

12ஆம் உலகத் தமிழ் இணைய மாநாடு

12th International Tamil Internet Conference 2013



மாநாட்டுக் கட்டுரைகள்
CONFERENCE PAPERS

உத்தமம்
INFITT

உலகத் தமிழ் தகவல் தொழில்நுட்ப மாநாடு
International Forum for Information Technology in Tamil



UNIVERSITI
MALAYA

The Leader in Research & Innovation

TI 2013 -12th International Tamil Internet Conference
Conference Proceedings

Computational Linguistics | கணினி மொழியியல்

1. A Knowledge-Engineering Methodology for Resolving Syntactic Level Ambiguities in Tamil Syntactic Parsing,
R.Padmamala
2. Computational approach to Word sense disambiguation in Tamil,
S.Rajendran, M. Anand Kumar and Soman K.P
3. A Rule Based Iterative Affix Stripping Stemming Algorithm For Tamil,
Damodharan Rajalingam
4. Developing online Sangam Corpus and Concordance,
A. Kamatchi
5. An Electronic Dictionary for Pathinenkilkanakku Literature,
K.Umaraj
6. தமிழ் தரவுத் தளங்கள்,
கோ.சக்கரபாணி
7. TagSets for Tamil,
V. Renuga Devi
8. Tamil Verb Conjugation Using Data Driven Approach,
V.Dhanalakshmi, M.Anand Kumar, K.P.Soman
9. Recognition of Kalippa Class of Tamil Poetry,
Rajeswari Sridhar, Rajkiran R, Narendra Kumar N J, Giridhar M
10. Programming Like Language In Tamil,
Yasasvi Sridharan

Natural Language Processing | இயற்கை மொழி ஆய்வு

11. தமிழை அறிய கணினிக்கு எத்தனை விதிகள் வேண்டும்?,
வாசு அரங்கநாதன்
12. Written to Spoken Tamil : Conversion Rules (A Study in Speech Synthesis),
K. Karunakaran and R. Krishnan
13. A Three-level Genre Classification for Tamil Lyrics,
Karthikeyan, Nandini Karky, Elanchezhiyan, Rajapandian & Madhan Karky
14. Lyric Objects & Spot Indices for Paadal, a Tamil Lyric Search Engine,
Karthikeyan, Elanchezhiyan, Rajapandian & Madhan Karky
15. An Avatar Rendering Hand Gesture For Tamil Words,
D.Narashiman, Bavatharani Suriyan, T.Mala
16. E-Content To Animation Conversion System (Etacs),
Chakkaravarthy.E, Narashiman.D
17. Lyric Visualization, Rajapandian,
Nandini Karky, Elanchezhiyan, Karthikeyan & Madhan Karky
18. Scoring Models for Tamil Lyrics, Elanchezhiyan,

- TamilSelvi E, Karthikeyan, Rajapandian, Madhan Karky
19. Modeling basic emotions for Tamil speech synthesis,
A G Ramakrishnan and Lakshmi Chithambaran
 20. Conceptual Based Search Engine (Cbse) System For Tamil And English,
Saviya shree K V, Umamaheswari E, Balaji J, T V Geetha, Ranjani Parthasarathi
 21. Morphological Analyzer for Classical Tamil Texts – A Rule based approach:
special Reference to Plural Markers,
R.Akilan, Prof E.R.Naganathan, G.Palanirajan
 22. Development of Hints in Tamil using the Techniques of Natural Language
Generation,
Rajeswari Sridhar, V. Janani, G. Rasiga, G. Monica
 23. Paeri : Evolving Tamil Name Generation Algorithm,
Rajapandian, Karthikeyan, Elanchezhiyan & Madhan Karky
 24. Semi-Automatic Query Testing Tool, Ganesh J,
UmaMaheswari E, T V Geetha, Ranjani Parthasarathi
 25. Patti Vaithiyam — An Information Extraction System for Traditional Tamil
Medicines,
J. Betina Antony, G.S.Mahalakshmi

Language Tools | மொழிக் கருவிகள்

26. தமிழ் மென்பொருள் உருவாக்கம்: தேவைகளும் திட்டங்களும்,
ப.அர.நக்கீரன்
27. Standardize fonts to catapult Tamil to the digital forefront,
A G Ramakrishnan
28. Notes on Tamil Orthography – puLLi, kAl, ai, ja, etc.,
Mani M. Manivannan
29. பல்வகை உணர்விகளில் மொழி மாதிரியங்களின் பயன்பாடு : ஒரு புதிய
கண்ணோட்டம்,
ஆ. க. ராமகிருஷ்ணன், ப. அருள்மொழி
30. A tool that converted 200 Tamil books for use by blind students,
Shiva Kumar H R, A G Ramakrishnan
31. A Multimodal Framework for the Recognition of Ancient Tamil Handwritten
Characters ...,
E.K.Vellingiriraj, P.Balasubramanie
32. Analysis of Statistical Feature Extraction Approaches Used in Tamil
Handwritten OCR,
Antony Robert Raj.M, Abirami.S, Murugappan.S
33. Offline Tamil Handwritten Character Recognition,
Jegadheeswaran J, Abirami S, Vijayalakshmi M,
34. Tamil Script Recognition using Smart Phone,
Rex Aantonny. P, Iniya Nehru .E
35. Use of Tamil Grammar Rules for Correcting Errors in Optical Character

Recognised Document,

Rajeswari Sridhar, Remitha Rathi L, Rithya P, Nivrutha P

36. Identification of Tamizh script on Tablet PC,

Bhargava Urala, A G Ramakrishnan

37. Recognition System for Tamil Sign Language Using Hand Gestures,

Vijay Jeyakumar, K.Bommanna Raja

Internet and software | இணையமும் மென்பொருள்களும்

38. Usage of Tamil in Internationalized Domain Names,
Balaji Vasudevan

39. வலைமச் சொற்கள்,

இலக்குவனார் திருவள்ளுவன்

40. திறவூற்றுத் தமிழ் மென்பொருள்கள்: மதிப்பீடும் எதிர்காலத் தேவையும்,

இரா. பன்னிருகைவடிவேலன்

41. தமிழ் இணையத்தில் புதிய அலை,

அண்ணாகண்ணன்

42. மின்தமிழ் வரலாற்றில் தமிழ் மரபு அறக்கட்டளையின் சுவடுகள்,

நா.கண்ணன், சுபாஷினி ட்ரம்மெல்

43. பார்வை மாற்றுத் திறனாளிகள் பயன்பாட்டில் தமிழ் கணினியும் செல்பேசியும்,

சே. பாண்டியராஜ், வா.மு.சே. முத்துராமலிங்க ஆண்டவர்

44. 'Raspberry PI' Future Outlook for Tamil,

Saravanan Mariyappan

45. முகநூலில் (facebook) தமிழர்கள் பயன்படுத்தும் மொழிவகைகள்,

மலர்விழி சின்னையா, பரமசிவம் முத்துசாமி, கிருஷ்ணன் இராமசாமி

46. கலாசாரப் பெயர்ப்பு,

Ilangkumaran Sivanadhan

47. E-Governance Initiatives in Tamil Nadu,

E.Iniya Nehru

48. மலேசியத் தமிழ் இளைஞர்களிடையே மொழித் தாவலும் புதிய

தொழில்நுட்பங்களின்வழி மின்னணுக் கலாசாரமும்,

செல்வஜோதி இராமலிங்கம்

49. தகவல் தொழில்நுட்பத்தில் தமிழ் – ஒரு பார்வை,

தங்கராஜா தவருபன்

50. திறந்தமூல தமிழ் மென்பொருட்கள் மற்றும் கணினிக்கு தமிழ் மொழியறிவு

ஊட்டல் பற்றிய ஆய்வுகள்,

ஜெயரஞ்ஜன் யோகராஜ்

Education | கல்வி

51. TEDU-Tamil Educational Cloud for anytime anywhere Tamil Learning,
G.Malini, T.Mala

52. Challenges in Online learning towards better learning environment ...,
Rathirane, Y

54. கற்றல் கற்பித்தலில் தகவல் தொழில் நுட்பத்தின் பயன்பாடு UK,
Siva Pillai
55. Effectiveness Of Computer Assisted Learning In Tamil Grammar,
G.Singaravelu
56. Using Augmented Reality as a Visual Aid to Teach Tamil Language in
Malaysian Tamil Primary Schools,
Kumaraguru
57. Bridging the Digital Divide for the Differently-Abled (Tamil) Students in Higher
Education,
J. Jerald Inico, T. Edwin Prabakaran
58. தமிழ்மொழி கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தின் பயன்பாடு -
மலேசியத் தமிழாசிரியர்களிடம் ஓர் ஆய்வு,
சாமிக்கண்ணு ஜெபமணி ஈசாக்கு சாமுவேல்
59. Computer-based and Paper-based Vocabulary Games for Tamil Learners,
Paramasivam Muthusamy, Kanthimathi Letchumanan, Tan Bee Hoon
60. மொழி விளையாட்டுகளும் சொற்கள் வளப்படுத்தலும்,
அ.பரிமளகாந்தம்
61. The Potential Role of YouTube in the process of 21st century approach to
Classical Tamil Literature,
M.Semmal
62. Role of Information Technology in Teaching and Learning of Classical
Literature,
Saravanan Raja
63. காணொளி பேசும் தமிழ்,
சங்கீதா இராமகிருஷ்ணன்
64. கிராமப்புற ஆசிரிய கல்விப் பட்ட மாணவர்களின் அறிவியல் கற்பித்தல் திறனை
மேல்புல கற்பித்தல் ...,
ச. அமுதா
65. பலகைக் கணினிகளின் வழி தமிழ் வாசிப்பதிலும் சமூகக் கட்டமைப்பு வலைகளில்
தமிழ் பதிப்பதிலும் ...,
இல.வாசுதேவன்
66. தமிழ்மொழி கற்றல் கற்பித்தலில் இணையத்தின் பயன்பாடும் பங்களிப்பும்,
சுப.நற்குணன்
67. கணினி சார்ந்த கற்றல்கற்பித்தலில் தமிழ்மொழிப் பயன்பாடும் நன்மைகளும்,
கல்பனா
68. PDF வடிவத்தில் பல்லாடகப் பனுவல் தயாரித்தல், இதன் மூலம் கற்றல்,
கற்பித்தல்.,
காசி வீராசாமி, சாத்தப்பன் மெய்யப்பன்
69. பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பப் பயன்பாட்டின் வாயிலாக
திருக்குறள் பயன்பாட்டினை மேம்படுத்துதல்.,
சிவகுமாரி
70. பன்முகத்தன்மை வாய்ந்த முகநூல், வாட்ஸ்ஆப் வழி ஒருங்கிணைந்த மொழித்
திறன்கள் கற்றல் கற்பித்தல்,
கோவிந்தசாமி சந்தானராஜ்

71. Let Us 'Flip' the Classroom!,
Seetha Lakshmi
72. முகநூல் வழி தமிழ் கற்றல் கற்பித்தல்,
கி.விக்கினேசு.
73. பள்ளிகளுக்கான இணையவழிக்கல்வி முறையும் தமிழ் அகராதி உருவாக்கமும்,
S.Maniam
74. Computer Assisted Tamil Teaching and Learning,
Louis Isack Kumar
75. மொழி கற்பித்தலுக்கான இணையத் தகவல் மூலங்களை மதிப்பிடுதலும்
தெரிவுசெய்தலும் -ஒரு கண்ணோட்டம்,
மணிமாறன் கோவிந்தசாமி
76. இணையம் வாயிலாக தமிழ்க்கல்வி கற்க பிற மொழி மாணவர்கள்
எதிர்கொள்ளும் சிக்கல்கள்,
கங்காதுரை கணேசன்
77. The Need for Integration Of ICT And Physical Infrastructure of School
Development for The Growth of The Tamil.,
Saravanan Mariappan
78. Sustainability of Teaching Information and Communication Technology (ICT)
Skills in Malaysia Tamil Schools ... ,
Gunasegaran Kandaswamy

Mobile and Handheld | செல்பேசிகளும் கைக்கருவிகளும்

79. Proactive Location Based Services In Tamil,
Karthikeyan K, Vinayak U S M P, Vijayalakshmi M
80. Social Interaction and Content Sharing in Mobile P2P Environment using
Tamil Language,
L.Libin Lougine, M.Vijayalakshmi
81. Android Based Mobile Application in Tamil for Danger Assistance,
Gowtham.N.R, Dhanraj.R, Lokesh Kumar.K, Tamizhselvi S P, Vijayalakshmi
M
82. Mobile Based Rural Road Monitoring System,
Jeyabalan.K, Iniya Nehru.E, Rex Aantonny. P
83. ஒரு முழுமையான தட்டல் மற்றும் கையெழுத்து உள்ளீட்டு மென்பொருளை
நோக்கி,
வெ. கிருஷ்ணமூர்த்தி
84. Translating Tamil Speech (SL) as English Text Message (TL) in Android
Mobile Phones,
Vikram Srinivasan
85. தன்னார்வ முனைந்து கற்றலுக்கான தமிழக ஆறாம் வகுப்பு தமிழ் பாடக்
குறுஞ்செயலி,
Abraham Meshach Ponraj, T.Priyanka, R.R.Jeyshree & B.Padma
86. செல்பேசியில் தமிழ்க் கலைச்சொல்லாக்கம், தரப்படுத்தம், செந்தரப்படுத்தம்,
நா.ஜானகிராமன்

87. அலைபேசிகளிலும் கைக்கணினிகளிலும் குறுஞ்செயலிகளுக்கான இணையத் தொழில்நுட்ப நியமனங்கள்,
சுகந்தி வெங்கடேஷ்
88. கைக்கணினியும் தமிழும்,
வெ.இராமன்
89. தமிழ்க் குறுஞ்செயலிகள் உருவாக்கம், உள்ளடக்கம், பயன்பாடு – ஒரு மதிப்பீடு,
சி. சிதம்பரம்
90. மென்பொருள்கள், தளங்கள் மற்றும் தொழில்நுட்பங்கள் வழி கையடக்க கணினி சாதனங்களில் தமிழ் மின் புத்தகங்கள்
சி.குணசேகரன்
91. பயன்பாட்டு நோக்கில் தமிழ்க் குறுஞ்செயலிகள்,
துரை. மணிகண்டன்

A Knowledge-Engineering Methodology for Resolving Syntactic Level Ambiguities in Tamil Syntactic Parsing

R.Padmamala M.C.A. Ph.D, Research Scholar, University of Madras, Chennai, India

Abstract

This paper discusses the ambiguities that arise in developing a Syntactic Parser for Tamil and proposes a knowledge-engineering methodology to resolve those ambiguities. Syntax is the scientific study of sentence structure. The sentences are hierarchically structured. The sentence is not just a linear string of symbols. The position of the words relative to one another, makes a difference in the meaning of the sentence. Such a feature gives rise to certain ambiguities. This paper gives a guideline to develop a word net for Tamil which will assist in word sense disambiguation.

Keywords: Syntax, ambiguities, knowledge – engineering, word net.

1. Introduction

After morphological parsing, morpheme-labelling and word-class tagging are completed, syntactic parsing is to be done. Various ambiguities arise at this level too. Let us discuss the ambiguities that arise at this level.

The ambiguities that arise at syntactic level are:

1. Lexical ambiguity
 - Category ambiguity
 - Homograph and polysemy
2. Structural Ambiguity
3. Anaphora resolution

This paper suggests a method involving synsets to resolve category and structural ambiguities.

2. A Knowledge-Engineering Methodology

For resolving the above said ambiguities, certain contextual information is also needed. Also semantic information is to be provided. These details can be provided using a WordNet, a lexical database, specially developed for Tamil.

WordNet is a lexical database for the English language. It groups English words into sets of synonyms called *synsets*, provides short, general definitions, and records the various semantic relations between these synonym sets.

3. WordNet for Tamil

WordNet differentiates words into nouns, verbs, adjectives and adverbs as each of the categories follow different grammatical rules. Usually prepositions, determiners etc. are not included in a synset. Each synset has a group of synonymous words or collocations. A collocation is a sequence of words that join together to form a specific meaning. E.g. *pachai kuzanthai* ‘infant baby’. Different senses of a word are in different synsets. The meaning of the synset is also attached to a synset with short definition and / or example sentences called gloss.

A typical example synset with gloss is:

cariyaaNa, nalla, kaNintha, thoothaaNa, - (suitable for a particular purpose; “thankam vaanka ithu cariyaaNa neram” ; “ithu oru nalla mudivu” ; “thirumaNnam ceiyya kalam kaNinthathu” ; “payaNnam ceiyya vimaaNam thoothaaNathu”).

The synonym sets are connected to other synsets through a number of semantic relations. The following are some of the relations for each category of words:

Nouns

- **Hypernyms:** Y is a hypernym of X if every X is a type of Y (*maram* is a hypernym of *maamaram*)
- **Hyponyms:** Y is a hyponym of X if every Y is a type of X (*maram* is a hyponym of *maamaram*)

- **Coordinate terms:** Y is a coordinate term of X if X and Y share a hypernym ()
- **Holonym:** Y is a holonym of X if X is a part of Y (*utal* is a holonym of *kai*)
- **Meronym:** Y is a meronym of X if Y is a part of X (*kai* is a meronym of *utal*)

Verbs

- **Hypernym:** The verb Y is a hypernym of the verb X if the activity X is a kind of Y (*aRithal* is an hypernym of *patiththal*)
- **Troponym:** The verb Y is a troponym of the verb X, if the activity Y is doing X in some manner (*vizunku* is a troponym of *caapitu*)
- **Entailment:** The verb Y is entailed by X, if by doing X you must be doing Y (*thuungu* is entailed by *kurattai vitu*)
- **Coordinate terms:** The verbs that share a common hypernym are termed as coordinate terms. (*noNntu* and *thaavu* share the hypernym *nata*)

Adjectives:

- Related nouns
- Similar to
- Participle of verb

Adverbs

- Root adjectives

3.1. A model synset for Tamil that assists in eliminating ambiguities

3.2. Syntax level ambiguities

a. Lexical ambiguity

The morphological problems are concerned with the ambiguities involved in analysing the words in more than one way. Lexical ambiguity involves interpreting words in more than one way. Lexical ambiguities are of three types: category ambiguities, homographs and polysemes, and transfer (or translational) ambiguities. In this paper, only category ambiguities and homographs and polysemes are considered.

- **Category ambiguity**

The most straight forward type of lexical ambiguity is that of category ambiguity: a given word may be assigned to more than one grammatical or syntactic category (e.g. noun, verb or adjective) according to the context. There are numerous examples for this in Tamil:

1. *ootu* – Roof tiles (Noun), Run (Verb)
2. *utai* – Dress (Noun), Break (Verb)
3. *pai* – Step (Noun), Read (Verb)
4. *aatu* – Goat (Noun), Dance (Verb)

Category ambiguity can be resolved by considering the inflectional morphology. For example, verb will take tense. From that words can be categorised. However, the problems increase when several categorically ambiguous words occur in the same sentence. A good example is the following sentence.

I. *Aattai kathial¹ vetta kathi² iranthadhu.* ‘as the goat was cut with a knife, it screamed and died’

Here *kathi¹* means knife (Noun) and *kathi²* is the Verbal participle form of the verb *kathu* which means Shout. Resolving such an ambiguity is quite a problem.

Consider a similar sentence.

II. *Velai¹ vaNnankuvathu yen velai².* ‘Worshipping the spear is my duty’

In the above seen II, *velai¹* means the spear (Noun+Associative case) and *velai²* means job (Noun). Here arises confusion what meaning to assign in each position. The meaning can be finalised based on the context in which it is occurring or on the basis of the properties of the verb that is occurring in the sentence. To give the contextual information, synsets can be used effectively. The following synsets can be considered:

Verb:

S1 - *vaNnanku, kumpitu, thožu, thuthi* – different forms of worship (*iRaivaNai vaNnanku, vizunthu kumpitu, thiruvatikaLaith thožu, iRai naamaththaith thuthi*)

Noun:

S2 – *veel, vil, thiruneeRu, kunkumam, thirisuulam, canku, cakkaram, karutaN, aathi seeshaN, cakrayutham* – things associated with Gods.

S3 – *yeesu, murukaN, civaN, vishNu, thirumaal, perumaal, allaa* – Names of Gods of various religions.

Here a relation within noun synsets itself will not work. Hence a cross relationship between noun synsets and verb synset is to be defined. That is, a relationship between the synsets S1, S2 and S3 is to be defined. It can be done as follows:

S1 < **Done On** > S2, S3

This relation will clearly finalise the meaning of all the constituent words in the sentence II. As the verb *vaNnanku* is in S1 and it is related to S2, where the noun *veel* is present, it can be ascertained the meaning ‘the spear’.

b. Structural Ambiguity

Structural ambiguity involves problems with the syntactic structures and representations of sentences. Ambiguity arises when there is more than one way of analyzing the underlying structure of a sentence according to the grammar used in the system. There are two types of structural ambiguities. They are

1. Real structural ambiguity
2. Accidental structural ambiguity

Synsets can be used to resolve Accidental structural ambiguity as follows.

Accidental structural ambiguity

The fact that a single word may serve in a different function as noun, verb etc. within the same syntactic context will lead to system ambiguity but will be unambiguous to the human reader. For example, *avaN pati thaaNntiNaaN* – ‘He crossed the steps’. Here *pati* is unambiguously ‘Steps’ for the human reader. But for the system it is ambiguous which category to assign to it, whether verb or noun, where the verbal meaning is ‘Study’. Such ambiguities which are not actually ambiguous for the human readers but ambiguous to the system alone are termed as Accidental structural ambiguity.

Resolution of structural ambiguity using synsets

When syntactic analysis produces more than one interpretation of a given sentence, a method is needed to choose the correct one. The following synsets can be considered:

Verb

S3 – *pati, ezuthu, vaaci, varai* – academic works

S4 – *thaaNntu, kuthi, nuzai, puku, nata* – various types of movements

Noun

S5 – *peeNaa, thaaL, puththakam, vaNnNnappencil* – stationery items

S6 – *pati, vaacal, thiNnNnai, viitu, aRai* – part of a house

Relationships

S5 <Needed for> S3

S4 <Can be done on> S6

Thus in the example sentence, *avaN pati thaaNntiNaaN*, the verb *thaaNntu* is in the synset S4 and *pati* is both in S4 and S6. A relationship exists between only S4 and S6. Therefore *pati* comes under part of a house. Hence it can be assigned the meaning ‘Steps’.

4. Conclusion

A full-fledged knowledge base can be developed involving a larger number of synsets. With minute categorization, it may be possible to resolve all the syntactic level ambiguities.

5. Bibliography

- Andrew Carnie. Modern Syntax A Course book, 2011. Cambridge University Press.
- Andrew Radford, English Syntax – An Introduction, 2004, Cambridge University Press, Cambridge
- Anne Abeillé (ed), Treebanks – Building and using parsed corpora, Kluwer Academic Publishers
- Bonnie Jean Dorr, Machine Translation : A view from the Lexicon, 1993, Publisher: The MIT Press, Cambridge, Massachusetts, London
- Brian Roark and Richard Sproat. Computational Approaches to Morphology and Syntax, 2007. Oxford University Press.
- Christopher D.Manning, Hinrich Schutze, Foundations of Statistical Natural Language Processing, MIT Press, England
- Daniel Jurafsky & James H.Martin, Speech and Language Processing, 2003, Pearson Education Inc.
- Ontology Development 101: A Guide to Creating Your First Ontology". Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.
- Pon. Kothandaraman. A grammar of contemporary literary Tamil, 1997. International Institute of Tamil Studies.

Computational approach to Word sense disambiguation in Tamil

Dr. S.Rajendran, Dr. M. Anand Kumar and Dr. Soman K.P

Amrita Vishwa Vidyapeetham, Coimbatore, India.

rajushush@gmail.com, m_anandkumar@cb.amrita.edu, kp_soman@amrita.edu

Native speakers of any natural language have an implicit knowledge of their language usually referred to as competence. Part of this tacit knowledge is concerned with interpretation, i.e. native speakers are capable of assigning an interpretation to any structure they generate. Native speakers not only assign an interpretation to every structure in their language, but also know that there are structures that may have more than one semantic interpretation. These structures are usually referred to as ambiguous structures. When an utterance has more than one interpretation, it is usually referred to as ambiguous. Ambiguity means that utterances may differ semantically but not phonetically, i.e. they differ in their interpretation but not in their form. Ambiguity may result from two homonyms occurring in the same structural position, as in the following example.

அவன் கால் பகுதியைச் சாப்பிட்டான் .

avan kaal pakutiyaic caappiTaan

‘He ate quarter of something’/‘He ate the leg part of something’

It may also occur when constituents in larger structures have more than one interpretation according to their internal structure and syntactic position.

வெள்ளை மருந்து குப்பி .

veLLai marundtu kuppi

‘medicine bottle which is white in colour/a bottle with white medicine’

The first one is called lexical ambiguity and the second structural ambiguity. Lexical ambiguity refers to the type of ambiguity those results from the occurrence of homonyms.

Here in this paper we are concerned with the lexical ambiguity only.

1 Lexical ambiguity

The lexical ambiguity is very common. It includes, for example, the nouns such as படி/*paTi*, குடி/*kuTi*, மடம்/*maTam*, etc, verbs such as பிடி/*piTi*, அறை/*aRai*, முடி/*muTi*, சுடு/*cuTu*, etc and the adjectives such as விரிந்த/*virinta*, குறைந்த/*kuRainta*, வெளுத்த/*veLutta*, கறுத்த/*kaRutta*, etc. There are tests for establishing lexical ambiguity. One test is that for the word கடினமான/*kaTinamaana* there are two opposite words, மெதுவான/*metuvaana* and எளிதான/*eLitaana*. Consider the following example,

கடினமான மிட்டாயை கடிக்க முடியாது .

kaTinamaana miTTaayaik kaTikka mutiyaatu

‘You cannot bite hard seet’

கடினமான சொல்லுக்கு பொருள் கூற இயலாது .

kaTinamana collukkup poruL kuuRa iyalaatu

‘you cannot give meaning to hard word’

The reason for this ambiguity is that the word has more than one meaning. But it is not clear when there is only one word involved in ambiguity.

Though the noun *படி/paTi* and the verb *படி/paTi* have same spelling/pronunciation they are two different words. They are examples of homophones. One may wonder whether the noun *கடி/kaTi* and the verb *கடி/kaTi* are examples for homonyms or not. Doubt may arise whether the word *mutal* in *முதல் மாணவன்/mutal maaNavan* and *ஐந்து முதல்/aindu mutal* are one and the same or not. To tell that one shows lexical ambiguity and the other is homonymous is not correct for all. This may be accidental.

There are three basic types in lexical ambiguity: category ambiguity, ambiguity due to homography and ambiguity due to ploysemy.

1.1 Category ambiguity

Category ambiguity is the most straight forward type of lexical ambiguity. This happens when a given word may be assigned to more than one grammatical or syntactic category as per context. One can find a number of such examples in Tamil. For example the word *paccai* ‘green’ can be both noun as well as adjective. Similarly the word *cuTu* can be both verb as well as an adverb. *kaTi* could be both verb as well as noun. As shown in the following example *paTi* can be noun as well as a verb denoting two different meanings and thus shows ambiguity.

அவன் தந்தை அவனிடம் நீ நன்றாக படி என்று கூறினார் .

avan tantai avaniTam nii nanRaakap paTi eRu kuuRinaar

‘His father asked to study well’

அவன் மாடி படி வழியாக மேலே ஏறினான் .

Avan maaTip paTi vaziyaaka meele eeRinaan.

‘He climbed up through stair case’

The words like *melee* and *kiizee* could be adverbs and postpositions.

அவன் மேலே இருக்கிறான் .

avan meele irukkiRaan. (noun)

He is at the top

அவன் மேலே சென்றான் .

avan meele cenRaan (adverb)

He went up

அவன் மேலே நிிற்கின்றான் .

Avan meecai meele niRkiRaan. (postposition)

Category ambiguities can be often be resolved by morphological inflection. For example *aTi* in *avan aTikkiRaan* ‘he is beating’ is a verb and *அடி/aTi* in *அவனால் அந்த அடியைத் தாங்க முடியவில்லை/avanaal andta aTiyait taangka muTiyavillai* ‘He could not bear that beating’ is noun. Frequently ambiguity can be resolved by syntactic parsing. However, the problem increases when several categorical ambiguous words occur in the same sentence, each requiring being resolved syntactically.

1.1.1 Categorical ambiguity due to historical functional reorganization

The inflected forms of nouns or verbs will denote different word category or functional category due to historical meaning change. For example many of the postpositions in Tamil are historically the inflected forms of verbs. The inflected forms *இருந்து/irundu* ‘from’, *பற்றி/paRRi* ‘about’, *குறித்து/kuRittu* ‘about’, *ஒட்டி/oTTi* ‘about’, *கொண்டு/koNTu* ‘by (means of)’, *வைத்து/vaittu* ‘by (means of)’, *சுற்றி/cuRRi* ‘around’, *நோக்கி/ndookki* ‘towards’, *முந்தி/mundti* ‘before’, *விட/viTa* ‘than’, and *கூட/kuuTa* ‘along with’ are the inflected forms of the verb *iru* ‘be’, *பற்று/paRRu* ‘catch’, *குறி/kuRi* ‘aim’, *ஒட்டு/oTTu* ‘stick’, *கொள்/koL* ‘have’, *வை/vai* ‘keep’, *சுற்று/cuRRu* ‘go around’, *நோக்கு/ndookku* ‘look at’, *mundtu* ‘over take’, *viTu* ‘leave’, and *kuuTu* ‘assemble’ respectively.

1a. அவன் வீட்டில் இருந்து வெளியேறினான். (முன்னுருபு)

avan viiTTilirundtu veLiyeerinaan

'He went out from the house'

1b. அவன் வீட்டில் இருந்து வந்தான் (வினைத்திரிபு வடிவம்)

avan viiTTil irundtu vandtaan

'He was in the house (habitually/continuously)'

2a. அவன் அவளைப் பற்றி பேசினான். (முன்னுருபு)

avan avaLaip paRRi peecinaan

'He talked about her'

2b. அவன் அவள் கையைப் பற்றி முத்தமிட்டான். (வினைத்திரிபு வடிவம்)

avan avaL kaiayip paRRi muttamiTTaan

'He catch hold of her hand and kissed it'

3a. அவன் அவளைக் குறித்து பேசினான். (முன்னுருபு)

avan avaLaik kuRittu peecinaan

'He talked about her'

3b. அவன் அவள் சொல்வதைக் குறித்து வந்தான். (வினைத்திரிபு வடிவம்)

avan avaL colvataik kuRittu vandtaan

'He was noting down what she was telling'

4a. அவன் அந்த தலைப்பை ஒட்டி பேசினான். (முன்னுருபு)

avan andta taippai oTTi peecinaan

'He talked about that title'

4b. அவன் போஸ்டர் ஒட்டி பிழைக்கின்றான். (வினைத்திரிபு வடிவம்)

avan poosTar oTTi pizaikkinRaan

'He eke his livelihood by pasting posters'

5a. அவன் கத்தி கொண்டு அதை வெட்டினான். (முன்னுருபு)

avan katti koNTu atai veTTinaan

'He cut it with a knife'

5b. அவன் பென்சிலைச் சீவிக் கொண்டு பேசினான். (வினைத்திரிபு வடிவம்)

avan pencilsaic ciivik koNTu peecinaan

'He was speaking while sharpening the pencil'

6a. அவன் கத்தி வைத்துப் பழம் வெட்டினான். (முன்னுருபு)

avan katti vaittup pazam veTTinaan

'He cut the fruit with the knife'

6b. அவன் பணம் வைத்துக் கொண்டு சூதாடினான். (வினைத்திரிபு வடிவம்)

avan paNam vaittuk koNTu cuutaaTinaan

'He gambled by keeping the money at hand'

7a. அவன் வீட்டைச் சுற்றி மரங்கள் நிற்கின்றன. (முன்னுருபு)

avan viiTTaic cuRRi marangkaL ndiRkinRana

'The house is surrounded by the tree'

7b. அவன் அவளையே சுற்றி வருகின்றான். (வினைத்திரிபு வடிவம்)

avan avaLaiyee cuRRi varukinRaam

'He is going after her'

8a. அவன் அவளை நோக்கி நடந்தான். (முன்னுருபு)

avan avaLai ndookki ndaTandtaan

He went towards her'

8b. அவன் அவள் முகத்தை நோக்கிச் சிரித்தான். (வினைத்திரிபு வடிவம்)

avaL mukattai ndookkic cirittaan

'He smiled looking after her face'

9a. அவன் அவளுக்கு முந்தி அங்கு வந்தான். (முன்னுருபு)

avan avaLukku mundti angku vandtaan

'He came there before her'

9b. அவன் அவளை முந்தி நடந்துகொண்டிருந்தான். (வினைத்திரிபு வடிவம்)

avan avaLai mundti ndaTandukoNTirundtaan

'He is walking overtaking her'

10a. அவன் அவளை விட நல்லவன். (முன்னுருபு)

avan avaLai viTa ndallavan

'He is better than her'

10b. அவன் அவளை விட விருப்பவில்லை. (வினைத்திரிபு வடிவம்)

avan avaLai viTa virumpavillai

'He does want to leave her'

11a. அவன் அவள் கூட வந்தான். (முன்னுருபு)

avan avaL kuuTa vandtaan

'He came with her'

11b. அவன் அவர்களுடன் கூட விரும்பினான்

avan avarkaLuTan kuuTa virumpinaan

'He wanted to gather together with them'

The word என்று/enRu which is the inflected from the verb en 'say', show two different grammatical functions thus showing ambiguity.

a. அவன் நல்லவன் என்று நினைத்தேன் (என்று/enRu functions as complementizer)

avan ndallavan enRu ndinaitteen

b. அவன் திடீர் என்று வந்தான் (என்று/enRu functions as adverbializer)

avan tiTiir enRu vantaan

The word என்று/enRu is having homographic relation with the inflected verbal form என்று/enRu.

avan enRu varukiRaam

அவன் என்று வருகிறான்

'When does he come?'

The inflected verbal forms which can be analyzed verb root+um (future suffix) can be interpreted at least in two ways.

Atu naaLai varum
'It will come tomorrow'
Atu varum naaL enakkut teriyaatu
'I don't know the date of its coming'

The inflected verbal form which can be analyzed as verb + tense + அது/atu can be interpreted in three ways.

atu vandtatu 'it came' (*vandtatu* is the finite verbal form)
atu vandtatu enakkut teriyaatu (*vandtatu* is the gerundival form)
"I did know that it had come"
andta ceytittaal neeRRu vandtatu (*vandtatu* is the participial noun)
'That newspaper is yesterday's one'

The ambiguity can be resolved by selection restriction, context, collocation, co-occurrence, etc.

1.2 Ambiguity due to Homography

If two entirely different words having same form have different meanings, the ambiguity arises due to homography. In the following examples the word அட்டை/aTTai shows homography; அட்டை/aTTai can denote 'leach' as well as 'binding'.

அவன் புத்தகத்தின் அட்டையைக் கிழித்து எறிந்தான்.
Avan puttakattin aTTaiyaik kizittu eRintaan
'He tore away the binding'
அவன் அட்டையைக் கொன்றான்.
Avan aTTaiyaik konRaana
'He killed the leach'

Sometimes among the homographs, the use of one may be greater than the other. In that case the ambiguity can be resolved on the basis of text. This is done by setting aside the unusual meaning from the dictionary unless it is required for translation.

1.2.1 Homography in inflected words

Homography may arise when two different words get inflected differently resulting in homographic forms. The homography can be resolved by different morphological analysis. The following examples will reveal this.

அவன் கடலை தின்று மகிழ்ந்தான்.
avan kaTalai tinRu makizntaan
'He enjoyed eating pea nut'
அவன் கடலை கண்டு மகிழ்ந்தான்.
avan kaTalai kaNTu makizntaan
'He enjoyed seeing the sea'

In the first sentence the noun கடலை/kaTalai denotes 'pea nut' and in the second case கடலை/kaTalai has to be analysed as கடல்/kaTal + ஐ/ai (accusative case marker) and interpreted as kaTal 'sea'.

As in the case of the following example, the inflected word of one type of morphological analysis resembles an inflected word form of another morphological analysis, thereby showing homography.

அவன் (துணி) நெய்தான்.
avan (tuNi) ndeytaan.
'He weaved (cloth)'

அவன் நெய்தான் விரும்புகிறான் .

avan ndeytaan virumpukiRaan

'He likes gee only'

In the first sentence the word நெய்தான்/*ndeytaan* has to be interpreted after analyzing it into நெய் + த்/*ney+t* (past tense)+ஆன்/*aan* (third person masculine singular) and in the second sentence *neytaan* has to be interpreted as நெய்/*ndey* 'ghee' + தான்/*taan* 'only'. Even the two root words *ney* 'weave' and *ney* 'ghee' are homographs showing categorical lexical ambiguity.

1.3 Ambiguity due to ploysemy

If a word has two or more meanings it can be said that the ambiguity is due to polysemy.

Polysemy expresses extension of meaning. The polysemous words may express new meaning by metaphoric and metonymic extensions. For example the word கிளை/*kiLai* 'branch' may denote branch of a tree as well as a branch of a bank. நட/*ndaTa* can denote the action of walking as well as happening or functioning of something.

அவன் தினமும் காலையில் பள்ளிக்கு நடந்து செல்கிறான் .

avan tinamum kaalaiyil paLLikku ndaTantu celkinRaan

'He goes to school daily by walking'

அந்த நிறுவனம் நன்றாக நடந்துக்கொண்டிருக்கின்றது

Anta niRuvanam nanRaaka ndaTandukoNTirukkinRatu

'That organization is functioning well'

அந்த தியேட்டரில் சினிமா நடந்துக்கொண்டிருக்கின்றது .

Anta tiyeeTTaril cinimaa ndaTandukoNTirukkinRatu

'A cinema is running in the theatre'

ஓடு/*ooTu* can denote human action of running as well flow of a river.

அவன் விரைவாக ஓடுகிறான் .

Avan viraivaaka ooTukiRaan

'He is running fast'

தஞ்சாவூர் வழியாகக் காவேரி ஓடுகிறது .

Tanjaavuur vaziyaakak kaaviriyaaRu ooTukiRatu

'The river Kaviri runs through Thanjavur'

கண்/*kaN* may denote the eye of animate beings as well as the eye like spot in the coconut.

அவன் தன் கண்களை மூடினான் .

avan tan kaNkaLai muuTinaan

'He closed his eyes'

தேங்காய்க்கு மூன்று கண்கள் உண்டு .

Teengkaaykku muunRu kaNkaL uNTu

'there are three eye like spots in the coconut'

Similarly many words denoting body parts show polysemous extension thereby denoting ambiguity. For example வாய்/*vaay* can be a human mouth as well as the mouth of a bottle; கால்/*kaal* can be a human leg or leg of furniture. In the following example the ambiguity is due to metonymic extension.

ஊர் சிரித்தது

uur cirittatu

'The people (of the village) laughed'

Here in this sentence ஊர் /uur 'people' is used as a metonymic extension of ஊர் /uur 'village'.

In the following sentence the word கேள்/keeL denotes both the perception through ears as well as 'asking'.

ராதை ராஜா கேட்டதை அவனிடம் கூறினாள் .

raatai raajaa keeTTatai avaniTam kuuRinaaL

'Radha told him what Raja has asked her'

'Radha told him that Raja has heard that'

This sentence is ambiguous giving at least two interpretations.

ராதை ராஜா தான் காதால் கேட்டதை அவனிடம் கூறினாள் .

Raatai raajaa tan kaataal keeTTatai avaniTam kuuRinaaL

'Radha told him what Raja has heard'

ராதை ராஜா வினவியதை அவனிடம் கூறினாள் .

Raatai raajaa vinaviyatai avaniTam kuuRinaaL

'Radha told him that Raja had asked her'

As for machine translation is concerned both the homography and ploysemy will be treated alike as the aim is to finding out the meaning by context. The homographs belonging to different grammatical categories can be resolved as explained before. But they belong to same grammatical categories syntactic parsing may not be enough. One common approach is to assign semantic features such as 'human', 'female', 'liquid' etc and to specify which features are compatible in the given syntactic constructions, via selection restrictions. For example it might be specified that the verb குடி/kuTi 'drink' have an 'animate' subject.

There are difficulties in finding semantic features that can be used consistently and specifying the selection restriction for nouns and verbs based on these features. Even then these are widely used in Machine translation system often in combination with case roles. But the semantic features cannot solve all the problems, even in situations for which they have been devised. For example let us take the word அட்டை/aTTai. As we have found out it is used in the senses of 'binding' and 'leach'. These two meanings can be differentiated explaining the relevant co-occurrence restrictions we find out in the following sentences in which அட்டை/aTTai is used.

புத்தகத்தின் அட்டை கிழிந்துவிட்டது .

Puttakattin aTTai kizintuviTTatu

'The binding of the book is torn'

அட்டை ஊர்ந்து செல்கிறது .

aTTai uurndtu celkinRatu

'The leach is crawling'

The verbs like கிழி/kizi will take the objects like அட்டை/aTTai which can be torn as their subjects and the verbs like ஊர்ந்துசெல்/uurndtucel takes objects like அட்டை/aTTai which can crawl.

2. Word Sense Disambiguation

Word sense disambiguation is finding out the correct senses of a word in its context though computational methodology. Word sense disambiguation (WSD) is considered as AI-complete problem. This is equivalent to very hard problem in AI.

2.1 External knowledge resources.

The important component of word sense disambiguation is knowledge. Knowledge resources gives data which is needed to relate words with meanings. They include unlabeled or annotated corpora of texts, machine readable dictionaries (MRDs), thesauri, glossaries, ontologies and others. Structured resources include thesaurus, Machine readable Dictionaries (MRDs), and ontologies. Unstructured resources include corpora, raw corpora, sense-annotated corpus, collocational resources and other resources.

2.2 Representation of context

As text is an unstructured source of information, to make it a suitable input to an automatic method it is usually transformed into a structured format. To this end, a preprocessing of the input text is usually performed, which typically (but not necessarily) includes the following steps:

tokenization, a normalization step, which splits up the text into a set of tokens (usually words);

part-of-speech tagging, consisting in the assignment of a grammatical category to each word (e.g. “அந்த/anta_DT viiTu/வீடு_NN azakaaka/அழகாக_ADV irukkiRatu/இருக்கிறது_VBD, ” where DT, NN, ADV, and VBD are tags for determiners, nouns, adverbs and verbs respectively);

lemmatization, that is, the reduction of morphological variants to their base form

(மரங்கள்/marangkaL > மரம்/maram, வந்தான்/vandtaan > வா/vaa);

chunking, which consists of dividing a text in syntactically correlated parts (e.g., [அந்த வீடு/anta viiTu]NP [அழகாக இருக்கிறது/azakaaka irukkiRatu]VP, respectively the noun phrase and the verb phrase of the example); and

parsing, whose aim is to identify the syntactic structure of a sentence (usually involving the generation of a parse tree of the sentence structure).

2.3 Choice of a Classification Method

The final step is the choice of a classification method. Most of the approaches to the resolution of word ambiguity stem from the field of machine learning, ranging from methods with strong supervision, to syntactic and structural pattern recognition approaches. We can broadly distinguish two main approaches to WSD:

supervised WSD: these approaches use machine-learning techniques to learn a classifier from labeled training sets, that is, sets of examples encoded in terms of a number of features together with their appropriate sense label (or class);

unsupervised WSD: these methods are based on unlabeled corpora, and do not exploit any manually sense-tagged corpus to provide a sense choice for a word in context.

We further distinguish between knowledge-based (or knowledge-rich, or dictionary based) and corpus-based (or knowledge-poor) approaches. The former rely on the use of external lexical resources, such as machine-readable dictionaries, thesauri, ontologies, etc., whereas the latter do not make use of any of these resources for disambiguation.

Finally, we can categorize WSD approaches as token-based and type-based. Token based approaches associate a specific meaning with each occurrence of a word depending on the context in which it appears. In contrast, type-based disambiguation is based on the assumption that a word is consensually referred with the same sense within a single text. Consequently, these methods tend to infer a sense (called the predominant sense) for a word from the analysis of the entire text and possibly assign it to each occurrence within the text. Notice that token-based approaches can always be adapted to perform in a type-based fashion.

The methodology we have chosen works in two phases. At the beginning of the training phase the system identifies the ambiguous words in the text using the ambiguous words list. For each occurrence of an ambiguous word a context is identified from the text. Different researchers have used different sizes of contexts, i.e. different window sizes. The window size influences the performance of the system considerably. A larger window size brings in more unrelated words to the context, while a smaller window size misses some important collocations. The frequent words also called *stop-words* are removed from the window. The context is then morphologically analysed to get the root form of the contextual words, which are used, in clustering. Each context is represented as a context or occurrence vector. Using the K-means approach these vectors are collected into different clusters. Collocations are then collected from these clusters automatically and senses are assigned to these collocations by human-annotators thereby developing a sense-collocation dictionary.

2.4 Significance of Collocations

Collocations are nearby words that strongly suggest the sense of the ambiguous word in a given sentence. In general, the term collocation refers to a quantifiable position specific relationship between two lexical items. Collocations encode information about words that are semantically *very* close to the ambiguous word. For example consider the word ‘plant’ having the two following senses.

plant1 – a living organism

plant2 – a manufacturing place

Collocations for the word ‘plant’ will include the words that frequently occur with ‘plant’ in texts. Some typical collocations for the above two senses of ‘plant’ are listed below.

plant1 – growth, height, flower, fruit, species, leaves 2.3

plant2 – car, union, equipment, assembly, nuclear, job, worker 2.4

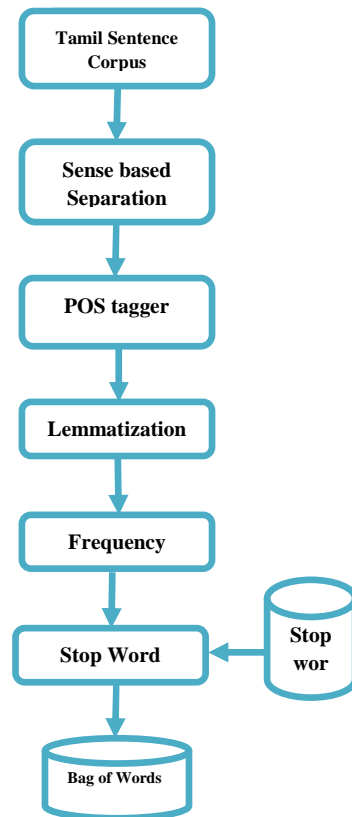
Generally collocations are sensitive to the distance from the ambiguous word. Words that are at a short distance from the ambiguous word strongly indicate its sense; as the distance increases the relevance of the word to that of the ambiguous word decreases. The distance is commonly referred to as window size. However, there is no fixed distance from the ambiguous word within which a word can be considered as a collocation. Sometimes, collocations may even occur far away from the ambiguous word.

2.5. Building of the present system.

The system has two phases: the training phase and the testing phase. In the training phase the bags of words or cluster of words occurring along with the ambiguous word will be obtained. The number of bags or cluster depends on the number of ways a particular word is ambiguous. For example the ambiguous word *aTTai* has two senses, and then we have to collect two bags of words or cluster of words.

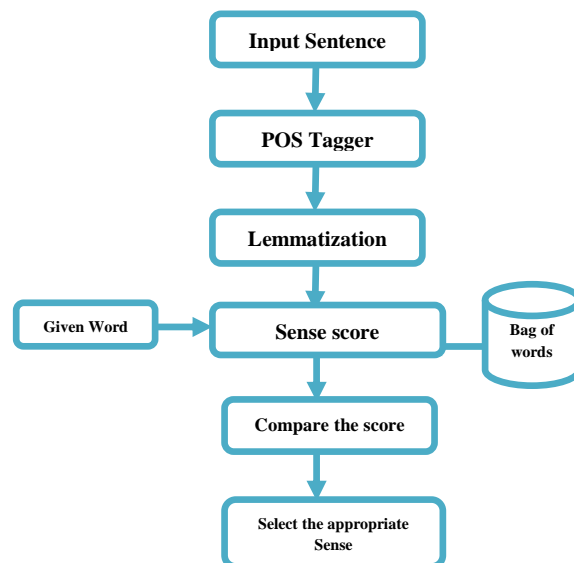
The ambiguous words will be identified and the sentences in which they occur will be extracted from the corpus. For example, the Tamil word அட்டை/*aTTai* at least has two senses: அட்டை/*aTTai* ‘leach’ and அட்டை/*aTTai* ‘binding of a book’. From the corpus the sentences with the target ambiguous word (அட்டை/*aTTai*) will be extracted and manually separated into two sets. Each set contains twenty sentences each. The sentences will be post-tagged and lemmatized. The occurrence of words along with the ambiguous word will be counted for each set of sentences. The functional words or words which do not contribute to the meaning of the ambiguous word are known as stop words. Such words need to be removed from sets of contextual words. After the removal of stop-words we will get a finite number of bags or cluster of words each belonging to the different senses of the ambiguous word. Now the system is ready for testing phase

Diagram representing the training phase



In the testing phase, the new sentence with the ambiguous word will be sense marked using the learning phase, i.e. using the bag of words. The input sentences will be post-tagged and lemmatized. The ambiguous word will be calculated sense score using the bag of words identified for the different senses. Using the sense score calculation, the appropriate sense of the ambiguous word will be selected.

Diagram representing the testing phase

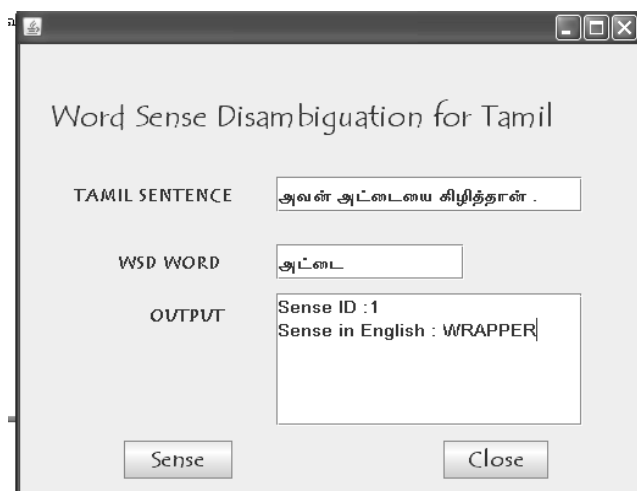


Using this method the ambiguity of the following ambiguous words such as மாலை *maalai*, நூல்/*nduul*, விலங்கு/*vilangku* can be disambiguated. The following table shows the bag of words associated with each senses of the ambiguous words.

Word	Meaning	Word occurring together
மாலை1	evening	காலை, பகல், சூரியன், மறை, நேரம், நாள், மாதம், மதியம், இரவு
மாலை2	garland	பூ, மணம், ரோஜா, மல்லிகை, மலர், பூஜை, மகளிர், கொன்றை
நூல்1	thread	இழை, தறி, சாயம், நெய், கதர், நெசவு, ஊடு, பாவு, துணி, ஊசி
நூல்2	book	ஆசிரியர், பதிப்பகம், அச்சு, பக்கம், தாள்
விலங்கு1	animal	காடு, சிங்கம், புலி, தோல், யானை, தந்தம், நாய், பூனை
விலங்கு2	Hand-cuff	குற்றவாளி, கோர்ட், போலீஸ், கைதி, விசாரணை, தண்டனை

Graphical User interface

The following is the graphical user interface for the system. the ambiguous words can be disambiguated using the GUI.



3 Conclusion

There are various methods for WSD. It can be said that for WSD unsupervised approach is economical than supervised method. The knowledge based approaches are welcome thing. The clustering method is a non-supervised method. The present approach is giving encouraging result.

References

- இராசேந்திரன், ச. 1999. பொருட்புல வகைப்பாடும் சொற்களஞ்சியமும் [Classification into semantic domains & Thesaurus]. புலமை தொகுதி 25, எண்2, டிசம்பர், 1999, 47-66.
- இராசேந்திரன், ச. 2001 தற்காலத் தமிழ்சொற்களஞ்சியம். தமிழ்ப்பல்கலைக் கழகம், தஞ்சாவூர்.
- இராசேந்திரன், ச. 2006. தமிழ் மின்சொற்களஞ்சியம். தமிழ்ப்பல்கலைக் கழகம், தஞ்சாவூர்.
- ARUN, M.C. "Evaluation of Word Sense Disambiguation methods applied on Tamil Corpus. PhD thesis.
- BASKARAN, S. 2002. Semantic analyser for word sense disambiguation. MS thesis. Madras Institute of Technology, Anna University, Chennai 2002.
- BASKARAN SANKARAN, VAIDEHI V, 2002, "Role of Collocations and Case-Markers in Word Sense Disambiguation: A Clustering-Based Approach". In Proceedings of IEEE International Symposium on Natural Language Processing and Knowledge Engineering 2002. Vol. I. pp. 625-630. Hammamet, Tunisia.
- BASKARAN S, VAIDEHI V, 2003, "Collocation based Word Sense Disambiguation using Clustering for Tamil". Communicated to International Journal of Dravidian Linguistics, Thiruvananthapuram, India.
- IDE, N. AND VERONIS, J. 1998. Word sense disambiguation: The state of the art. Computat. Ling. 24, 1, 1-40.
- NAVIGLI, R. 2009. Word Sense Disambiguation: A Survey. ACM Computing Surveys, Vol 41, No.2, Article 10, Publication date February 2009.
- PUSTEJOVSKY, J. 1995. The Generative Lexicon. MIT Press, Cambridge, MA.

- SOWA, J.F. 1984. *Conceptual Structures: Information Processing in Mind and Machine*. Reading, Massachusetts: Addison-Wesley Publishing Company.
- YAROWSKY, D. 1992. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In *Proceedings of the 14th International Conference on Computational Linguistics (COLING, Nantes, France)*. 454–460.
- YAROWSKY, D. 1993. One sense per collocation. In *Proceedings of the ARPA Workshop on Human Language Technology (Princeton, NJ)*. 266–271.
- YAROWSKY, D. 1994. Decision lists for lexical ambiguity resolution: Application to accent restoration in Spanish and French. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (Las Cruces, NM)*. 88–95.
- YAROWSKY, D. 1995. Unsupervised word sense disambiguation rivaling supervised methods. In *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (Cambridge, MA)*. 189–196.
- YAROWSKY, D. 2000. Hierarchical decision lists for word sense disambiguation. *Comput. Human.* 34, 1-2, 179–186.
- YAROWSKY, D. AND FLORIAN, R. 2002. Evaluating sense disambiguation across diverse parameter spaces. *J. Nat. Lang. Eng.* 9, 4, 293–310.

A Rule Based Iterative Affix Stripping Stemming Algorithm For Tamil

Damodharan Rajalingam

ABSTRACT

Stemming is an important step in many of the Information Retrieval (IR) and Natural Language Processing (NLP) tasks. Stemming reduces derived forms of words to a common root. When used in IR it increases the recall performance. Stemming algorithms are very specific to a languages as different languages have different rules for derivation. For a language to have better IR and NLP tools stemming algorithm is a basic necessity. There is currently no open implementation of a stemming algorithm available for Tamil. There might be proprietary products that include a stemming algorithm for their uses but having an openly available version will help in implementation of IR and NLP tools for Tamil.

This paper discusses about the implementation of a stemming algorithm that is available[1] as Open Source Software. The algorithm implemented is a rule based iterative affix stripping algorithm. The algorithm is implemented using Snowball[2], a string processing language specifically used for implementing stemming algorithms. The algorithm was tested against the Tamil WordNet data. The results of these tests are also presented in this paper.

Categories and Subject Descriptors

H.3.1 [Content Analysis and Indexing]: Linguistic processing, I.2.7 [Natural Language Processing]: Text analysis

General Terms

Algorithms, Design, Languages

Keywords

Stemming

1 INTRODUCTION

The problem of information storage and retrieval has been receiving more and more attention in the recent years. A more obvious example is how much people depend on web search engines like Google, Yahoo etc in their everyday life. With the growth of the World Wide Web the amount of information being generated is growing at a tremendous rate. This has created a need for faster and better information retrieval systems. The requirement for a good retrieval system is not limited to internet only. Users are having a lots of data in their personal computers that the old method of maintaining hierarchy of folders is not a viable way of finding the required information. So we have desktop search tools which index the data in one's personal computer. Information explosion and need to find relevant information from a huge collection is driving the improvements in Information Retrieval field.

One technique to improve the Information Retrieval performance is to provide the users with ways of finding morphological variants of search terms. If, for example, a user enters the term 'stemming' as a part of the query, it is possible that he/she is also interested in variants such as 'stem' and 'stemmer'. This increases the recall of the IR system. Stemming is also helpful in reducing the size of the index as we need not index all the morphological variants of a word. Most of the times it is good enough to index the stem of the word.

Most of the research on informational retrieval has been based on English as the reference language. Though most of the research in IR is language agnostic there are some areas which are specific to language. Stemming algorithms is one such area. Stemmers for different languages have been developed in the recent years. There is however no stemming algorithm publicly available for Tamil language. With the usage of internet reaching far corners of the world people are now able to create and share content in Tamil. Having good NLP tools will help in developing better Information Retrieval systems for

Tamil content and also in many other text and document processing tools. Stemmer is one such basic NLP tool and is part of other complex tools.

The paper is organised as follows: Sections 2 and 3 give a brief overview on the structure of Tamil words and Stemming algorithms respectively. Section 4 discusses the design of the stemming algorithm and section 5 give details about implementation. The evaluation of the algorithm is presented in section 6 followed by related work in section 7 and conclusion in section 8.

2 STRUCTURE OF A TAMIL WORD

From Wikipedia[4]: “Tamil employs agglutinative grammar, where suffixes are used to mark noun class, number, and case, verb tense and other grammatical categories. Tamil words consist of a lexical root to which one or more affixes are attached.

Most Tamil affixes are suffixes. Tamil suffixes can be derivational suffixes, which either change the part of speech of the word or its meaning, or inflectional suffixes, which mark categories such as person, number, mood, tense, etc. There is no absolute limit on the length and extent of agglutination, which can lead to long words with a large number of suffixes, which would require several words or a sentence in English. To give an example, the word *pōkamūṭiyātavarkalukkāka* (போகமுடியாதவர்களுக்காக) means 'for the sake of those who cannot go', and consists of the following morphemes:

<i>pōka</i>	<i>muṭi</i>	<i>y</i>	<i>āta</i>	<i>var</i>	<i>kaḷ</i>	<i>ukku</i>	<i>āka</i>
go	accomplish	word-joining letter	negation (impersonal)	nominalizer <i>he/she who does</i>	plural marker	to	for

The aim of the stemming algorithm is to strip these extra constituents and map them to a stem corresponding to the root word. Ideally all the words with same root words should be stripped to same stem.

3 STEMMING ALGORITHMS

A stemming algorithm is a computational procedure which reduces all words with same root (or if prefixes are untouched same stem) to a common form, usually by stripping each word of its derivational and inflectional suffixes[3]. Simply stated stemming algorithms are used to group words that arise from same stem or root. The result of a stemming algorithm need not be identical to the morphological root of the word; it is usually sufficient that related words map to the same stem, even if this stem is not in itself a valid root. The stemming process is also called as *conflation* sometimes. There are several types of stemming algorithms based on their approach, accuracy etc like: Brute force algorithms, Affix stripping algorithms n-gram based algorithms, Lemmatisation algorithms, Stochastic algorithms etc.

3.1 AFFIX STRIPPING ALGORITHM

Affix removal algorithms remove suffixes and/or prefixes from terms leaving a stem. These algorithms sometimes also transform the stem. A simple example of an affix stripping algorithm is the one that remove the plural forms by Harman [5]

If a word ends in "ies" but not "eies" or "aies"
Then "ies" -> "y"
If a word ends in "es" but not "aes", "ees", or "oes"
Then "es" -> "e"
If a word ends in "s", but not "us" or "ss"
Then "s" -> NULL

Affix removal algorithms can be simple removal or iterative. In an iterative algorithms affixes are removed until no more affixes can be removed.

4 STEMMING ALGORITHM DESIGN

4.1 WHY AN AFFIX STRIPPING ALGORITHM?

An affix stripping algorithm was chosen for the following reason:

1. An affix stripping algorithm does not require a dictionary. In Tamil the suffixes are attached in an order. So a stemming algorithm which stems most of the words satisfactorily can be designed without the help of a dictionary.
2. The algorithm is very fast. The algorithm need not lookup any dictionary or do complex statistical analysis based on any collected corpus. It just works on the string to be stemmed. So it is very fast.
3. Since it does not require any supporting data the algorithm can be run on any device. For example to port any dictionary based stemmer to a low memory device the dictionary might need to be trimmed down thereby reducing the accuracy of the stemmer. But an affix stripping algorithm does not have any such memory requirements and does not hold much data during its operation
4. There is lack of quality corpus to train statistical algorithms.

4.2 Overview of the algorithm

In Tamil suffixes are used for many things like tense, plurality, person etc. So the suffixes are grouped into categories and a routine is defined for each category to handle the removal the respective suffixes. After removal of suffix for each category there is routine to fix or recode the ending of the word to make it consumable for the next routine. Also before stripping the suffix every routine checks for the current size of the string.

As shown in Figure - 1, first the prefixes are removed followed by the suffixes. Every suffix stripping routine checks for the length of the string before proceeding and after removing a suffix calls the routine responsible for fixing the endings.

4.3 Prefix Removal

There are two routines in the algorithm to handle prefixes. One is for handling the prefix in the questions. Eg. எக்காலம் (which period?). - காலம். Another one is for removing the pronoun prefixes, அ, இ and உ. Eg. அக்காலம் (that period).- காலம்

After removing the prefixes another routine handles fixing the start of the word. The above prefixes introduce ஁ when the root word starts with a vowel. ஁ in the start of the word cannot combine with certain vowels. In such cases this routine substitutes with appropriate vowel as the starting.

Figure 1 - Flowchart of stemming algorithm for Tamil

4.4 Fixing the ending

When a suffix joins a root word one of the following can happen

1. New letters are introduced
2. Some letters are removed
3. The letters are transformed
4. Joins naturally without addition/removal

fix_ending routine tries to handle these modifications before the next suffix removal routine is called.

If the join had caused new letters to be introduced, this routine removes it. For example vallinam consonants appear as conjunctions in many cases. A normal word will not end with a vallinam consonant.

கடக்க	கடக்	கட
original word	suffix stripped	vallinam consonant removed

If the join has caused some characters to be removed it leaves it since it is possible for more than one valid character to be appropriate candidates.

If the join has transformed some of the characters it tries to recode it. It currently cannot recover all such transformations. Eg.

மரத்தின்	மரத்த்	மரம்
original word	suffix removed	end recoded

The fix_ending removes the conjunctions and recodes the transformed letters.

4.5 Suffix removal

The stemming algorithm handles different kinds of suffixes. They are discussed in the following sections

4.5.1 Question suffixes

This routine removes the suffixes. The suffixes are ஆ, ஏ, ஓ.

கண்ணா	கண்ணன்
Is it Kannan	Kannan

4.5.2 Conjunction suffix

This routine removed the suffix உம்

அவனும்	அவன்
Him and	Him

4.5.3 Common words

This algorithm tries to remove some of the common words that are attached to verbs or nouns. These are not suffixes and are proper words.

அவனில்லாத	அவன்
without him	Him

4.5.4 Case suffixes

Tamil case suffixes are attached to the ends of nouns to express grammatical relations (e.g., subject, direct object, etc.) as well as meanings typically expressed in English through pre-positions (e.g., 'in', 'to', 'for', 'from', etc.).

அவனிடம் (with him)	அவன் (him)
மரத்தில் (in tree)	மரம் (tree)

4.5.5 Plural suffix

The plural suffix in Tamil is கள்.

மரங்கள்	மரம்
Trees	Tree

4.5.6 Imperative suffixes

These are used to command a person.

காண்பி	காண்
show me	see

4.5.7 Tense suffixes

This routine removes tense indicating suffixes. It also include person suffixes.

பிரிகின்றன	பிரி
leaving	leave

Apart from the standard suffixes the routine also removed கொண்டு and similar words.

4.6 Minimum length criteria

Being a strong stemmer it has a tendency to overstem some words to single letters. To prevent this every routine checks for the length of the string. Currently the minimum length is set as 4 characters. These are not 4 characters exactly since in Unicode a meaningful character can be represented by more than one code points. So the check made in the implementation actually only verifies the number of codepoints in the string than the actual meaningful characters. Also the routine which fixes the ending does not check for the length of the string. So it is still possible to get a stem of length one character.

5 ALGORITHM IMPLEMENTATION

The algorithm described in the previous section was implemented using Snowball language. Snowball[2] is a small string handling language mainly designed to define stemming algorithms in a natural way. The language was created by Dr. Martin Porter when he saw various buggy implementations of his famous Porter algorithm [6] for English. The reasons for errors in the implementation can be grouped into following: misunderstanding of the original algorithm, errors in handling the encoding and the programmers urge to improve the algorithm. The language was mainly developed to avoid such implementation errors and it widely used now for developing stemming algorithms. Stemming algorithms for many languages like German, French, Turkish etc have been implemented using the Snowball language.

The code written in Snowball cannot be used with other programs as such. We use snowball compiler to convert the Snowball code to any other programming language. As of now Snowball supports C and Java output. Using the generated C code we can create bindings for many other programming languages. The algorithm implementor is now not bothered about implementing algorithms in various other programming languages. It is also possible to extend the snowball compiler to generate code in other programming languages also.

6 EVALUATION

6.1 Tamil WordNet

Tamil WordNet[8] is an attempt to build a lexical network for Tamil language along the lines of the English WordNet so that it can be used as a tool for enhancing the performance of MT systems involving Tamil. The wordnet data is available for free as a sql dump. It has more than 4lakh words with its morphological root. The data available is coded in english transliteration. A transliterator program was written to convert it to UTF8 data. Some of the issues with data are the typing errors and inclusion of many foreign words and non classical Tamil words.

Figure 3 - Database Schema of Tamil WordNet

6.2 Correctness of the algorithm

The correctness of the algorithm is usually measured by identifying the number of semantically related words that are correctly assigned to the same conflation class. Another measure is to see how close the stem is the morphological root of the word.

6.2.1 Variation with morphological root

The stemming algorithm was run against the collection in Tamil WordNet. The Hamming distance between the output of the stemming algorithm and the morphological root was measured. The results are listed in the table below.

Table 1 - Hamming distance from morphological root

Measure	Value
Mean	1.9237
25th percentile	0
Median	2
75th percentile	3

6.2.2 Stems per morphological root

In this test we measure the number of stems per morphological root. The test counts the number of stems created for the words derived out of same root. This describes how correctly we assign the conflation class for the inflated words. Higher number indicates that the algorithm is not stemming the semantically related words to same conflation class.

Table 2 - Stems per morphological root

Measure	Value
Mean	1.7383
25th percentile	1
Median	1
75th percentile	1

For 75% of the root words the number of stems is 1. So the stemmer is doing a good job of mapping similar words to same conflation class.

6.3 Strength of the algorithm

The amount of change the algorithm causes to the given string decides the strength of the algorithm. A strong algorithm tries to remove as many suffixes as possible. A light stemmer usually handles less cases and does not make much modifications to the provided string.

In their paper "Strength and Similarity of Affix Removal Stemming Algorithms", Frakes and Fox[7] propose the following metrics to measure the strength of affix removing algorithms:

1. Mean number of words per conflation class - average number of words that correspond to the same stem for a corpus.
2. Index compression factor - this is the fractional reduction in the index size achieved by stemming. This is given as
3. The number of words and stems that differ - stemmers may often leave words unchanged. This measures such words
4. Mean number of characters removed in forming stems
5. Median and mean modified Hamming distance between the words and their stems - Hamming distance between strings of equal length is the number of character they are differing at the same position. For the strings of unequal length the Hamming distance is the difference in their lengths are also added up.

Table 3 - Modified Hamming Distance Descriptive Statistics

Mean	2.76
Std. deviation	1.97
Minimum	0
25th percentile	2
Median	3
75th percentile	4
Maximum	18

Table 4 - Modified Hamming Distance Descriptive Statistics of some popular stemming algorithms for English Frakes and Fox [7])

	Lovins	Paice	Porter	S-removal
Mean	1.72	1.98	1.16	0.03
Std. deviation	1.64	1.92	1.40	0.19
Minimum	0	0	0	0
25th percentile	0	0	0	0
Median	1	2	1	0
75th percentile	3	3	2	0
Maximum	10	13	9	3

Table 5 - Strength description statistics for Tamil stemmer

Mean Modified Hamming Distance	Median Modified Hamming Distance	Mean Characters Removed	Compression Factor	Mean Conflation Class Size	Word and Stem Different
2.76	3	2.4	0.65	2.88	86.53%

Table 6 - Strength description statistics of some popular stemming algorithms for English (Frakes and Fox [7])

	Mean Modified Hamming Distance	Median Modified Hamming Distance	Mean Characters Removed	Compression Factor	Mean Conflation Class Size	Word and Stem Different
Lovins	1.72	1	1.67	0.29	1.42	69.4%
Paice	1.98	2	1.94	0.33	1.49	69.5%
Porter	1.16	1	1.08	0.17	1.20	56.2%
S-Removal	0.03	0	0.03	0.01	1.01	3.3%

From the metrics measured above it is evident that the designed stemmer is a strong stemmer and is decently accurate. The data for English stemmers is provided as an information and not for comparison. An apple-to-apple comparison cannot be made since the algorithm is for English and uses a different corpus.

6.4 Shortcomings

Because of the agglutinative nature of the language it is possible to form compound words which combine two or more stems into a single word. In such scenarios we will have to find the word boundaries and identify the individual stems. The algorithm proposed in this paper does not handle such scenarios and is only applicable to non-compound words.

7 RELATED WORK

There is a paper published by Vivek Anandan Ramachandran and Ilango Krishnamurthi [9] on an iterative suffix stripping stemming algorithm for Tamil. The paper is behind a paywall and no further details are available regarding the implementation and there is no openly available implementation of the algorithm.

8 CONCLUSION

This paper described the implementation of a rule based iterative affix stripping algorithm and its implementation using Snowball language. The evaluation of the algorithm was done using the Tamil WordNet corpus. The correctness of the algorithm was measured in following ways: 1. Hamming distance between the stem and morphological root. 2. Number of stems per root. The values obtained for these measures indicate that the designed algorithm is reasonably accurate. The strength of the algorithm was measure using the methods proposed by Frakes and Fox[7]. The values obtained in these tests indicate that the designed algorithm is a strong stemmer and it provides good index compression ratio. The The implementation of the algorithm is available in open source and can be used by other tools that require Tamil language stemming.

References

- <https://github.com/rdamodharan/tamil-stemmer>
- <http://snowball.tartarus.org/>
- Lovins, Julie Beth. "Development of a Stemming Algorithm." Mechanical translation and computational linguistics 11 (1968): 22-31.
- http://en.wikipedia.org/Tamil_grammar
- Harman, D. "How effective is suffixing?" Journal of the American Society for Information Science.42 (1991): 7-15
- Porter, M. F. "An Algorithm for Suffix Stripping." Program 14 (1980): 130-137.
- Frakes, William B. and Christopher J. Fox. "Strength and similarity of affix removal stemming algorithms". ACM SIGIR Forum 37 (2003): 26-30.
- http://www.au-kbc.org/research_areas/nlp/projects/tamil_wordnet.html
- Ramachandran, Vivek Anandan and Krishnamurthi, Ilango. "An Iterative Suffix Stripping Tamil Stemmer". Proceedings of the International Conference on Information Systems Design and Intelligent Applications (2012): Volume 132, 583-590

Developing Online Sangam Corpus and Concordance

Dr. A. Kamatchi, CAS in Linguistics, Annamalai University

Introduction

Corpus linguistics, a new method of language study, has emerged in recent years and it has generated a number of research methods, attempting to trace a path from data to theory. Further it is explained, as we all know, that corpus is a large collection of written/spoken materials in machine readable form. It provides, as much as possible, an authentic data for linguistic studies as well as for other related studies in the languages. Most lexical corpora, today, are part-of-speech-tagged (POS-tagged). According to Wikipedia, “a landmark in modern corpus linguistics was the publication by Henry Kucera and W. Nelson Francis of *Computational Analysis of Present-Day American English* in 1967, a work based on the analysis of the Brown Corpus, a carefully compiled selection of current American English, totalling about a million words drawn from a wide variety of sources. Around forty thousand lines are available in these *Sangam* texts, which comprise *eTTuttokai* – *naRRinai*, *kuRuntokai*, *aikkuRunuuRu*, *patiRRuppattu*, *paripaaTal*, *kalittokai*, *akanaa~~nu~~uuRu* and *puRa naa~~nu~~uuRu* and *pattup paaTTu* literature – *tirumurukaaRRuppaTai*, *porunaraaRRuppaTai*, *ciRupaaNaaRRuppaTai*, *perumpaanaaRRuppaTai*, *mullaippaaTTu*, *maturaikkaa~~ni~~*, *neTunal vaaTai*, *kuRi~~nc~~ippaaTTu*, *paTTinappaalai* and *malaipaTukaTaam* – which are earliest literary texts and dated to 3rd century B.C. to 2nd century A.D.

Developing the online corpus is the need of the hour

As far as Tamil language is concerned, the first corpus for modern written Tamil was started to be built in the Central Institute of Indian Languages (CIIL), Mysore in 1987. But its usage by the people is very less in number. The reason may be that it is only in the CD form but not posted in the internet. The other one which is now available in the internet is the Cre-A: Online Tamil Language Repository posted by Cre-A. These corpuses are, of course, concerned only with the modern Tamil, but not with the other period of languages.

Besides these corpora of living languages, computerized corpora have also been made of collections of texts in ancient languages. According to Wikipedia, “An example is the Andersen-Forbes database of the Hebrew Bible, developed since the 1970s. The Quranic Arabic Corpus is an annotated corpus for the Classical Arabic language of the Quran. On this line, the present study attempts to prepare a *Corpus and Concordance to Sangam Tamil*, through which one can search lexical items, rather than words, and their concordance available in these texts.

Useful for inter and intra language studies

The completion of this work definitely leads to develop software for *Sangam* Tamil and also for all other old texts. Ultimately, it is highly expected that once this work would be completed, it will, undoubtedly, help us to compare intra and inter languages and language families grouped in the world. Moreover it could be useful for the scholars/researchers working in the field of comparative language as well as historical linguistics studies. Further, it may be helpful for the quantitative analysis, too.

A maximum number of words in *Sangam* Tamil, as we all know, occupy the head entries in Dravidian Etymological Dictionary (DED), which was prepared five decades ago and was widely used by the scholars worldwide. So, posting this material in the website is necessary for the use of the scholars working in the areas of comparative linguistics, in general, and comparative Dravidian, in particular. In the same way, it is also very useful for the historical linguistics scholar in the world. Moreover, it may also use to glottochronological linguistics study all over the world. It is sure that this online *Sangam* Corpus and Concordance would represent the classic language in Tamil. After *Tolkaappiyam*, in which a few words have

been simply explained on the part *uriccol*, this would be the potential work with using the modern theories in linguistics and scientific methods in the process of preparing the collection of lexical items. This work may be very useful for utilizing for the school curriculum because most of the students and teachers cannot understand the old Tamil words in proper way. In this way, this study tries its level best to make classical literature easier for the teachers and students community in understanding the classical works in the school curriculum of Tamil living countries.

Morphological Parser

Approximately, ten thousand words are, perhaps, attested in these texts. Without parsing the poetic lines of these texts in these literatures, we are not able to make the machine to understand the materials. Therefore, every line should be parsed morphologically then only, the original root forms could be retrieved from the database. This parser would consist of not only nouns and verbs but also the possible grammatical items found in the texts. These words would be collected systematically and incorporated in the corpus. For developing this online *Sangam* Corpus and Concordance, the data would be collected from old Tamil *Sangam* texts. In fact, it is the pioneer attempt to develop the *Corpus* and *Concordance to Sangam Tamil*, which consists of not only lexical items but also the grammatical elements attested in these texts. When we click one word already given in the Drop-Down box in the window, we can, accurately, get a number of occurrences of that word, the line on which it occurs, its meaning in that particular line, literature name and poem number along with the line number. If it is successfully completed at the earliest, this work will be the model for other period of Tamil works.

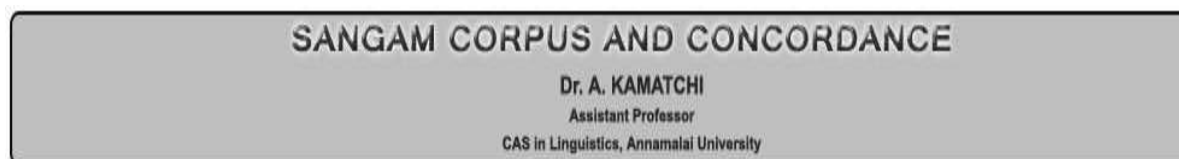
Available materials

There are indexes for some of the *Sangam* works, but not for all, of course. Among these *Sangam* texts, only for some of the anthologies, the indexes have been made by the scholars and published. Though the others have been indexed, they have not published but are only in the form of unpublished Ph. D. Dissertations. Of course, these indexes comprise the words with root form plus some other suffixes, as far as the verb is concerned. Further, it includes their occurrences only with poem and line number. In case someone wants to refer something, he needs of original text for the reference. Through these available published materials, it is too difficult for a scholar to locate such word in *Sangam* Tamil. But, if such an online work is completed, then we need not want of such textual materials. By a single click of a word, one can get all the occurrences in this literature with poem and line number of those occurrences. As we all know, accuracy is one of the unique features for computer. Through this program, one can attain the total number of occurrences in these *Sangam* works. Of course, it is the user friendly material, of course. One can easily find a word which is doubtful for the researcher working in *Sangam* Tamil. Further, such word that occurs in one text may or may not available in the other texts. For such situation, it is very helpful for the researcher or reader to find out such things in these areas.

Creation of data base

The data would be collected from the original texts of *Sangam* literature, using the Index works, which are available in various institutions in the nation as well as the commentaries of the literature. Using the computer software such as POS tagging, *Sangam* corpus, searching engines, the data would be collected from the *Sangam* texts and analyzed in the way of descriptive methods. The works of the eminent Tamil Linguistics Scholars such as Prof. Nida, Prof. V.I. Subramoniam, Prof. S. Agesthalingom, Prof. M. Israel, Prof. M. Elayapermal, Prof. A. Kamatchi, Prof. Rm. Sundaram Prof. S. N. Kandaswamy, Ms. Eva Wilden and so on, would mostly be utilized for this study.

The proposed online corpus model would be as follow:



Select Literature to search

குறு

Select Word

அணி-beauty

Select Grammaticalitem

அ1-வேற்றுமை6/உடமைப்பொருள்

Search Any Word / Grammaticalitem

அல்

Search

அலங்கு சினை மாஅ த்து அணி மயில் இரு க்க் உம்2

அலங்குசினை மாஅத் தணிமயி லிருக்கும் ஐங் 8 :4

அணி துறை ஊரன் மார்பு ஏ

மணித்துறை யூரன் மார்பே ஐங் 14 :3

சிலம்பு அணி கொள் ட்/த் அ2 வலம் சுரி மரா த்து

சிலம்புஅணி கொண்ட வலம்சுரி மராஅத்து குறு 22 :3

பூ குல் ற்த் உ2 எய்து இ2 அ2 புனல் அணி ஊரன்

பூக்குற் றெய்திய புனலணி யூரன் ஐங் 23 :2

பசப்பு அணி ந்த் அன் அ5 ஆன் மகிழ்ந என் கண் ஏ

பசப்பணிந் தனவான் மகிழ்நவென் கண்ணே ஐங் 45 :4

யாறு அணி ந்த் அன்று1 நின் ஊர் ஏ

யாறணிந் தன்றுநின் னூரே ஐங் 45 :3

புலம்பு அணி ந்த் அன்று1 அவர் மண ந்த் அ2 தோள் ஏ

புலம்புஅணிந் தன்றுஅவர் மணந்த தோளே. குறு 50 :5

துறை அணி ந்த் அன்று1 அவர் ஊர் ஏ இறை இற ந்த் உ

துறைஅணிந் தன்றுஅவர் ஊரே இறைஇறந்து குறு 50 :3

நல் அணி நய ந்த் உ2 நீ துற த்தல் இன்1

நல்லணி நயந்துநீ துறத்தலிற் ஐங் 55 :3

மலர் அணி வாயில் பொய்கை ஊர் நீ

மலரணி வாயிற் பொய்கை யூரநீ ஐங் 81 :3

Selected Keyword : அணி - beauty

No of Occurances : 11

Of course, the database also enables for programming of online dictionary of *Sangam Tamil*. There are a number of online dictionaries – online dictionary kids, online dictionary for students, medical online dictionary, legal online dictionary, etc. – which are developed neatly in many languages in the world. Even in Modern Tamil, many websites for online dictionary are available but there is no *Sangam Classical Tamil-English* dictionary although a number of indexes are available with lack of head entries. The scholars in Tamil have prepared indexes for individual literature but they have not consolidated in one platform. In fact, such a dictionary is necessitated for translation from one language to another.

Organization of dictionaries

As we all know, accuracy is one of the unique features for computer. The text of dictionary is organized under head words, which would be listed in alphabetical order. It is estimated that there may be more than 10, 000 words available in *Sangam Tamil*. All the words could be accommodated in the dictionary. Compound words will also be given separate entries in the alphabetical order. Certain words may have the same spelling but different meanings and different etymologies and such words are, as we all know, called as homonyms; they would be treated as separate head words, even when they have the same parts of speech. Homonym would be added with the words to distinguish identical headwords. There is a list of words which is alphabetically arranged and provided in the data base, from which the data would be retrieved for the online dictionary. A few words from the *Sangam texts* are given as sample here.

அணி
அம்
அம்பல்
அரசு
அல்லது
அலங்கு
அவினி
அறம்
ஆகு
ஆதன்
ஆர்
இதழ்
இரவலர்
இரு
இவண்
உழவர்
உளை
உளர்
உளரன்
உளறு
எந்தை
எம்
ஓது
கஞல்
கயலார்

References

- Agesthialingom, S. 1979. **A Grammar of Old Tamil (With special Reference to patiRRuppattu)** Annamalai University, Annamalaiagar.
- Andiappa Pillai, D. 1970. **Descriptive Grammar of kalittokai** Ph.D. Dissertation, University of Kerala, Trivandrum.
- Andronow, M. 1959. '**On the future tense base on Tamil**', Tamil Culture, Vol.3, Madras.
- Andronow, M. 1978. **Tense and Mood in Dravidian: A Comparative Study**, in (tamiḷiyal) Jornal of Tamil Studies.
- Elayapermal, M. 1958. '**The maar suffix in Early Tamil Literature**', Indian Linguistics, Ralfh Turner Jublee, Vol.1.
- Elayapermal, M. 1975. **Grammar of AṇḱuRuṇḱuRu with Index**, University of Kerala, Trivandrum.
- Eva Wilden, 2008. **Word Index of NaRRiNai** (vol.3) Tamil mann patippakam, Chennai.
- Israel, M. 1964. '**The finite verbs of 'ceyyum' pattern in Tamil**' Indian Linguistics, Vol.25, Poona.
- Kandaswamy, S.N. 1962. **paripaaTal- A Linguistic Study**, M.Litt. Dissertation, Annamalai University, Annamalaiagar.
- Krishnambal, S.R. 1974. **Grammar of kuRuṇḱokai with Index**, University of Kerala, Trivandrum.
- Subrahmanyam, P.S. 1971. **Dravidian Verb Morphology (A comparative study)**, Annamalai University, Annamalaiagar.
- Subramaniyan, S.V. 1972. **Grammar of AkāṇḱaanuuRu with Index**, University of Kerala, Trivandrum.
- Subramoniyam, V.I. 1962. **Index of puRaṇḱaanuuRu**, University of Kerala, Trivandrum.

An Electronic Dictionary for Pathinenkilkanakku Literature

Dr.K.Umaraj, Assistant Professor, Dept. of Linguistics, Madurai Kamaraj University, umarajk@gmail.com

Introduction:

Electronic dictionary is a machine readable dictionary, which provides search facilities to identify meaning and grammatical information of a particular word. There are number of Electronic dictionaries for Tamil have been web published by different Research Institute and Commercial organizations. For example, PAL organization published an English –English -Tamil electronic dictionary. It has 22, 000 heads words and 35, 000 sub words, Tamil Lexicon and Muthu Shanmugam pillai's Tamil- Tamil e_dictionary was available in the Tamil Virtual University website. Winslow and Lifco companies published online dictionaries. Cre-a's Tamil –Tamil –English dictionary was web enabled recently. This dictionary has 21, 000 headwords, 38, 000 citation, 1700 Srilankan Tamil words, 342 pictures and 1892 pages. Moli trust published 'Tarkalat Tamil maraputtotar akarathi' in electronic form on September 2009. An electronic dictionary for Scientific Technical Terms in Tamil has also been developed by Chellapan Radha. It has different kinds of retrieval and browsing facility. CoRpuaiyal is an online dictionary contains 20, 000 root words. Each entry in the dictionary includes the Tamil root word, its English Equivalent, different meaning of the word, and the associated syntactic category.

At international Level, the Institute of Indology and Tamil studies, Colone, Germany has produced Sanskrit, Tamil, Pahlavi dictionary. This dictionary has 1, 66, 434 entries. It was developed by Prof. Malton from University of Colone. "Core vocabulary for Tamil" is an online searchable dictionary published by the department of South Asia Regional Studies, University of Pennsylvania. At National level, Vijaya and Paul (2004) developed an electronic dictionary for Tamil named Multidimensional SMART dictionary, a project work of Central Institute of Indian Languages, Mysore. However, Most of these dictionaries are based on modern Tamil. So an exclusive dictionary for Pathinenkilkanakku Literature is the need of the hour and an Electronic dictionary exclusively for Pathinenkilkanakku is yet to be developed. While developing an electronic dictionary for pathinenkilkanakku a number of linguistic problems arise. Those problems are discussed in detail in this paper.

Uses of Electronic dictionary for pathinenkilkanakku Literature

E-dictionary for pathinenkilkanakku Literature helps to strengthen the reference section of Classical Tamil studies.

The electronic data can be used as an e_Library to enhance the knowledge of Post Sangam Literature and this will help to know the archaic words.

The attempt would result for an updating of existing Tamil Lexicon.

The study would also assist in the preparation of Historical Tamil Grammar.

The creation of dictionary would be of great use to all the research works from Post Sangam Tamil to Modern Tamil.

It is a supplement for updating the current Tamil Electronic corpus.

It also helps to develop Thesaurus for Tamil Language and concordance tool for Tamil.

The entire lexical data base (e_text and e_dictionary) can be used as a tool and that will serve the purpose for those who are involving to develop further the NLP system in Tamil.

Methodology

The Methodology of the e_dictionary for Pathinenkilkanakku Literature was given below. Preparation of texts activities will be done with the help of special software. After the text preparation, most commonly occurring head words from the Pathinenkilkanakku corpus should be selected. Transliteration, grammatical information and meaning with citations are given for each head word. Then the content will be put into data base and an advanced search engine will be developed. The data will be encoded in Unicode notation.

Architecture

Preparation of texts

|

Selecting most commonly occurring head words from the Pathinenkilkanakku Corpus

| Giving transliteration for each head word
 | Assigning grammatical information to the head word
 | Assigning meaning (English and Tamil) to the head word along with citation
 | Designing the data base
 | Put into the data base
 | Developing an e_dictionary

Issues while developing a dictionary for Pathinenkilkanakku Literature

1. Identifying what is a word? what is not a word ? in Pathinenkilkanakku Literature is a problem. .whether all the compound words should be written as one word or not, whether all postpositions and clitics should be part of the word or it should be a separate word? There was no clear guidelines for segmenting the Pathinenkilkanakku texts.
2. Assigning grammatical information to the head word is another issue in Pathinenkilkanakku Literature. The Tamil traditional grammarians classified the words into four types. Asher (1982:101, 102) classified words into 6 types. Lehmann (1989) classified words into 8 types and .Kothandaraman.R (1989) classified the words into 10 types. Due to, different approaches in classification of words by grammarians, each dictionary follows their own way of assigning the grammatical information to a particular word. The lexical entry 'aktaanru' is marked as adjective in Tamil Lexicon and it is marked as verb in Maree's Dictionary. Similarly the word 'akiya' is marked as verb in Maree's dictionary. But it is an adjective.
3. Identification of meaning for a particular word in Pathinenkilkanakku is another issue.

Consider the following sentences.

1. nilaiyaa ena uNarntaar (Nalatiyaar" 182-3)
2. tan makan caanRoon ena k eeTTa taay (Thirukural: 69:2)
3. pagai, paavam, accam, pali ena naanku (Thirukkural : 146-1)
4. pon ena aangee puramveraar (Thirukural 487-1)
5. olleena ooTum malai naataan (kainilai 7-3)

In the above sentences, the particles 'ena' has different senses based on the context of the text. In the same way the word 'aaga' in Pathinenkilkanakku Literature has different senses based on the context. For example in the following sentence "kanru aaga, iivaar aaga" (Nalatiyar 279-1), the word 'aaga' functions as a role (in the sense of "as") .But the following sentence the word 'aaga' functions as an adverb " paayntu aruvi aatinoomaka " (ainthinai:15-3).

Conclusion:

An Electronic dictionary for Pathinenkilkanakku Literature is yet to be developed. While developing an electronic dictionary for Pathinenkilkanakku Literature a number of linguistic problems arise. So we need to solve these problems before developing the e-dictionary for Pathinenkilkanakku Literature.

Bibliography

- Lehmann, Thomas (1989) A grammar of Modern Tamil. Pondicherry: Pondicherry Institute of Linguistics and Culture.
- Ramakrishnan, S. (Ed.) (2008) Kriyavin Tarkalat Tamil Akatati (Dictionary of contemporary Tamil) Chennai: Mozhi. Chennai: Cre-A.

தமிழ் தரவுத் தளங்கள்

கோ.சக்கரபாணி,

முனைவர் பட்ட ஆய்வாளர், தமிழ்ப் பல்கலைக் கழகம், தஞ்சாவூர்

முன்னுரை

வரலாறு என்பது பல ஆண்டுகளுக்கே முற்பட்டது என்பதல்ல. ஏன் நேற்றும் இன்றும் இந்த நேரத்திலும் நடப்பது கூட சரித்திரம் தான். இன்று நடப்பது முதற்கொண்டு பின்னோக்கி பயணிப்போமேயானால் மூல ஆதாரம் அல்லது ஆவணங்கள் கிடைப்பது வரை நடந்த சம்பவங்கள் அனைத்தும் சரித்திரமேயாகும். நிறைய விடயங்கள் குறுகிய காலத்தில் அழிந்து போகக்கூடியதாக இருப்பதால் ஆவணப்படுத்தலில் அனைவரும் அக்கறை காட்ட வேண்டியவர்களாக இருக்கின்றோம். இன்று பயனற்றவைகள் என ஒதுக்கும் காரியங்கள் நாளை பயனுள்ளவைகளாக மாறக்கூடும். சங்க இலக்கியங்கள் பலவற்றை நாம் ஒதுக்கியதை இன்று நினைத்து வருந்துகிறோம். குடும்பப் புகைப்படங்கள், வீட்டுப் பத்திரங்கள், பழைய திருமண அழைப்பிதழ்கள் கூட எதிர்காலத்தில் ஆவணமாகி விடுகின்றன. பாதுகாப்பது என்பது வேறு ஆவணப்படுத்துவது மற்றும் சேமிப்பது என்பது வேறு. தரவுகளை செம்மைப்படுத்தித் தரப்படுத்தி வடிவமைக்கப்பட்ட தளத்தில் சேமிப்பதினால் தரவுத் தளம் என்கிற வழக்காற்றுச் சொல்லை உடைத்தாயிற்று.

கணினியின் பயன்பாடு

உலகம் முழுவதிலும் கணிப்பொறியின் பரவல் அதிகரித்துள்ளது. சாதாரண மக்கள் தங்களுக்கேற்ற பயன்பாடுகள் தத்தம் தாய்மொழியிலேயே இருக்கவேண்டும் என எதிர்பார்க்கின்றனர். கணினி பயன்பாடும் அவரவர் தாய்மொழி எழுத்துருக்களில் இருக்கவேண்டும் என எண்ணுகின்றனர். அதற்கேற்றவாறு விசைப்பலகைகள் உருவாக்கப்படவேண்டும். எழுத்துருக்கள் ஒருங்குறியீடாக (Unicode) இருந்தால் மட்டுமே உலகமுழுவதிலுள்ள மக்கள் அம்மொழி தெரிந்தவர்கள் பயன்படுத்திப் பயன்பெற முடியும்.

தரவுத் தளங்கள்

ஆவணத் தரவுகளை சேமித்து வைத்துப் பயன்படுத்துவதில் காட்டும் அக்கறை தரவுத் தகவல்களை தரவுத் தளத்தில் சேமித்து வைப்பதிலும் காட்டவேண்டும். தகவல் தளங்களை எண்ணி அல்லது மின்படிவ நூலகத் தளங்கள் [Digital Libraries], இணையத் தளங்கள் [Internet sites], தரவுத் தளங்கள் (databases) என வகைப்படுத்தலாம். தமிழ் மொழி சார்ந்த தகவல்களை மற்றும் பட்டியல்களை ஆவணப் படுத்தியிருப்பதால் இவைகளை தமிழ் மின்நூலகத் தளங்கள்[Tamil Digital Libraries], தமிழ் இணையத் தளங்கள் [Tamil Internet sites] மற்றும் தமிழ் தரவுத் தளங்கள் (Tamil databases) என்று அழைக்கிறோம்.

தமிழ் எண்ணிம நூலகங்கள்

பழையன போக்குகிறோம் என்கிற பெயரில் போகி பண்டிகையில் நாம் ஏடுகளை எரித்திருக்கிறோம். எட்டுச் சுவடிகளைப் பெயர்த்தெழுதும் பழக்கத்தை விட்டிருக்கிறோம். பழைய ஆவணங்களை, ஓலைச்சுவடிகளை எண்ணிமப் படுத்தி பேணிப் பாதுகாக்க வேண்டும். கணிப்பொறி மூலம் ஓலைச்சுவடிகளின் தகவல்களை பாதுகாத்து எளிதில் தொகுக்க இயலும். ஓலைச்சுவடிகளின் தகவல்களை தட்டச்சு மூலம் கணிப்பொறியில் சேமிப்பதனால் பல்லாயிரக்கணக்கான ஓலைச்சுவடிகளிலுள்ள செய்திகள் இலக்கமாக்கப்படலாம். தனியார் வசம் உள்ள ஓலைச்சுவடிகளை தேடி கண்டுபிடித்து மின்படிவமாக்கிய பின் அவற்றை அவர்களிடமே ஒப்படைத்து விடலாம். பல்வேறு இடங்களிலிருந்து பெறப்பட்டு மின்படிவமாக்கிய தகவல்களை இணையத்தின் மூலம் பெற்று ஒரே தளவுத் தளத்தில் சேமித்து வைத்துக்கொள்ளலாம். அரசே முன்வந்து ஓலைச்சுவடிகளிலுள்ள அறிவுச் செல்வங்களை எண்ணிமப் படுத்தவேண்டும். தகவல்களை தட்டச்சு செய்வதற்கும் இணையத்தின் வழியே பெறுவதற்கும் தரவுத் தளத்தில் சேமித்து வைப்பதற்கும் ஏற்ற நிரல்களை தயார் செய்து வைத்துக் கொள்ள வேண்டும். ஏற்ற தகவல்களை வினவி பெறுவதற்கும், பெற்ற தகவல்களை அட்டவணைப் படுத்துவதற்கும், ஆவணப்படுத்துவதற்கும் ஏற்ற நிரல்களை விரிவான முறையில் திட்டமிட்டு தயாரிப்போமேயானால் தரவுத் தளத்திலிருந்து தங்கு தடையின்றி தகவல்கள் கொட்டும். வருங்காலத்தில் அடிப்படையான தகவல் தொடர்புகள் அனைத்தும் டிஜிட்டல் வடிவத்தில், இணையம் போன்ற ஏதோ ஒரு வலைப்பின்னலில் தான் நிகழப் போகிறது என்பதால் அந்தத் தொழில்நுட்ப வளர்ச்சிக்கு ஈடுகொடுக்கும் வகையில் தன்னை வளர்த்துக் கொள்ளத் தமிழ் தயாராக வேண்டிய நிர்ப்பந்தம் இருக்கிறது.

தமிழ் இணையத் தளங்கள் மற்றும் வலைப்பக்கங்கள்

தமிழ் வலைப்பக்கங்களில் தகவலைத் தேடிப் பெறுவதற்கான சிறந்த தீர்வு நெறிகள் உருவாக்கப்பட வேண்டும். சொற்களின் பொருளையும் அடிப்படையாகக் கொண்டு தேடும் தேடுபொறிகள் உருவாக்கப்பட வேண்டும். 'மலர்கள்' எனத் தேடும்போது 'பூக்கள்' இடம்பெற்ற ஆவணங்களும் தேடுபொறியில் சிக்க வேண்டும். சுய உதவிக் குழுமங்கள் தன்னார்வத் தொண்டு நிறுவனங்கள் சமயச் சார்புடைய நிறுவனங்கள் வணிக மையங்கள் மைய அரசு நிறுவனங்கள் தங்கள் மின் வளங்களைத் தமிழில் மொழிமாற்றம் செய்து இணையத்தில் வெளியிட அரசு ஊக்கம் அளிக்க வேண்டும்.

தமிழ் தரவுத் தளங்கள்

தரவுத் தளம் (databases) என்பது, தரவுகளை (data) கணிப்பொறியில் சேமித்து வைக்கப்படும் ஒரு இடம். எடுத்துக்காட்டாக ஆரக்கிள், மை-எஸ்கியூஎல், மைக்ரோசாப்ட் எஸ்கியூல், டிபி2, போஸ்ட்கிரஸ் போன்றவற்றை கூறலாம். இதை உருவாக்குவதற்கு வகைப்படுத்தல் பற்றிய செந்தரமொன்று (standard) தேவை. பல்வேறு பணிகளுக்கு நாம் அன்றாடம் பயன்படுத்தக்கூடிய சொற்களை கணினியில் உள்ளீடு செய்யும்பொழுது அதனையே ஒரு தரவுத் தளமாக மாற்றி தேவைக்கேற்ப பயன்படுத்திக் கொள்ள ஏதுவாக ஒரு மென்பொருள் (Software) வடிவமைத்தால், தமிழில் தரவுத் தளம் உருவாவது மாத்திரமல்ல தமிழ்க் கணிப்பணியை மிக எளிதாகவும் சிறப்பாகவும் செய்யமுடியும்.

அரசு ஊழியர்கள் குறித்த விவரங்கள் அடங்கிய மைய தரவுத் தளம் உருவாக்கப்படும்' என, தமிழக அரசு அறிவித்து உள்ளது. அரசு ஊழியர்கள் தொடர்பான விவரங்கள் அடங்கிய மைய தரவுத் தளத்தின் மூலம், மேலாண்மை தகவல் முறையில், பணியாளர்களின் விவர எண்ணிக்கை, ஊதியம் கணக்கிடல், வங்கிகள் மூலம் ஊதியம் ஒப்படைத்தல், பல்வேறு வகையான விடுப்புகள், ஊழியர்களின் தண்டனை குறித்த விவரங்கள், நன்னடத்தை பற்றிய தகவல்கள், பல்வேறு வகையான முன்பண பறிமாற்றங்கள், கடன் நிலுவை, குடும்ப ஓய்வூதியத்திற்காக குடும்ப நபர்கள் பற்றிய விவரங்கள், சேமநலநிதி விவரங்கள், ஊழியர் பணியிட மாற்றம், பதவி உயர்வு, பணியிட தேவை, பணி பகிர்ந்தளித்தல் மற்றும் அதை சார்ந்த பிற விவரங்களை பெற தரவு அட்டவணை (Database Table) ஒன்று தயாரிக்கவேண்டும். அதற்கேற்றவாறு வலைதள பட்டியல் மென்பொருள் (Software)உருவாக்கப்படும். இத்திட்டம், 2013-14ம் ஆண்டில் நடைமுறைப்படுத்தப்படுகிறது. இதற்காக வலைதள பட்டியல் மென்பொருள் தயாரிக்கும் நடவடிக்கைகள் மேற்கொள்ளப்படும். இது மனித வள மேலாண்மை தொடர்பான திட்டமிடலுக்கு, உகந்த கருவியாக அமையும் என, பொதுத்துறை கொள்கை விளக்க குறிப்பில் தெரிவிக்கப்பட்டு உள்ளது. தமிழக அரசின் துறை தொடர்பான மின் வடிவ வளங்கள் இணையத் தளங்கள் தரவுத் தளங்கள் இரு மொழிக் கொள்கை அடிப்படையில் ஆங்கிலத்திலும் தமிழிலும் இருக்க அரசு ஆணை பிறப்பிக்க வேண்டும். இவை மாத்திரமல்ல, மக்கள் கணக்கெடுப்பு, குடியிருப்பு கணக்கெடுப்பு, பல்வேறு வகையான பள்ளிகள், கல்லூரிகள், நூலகங்கள் தொடர்பான தகவல்கள், பல்வேறு மதம் சார்ந்த கோயில்கள், சமூகக்கூடங்கள், சமுதாய மையங்கள் ஆகியவற்றின் தகவல்களும் சேகரிக்கப்பட்டு தரவுத் தளத்தில் சேமித்து வைக்கப் படவேண்டும். குடிமக்கள் சாசனம், குடிமக்கள் அரசாளுமையில் பங்குபெறும் வாய்ப்புகள் அனைத்தும் முழுக்க முழுக்கத் தமிழ் வயப்பட அரசு பரிந்துரைக்கவேண்டும்

இணைய வழியில் எண்ணிம நூலகங்கள், வலைத் தளங்கள் மற்றும் தமிழ்த் தரவுத் தளங்கள்

அரசு அலுவலகங்களில் உள்ள கணினி வலைத் தளங்கள் ஒருங்கிணைக்கப்பட்டுப் பல நிலைகளில் தமிழ் வடிவக் கோப்புகள் ஆணைகள் விண்ணப்பங்கள் உருவாக்கம் செய்யப்பட்டுப் பயன்பாட்டில் இருக்கத் தேவையான செயல்களை அரசு மேற்கொள்ள வேண்டும். விடுப்பு கோரல், நிதி கோரல், அரசு ஆவணங்கள் வேண்டுதல், சான்றுகள் அனைத்தும் தரப்படுத்தப்பட்டு ஆன்லைனில் நிரப்பி வழங்க ஏற்பாடு செய்ய வேண்டும். இணைய வெளியில் இயங்கும் பல்லாயிரக்கணக்கான இணையத் தளங்களும் வலைப்பூக்களும் சமூக வலைத்தளங்களும் ஒருங்குறியீட்டு வரிவடிவில் இயங்க வலியுறுத்த வேண்டும். மேலும், பேரகராதி வளர்ச்சிப் பணிக்கும் பேருதவியாக இருக்கும். ஆங்கிலப் பேரகராதியில் பல பொருள் தரக்கூடிய அடிப்படைச் சொற்கள் மொத்தம் 3000 எனக் கணக்கெடுத்து பட்டியலிட்டுக் காட்டியிருக்கிறார்கள். ஆனால் இதுவரையில் தமிழ் அகராதியில் அவ்வாறு பல பொருள் தரும் அடிப்படைச் சொற்கள் என்று எதுவும் பட்டியலிட்டுக் காட்டப்படவில்லை. இத்தகைய பணியெல்லாம் சிறப்பாக நடைபெற ஏற்கனவே அகராதியில் இருக்கும் சொற்களைத் தொகுத்து, திருத்தி விடுபட்ட சொற்களைச் சேர்த்துக் கணினியில் உள்ளீடு செய்து, ஒரு தரவுத் தளத்தை உருவாக்க வேண்டும். இவ்வாறு செய்தால் தமிழ்-ஆங்கிலம் போன்ற அகராதிகளை அமைக்க முடியும்

தரவுத் தளங்களில் (databases) தேடல் விரைவு படுத்தப்பட வேண்டுமெனில் தரவுகள் (data)அகர வரிசையில் வரிசையாக்கப்பட (sorting) வேண்டும். எழுத்துருக் குறியாக்க முறைகள் வரிசையாக்கத்துக்கு ஒத்துழைக்க

வேண்டும். யூனிகோடு குறியாக்க முறையிலேயே தமிழ் எழுத்துகள் அகர வரிசைப்படி அமைக்கப்படவில்லை. எனவே வரிசையாக்கத்துக்கான சிறந்த தீர்வு நெறிகள் (sorting algorithms) கண்டறியப்பட வேண்டும்.

ஆவணங்களை எண்ணிம வடிவத்தில் பாதுகாப்பது முதல் கட்டவேலை. இரண்டாவது கட்ட வேலை, ஆவணங்கள் பற்றிய விபரத் தரவுகளை (Meta data) இணையத்தில் இணைப்பது. கடைசியாகப் பதிவேற்றம் செய்வது. ஆவணப்படுத்திய அனைத்தையும் உடனடியாகப் பதிவேற்றம் செய்வதில்லை. பலருக்கும் இணைய நூலகம் என்பது புதிது.

தகவல் தளங்களின் பயன்கள்

தகவல் மூலகங்களான அகராதிகள், கலைக்களஞ்சியங்கள் சாராம்ச சொல்லடைவு ஆகியவைகள் சேமித்து வைக்கப்பட்டிருந்தால், **இலக்கண சொல் / இலக்கணப் பிழை திருத்தம்** (Spell check / grammar check) **செய்ய இயலும்**. **தகவல்கள்** ஒருங்குறியீட்டு முறையில் இருந்தால் தரவுத் தளங்களில் வரைசையாக்கமும் தேடலும் (Sorting and searching) எளிதாகும். அச்சிட்ட ஆவணத்தை பிட்மேப் ஃபைலாக ஸ்கேன் செய்து கணிப்பொறித் தகவலாக (digital information) மாற்றி சேமித்து வைத்துக்கொண்டால் அதிலிருந்து எந்த தகவலையும் தேடிப் பெற முடியும். மின் எழுதுகோலால் பலகைக் கணினியின் திரையில் எழுதப்பட்ட கையெழுத்து ஆவணங்களை கணிப்பொறித் தகவலாய் மாற்றி **கையெழுத்து அறிதல்** (Handwriting recognition) மூலம் ஆவணமாகச் சேமிக்க முடியும். கணிப்பொறி முன் அமர்ந்து டிக்டா ஃபோன் என்னும் கருவியில் பேசப் பேசக் கணிப்பொறி அதனை உரை ஆவணமாக மாற்றி சேமித்து வைத்துக் கொள்ளலாம். இதற்கு **குரல் அறிதல்** (Voice recognition) என்று பெயர். கணிப்பொறியில் உள்ள ஓர் ஆவணத்தைத் திறந்தால் அந்த ஆவணத்தில் உள்ள விவரங்களைக் கணிப் பொறியே படித்துக் காட்டும் இதற்கு **உரையைக் குரலாக்குதல்** (Text to voice) என்று பெயர். ஓர் அலுவலகத்தின் தகவல் மேலாண்மையும் தகவல் பரிமாற்றமும் தாள்கள் மூலமாக நடைபெறாமல் முற்றிலும் கணிப்பொறி வழியாகவே நடைபெறும் 'தாளில்லா அலுவலகம்' (paperless office) சாத்தியம். அதுபோலவே, அரசு அலுவலகங்களைக் கணிப்பொறிப் பிணையங்கள் (computer networks) மூலமாகப் பிணைத்து அனைத்து அரசு நடவடிக்கைகளையும் தகவல் பரிமாற்றங்களையும் கணிப்பொறி மூலமாக மேற்கொள்ள முடியும். அரசின் திட்டப் பணிகள் விரைவாக நிறைவேறும். அவற்றின் பலன்கள் நேரடியாக மக்களைப் போய்ச் சேரும். ஊழல் குறையும். இத்தகைய அரசாட்சியை 'மின்-அரசாண்மை' என்கிறோம். தமிழ்நாடு அரசுப் பள்ளிகளில் பிளஸ்-1, பிளஸ்-2 மாணவர்களுக்கான கணிப்பொறியியல் புத்தகங்கள் தமிழில் வெளியிடப்படுகின்றன.

நிறைவுரை

பள்ளிகள், கோயில்கள், சமூக நிறுவனங்கள், முக்கிய மனிதர்கள் பற்றிய தரவுகளை ஆவணப்படுத்துவதோடு, ஆவணப்படுத்திய விடயங்களை மக்களுக்கு பயன்படுத்தக் கொடுப்பதும், பயனாளர்கள் கேட்கும் கேள்விகளுக்கு பதிலனுப்புவதும் பயனுடையதாகும்.

TagSets for Tamil

Prof. V. Renuga Devi

*Chairperson & Head, School of Linguistics and Communication Sciences,
Department of Linguistics, Madurai Kamaraj University, Madurai – 625 021.*

Tamil is a classical language spoken by millions people all over the world, but it still lacks significant researchers in the area of Natural Language Processing.

Part-of-speech (POS) tagging is a technique for assigning each word a text with an appropriate parts of speech tag. The significance of parts-of- speech also known as POS, word classes, morphological classes or lexical tags, for language processing is the large amount of information they gave about a word and its neighbour.

POS tagging can be used in text to speech, information retrieval, shallow parsing, information extraction, linguistic research for corpora and (2) also as intermediate step for higher level NLP tasks such as parsing, semantics, translation and many more (3) POS tagging, thus, is a necessary application for advanced NLP application in Tamil or any other languages.

POS varies language to language i.e. interlingually and also within the language i.e. intralingually. Each word in every language belongs to a category. ‘naay’ is a noun; ‘ooTu’ is a verb, ‘azakaana’ is an adjective; ‘ai’ is a case marker, ‘meelee’ is a postposition and so on. A word such as ‘naay’ shows various properties with the word ‘maram’. The plural suffix –kaL can be attached to each of these words to form plural ‘naaykaL’ and ‘marankaL’. The suffix attaches to words are classified as nouns and produces plural nouns. There are exceptions- for example mass nouns ‘makkaL’ cannot be pluralized and the word ‘ooTu’ cannot be pluralized in this fashion. Thus, these morphological evidences exist for distinguishing nouns from words belonging to other categories.

Morphological evidence also exists that differentiates the other categories from one another. On the basis of the morphological evidences which are used to differentiate each word from other, English has eight parts of speech viz., Noun, Pronoun, Verb, Adjective, Adverb, Preposition, Conjunction and Interjection whereas in Tamil there are four parts of speech Noun, Verb, Uriccol and Itaiccol according to Traditional Grammar. There are five parts of speech viz., Noun, Verb, Adjective, Adverb and Particles according to Modern Theories. It also varies from five to ten.

Verbs in Tamil take the tense markers kiRu, kinRu in Present tense, t, T, R in Past tense and p, v in Future tense. A verb also inflects for Person, Number and Gender markers.

Adjectives can usually take the suffixes: -aana and some adjectives occur without any suffix

Azakaana - Alli oru azakaanap puu

Arumaiyaana - Avan arumaiyaana kaRcilaiyaic ceytaan

Nalla - Avan nalla paiyan

niRaiya - avan niRaiya puttakangaL paTittaan

Adverbs in Tamil inflect for the suffix –aaka.

Veekamaaka - Avan veekamaaka vaacittaan

Postpositions are invariants No affixes can be attached to them.

Meelee - Puunai meelee niRkinRatu.

Kiizee - Avan kiizee vizuntaan

Particles in Tamil are functors which are not inflected any affixes.

The question now arises that all these categories are found in all languages or just in a few languages. The answer is by no means simple. However, linguists generally assume that certain ‘major’ categories in particular nouns and verbs exist in most if not all languages.

The grammatical properties of a given parts-of-speech class are quite specific to a given language or small group of languages. For example, the property of nouns taking case markers, which defines Tamil nouns obviously cannot be used as a general defining property for nouns across languages.

Though it may be true that most, if not all languages share the categories noun and verb, it is also clear that other categories are found in some languages but not others. Tamil has a class of bound morphemes known as particles, which are attached to noun or verb phrases to indicate grammatical functions.

Whether or not all languages share part of speech categories, one can expect to find groups of words within any given language that share significant grammatical properties. The words sharing significant properties all belong to the same category. Such categories traditionally labeled as noun, verb, adjective, adverb and so on. But one must remain open to the possibility that a given language may have a grammatical category not found in other languages. The existence of parts-of-speech categories shows that the lexicon of a language is not simply a long, random list; rather, it is structured into special subgroup of words.

IIIT-tagset

Tagset is the set of tags from which the tagger is used to choose and attach tag to the relevant word. For achieving POS tagging, deciding and creation of tagset is very important. There are several POS tagsets for Tamil.

Sankaran Baskaran et al., 2008 proposed a Common Parts-of-Speech Tagset Framework for Indian Languages. They have identified partial 12 tagsets as universal categories for any tagset. The following tagsets are

1. [N] Nouns
2. [V] Verbs
3. [PR] Pronouns
4. [JJ] Adjectives
5. [RB] Adverbs
6. [PL] Participles
7. [PP] Postpositions
8. [DM] Demonstratives
9. [QT] Quantifiers
10. [RP] Particles
11. [PU] Punctuations
12. [RD] Residual

The tagsets also varies within a language. LDC-IL POS Tagset has 14 tags,

LDC-IL-TAMIL POS TAGSET		
SL No.	CATEGORY	TYPE
1	NOUN	Common (NC)
		Proper (NP)
		Verbal (NV)
2	PRONOUN	Pronominal (PPR)
		Reflexive (PRF)
		Reciprocal (PRC)
		Wh-pronoun (PWH)
3	DEMONSTRATIVE	Absolutive (DAB)
		Wh-demonstrative (DWH)
		Adjective (JJ)
4	NOMINAL MODIFIER	Quantifier (JQ)
		Intensifier (JINT)
5	VERB	
6	ADVERB	Manner (AMN)
		Location (ALC)
7	PARTICIPLE	Relative (LRL)
		Verbal (LV)
		Nominal (LN)
		Conditional (LC)
8	POSTPOSITION	
9	PARTICLES	Co-ordinating (CCD)
		Subordinating (CSB)
		Interjection (CIN)
		(Dis)Agreement (CAGR)
		Confirmative (CON)
		Delimiting (CDLIM)
		Others (CX)
10	NUMERAL	Real (NUMR)
		Serial (NUMS)
		Calendric (NUMC)
		Ordinal (NUMO)
11	REDUPLICATION	
12	RESIDUAL	Foreign Word (RDF)
		Symbol (RDS)
13	UNKNOWN	Unknown (UNK)
14	PUNCTUATION	Punctuation (PU)

Dhanalakshmi et al., developed an Amrita POS Tagset for Tamil consists of 32 tags,

Amrita POS tagset		
Sl. No.	POS Label	Description
1.	<NN>	NOUN
2.	<NNC>	COMPOUND NOUN
3.	<NNP>	PROPER NOUN
4.	<NNPC>	COMPOUND PROPER NOUN
5.	<ORD>	ORDINALS
6.	<CRD>	CARDINALS
7.	<PRP>	PRONOUN
8.	<PRIN>	PRONOUN INTROGATIVE
9.	<PRID>	PRONOUN INDEFINITE
10.	<ADJ>	ADJECTIVE
11.	<ADV>	ADVERB
12.	<VNAJ>	VERB NON FINITE ADJECTIVE
13.	<VNAV>	VERB NON FINITE ADVERB
14.	<VBG>	VERBAL GERUND
15.	<VF>	VERB FINITE
16.	<VAX>	VERB AUXILARY
17.	<VINT>	VERB INFINITE
18.	<CNJ>	CONJUNCTION
19.	<CVB>	CONDITIONAL VERB
20.	<QW>	QUESTION WORD
21.	<COM>	COMPLEMENTIZER
22.	<NNQ>	QUANTITY NOUN
23.	<QTF>	QUANTIFIERS
24.	<PPO>	POSTPOSITIONS
25.	<DET>	DETERMINERS
26.	<INT>	INTENSIFIER
27.	<ECH>	ECHO WORDS
28.	<EMP>	EMPHASIS
29.	<COMM>	COMMA
30.	<DOT>	DOT
31.	<QM>	QUESTION MARKS
32.	<RDW>	REDUPLICATION WORDS

Kongu Pos Tag-Set (more than 500 tags)

POS Forms and Morphological Inflections		
POS & Others	Forms	Morphological inflections
Noun	Simple Noun	Number
	Proper Noun	Singular
	Participle Noun	Plural
	Adjective Noun	Gender
	Positive Tensed Verbal	Male
	Noun	Female
	Negative Tensed Verbal	Neutral
	Noun	Common
	Un-tensed Verbal Noun	Oblique
	Simple Verb	
Verb	Transitive Verb	Person
	Intransitive Verb	First
	Causative Verb	Second
	Infinitive Verb	Third
	Imperative Verb	Number
	Reportive Verb	Singular
		Plural
		Gender
		Male
		Female
		Neutral
		Common

Kongu POS Tag set has more than 500 tags and there are other tag sets also.

		Tense
		Present
		Past
		Future
		Passive
		Honorific
		Negative
		Interrogative
		Suffix
Adverb	Simple Adverb	Tense
Adjective	Simple Adjective	Present
	Participle Adjective	Past
		Future
		Negative
Preposition	Simple Preposition	Cases
	Noun + cases	Accusative
		Dative
		Instrumental
		Sociative
		Locative
		Ablative
		Benefactive
		Genetive
		Vocative
		Clitics
		Selective

		Negative
Conjunction	Simple Conjunction	Wh-words
	Coordinating	What
	conjunction	Who
		Whose
		When
Participle		
		Where
		Whom
		Which
		How
	Verbal Participle	Positive
	Conditional participle	Negative
Interjection	Simple Interjection	
Others	Echo words	Same
		Different
	Determiner	
	Quantifier	
	Complementizer	
	Ordinal	
	Optative	

IL - ILMT Tagset
Noun NN
Pronouns PRP
Locative Nouns NST
Verbs Finite VBF
Verbs Auxiliary VAUX
Adjectives JJ
Conjunctions CC
Interjections INJ
Question Words QW
Echo Words ECH
Particles RP
Numbers Ordinals QO
Question Words QW
Symbols SYM
Noun NN
Pronouns PRP
Locative Nouns NST
Verbs Finite VBF
Verbs Auxiliary VAUX
Adjectives JJ
Conjunctions CC
Interjections INJ
Question Words QW
Echo Words ECH
Particles RP
Numbers Ordinals QO
Question Words QW
Symbols SYM

TABLE I.

LIST OF TAGS FOR TAMIL AND THEIR DESCRIPTION

Tag Description

1. N → Noun
2. NP → Noun Phrase
3. NN → Noun + noun
4. NNP → Noun + Noun Phrase
5. IN → Interrogative noun
6. INP → Interrogative noun phrase
7. PN → Pronominal Noun
8. PNP → Pronominal noun
9. VN → Verbal Noun
10. VNP → Verbal Noun Phrase
11. Pn → Pronoun
12. PnP → Pronoun Phrase
13. Nn → Nominal noun
14. NnP → Nominal noun Phrase
15. V → Verb
16. VP → Verbal phrase
17. Vinf → Verb Infinitive
18. Vvp → Verb verbal participle
19. Vrp → Verbal Relative participle
20. AV → Auxiliary verb
21. FV → Finite Verb
22. NFV → Negative Finite Verb
23. Adv → Adverb
24. SP → Sub-ordinate clause conjunction Phrase
25. SCC → Sub-ordinate clause conjunction
26. Par → Particle
27. Adj → Adjective
28. Iadj → Interrogative adjective
29. Dadj → Demonstrative adjective
30. Inter → Intersection
31. Int → Intensifier
32. CNum → Character number
33. Num → Number
34. DT → Date time
35. PO → Post position

A Comparison of various tagsets:

Category	LDC-IL	Amrita	Kongu	IL-ILMT	IIIT-tagset	Description
Noun	Common, Proper, Verbal	Compound, Proper, Compound Proper, Quantity	Simple, proper, Participle, Adjective Noun, Positive Tensed, Verbal, Negative Tensed Verbal, Un-tensed VerbalNoun	Locative Nouns	Nouns	Noun Phrase, Noun +Noun, Noun + Noun Phrase, Interrogative Noun, Interrogative Noun Phrase, Pronominal Noun, Verbal Noun, Verbal Noun Phrase Nominal Noun, Nominal Noun Phrase
Pronoun	Pronominal, Reflexive, Reciprocal WH-Pronoun, Absolutive	Pronoun, Pronoun Interrogative, e, Pronoun Indefinite	---	Pronouns	Pronoun	Pronoun, Pronoun Phrase
Demonstrative	Wh- demonstrative, Adjective	Adjective	----	---	Demonstrative	Adjectives
Nominal modifiers	Quantifiers, Intensifier	Quantifiers, Intensifiers, Determiners Complementizer	---	---	---	---
Verb	<i>Adverb:</i> Manner Location <i>Participle:</i> Relative Verbal	Non finite Adjective, Non finite Adverb, Verbal Gerund, Verb finite,	---	Verb finite, Verbs Auxiliary	Verb	Verb, verbal phrase, verb infinitive, Verb verbal participle, Verbal relative participle, auxiliary verb, Finite verb

	Nominal	Verb				
	Conditional	auxiliary, Verb infinite, Conditional Verb				
Postposition	Postposition	Postposition	Postposition	Postposition	Postpositi on	Postposition
Particles	Coordinating, Subordinating, Interjection, Agreement, Confirmative, Delimiting, Others	---	---	Particles	Particles	Particles
Numeral	Real, Serial, Calendric, Ordinal	Ordinals, Cardinals	---	Numbers Ordinals	---	Number, Character number, Date, time
Adverb		Adverb	Simple Adverb	Adverb	Adverb	Adverb
Adjective		Adjective	Simple Adjective, Participle Adjective	Adjective	Adjective	Adjective, Interrogative adjective, Demonstrative adjective
Reduplication	Reduplication	---	---	---	---	---
Residual	Foreign Word, Symbol	---	---	---	Residual	---
Unknown	Unknown	---	---	---	---	---
Punctuation	Punctuation	Question marks, Comma, Dot	---	---	Punctuatio n	Punctuation

Conjunction	--	Conjunction	Simple Conjunction, Coordinating conjunction	Conjunction	--	Subordinate clause, subordinate clause conjunction phrase
Participle	--	---	Verbal, Conditional	---	---	---
Interjection	---	---	Simple Interjection	Interjection	--	Interjection
Others	---	---	Echo words, Determiner, Quantifier, Complementizer, Ordinal, Optative	---	---	---
---	---	---	---	---	Quantifier s	---
---	---	---	---	---	---	Intensifiers
---	---	Echo words	---	Echo words	---	---
---	---	Emphasis	---	Symbols	---	---
---	---	Question word	---	Question word	---	---

Though these tagsets are useful to categorize words very minutely so as the machine can recognize and set the tag of the particular word. There is a need to bring uniformity among the tag sets. Single software is standardized for English which helps the user all over the world to work easily. But in Tamil various types of fonts are available. Different schools use different type of fonts which make the retrieval of the message hard. To avoid this now Unicode is introduced for Tamil. That makes more people to work with the computers and retrieval of information is made easy. Likewise the various tagsets used by different people gives trouble to the user. Bringing uniformity among the tagsets make many people to work in NLP. Uniform tagset make the machine translation easy and the performance level of that tagsets should be high.

Tamil Verb Conjugation Using Data Driven Approach

Dr.V.Dhanalakshmi
HOD, Department of Tamil
SRM University
Kattankulathur-603 203

Dr.M.Anand Kumar
Assistant Professor
CEN, Amrita University
Coimbatore- 641 112

Dr.K.P.Soman
HEAD, CEN
Amrita University
Coimbatore- 641 112

ABSTRACT:

This Paper presents a novel methodology adopted for developing Tamil Verb Conjugation system. The Verb Conjugation system implemented here is a new data driven approach which is simple, efficient and does not require any rules and morpheme dictionary. There are two options available in this verb conjugator. As a first option, this tool will produce the intended word form corresponding to the user's input. The second option available will generate all the word forms of a particular word. If the user gives the root word or lemma as an input this tool it will automatically generate all the possible word forms (10, 000 verb forms for a single root). This tool is also used for Sentence Generator. So this can be very useful for teaching Tamil verb forms which is highly rich, agglutinative nature.

KEY WORDS:

Verb Conjugation, Tamil Morphology, Agglutinative Language, Word-forms, Morpho-lexical information, Data Driven Approach

INTRODUCTION:

Conjugation is the regular arrangement of the forms of the verb that shares a similar conjugation pattern. Verb conjugator conjugates a verb root and yields all the word forms of that particular verb root. Tamil is a momentous classical Dravidian language. It is a morphologically rich language with agglutinative and disparate morphological structure. Tamil verbs are highly inflected for tense, person, number, gender, mood and voice. Even though the morphotactic order of morphemes are fixed there is some flexibility to allow the insertion of certain morphemes like A, E, O, thAn etc. This nature of Tamil verb increases the number of verb forms for each verb. Our proposed data driven approach based Tamil verb conjugation system could conjugate finite, infinitive, adjectival, adverbial and conditional forms of verbs along with auxiliaries and clitics which could yield more than 10, 000 forms for each verb.

A few attempts have been made to develop Tamil conjugation system. VERBIX¹ on-line verb conjugator contains verb conjugations for hundreds of languages, ranging from national and international languages (including Tamil) to regional and even extinct languages. Tamil verb pattern has been studied and conjugated by CREA². Dr.V.S.Rajam has proposed Conjugation by Verb-Stem Ending³.

The Verb Conjugation system implemented here is a data driven approach which is simple, efficient and does not require any rules and morpheme dictionary. The system functions efficiently on the novel algorithm which is implemented using java program. Morphological Generator have already developed using this novel algorithm [Anand Kumar M e.tal.,] . Morphological Generator is not an end-user tool since the user should know how to feed the morphological information. But Verb Conjugation system is an end-user tool which could give all the verb forms with its Morpho-lexical information. Three different modules are developed to build this system. The first module takes the verb root as input and gives the verb's paradigm number. The second module gives the verb's stem (Stemming process) as output. The second module also contains

¹ <http://www.verbix.com>

² <http://www.crea.in/verb-table-downloads/Tamil-Verb-Pattern.pdf>

³ <http://www.crea.in/verb-table-downloads/Tamil-Verb-Pattern.pdf>

all the morpho-lexical information. In third module a suffix-table is used to conjugate the verb with all the morpho-lexical information. This tool is also used for generating simple sentences which would be useful for pedagogic purpose.

ARCHITECTURE OF THE SYSTEM:

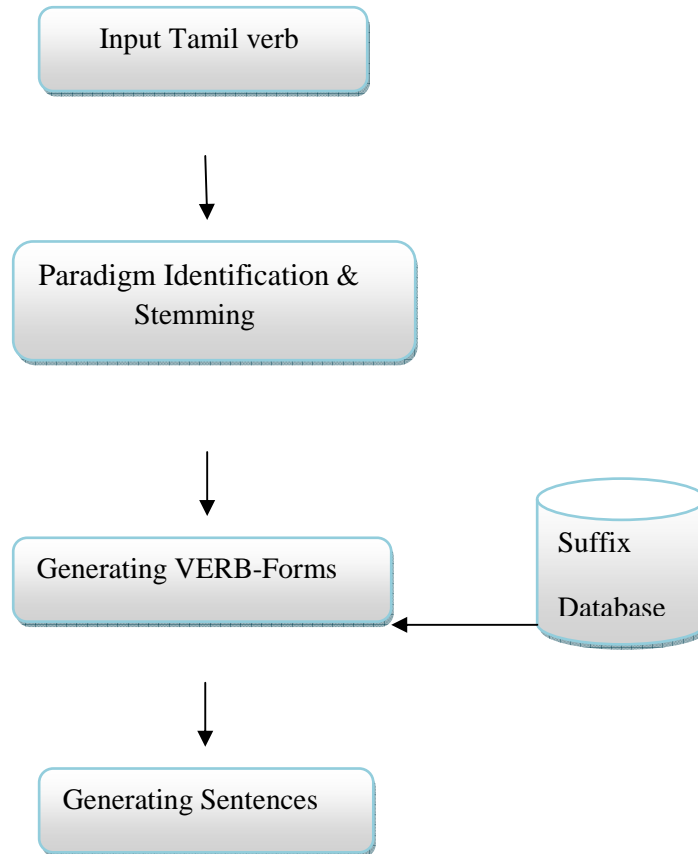


Fig:1 Architecture of the System

IMPLEMENTATION:

Tamil verb conjugation system is implemented by building three modules. The first module takes the verb root as input and gives the verb's paradigm number (**Paradigm Identification**). The second module gives the verb's stem (Stemming) as output.. The second module also contains all the morpho-lexical information. In third module a suffix-table is used to conjugate the verb with all the morpho-lexical information.

Paradigm Identification: The input verb root word is romanized using Tamil Unicode to roman mapping file. This romanized form is compared with end suffixes in paradigm classification file. If an end suffix is matched with the end characters of the root word then the paradigm number is identified. End suffixes are created based on the paradigms and sorted according to their character length. The algorithm for paradigm classification is given below.

```
Root word is Romanized
For all End Suffix
If End Suffix is matched with root word
Then, Paradigm number is identified
End if
End for
```

Stemming : Stemming is the process for reducing a words to its stem, base or root form. The stem need not be identical to the morphological root of the word. So for stemming we have identified certain characters for each verb paradigm, based on this our system would reduce the root verb to its stem form. For Example: the root word “pAdu” falls in fourth paradigm, in stemming process the suffix “du” will be deleted and the stem “ pA” will be restored to join with the other word forms in the suffix table. In some cases no deletion would take place so that the root would be joining with the other morpho-lexical forms. For example: the root word “padi” falls in the first paradigm, here no stemming is done, so the root word “padi” would join with the other morpho-lexical forms present in the suffix table.

Suffix table creation: The **suffix table** is the most essential resource for this verb conjugation algorithm. It is a simple two dimensional (2D) table where row corresponds to the morpho-lexical information and column corresponds to the paradigm number. Verb suffix table has two suffix tables, first table is made without auxiliaries and the next is designed with auxiliary forms. First table contains 164 rows and the next has 67 rows. Both the tables contain 32 columns (total number of paradigms). Table1: shows the number of paradigms and number of conjugation for a single verb. The verb stem is conjugated with the all forms present in the suffix table. Table:2 shows the model of the suffix table.

VERB	Paradigms	Simple forms	Inflection	Auxiliary forms	Total
	32	164		67	10988

Table1: Paradigms & Inflections

	P-1	P-2	P-3	P-4	P-5	
MLI-1	ththAn	wAn	inAn	thAn	Ran
MLI-2	ththAL	wAL	inAL	thAL	RAL
MLI-3	ththAr	wAr	inAr	thAr	RAr
MLI-4	kinRAn	kinRAn	kinRAn	kinRAn	kinRAn
MLI-5					

Table:2: Suffix Table

Conclusion:

An user friendly GUI is created to operate the verb conjugation system. Separate columns are created to show the simple finite verb forms, participle forms, primary auxiliary forms, secondary auxiliary forms, etc. The verb conjugation system for simple finite verb forms would conjugate along with the sentence. For example the verb “pAdu” would conjugate as “pAdukiRen” (“wAn pAdukiRen) for the morphological form “1S” (1st person singular). This is sentence creation is created for all the verb simple forms. As it is used for Sentence Generator, this system could be very useful for teaching Tamil verb forms to the students in a non-conceptual method.



Figure :2 : GUI of the system

References:

- Anand Kumar M, Dhanalakshmi V, Rekha R U, Soman K.P, Rajendran S, “A Novel Data Driven Algorithm for Tamil Morphological Generator” International Journal of Computer Applications (0975 – 8887) Volume 6– No.12, September 2010.
- Annamalai, E. 1985, “Dynamics of verbal extension”, Trivandrum: Dravidian Linguistics Association.
- Annamalai E., (Chief Editor), 2009, “A Handbook of Tamil Verbal Conjugation”, McNeil Technologies, Inc.
- Rello, Luz, and Eduardo Basterrechea. "Onoma: a linguistically motivated conjugation system for spanish verbs", Computational Linguistics and Intelligent Text Processing. Springer, Berlin, Heidelberg, 2011. 227-237.
- Rajendran S, Viswanathan S, Ramesh Kumar, 2003. “Computational morphology of verbal complex” Language in India: Volume 3 : 4 April 2003.

Programming Like Language In Tamil

Yasasvi Sridharan

s.yasasvi@gmail.com

ABSTRACT

Most of the programming language use English as the medium for instruction. Understanding the concepts in English is difficult for rural people, so to gain interest towards programming the programming must be taught to them in their native language.

Any concept when taught in their native language is easy for a person to understand than that taught in any other language which is known to him.

This new programming like language - medium in Tamil, is an Open-Source, available to everyone and it will create interest in young minds towards programming. This Tamil like programming language serves as a teaching guide for students to learn the concepts of Programming Language with hands on session and for better understanding the bilingual keyword list is provided. This project called Pachondhi (pacch) has been created using java-script.

Keywords: programming like language, javascript, jquery, html, css, console environment

1. INTRODUCTION

Like Mathematics, computing is a concept, and can be introduced through any native language. Our motive is to create programming language which is easier for everyone. Once they are strong with their basic concepts and logical thinking they can implement their own ideas in programming.

Generally every person first starts to learn new things in their native language only. So if things are provided in the native language they could learn it fast.

This project is designed like a dynamically typed (i.e. Declaration free) procedural language. And also that this project can act as a teaching guide (i.e. tutorial) to help people understand the flow of a programming language.

The entire project revolves around the concept of the control flow of any procedural language, so it is best to use scripting language. The project is based on java-script to describe the flow of any procedural language.. The details about how to use java-script is given in this project to act as a teaching module. For teaching purpose, we use the concept of java-script.

This project is bilingual i.e. we can use both Tamil and English code so as to make computing easy. This way we can easily identify how code in one language is associated with that of other language. The variety of codes dealing in these two languages (i.e. English and Tamil) can be implemented either way and helps to understand the concepts.

In this paper we will discuss the Tamil keywords used, their relationship with that of English keywords in javascript, and how it is implemented in the project.

2. PREVIOUS ATTEMPTS ON TAMIL PROGRAMING LANGUAGE

SWARAM & EZHIL

Many solutions to use Tamil as programming language have been suggested earlier.

A static-typed Tamil-language system called Swaram was introduced in 2003. A static-typed Tamil-language system Ezhil, introduced in 2008. Swaram is a full-fledged static-typed programming language, with a feature set resembling C-programming language. Swaram has inspired some of the choice of keywords in the Ezhil language. To its credit Swaram is the first programming language in Tamil, in the true sense with a JIT compiler from source and a virtual machine (VM). In reporting Swaram, the authors justify the need for a complete language rather than plain pre-processors, and other syntactic sugar. At time of writing, Swaram is not publicly available, which severely limits language development, system use, community support and improvement. From experience of Python one generally thinks that interpreted languages are easier for students to pick up and enjoy programming with.

Swaram is strongly typed, and allows mixed English & Tamil identifiers. This useful feature has been incorporated to provide some form/type of access to allow external libraries which are (invariably) written in English. Now this will mean the language has to accept identifiers & functions / variables in a mix of English & Tamil. Keywords can still be Tamil-only.

Swaram is the more complete of the previous attempts at providing a Tamil programming language. Ezhil system borrows some of the better ideas freely from the Swaram system.

The Ezhil system incorporates most of the concepts from Swaram. In the initial stage both Swaram and Ezhil are not freely available. But at a later stage Ezhil became like opensource.

3. Pachondhi KEYWORDS AND ITS MAPPINGS

Some of the keywords used in this project are listed below. In order to make the keyword list more understandable, their corresponding English keyword which is used in javascript are also mapped here.

நிரல்பாகம்	function
தொடர்ச்சி	continue
மறுபடியும்	repeat
ஆனால்	if
இல்லை	else
நிறுத்து	break
தேர்ந்தெடு	select
ஆக	for
பின்கொடு	return

தேர்வு	Case
செய்	Do
ஏதேனில்	While
பதிப்பி	Print
முடியேனில்	Until
முடிவு	End
காட்டு	Show
கேள்	Prompt
எண்	Number

The keyword mapping in this project is made unique because there are no specific global keywords for Tamil language. The keyword specified here are just samples. There are more keywords specified in the teaching module.

4. THE PROGRAMMING LANGUAGE SYSTEM

The programming language system used in this project is quite simple. The keywords and statements are closely chosen to represent the computer programs, the same chain of reasoning and logic followed in Tamil language. The conditional statements like *IF – ELSEIF – ELSEIF - ... - ELSE* can be represented in a similar way of replacing IF and ELSE with Tamil keywords i.e. the statement would look like:

ஆனால் - இல்லை ஆனால் - இல்லை ஆனால் - ... - இல்லை

The normal syntax of javascript is being used which includes the use of { } for a block, the use of [] for arrays and all the similar syntax is being followed.

The program will look like this:

Tamil code with Tamil keywords	Tamil Code with English keywords
நிரல்பாகம் பெருக்கல் (அ, ஆ({ காட்டு (அ*ஆ(} பெருக்கல் (2, 41);	function பெருக்கல் (அ, ஆ({ show (அ*ஆ(} பெருக்கல் (2, 41);

Example: Using loops

நிரல்பாகம் கூட்டு () { இ=0; ஆக (ஈ=1;ஈ<=10;ஈ++) { இ=இ+ஈ; } காட்டு (இ); } கூட்டு ();	functionகூட்டு () { இ=0; for (ஈ=1;ஈ<=10;ஈ++) { இ=இ+ஈ; } show (இ); } கூட்டு ();
---	---

5. System Implementation

This project is implemented in an object oriented fashion in the web page. For the web page the use of html and java-script are of the foremost importance. The content to be available in Tamil needs Unicode keyboard.

Systems with windows 7 or higher has Tamil keyboard in Unicode format in the language bar. But for systems with different configuration, a virtual keyboard is a must for this project to work.

In order to create a virtual keyboard we use the concept of JQuery which helps in mapping every English alphabet to that of the Tamil alphabet. In Tamil we also have to keep in mind the combination of alphabet resulting in a new alphabet. For the purpose of typing in Tamil we use the Unicode characters in the project. The Unicode format which we use is UTF-8. We can create the virtual keyboard by assigning a unique value for nearly all the keys of the keyboard by which we can map Tamil letters.

The 'Esc' key will be used to toggle between Languages (Tamil and English).

This project contains a console environment where in a person can type a code and check for its flow and correctness. In order to create this console environment the use of a light weight java-script library like *mochikit* is used. This mocha kit acts like a framework and helps in creating a java-script console environment. In the right side of the console environment we type the code and in the left side we get the result.

The future work is to implement the same concept for all the available programming languages. So that this teaching site would enable in creating interest for students in programming.

6. Conclusion

The procedural language whose medium is Tamil, will be a stepping stone that creates an interest for young minds towards programming. Tamil as a medium of instruction would help in developing the programming knowledge for rural people. (This can also be used for helping the students to understand the programming concepts in Tamil.)

7. REFERENCES

- [¹] David Flannagan, (1997), JavaScript: The Definitive Guide, Second Edition,
- [2] Eloquent JavaScript <http://eloquentjavascript.net>
- [3] Laurel Peterson, (1999) Principles of Compiler design, Pearson Publication.
- [4] Malcolm McLean, 'MiniBasic - how to write a script interpreter', online ebook.
- [5] Mutthaiya Annamalai, (2008) Ezhil Project: <http://students.uta.edu/mx/mxa6471/download.html>.
- [6] S.G.Ganesh, G.R.Prakash, K.K.Ravikumar (2003), Proceedings of Tamil Internet Conference, an overview of 'Swaram', a programming language in Tamil, .
- [7] Unicode 5.0 Tamil code chart: <http://www.unicode.org/charts/PDF/U0B80.pdf>

தமிழை அறிய கணினிக்கு எத்தனை விதிகள் வேண்டும்?

வாசு அரங்கநாதன்

பென்சில்வேனியாப் பல்கலைக்கழகம், vasur@sas.upenn.edu

இக்கட்டுரையில் கணினிக்குத் தமிழைப் புரிந்துகொள்ளச் செய்வதற்கான ஒரு எளிய வழிமுறை பற்றி விளக்கப்படுகிறது. எந்த மொழியும் தன்னகத்தே பல சிக்கலான விதிகளைக் கொண்டுள்ளதாக இருக்கும் என்பது உண்மை. ஆனால் மொழியின் கட்டமைப்பு என்பது ஒரு வரையறைக்குட்பட்டதாகத் தான் இருக்கும் என்றும் இவ்வரையறையைக் கொண்டு மொழி போகவோர் மிகவும் சிக்கலாக மொழியைப் பயன்படுத்தும் வழியறிவர் என்பது பொதுவான உண்மை. இல்லையெனில் யார் வேண்டுமானாலும் எப்படி வேண்டுமானாலும் மொழியை அமைக்கலாம் மற்றும் பேசலாம் என்ற சூழல் வரலாம். இந்த ஒரு குறிப்பிட்ட வரையறையை மனதில் கொண்டு ஒரு புதிய நோக்கில் புரோலாக் என்னும் கணிப்பொறி மொழியின் கட்டமைப்பு விதிகளைப் பயன்படுத்தி தர்க்கவிதிகளின் (logic rules) கோட்பாட்டின் அடிப்படையில் கணினிக்குத் தமிழின் அறிவைப் புகட்டும் ஒரு முயற்சி இது. இக்கட்டுரையில் கொடுக்கப்பட்டிருக்கும் எடுத்துக்காட்டுகளை பயன்படுத்திப் பார்க்க <http://www.thetamilnlanguage.com/tamilnlp/tagit.html> என்னும் வலைப்பக்கம் உருவாக்கப்பட்டிருக்கிறது.

இயற்கை மொழி என்னும் போது நமக்கு மனதில் வருவன ஒலியன் மற்றும் உருபன் விதிகள், தொடரியல் விதிகள், பொருளியல் விதிகள் மற்றும் உலகியல் கோட்பாட்டு விதிகள் என்பனவாகும். இத்தோடு இலக்கியவிதிகள் பண்பாட்டு விதிகள் அரசியல் மற்றும் மதம் சார்புள்ள விதிகள் இப்படி பலவிதிகளின் சிக்கலான பின்னலில்தான் மனித சமூகம் மொழியைப் பயன்படுத்தி இயங்குகிறது. இவ்வகையில் இவ்வனைத்து விதிகளையும் மொழி வழியாகத்தான் நாம் செயல்படுத்துகிறோம். அத்தகைய சிக்கலான மொழியைக் கணினிக்கு எங்கனம் கொடுக்க இயலும் என்னும் வினா எழலாம். இங்கு நாம் ஒரு உண்மையை உற்று நோக்கவேண்டியிருக்கிறது. நாம் மேற்சொன்னவாறு மொழி என்பது ஒரு குறிப்பிட்ட வரையறைக்குள் அமைக்கப்பட்ட ஒரு கட்டமைப்புதான் என்பது பற்றி எண்ணிப்பார்க்கவேண்டும். அதாவது மொழியைப் பொறுத்தவரையில் அவற்றின் விதிகள் ஒரு குறிப்பிட்ட எண்ணிக்கைக்குள்ள்தான் இருக்கும். அவ்வரையறையை நாம் பலவாறு பல கோணங்களில் பல சிக்கல்களோடு பயன்படுத்தும் போது அது நமக்கு மலைப்பாகத் தோன்றும். அந்த விதிகளின் ஒரு குறிப்பிட்ட வரையறையில் அவ்விதிகளைக் கணினியில் உள்ளிடும் தருணத்தில் நாம் எப்படி மொழியைப் பயன்படுத்துகிறோமோ அதே போன்ற ஒரு சூழலைக் கணினிக்குக் கொடுக்கும் வாய்ப்பை ஏற்படுத்தலாம் என்பதே இக்கட்டுரையின் உட்கருத்து. இங்கு சோம்சுக்கி என்னும் மொழியியலாளர் கூறிய கருத்தை நோக்கலாம். ஒரு குறிப்பிட்ட எண்ணிக்கையிலான விதிகளைக் கொண்டு நாம் பல சிக்கலான வாக்கிய அமைப்புகளை ஏற்படுத்திக்கொண்டிருக்கிறோம் என்று அவர் தனது மொழியியல் கோட்பாட்டு நூலில் விளக்கியுள்ளார் (Chomsky 1965).

இவ்வகை விதிகளை முறைப்படி கணினியில் உள்ளிட்டு அவை கொண்டு மனிதன் மொழியைப் பயன்படுத்துவது போன்ற ஒரு சூழலை கணினிக்குக் கொடுத்தல் என்பது செயற்கை நுண்திறன் ஆராய்ச்சிகளில் ஒன்றாகும். இவ்வழிநிரலில் தமிழ் உருபன் மற்றும் சொற்றொடர்களின் தொடரியல் பிரிப்புப் பகுதிகள் அமைக்கப்பட்டுத் தமிழ் மொழி வாக்கியங்களைக் கணினியில் ஆய்ந்தறியும் வழியில் ஒரு தொடரான கட்டமைப்பு வடிவத்துக்கு (list form) மாற்றும் ஒரு முறையில் விதிகள் உட்படுத்தப்பட்டிருக்கின்றன. இவ்வகைக் கட்டமைப்புகளைக் கூட்டமைப்பு விதிகளின் (set theory) அடிப்படையில் கணினி தமிழை அறிந்துகொள்ளும் ஒரு புது உத்தியை விளக்குகிறது இக்கட்டுரை. கிட்டத்தட்ட பத்தாயிரம் சொற்களைக் கொண்ட ஒரு மின்னகராதியும் இவ்வழிநிரலில் உட்படுத்தப்பட்டிருக்கிறது.

தமிழ்ச் சொற்களைப் பற்றி அறியும் உருபன் விதிகள்

இக்கணினி வழிநிரலில் தமிழ்ச் சொற்களைக் கணினி வழியில் புரிந்துகொள்ளும் ஒரு வழிமுறை முதற்கட்டமாக அமைக்கப்பட்டுள்ளது. நாம் "நீங்கள் சென்னைக்குப் போனீர்களா?" என்ற ஒரு கேள்வி வாக்கியத்தைப்

பயன்படுத்தும் போது நமக்கு இச்சொற்களில் உள்ள விகுதிகள், காலம் முதலியன பற்றி தெரிகிறது. இவற்றில் உள்ள விகுதிகளை பிரித்து நாம் சிந்திப்பதில்லை. ஆனால் மொழிக் கூறுகளை ஆய்ந்து நோக்கும் தருணத்தில் இத்தகைய வாக்கியங்களுக்கான விகுதிகள் மற்றும் பல விளக்கங்களைத் தனித்தனி இழையாகப் பிரித்தறிய வேண்டிய அவசியம் ஏற்படுகிறது. அவ்விழைகளைக் கீழ்க்கண்டவாறு விளக்கலாம்.

முதற்கட்ட இழை

[[["nom", "ndii", "noun", "pl"], ["dat", "cennai", "noun"], ["pa", "poo", "2pl", "inte"], ["q", "?"], ["nom", ".", "period"]]]

அதாவது முதற்கட்ட இழையில் எல்லாச் சொற்களையும் பிரித்து வேற்சொற்களைக் கண்டுபிடித்து ஒவ்வொரு விகுதிக்கும் ஒரு பெயர் கொடுத்து அதை எந்த வரிசைக்கிரமத்தில் கேட்டோமோ அதே வரிசைக்கிரமத்தில் கொடுக்கவேண்டும். இந்த இழையை விகுதிகளின் பகுத்தறிதல் (morphological tagger) என்று கூறுவார்கள். விகுதிகளுக்கு நாம் கொடுக்கும் பெயரை விகுதிக் கூறுகள் (tag sets) என்றும் கூறுவார்கள். இது போன்ற ஆராய்ச்சிகள் தமிழில் நடந்த வண்ணம் இருக்கின்றன. ஆனால் தமிழ்ச் சொற்களை நாம் எப்படி புரிந்துகொள்கிறோமோ அதே போல் கணினியும் புரிந்து கொள்ளும் வகையில் ஒரு அமைப்பை யாரும் ஏற்படுத்தியதில்லை. இந்த வகையில் இந்த ஆராய்ச்சி அதற்கு ஒரு முன்முயற்சி எனக்கொள்ளலாம். எடுத்துக்காட்டாக ["pa", "poo", "2pl", "inte"] என்னும் சொற்பிரிவை நோக்கினால் ஒன்று விளங்கும். இது "போ" என்னும் இறந்தகால வினை. இத்தோடு முன்னிலைப் படர்க்கை விகுதியும் வினாவுக்கான விகுதியும் இணைக்கப்பட்டிருக்கின்றன என்பது தெரியும். இப்படித்தான் நாம் "போனார்களா" என்னும் சொல்லைப் புரிந்துகொள்கிறோம். இந்த இழையில் உள்ள விதிகள் அனைத்தையும் புரோலாக் என்னும் கணினி மொழியின் கட்டமைப்பு அடிப்படையில் பகுக்கும் முறையில் (list processing) கணினிக்குக் கொடுத்திருக்கிறோம். இந்த வகையில் மேற்படி வாக்கியத்தை நாம் உள்ளீடு செய்யும் போது நாம் எப்படி புரிந்துகொள்கிறோமோ அதே வழியில் கணினியும் புரிந்துகொள்ள ஒரு வழியை நாம் இங்கு ஏற்படுத்துகிறோம். இந்தக் கூட்டமைப்பு வடிவத்தின் வசதி என்னவென்றால் கணினி வழிநிரலில் சொற்களைப்பற்றி மிக எளிதாக மொழியைப் பகுத்தறிய வழிமுறைகளைச் செய்யலாம். எல்லாக் கடந்தகால வாக்கியங்களைக் கொடு என்றோ "போ" என்னும் வினை கொண்ட வாக்கியங்களைக் கொடு என்றோ கேட்டால் இம்முறையில் சேமிக்கப்பட்ட தரவிலிருந்து இத்தகைய செய்திகளை எளிதில் கண்டுபிடிக்கலாம். இம்முறையின் மற்றொரு வசதி இக்கட்டுரையில் விளக்கப்படும் செய்திக்கொணரல் முறையாகும். மேலும் ஒரு சொல்லிலிருந்து உருபுகளை முறையாகப் பிரித்து அவற்றை மேற்படி காட்டிய அமைப்பில் மாற்ற Lexical Phonology என்னும் கோட்பாட்டை வெகுவாக இவ்வழிநிரல் பயன்படுத்துகிறது (காண்க Renganathan 1997).

இரு வழி சொல் பகுக்கும் விதிகள்

இவ்வழிநிரலில் "பார்த்துக்கொண்டேயிருக்கிறதனால்" என்னும் சொல்லை உள்ளீடு செய்யும் போது அச்சொல்லை ["pr_prog", "paar", "neut.sg", "causal"] எனப் பகுத்தறிகிறது. அதே நேரத்தில் இவ்வகை அமைப்புகளைக் கொண்டு அதே சொற்களை மீண்டும் உருவாக்கவும் ஏற்பாடுகள் செய்யப்பட்டுள்ளன. இந்தவகையில் தமிழ்ச்சொல்லைப் பிரிக்கவும் பிரித்ததை மீண்டும் சொல்லாக உருவாக்கவும் இவ்வழிநிரலில் விதிகள் கொடுக்கப்பட்டுள்ளன. முக்கியமாக, மேற்படி சொல்லை "சொல்லிக்கொண்டிருக்காததால்" என் மாற்றவேண்டுமானால் மேற்படி கட்டமைப்பு வடிவில் "neg_causal" என்னும் சொல்லையும் "col" என "சொல்" என்ற வினைக்கும் ["pr_prog", "col", "neut.sg", "neg_causal"] என இணைக்கவேண்டும். அதே போல் "சொல்" என்னும் வினைக்குப் பதிலாக "கொடு" என்னும் வினையை மாற்றினால் "கொடுத்துக்கொண்டிருக்காததனால்" என்னும் வடிவைக் கொடுக்கும். இவ்வகையில் இந்த வடிவம் கணினி தமிழை எளிதில் பகுத்தறிய நாம் அமைக்கும் ஒரு இடைநிலை வடிவம் ஆகும்.

வாக்கியங்களைப் பற்றி அறிய உதவும் இலக்கண விதிகள்

இரண்டாவது கட்டமாக வாக்கியத்தில் உள்ள சொற்றொடர்களைப் பிரித்தறியவேண்டும். அதாவது பெயர்த்தொடர் எது? வினைத்தொடர் எது? என்பதைப் பிரிக்கவேண்டும். பெயர்த்தொடரில் பெயரடைகளையும்

உடைமைத் தொடர்களையும் உட்படுத்த வேண்டும். வினைத்தொடரில் வினையடையை இணைக்கவேண்டும். இந்த அமைப்பையும் புரோலாக் மொழியின் கட்டமைப்பு விதிகளுக்குள் அமைக்கவேண்டும். இதை இங்கு இரண்டாவது இழையாகக் கணினிக்கு கொடுத்துள்ளோம். கீழ்க்காணும் இரண்டாம் கட்ட இழையில் ஒவ்வொரு தொடரும் தனித்தனி கூட்டமைப்புக்குள் கொடுத்துள்ளதைக் காணலாம்.

புரோலாக் மொழியின் கட்டமைப்பு விதியறைக்குள் கொடுக்கப்பட்ட இரண்டாம் கட்ட இழை

```
[[["nom", "ndii", "noun", "pl"]], [{"dat", "cennai", "noun"}], [{"pa", "poo", "2pl", "inte"}], [{"q", "?"}], [{"nom", ".", "period"}]]
[[["nom", "ndii", "noun", "pl"]], [{"dat", "cennai", "noun"}], [{"pa", "poo", "2pl", "inte"}], [{"nul", ""}], [{"nom", ".", "period"}]]
```

இவற்றில் பெயரடை கொண்ட கட்டமைப்பில் பெயரடைகளும் பெயரும் ஒன்றாகக் கொடுக்கப்பட்டிருக்கும். எடுத்துக்காட்டாக "அழகான உங்களுடைய வீட்டுக்கு" என்னும் தொடர் பின்வருமாறு பிரித்தரியப்படும்.

```
[[["adj", "azhakaana"], ["gen", "ndii", "noun", "pl"], ["dat", "viiTu", "noun"]]]
```

அதாவது உள்ளிடும் வாக்கியங்கள் அனைத்தையும் முதற்கட்டமாகத் தமிழின் உருபனியல் அடிப்படையில் பிரித்து ஒவ்வொரு சொல்லையும் விகுதிகளின் அடிப்படையில் விகுதிகளுக்கான கட்டமைப்பைக் கொடுக்கிறது இவ்வழிநிரல். இரண்டாவது கட்டமாக தொடரமைப்பின் அடிப்படையில் தொடர்களின் கட்டமைப்பை ஏற்படுத்துகிறது. இவ்வழிநிரலில் பெயரடைத்தொடர், கூட்டுவாக்கியங்கள் போன்ற சில சிக்கலான வாக்கியங்களையும் பகுத்தறிய வழிவகை செய்யப்பட்டிருக்கிறது. தமிழ் இலக்கண அமைப்பை முறையாக விளக்கவேண்டுமானால் மிக எளிய வாக்கிய அமைப்பான பெயர்-பெயர் அமைப்பு, கட்டளை வாக்கிய அமைப்பு, நிகழ்கால வாக்கிய அமைப்பு என்று தொடங்கி மிகச்சிக்கலான வாக்கிய அமைப்பான பெயரடைத்தொடர் வாக்கியங்கள் மற்றும் கூட்டு வாக்கியங்கள் என எளிமையான வடிவத்திலிருந்து கடினமான வாக்கிய அமைப்பில் ஒரு வரிசைக்கிரமத்தில் விளக்கலாம். இவ்வகையில் விளக்கப்பட்ட இலக்கணத்தை Renganathan (2011) நூலில் காணலாம்.

கட்டமைப்பு விதிகளும் தருக்க விதிகளும்

மற்றக் கணினி மொழிகளைக் காட்டிலும் புரோலாக் என்னும் கணினி மொழி சில பண்புகளில் தனிப்பட்டதாக இருக்கும். தருக்க விதிகளைப் பயன்படுத்தும் முறை மற்றும் கட்டமைப்பு முறையில் தரவுகளைப் பகுத்தறியும் முறை ஆகியன இம்மொழியின் தனிச் சிறப்பு எனலாம். இவ்வகையில் இந்த இரண்டு முறைகளிலும் தமிழ் மொழித் தரவை வடிவமைத்துச் செய்யப்படும் முயற்சியே இந்தக் கட்டுரையில் விளக்கப்படும் தமிழ் மொழி அறிதல் மற்றும் செய்தி சேமித்தல் மற்றும் கொணரல் பற்றிய ஆராய்ச்சியாகும். கட்டமைப்பு மற்றும் தருக்க முறையில் பயன்படுத்தப்படும் புரோலாக் மொழியப்பற்றி அறிய காண்க Clocksin and Chris Mellish (1987) மற்றும் Gazdar and Chris Mellish. (1989). எடுத்துக்காட்டாக "நான் நேற்று சென்னைக்குப் போய் என்னுடைய மாமாவைப் பார்த்தேன்" என்ற தொடரை உள்ளீடு செய்யும் போது இத்தொடரை தொடர்களின் கட்டமைப்பில் பின்வருமாறு தனது தரவில் இணைக்கிறது.

```
[[["nom", "ndaan", "noun"]], [{"tm", "ndeeRRu"}], [{"dat", "cennai", "noun"}], [{"nom", "pooy", "tr"}], [{"gen", "ndaan", "noun"}, {"acc", "maama"}], [{"pa", "paar", "1sg"}], [{"nom", ".", "period"}]]
```

பின்னர் கணினியிடம் "நேற்று" பற்றியோ "என்னைப் பற்றியோ", "போனது பற்றியோ", "சென்னை பற்றியோ" கேட்டால் அது அந்த கேள்விகளையும் இதே கட்டமைப்பு முறையில் மாற்றிக்கொண்டு முன்னரே சேமிக்கப்பட்ட தரவுகளோடு ஒப்பிடுகிறது. அப்படி ஒப்பிடும்போது கேள்வியின் அமைப்புக்கு உட்பட்ட வாக்கியங்களை வெளிக்கொணருகிறது.

எடுத்துக்காட்டாக "நான் எப்பொழுது மாமாவை பார்த்தேன்" என்னும் வினாவை எழுப்பும் போது இவ்வினாவை பின்வருமாறு இவ்வழிநிரல் பிரித்தறிகிறது.

[[["nom", "ndiingkaL", "noun"]], [{"acc", "maamaa"}], [{"qtime", "eppozhutu"}], [{"pa", "paar", "2pl"}], [{"nom", ".", "period"}]]

இக்கட்டமைப்பைத் தன்னுடைய தரவுக்கோப்பில் ஒன்றன்பின் ஒன்றாக இணைத்துப்பார்க்கும் போது மேற்கூறிய வாக்கியத்துக்கு இவ்வாக்கியம் உட்பட்டுப்போவதால் அவ்வாக்கியத்தை இக்கேள்விக்கு பதிலாக "நீங்கள் சூலை மாதம் ஏழாம் தேதி உங்களுடைய மாமாவை பார்த்தீர்கள்" என்று கொடுக்கிறது.

இதன் அடிப்படையில் உருபன் மற்றும் தொடரியல் விதிகளோடு அமைக்கப்படும் வாக்கியங்களை ஒன்றோடு ஒன்று ஒப்பிட்டுப் பார்க்கும் திறனுக்கான விதிகளும் இவ்வழிநிரலில் உள்ளீடு செய்யப்பட்டிருக்கின்றன. இவ்வழிநிரலைக் கொண்டு செய்திகளைப் பகுத்தாயும் திறனைக் கணினிக்கு அளிக்க முடிகிறது. எத்தகைய சிக்கலான செய்திகளை இவ்வழிநிரல் பகுத்தறியும்? எந்த இடங்களிலெல்லாம் இவ்வழிநிரலைப் பயன்படுத்தலாம்? என்பது போன்ற பயன்பாடு குறித்த வினாக்களுக்கு இவ்வழிநிரலில் கொடுக்கப்பட்ட அகராதியின் தன்மையையும், கொடுக்கப்பட்டுள்ள தமிழ் மொழி குறித்தான விதிகள் ஆகியனவற்றையும் வைத்தே சரியான விடையைக் கொடுக்கமுடியும். மொழியின் பொருள் நுட்ப விதிகள் பற்றியோ உலகியல் கோட்பாடுகள் பற்றிய விதிகளையோ பற்றி அதிகமாகக் கவலைப்படாமல் சூழ்நிலை அடிப்படையில் அல்லாத வாக்கிய அமைப்புகளைக்கொண்டே சில செயல்பாடுகளைக் கணினியைக் கொண்டு செயல்படுத்தலாம் என்பதே இங்கு வலியுறுத்தப்படுகிறது.

"கங்கா ஸ்நானம்" என்னும் அகிலனது நாவலிலிருந்து சில வாக்கியங்களை இவ்வழிநிரல் பிரித்தறிவதற்கான எடுத்துக்காட்டுகள் சில இங்கு கொடுக்கப்பட்டுள்ளது.

அத்தனையையும் கேட்டுக்கொண்டுதான் இருந்தாள் பாகீரதி.

[["acc", "attanai", "noun", "conj"], ["avp_refl", "keeL", "empt"], ["pa", "iru", "3fem.sg"], ["nom", "paakiirati", "noun"], ["nom", ".", "period"]]

நாற்காலியில் ஒரு யோசனையும் இல்லாமல் அறைக்கு வெளியே உட்கார்ந்திருந்த சுந்தரத்திற்கு அந்தச் சத்தம் தலையில் அடித்தது போல் கேட்டது.

[["loc", "ndaaRkaali", "noun"], ["adj", "oru"], ["nom", "yoocanai", "noun", "conj"], ["neg_avp", "il"], ["dat", "aRai", "noun"], ["acc", "veLi", "noun"], ["pa_ajp_perf", "uTkaar"], ["dat", "cundtaram", "noun"], ["adj", "andta"], ["nom", "cattam", "noun"], ["loc", "talai", "noun"], ["pa", "aTi", "neut.sg"], ["nom", "pool", "lk"], ["pa", "keeL", "neut.sg"], ["nom", ".", "period"]]

அவன் அதற்கு மேல் தயங்கினான். அதற்கு மேல் அவன் சொல்லப் போவதை அவன் உற்றுக் கேட்கலானான். அவன் கண்களில் சோகம் கலந்த ஆவல் ஊடாடியது.

[["nom", "avan", "noun"], ["dat", "atu", "noun"], ["nom", "meel", "loc.part"], ["pa", "tayangku", "3mas.sg"], ["nom", ".", "period"], ["dat", "atu", "noun"], ["nom", "meel", "loc.part"], ["nom", "avan", "noun"], ["inf", "colla"], ["nul", "poovatai"], ["nom", "avaL", "noun"], ["adj", "uRRu"], ["nul", "keeTkalaanaaL"], ["nom", ".", "period"], ["nom", "avaL", "noun"], ["loc", "kaN", "noun", "pl"], ["nom", "cookam", "noun"], ["pa_ajp", "kala"], ["nom", "aaval", "noun"], ["pa", "uuTaaTu", "neut.sg"], ["nom", ".", "period"]]

அகராதி வழி கொடுக்கப்பட்டிருக்கும் விதிகள்

இவ்வழிநிரலுக்கு அகராதியின் பங்கு மிக முக்கியமானதாகும். அகராதியில் உள்ள சொற்களைப் பயன்படுத்தியே இவ்வழிநிரல் வாக்கியங்களில் உள்ள சொற்களைப் பகுத்தறியும். அகராதியில் உள்ளீடு செய்யப்பட்ட சொல் இல்லையெனில் அச்சொல்லைப் பகுத்தறிய முடியாத நிலையில் nul என்று குறித்து முழுச் சொல்லையும் கட்டமைப்புப் பகுதியில் தருகிறது. இவ்வகை வாக்கியங்களையும் இவ்வழிநிரல் தன்னுடைய தரவோடு சேர்த்துக்கொள்கிறது என்றாலும் இத்தகைய பகுக்கப்படாத சொற்கள் வாக்கியங்களோடு இருக்கும் தருணத்தில்

நாம் எதிர்பார்க்கும் பலன் கணினியிலிருந்து கிடைக்காது. அகராதியில் பெயர், வினை, பெயரடை மற்றும் நேரப்பெயர்கள் தனித்தனியாகக் கொடுக்கப்பட்டிருக்கின்றன. இவ்வகை பெயர்களின் எடுத்துக்காட்டுகள் சிலவற்றை இங்கே கொடுத்துள்ளோம்.

பெயருக்கான அகராதிக் தரவு

பெயர்களை "noun" என்று காட்டி அதோடு அதன் பால் மற்றும் எண் காட்டும் விசுவசைத் தனியாகக் கொடுத்திருக்கிறோம். மேலும் கூட்டுப்பெயர்களை எளிமை கருதி ஒன்றாகக் கொடுத்துள்ளோம். கடைசியில் கொடுக்கப்பட்டிருக்கும் [] என்னும் பகுதி புரோலாக் கட்டமைப்புப் பகுதியாகும். இதில் அந்த பெயர் குறித்தான வேறு தகவல்களை [animate, domestic] என மற்ற செய்திகளைச் சேர்த்து அதை வழிநிரலில் உள்ள விதிகளோடு இணைத்துக்கொள்ளலாம். இவ்வகையில் பொருளடிப்படையில் மொழியைப் புரிந்துகொள்ளும் முயற்சியைச் செய்யும் போது இப்பகுதி பயனுள்ளதாக இருக்கும்.

words ("kuLirkaalam", "noun", "atu", [])

words ("kanakacapai", "noun", "aar", [])

words ("aaciriyar", "noun", "aarkaL", [])

இங்கு கொடுக்கப்பட்டிருக்கும் atu, aar, aarkaL போன்ற விசுவசை வாக்கியத்தில் இவ்வகை பெயர்கள் எழுவாயாக வரும் போது வினையில் இணைக்கப் பயன்படுவதுடன் அவை அப்பெயர்களின் மற்ற குணங்களான உயர்திணை, அஃறிணை போன்ற செய்திகளையும் அறியப் பயன்படுகின்றன.

வினைக்கான அகராதிக் தரவு

வினைகளை அவற்றின் ஏழு வகைக்குள் உண்டான வினைப்பிரிவின் எண்ணோடும் அவற்றின் செயப்படுபொருள் எடுக்கும் தன்மை பற்றிய செய்தியோடும் கொடுக்கிறோம். பெயர்களைப்போலவே இங்கும் கூட்டுவினையை ஒன்றாகக் கொடுக்கிறோம்.

words ("kalviicu", "intr", "3", [])

words ("aTiyeTu", "tr", "6", [])

இவ்வழி நிரலில் வினைகளை ஏழு முழு வகையாகப் பிரிக்கப்பட்டு சில வகைகளில் உட்பகுதியாகவும் பிரிக்கப்பட்டுள்ளது. இந்த வினை வகைகளை வைத்துதான் இவ்வழிநிரல் வினைகளுக்கான விசுவசைப் பிரிப்பதுடன் அவற்றை இணைத்து வினைகளை உருவாக்கவும் செய்கிறது.

பெயரடைக்கான அகராதிக் தரவு

தனிப் பெயரடைகள் adj என்னும் குறிப்போடு கொடுக்கப்படுகின்றன. பெயரிலிருந்து உருவாக்கப்படும் "அழகான", "உயரமான" போன்ற பெயரடைகள் விதிகள் மூலம் பகுத்தறியப்படும்.

words ("andta", "adj", "", [])

words ("ciRiya", "adj", "", [])

இவற்றோடு காலம் மற்றும் எண்கள் தனியாகக் கொடுக்கப்படுகின்றன.

words ("mee", "mnth", "5", [])

words ("cuun", "mnth", "6", [])

மேற்சொன்னவாறு இவ்வழிநிரலின் சரியான பயன் அகராதியைச் சரியான முறையில் கொடுப்பதில்தான் இருக்கிறது.

முடிவு

இந்த வழிநிரலே தமிழ் உருபுகளைப் பிரிக்கும் திறன் கொண்டிருப்பதால் அகராதியில் சொற்கள் இருக்கும் தருணத்திலும் மேலதிகமான விதிகளை உள்ளடக்கும் தருணத்திலும் இந்த வழிநிரலைக் கொண்டு பல

செயல்பாடுகளை மேற்கொள்ளலாம். கதைகள், செய்திகள், மற்றும் கட்டுரைகள் போன்ற தரவுகளிலிருந்து குறிப்பிட்ட செய்திகள் பற்றி விரைவில் தெரிந்து கொள்ள வாய்ப்பு இருக்கும். மேலும் தரவு மென்பொருள்கள் போன்று செய்திகளைச் சேமிக்கவும் அவற்றிலிருந்து ஒரு குறிப்பிட்ட வகை செய்திகளை பெறவும் இவ்வகை வழிநிரல்கள் பயன்படும். தமிழ் வாக்கியங்களை கணினி பகுத்தறியும் வகையிலான இந்த ஆராய்ச்சி இதனுடைய தொடக்கநிலையில்தான் இருக்கிறது எனலாம். எனினும் தமிழ் மொழியின் விதிகளை கணினிக்கு முறைப்படி கொடுப்பதற்கான ஒரு உத்தியை இக்கட்டுரை விளக்குகிறது. நாம் தமிழ் மொழியை எங்ஙனம் புரிந்து கொள்கிறோமோ அதே முறையில் புரோலாக் கணினி மொழி வாயிலாக கட்டமைப்பு முறையைப் பயன்படுத்தியும் தருக்க விதிகளைப் பயன்படுத்தியும் முறைப்படி கணினிக்கும் மொழியைப் புரிந்துகொள்ள வைப்பது ஒரு இயலக்கூடிய ஆராய்ச்சியாகும். மொழிச்சிக்கல்களைத் தருக்கவிதிகளுக்கு உட்பட்டு எவ்வளவுக்கு எவ்வளவு நம்மால் தருவிக்கமுடியுமோ அந்த அளவிற்கு மொழியைக் கணினி புரிந்துகொள்ளும் திறனும் கூடும் எனலாம். தமிழ் மொழிச் சொற்களை மிகுந்த அளவில் மின்னகராதிக்குக் கொடுக்கும் தருணத்தில் உருபன் விதிகளை எளிதில் பிரிப்பதற்கான வாய்ப்புகள் அதிகமாக இருக்கும். உருபன் விதிகளைப் பிரித்து வாக்கியங்களை இக்கட்டுரையில் கொடுக்கப்பட்டுள்ளவாறு ஒரு கட்டமைப்பு முறையில் கொடுத்துவிட்டால் அவற்றை கணினி வழி பகுத்தறிவது மிகவும் எளிதாகும். இவ்வகையில் இதற்கான வழிமுறைகளை முறையாக இக்கட்டுரை விளக்குகிறது. மேலும் <http://www.thetamilanguage.com/tamilnlp> என்னும் பக்கத்தில் இவ்வழிநிரலில் கொடுக்கப்பட்டிருக்கும் விதிகளைக் கொண்டு ஆங்கிலம் - தமிழ் மொழிபெயர்ப்புக்கான ஒரு வழிநிரலையும் தருவிக்கப்பட்டுள்ளது.

மேற்கோள் நூற்களும் கட்டுரைகளும்

- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. MIT Press, Boston.
- Clocksin, W.F. and Chris S. Mellish (1987). *Programming in Prolog*. (3rd ed.) Springer-Verlag: New York.
- Gazdar, Gerald. and Chris S. Mellish. (1989). *Natural Language Processing in Prolog*. Addison-Wesley Publishing Company: Wokingham.
- Trost, Harald. (1991). "Recognition and Generation of Word forms for natural language understanding systems: integrating Two-level morphology and feature Unification". *Applied Artificial Intelligence*. 5:411-457
- Renganathan, Vasu. (1997). "A Lexical Phonology Approach to Processing Tamil Words by Computer." *IJDL* Vol. XXVI No. 1, January 1997.
- Renganathan, Vasu. (2011). *Tamil Language in Context: A Comprehensive approach to learning Tamil*. Thetamilanguage.com publication, Cherry Hill, New Jersey, USA (<http://www.thetamilanguage.com/book>).
- Winograd, T. (1983). *Language as a Cognitive Process*. Addison-Wesley.

Written to Spoken Tamil : Conversion Rules (A Study in Speech Synthesis)

K. Karunakaran and R. Krishnan
University of Malaya, Kuala Lumpur

Abstract:

Tamil is a diglossic language for centuries. The two varieties which participate in the diglossic situation namely the written and spoken varieties show considerable change both in language structure and language use. These two varieties are taught at different levels for different purposes. There are learners interested in learning both the varieties, either one of the varieties, depending upon their needs.

When we try and used to introduce these two varieties in the teaching of Tamil as a second language/foreign language, there arises the need to convert one variety from the other. So, this paper focuses on written to spoken conversion by formulating a set of rules called conversion rules. The conversions start right from the phonetic level itself and extended up to the level of discourse. There are phonetic and phonological conversions of which some are significant; morphological conversions which are conditioned in many environments; there are morphophonemic rules which are exclusively found in each of these varieties and there are phrase and sentence formations found to be little different in them. So, the written to spoken conversions as found in the above mentioned levels are considerably ordered and systematically explainable. Hence, the aim paper is to demonstrate the way the conversions take place along with occurrences as well as conditioning factors. Though there are differences in the occurrences of the final phones-phonemes-morphemes, their occurrences show similarities also. The conversions show a pattern also, like word-final occurrences (before #), when followed by a suffix, phonetic and phonological significance, the phonetic realization and so on. The conversion rules include all such patterns in word formation, inflectional, derivational processes vice-versa process can also be explained, but needs.....

(Key words: diglossia, conversion, varieties, conversion rules)

1.0 Introduction

In the modern period dialectal words have found their way into the vocabulary domain of standard Tamil either by the virtue of their intrinsic worth or by the authority of the poets who pressed them into service (Sethupillai, 1953). The two distinct varieties had parallel development and functional significance since then and continue to be so even today. The two varieties not only have structural variation at different linguistic levels but also functional variation in the use of language in different socio cultural domains such as education, administration, mass media science and technology, home language and so on. That is each variety has different sets of social functions in formal and informal levels. These varieties do not overlap or merge much in the modern use, but for the movie language and dialogues and conversations among characters in novels and short stories in modern literature as well as some of the socio cultural usages in writings.

So, a study pertaining to the conversion of written to spoken and the vice versa is needed in order to make appropriate selection of usages for different purposes. With reference to the conversion written to spoken Tamil seems to be more systematic based on rule formation.

2.0 Aims of the Study

The following are the aims of the present study:

- 2.1 To generalize the regular written to spoken conversion at the word level.
- 2.2 To explain the conditions in the form of conversion rules at the word final and medial positions.
- 2.3 To identify and explain all those changes found in the morphological structures of Tamil in the form of conversation rules.

3.0 Research Questions

The following questions are put forward:

- 3.1 What are all the different word final and medial conversions to be generalized at the word level?
- 3.2 What type of conditions are needed to explain the word level conversions?
- 3.3 What are all the conversions found in the morphological structures and how to explain the conversions in the form of rules?

4.0 Methodology

Qualitative methodology is adapted in this study in order to present the exact variations and presenting the variations in the form of conversion rules from written to spoken Tamil. The rules are presented under two headings (i) Word Level Conversion Rules and (ii) Grammatical Level Conversion Rules.

The data for analysis includes words /lexical items found in modern spoken Tamil in the Malaysian contexts. However, those distinct variations found in the Tamilnadu spoken Tamil also have been taken into account, as this would help to generalize the conversion rules as far as possible. So, the frame work has two segments the use of spoken Tamil in different contexts and their functional significance on the one hand and that of making rules for systematic conversions.

5.0 Data Analysis

5.1 Word Level Variation and Conversions

- 5.1.1 Variations at the word final positions
- 5.1.2 Variations at the word medial positions
- 5.1.3 Some of the explainable variations at word initial position

5.2 Sample analysis

CR1 - ai # → -ε#

CR2 - am / - a:m / - a:n / - um / -o:m etc. → nasalized vowels
followed by #

CR3 - C # → Cu # (phonetically í)

CR4 - y # → -yi # / -yyi # (the addition of -y or doubling of -yy has
to be explained by a morphophonemic rule in the process of
conversion.

CR5 - this conversion rule takes care of word medial conversion
especially -rr → -tt- (spoken Tamil does not show use of R except in Kanyakumari Tamil but
in the case of words with -rr- they are pronounced as -tt- with or without case suffix)

CR6 - this conversion rule takes care of word medial conversion
especially -nr → nn / -n (conditioned by initial syllable CV /CV)

Conversions at the Grammatical Level

5.3.1. Conversions in the Noun Structures

The conversions include : (i) Number Suffixes (ii) Gender Number Suffixes (iii) Link morphemes (iv) Case suffixes (v) Derivative Suffixes such as adjectival suffix, adverbial suffix, common human noun suffixes and so on.

There are numeral forms both cardinal and ordinal which show variations and they have to be explained in the form of conversion rules. The cardinal bases also show variations especially from 1-9.

So, at least there would be 15 conversion rules with reference to noun structures.

5.3.2 Conversions in the Verb Structures

The conversion rule includes all those variations found in the finite and non-finite verb forms in spoken Tamil.

5.3.2.1 Finite Verb Conversions

There are at least three types of finite verb forms namely (i) with tense suffixes, (ii) without tense suffixes, (iii) imperative permissive, hortative type of finite forms.

5.3.2.2 Non Finite Verb Conversions

The following non finite verb forms have to be explained in the form of conversion rules.

- (i) Infinitive Forms
- (ii) Verbal Participle Forms
- (iii) Adjective Participle Forms
- (iv) Conditional Participle Forms
- (v) Temporal Participle Forms

5.3.2.3 Person – Gender – Number Suffixes which occur in finite Verb Forms after the tense suffixes

5.3.2.4 Negative suffixes, Negative words and Affirmative Negative differences in spoken Tamil

5.3.2.5 Participial Nouns and Verbal Noun Formations

5.3.2.6 Verbal Derivative Nouns (Verb based + Link Morpheme + Derivative Suffix)

There are complex verb forms as well as compound verb forms in spoken Tamil which show distinctive variations. So, there is a need to explain all such variations in the form of conversion rules.

Type 1 : Complex Verb Forms

There are two divisions under this namely aspectual forms and modal forms which show distinct conversion.

Type 2 : Compound Verb Forms (Verb form 1 + verb form 2 : Noun form + verb form etc.)

There are dependent parts of speech like adjectives, adverbs and particles which occur with noun forms, verb forms, noun and verb forms respectively. All these formations have to be explained in the form of conversion rules. So, all together there would be around 30-35 rules needed. The speech synthesis study with reference to modern Tamil have to be modernized by explaining conversions which are needed when one goes from written to spoken Tamil. The NLP study should try to formulate a series of ordered set of rules (conversions or correspondences).

This kind of standardization could be done with the help of data using modern methodology in NLP studies.

A Three-level Genre Classification for Tamil Lyrics

Karthikeyan, Nandini Karky, Elanchezhiyan, Rajapandian & Madhan Karky
Karky Research Foundation, Chennai, India.

Abstract

A song is a merging point of two art forms, music and lyrics. Thousands of songs are created every year in Tamil in various forms. Presently, a song's genre is determined by the genre of the music. To the best of our knowledge, we do not have a genre classification for lyrics. In this paper, we propose a three-level genre classification system for Tamil lyrics. The first level, which we call the base classification, classifies Tamil lyrics into 10 base categories that would abstract the theme of the song. The second level classifies the song based on the mood of the song. The third level classifies the song based on the language style. Based on this three-level classification, we have identified 240 buckets in which a Tamil lyric can be placed. In this paper, we describe each level in detail and present genre classification statistics for a collected sample of 1200 Tamil lyrics.

Keyword

Lyric Genre, classification, analysis

1. Introduction

A song is a culmination of two art forms, music and lyrics. From lullabies to rhymes to jingles to original sound tracks, songs travel with us from birth. Tamil, a classical language with rich literature, produces thousands of songs every year in the form of jingles, private albums and original sound tracks of movies. All over the world, music has been well studied and classified in various contexts. In today's world, a genre of a song would denote the genre of the music and not that pertaining to lyrics. Jazz, blues, heavy metals, classical are a few examples of music genres. To the best of our knowledge, a classification of songs based on lyric content does not exist. In this paper, we propose a three-level genre classification for Tamil lyrics, which we believe can be extended to other languages.

In our attempt to classify Tamil lyrics, we identified that a single level classification was insufficient. A lyric has multiple dimensions such as the base concept of the song, the mood conveyed by lyrics and the language style. In this paper, we propose ten base genres, six mood genres and four style genres. We have applied this classification to a collection of 1200 songs in Paadal[1], a lyric portal that facilitates search and ranking of lyrics. We also present a statistical analysis of genre distributions.

This paper is organized into five sections. The following section presents a brief survey of literature relevant to genre classification in general. The third section describes the three levels of genre classification in detail. We present a statistical analysis of genre distribution in section four. The final section summarizes the paper and discusses the work in progress and further extension of this work.

2. Background

In this section, we review literature relevant to this paper. Most of the genre classification research happens in the area of music.

In [1] the authors present a system to extract more than hundred features from music files. The features are used to classify the songs. Features such as instrumentation, texture, rhythm, etc are used to classify the music. The ambiguities in genres are discussed in [2].

In [3] a hybrid classification system is proposed for hierarchical, flat and round robin classification. This again is an automated classification system that mimics human classification. In [4] authors take POS information from lyrics to classify

the music genre and mood. Interestingly in [5] rhyme feature from lyrics is used to classify the music. None of these works attempt to classify lyrics.

The primary objective of this paper is to define a genre set for Tamil lyrics. The secondary objective is to create a training set for an automated classifier, by manually categorizing a set of lyrics using the proposed classification.

3. Three-Level Genre Classification

In this section we define the three levels of genres namely Base, Mood and Style. Each of the levels and the sub classes are defined in this section.

3.1 Base

Every lyric has at least a base concept for which the lyric was created. A base genre would abstract the theme of the song based on the purpose. We have classified the base genre into the following ten classes.

3.1.1 Character

Character class comprises of songs created for the purpose of describing the character of a person. Physical features, capabilities, strengths and weaknesses of a person are usually described in this category.

3.1.2 Festival

Songs celebrating a common event such as new year, festival of lights, holi, etc. come under this category. Personal events such as birthdays and weddings are not classified under this category.

3.1.3 Nature

Songs on nature or on love for nature fall under this base genre. The purpose is to describe nature as it is and as a person's love or thoughts on natural elements such as hills, sun, sky, moon, trees, flowers, etc.

3.1.4 Philosophy

Songs that focus on a deeper meaning in various topics come under this category. Philosophical songs usually talk about human relationships, emotions and God from a worldly view.

3.1.5 Relationship

Songs on motherhood, fatherhood, relationship with siblings and friends fall under this category. Songs on romantic relationship alone are excluded from this category.

3.1.6 Romance

Songs related to single and multi-sided romantic relationships between two or more persons. Themes such as falling in love, state of love and breaking up in love fall under this class.

3.1.7 Occasion

Songs on birth, wedding, family events and death come under this category. This is not to be confused with festivals category, which are celebrated by large communities based on religion, language or country. Songs in the occasion category are usually specific to a family or a person.

3.1.8 Spiritual

Songs on religious beliefs, religions and gods come under this category. Religious festivals are categorized under festivals.

3.1.9 Patriotic

This category groups songs on a country, rendered with patriotic fervor.

3.1.10 Miscellaneous

Songs that do not fall under any of the above nine categories are grouped under this category. Songs that are ambiguous can also be categorized under miscellaneous.

3.2 Mood

Irrespective of the base genre, a song can be categorized based on its mood. In this section we present six mood classes. The classes have been based on the emotions described in [4]

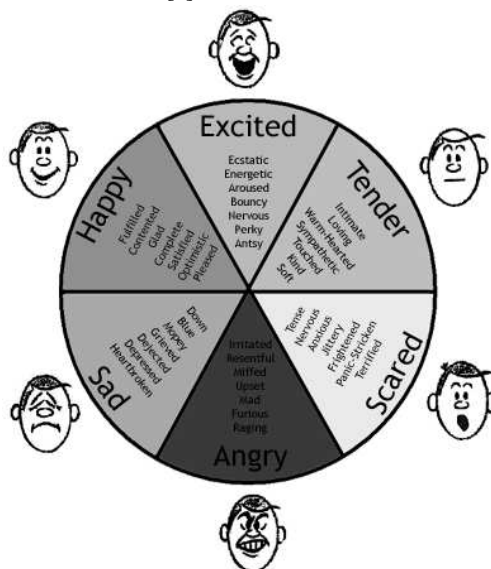


Figure 1. Six Basic Moods [4]

3.2.1 Happy

Lyrics that talk about fulfillment, satisfaction, contentment and optimistic thoughts fall under happy category.

3.2.2. Excited

Songs with a bouncy, energetic and ecstatic feeling are categorized under Excited.

3.2.3. Tender

Lyrics depicting intimate and soft feelings directly or indirectly as well as those portraying a kind, sympathetic or humane nature fall under tender category.

3.2.4. Scared

Lyrics expressing nervousness, tension and jittery, frightened feelings are classified as scared.

3.2.5. Angry

Irate feelings, fury or rage in a person, expressed in lyrics are categorized as angry.

3.2.6. Sad

Songs on the depressed state of a person, songs expressing dejection, grief or heartbroken feelings fall under sad.

A lyric is classified into one of these categories. If a lyric expresses multiple moods, the predominant mood is tagged. A song with romance as the base category may have an excited, angry or sad mood associated to its mood.

3.3 Style

The third level of the classification is based on language style. Tamil language has evolved over many thousand years and is characterized by innumerable style variations. For the sake of simplicity, we categorize the language styles into four simple classes.

3.3.1 Traditional

Lyrics with non-colloquial Tamil, without use of words from other languages, are classified as Traditional style.

3.3.2. Folk

Lyrics with conversational-style Tamil, specific to some dialects, are classified under folk style.

3.3.3. Contemporary

Lyrics that use words from other languages mixed with Tamil words or those that coin or introduce new Tamil words or sounds are classified as contemporary.

3.3.4. Mixed

A song may have a mix of traditional, folk and contemporary lyrics in it. Such songs are classified as Mixed.

4. Analysis

This lyric classification is part of a Paadal[7], a lyric portal. Around 1000 songs in Paadal portal were classified manually applying the three-level genre classification system. Ambiguities that arose in the manual classification process were resolved by assigning the most appropriate tag. For instance, if a song talks partly about nature and partly about a romantic relationship, human judgment was used to classify, depending on which class the song tended more towards.

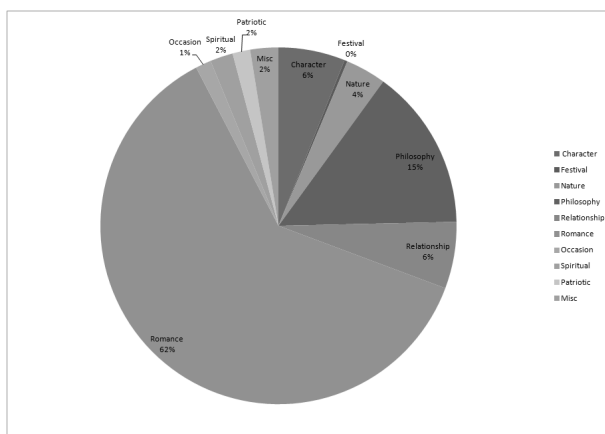


Figure 2. Base Genre Distribution

Of the base genre category, romance takes a huge slice of the pie. 62% of the lyrics belonged to this category. The sample of 1000 lyrics may be very small over 70 years, but we believe the pie would not change drastically when more songs are added. Philosophical songs and songs on relationships take the second and third place respectively in this distribution.

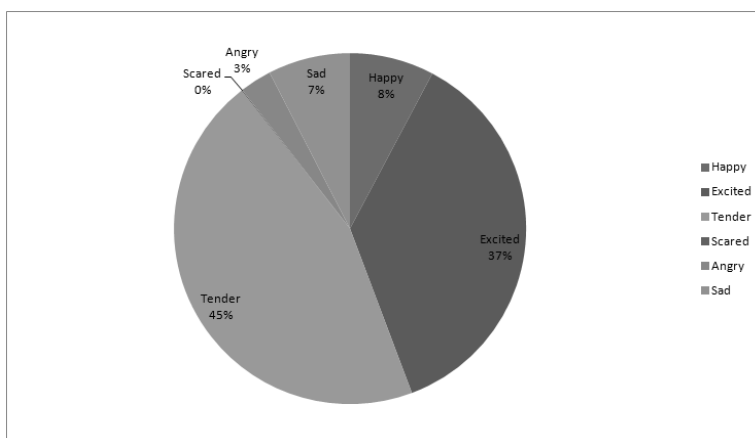


Figure 3. Mood Genre Distribution

Songs with tender feelings rule the mood distribution. Over 45% songs in the collection belong to the tender mood. Excited and Happy songs together share almost the other half of the pie.

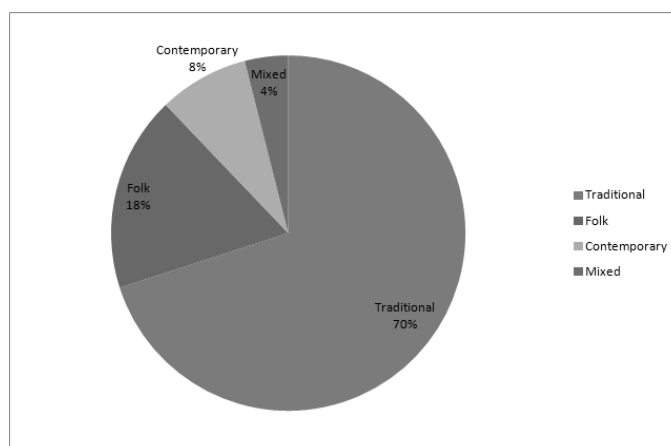


Figure 4. Style Genre Distribution

Over 70% songs in the set belong to traditional style lyrics. Around 18% songs belong to the folk category. On the whole ‘Romance.Tender.Traditional’ is the most popular song category, out of 240 possible classes, in this three-level genre classification.

5. Summary and Future of the work

In this paper, we proposed a three-level genre classification scheme for Tamil lyrics. Each level and category was explained. The lyric distribution statistics over these genres were discussed. The manual classification applied to over 1000 Tamil songs can be used as training set for an automated classifier, which is our work in progress. The base genre classification may not be complete. Even though there is a Misc category, a few popular categories might be missing owing to lack of samples. This three level classification system is open to modifications. New base genres can be added in the future. Searching for lyrics and lyric suggestions based on this classification will also be useful features in a lyric portal. This system also gives scope to language research in lyrics where lyric styles of various lyricists can be analyzed.

6. References

- Automatic Genre Classification Using Large High-Level Musical Feature Sets, Cory McKay Ichiro Fujinaga, McGill University. Faculty of Music Montreal, Quebec, Canada cory.mckay@mail.mcgill.ca, ch@music.mcgill.ca,
- Musical Genre Classification, Cory McKay and Ichiro Fujinaga Music Technology, Schulich School of Music, McGill University Montreal, Quebec, Canada cory.mckay@mail.mcgill.ca, ich@music.mcgill.ca
- Automatic Genre Classification of MIDI Recordings, Cory McKay Music Technology Area Department of Theory Faculty of Music McGill University, Montreal Submitted June 2004
- Teh Chao Ying; Doraisamy, S.; Abdullah, L.N., "Genre and mood classification using lyric features, " Information Retrieval & Knowledge Management (CAMP), 2012 International Conference on, vol., no., pp.260, 263, 13-15 March 2012
- Rudolf Mayer, Robert Neumayer, and Andreas Rauber, Rhyme And Style Features For Musical Genre Classification By Song Lyrics, ISMIR 2008 on Content Based Retrieval, 340 – 342.
- Emotion, Examples of basic Emotions <http://en.wikipedia.org/wiki/Emotion>, Last Accessed 08th July 2013
- Paadal – Lyric Portal, <http://www.karky.in/labs/paadal>, last accessed 8th July 2013.

Lyric Objects & Spot Indices for Paadal, a Tamil Lyric Search Engine

Karthikeyan, Elanchezhian, Rajapandian & Madhan Karky

Karky Research Foundation, Chennai, India.

Abstract

Tamil lyrics play a vital role in language literacy. We have estimated that over one lakh lyrics exist in Tamil in the form of original soundtracks of movies. The traditional way of organizing text, for search and ranking, will not be efficient for lyrics. Structure of lyric, frequent use of compound words as metaphors, and the difference in user perspective of search makes the existing indexing methods inefficient. In this paper, we propose LOB (Lyric Object), an object oriented structure for representing lyric and spot indices for LOBs. Spots are tiny locator objects that pin-point the exact location of a root word in an LOB. In this paper, we explain in detail how the spots are stored in word objects and how intersection of spots are used in multi-word search to produce efficient results in Paadal, a Tamil lyric search engine.

Keyword

Spot Indices, Lyrics, Indexing, LOB (Lyric Object)

1. Introduction

Lyrics play a vital role in carrying cultural elements across time. They play a pivotal role in language literacy. Many research results in the field of education suggest that songs are the best tools for effective teaching. In Tamil language, we have numerous words that are not used in day to day conversation but can be found only in lyrics and poetry. With the advent of UNICODE, search engines now can index Tamil documents. A few search engines, such as Google, use a basic stemmer to find the root words for indexing. In Tamil we estimate over one lakh lyrics to exist in the form of movie sound tracks, television serial sound tracks, jingles and more. Techniques employed in searching a web page news article does not prove efficient in searching lyrics, as lyrics have numerous features such as similes, metaphors, compound words and they have a different structure. Apart from this, the search requirement of a user is totally different when it comes to lyrics. Stop words, which make sense in traditional search, does not make sense here.

In this paper, we define Lyric Object (LOB), an object-oriented structure to represent a lyric. A lyric is comprised of letters, words, lines, stanzas, and markers for male or female perspective, in addition to meta-data such as album, year etc.

We also propose Spot Indices for LOBs. We define spots as locators of keywords in a lyric. Every keyword would have one or multiple spots associated with it. A spot would pinpoint the exact location of that keyword in an LOB.

We have represented 1064 Tamil lyrics as LOBs. Keyword objects were created containing their corresponding spots. The size of the keyword objects vary based on the number of spots contained within the object. This method, of storing spots within objects, enables us to store the individual keyword objects in hard disk rather than having the entire lyric index in memory.

We also show how multi-word search is simplified to an intersection of spots contained in their respective keyword object.

This paper is organized into five sections. The second section provides background study on indexing strategies and object representations. We present the LOB structure, the keyword object and spot structures in the third section. The fourth section presents implementation and results. The analysis of spot indices for lyric search in Paadal, a lyric portal, is given in the fifth section. The final section summarizes the paper and discusses future directions of this research.

2. Background

Indexing of web articles has gone through a sea change in the past few decades. In this section we review a few articles relevant to this paper. In [1] a morphological analyser is presented. We use the stemmer module of this analyser for our queries and keyword extraction. A concept based indexing is proposed in [2] where a tree structure index is presented. This indexing involves applying UNL concepts and relations between words.

In [3] authors discuss compaction strategies for indices. With number of indices growing rapidly as new documents are added for indexing, this provides means of making the subsystem memory efficient. In this paper we propose an object structure, LOB, designed specifically for Tamil lyrics and introduce spots, pre-computed locator information for indexing LOBs.

Nested indexing in an Object oriented system is discussed in [4] whereas [5] discusses efficient keyword search in relational database. Both these works consider structured data and not a semi-structured data such as a lyric. [6] Investigates web page importance ranking. The modeling of web page importance discussed in this paper inspired us to design the ranking described in section 4.

3. Methodology

In this section, we introduce Lyric Object (LOB), a structure to represent Tamil lyrics along with meta-data. We also define the structure of spots based on LOBs and describe in detail, how spots are associated with keywords. We use Atchayam morphological analyser [6], for all our online and offline stemming process.

3.1 Lyric Object (LOB)

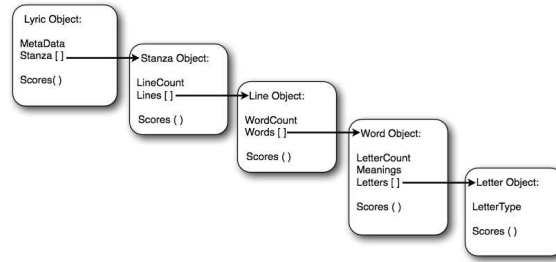


Figure 1. Lyric Object Structure

The structure of lyric object (LOB) is depicted in figure 1. The natural hierarchy of a lyric is used to represent a lyric as an object. Each letter/alphabet in the language is treated as a letter object. Collection of letter objects constitutes a Word object. A word has associated meanings, synonyms and associated scores such as pleasantness and word popularity. A Line object is built as an order of Word objects. A line has associated scores and voice indicator such as male, female or chorus. Similarly a Stanza object is an order of lines. An LOB is built as an order of these stanzas. The primary purpose of this fine-grain representation is to facilitate an effective search mechanism. Even a particular morphological ending in a selected word can be located efficiently with the indexing mechanism proposed in the next section.

3.2 Keyword Object and Spots

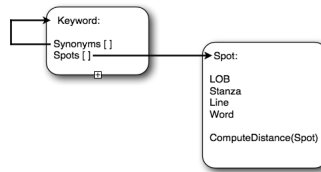


Figure 2. Keyword Object & Spots

A Keyword object as depicted in Figure 2 holds the pointers to synonymous keyword objects and spots. A keyword object would contain a minimum of one spot as the keyword objects are created from the LOBs by stemming the original words. There is no upper limit on the number of spots for a keyword object.

A Spot is a locator object that connects a keyword to all its locations. There exists a spot object for every occurrence of a word. A Spot object comprises of an LOB ID, Stanza, Line & Word position information.

4. Implementation and Analysis

An LOB Converter converts Tamil lyrics in UNICODE xml files to LOBs. As an offline process, the word objects are retrieved and stemmed using a Morphological Analyser [6]. The stemmed word is checked for existence in Keyword objects.

If such an object is not present, then a new keyword object is created with corresponding spot information. If not, the spot information is added to the existing keyword object.

4.1 Ranking

Ranking is divided into offline and online processes to increase the search efficiency.

4.1.1 Offline Ranking:

With the kind of frame work setup for search, ranking the lyrics for single keyword becomes effective with offline ranking.

Let Search be $S = \{kw1, kw2, \dots, kwn\}$ where kwi denotes keyword i .

Let Lyric objects be denoted by $L = \{l1, l2, \dots, lm\}$ where lj denotes an LOB j

Let **Oc** denote the occurrence of the keyword

$Oc(kw1, l1) \square Oc(kwi, lj)$ be an integer.

Fp \square First position of the keyword

$Fp(kwi, lj) \square fp(kwi, lj)$ be an integer.

4.1.2 Online Ranking:

Online ranking is carried out for multi-word search queries.

Let Ds \square Distance between two keywords.

$Ds(kwi, kwj) \square 0$, if kwi and kwj are not in same Stanza.

$(ABS(Oc(kwi) - Oc(kwj))) / Swc$, if kwi and kwj are in same Stanza.

Note: $Swc \square$ Stanza total word counts

POP (lj) \square Popularity of the lyric

Number of user views is used to compute popularity.

RAT (lj) \square Rating of the songs

Average of user ratings is used to compute Rating of song

4.1.3 Application

We employed spot indices for Paadal[7], a lyric portal comprising of 1064 lyrics. The entire framework was developed on Java platform. Total number of keywords identified after stemming was 5533, total number of spots identified was 115304 and total number of spots unique words was 15068.

Corresponding keyword objects of Stemmed keywords identified from user query were retrieved. The intersection of the spots of two keywords is identified. The ranking features described in section 4.1.1 is used with their corresponding weightage values to compute the rank of each LOB.

Figure 3 gives a snapshot of the search results. The lyric title, album title and snippet containing keyword (s) are displayed in result boxes. Clicking the result boxes would direct the browser to the corresponding lyric page.



Figure 3 : Search Results in Paadal Lyric Portal

5. Analysis

We compare the spot indices with regular tree based indices in this paper. The graph provided in Figure 4 compares the memory requirements for single word and multi word queries. The maximum size of a keyword object is 225KB, which alone is needed in memory for single keyword search. For n-keyword search, a maximum of $225*n$ KB memory alone is used. This in contrast to traditional tree based search where 80 MB of memory will be used for a small collection of 1064 lyrics.

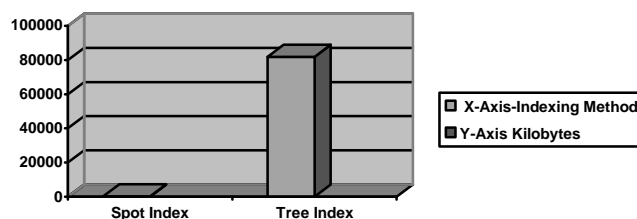


Figure 4: Single word memory occupation

6. Future Work and Conclusions.

In this paper, we discussed the need for an effective search framework for lyrics as lyrics have special features such as rhyme, similes, metaphors and compound words. Apart from that, requirements of users searching lyrics are different from that of users searching articles. We proposed an object-oriented structure, LOBs for lyrics. Structures of keyword objects and Spot locators were discussed. With detailed description of the implementation and setup for ranking, we concluded this paper with analysis of spot based indices with traditional tree based indices. Spot indices prove to be both memory and time efficient in the comparison.

Compacting spot indices by grouping multiple spots for same keyword in one LOB, analyzing the search results to evaluate ranking, personalizing ranking for lyrics can be interesting avenues of further research.

7. References

- Anandan, R. Paarthasarathi, and T.V.G. Geetha, Morphological Analyser for Tamil, ICOM 2002, 2002.
- Subalalitha, R. Paarthasarathi, T.V.G. Geetha and Karky, M. 2008. Concept based semantic indexing technique. Swm-08
- Index Compression for Information Retrieval Systems. <http://www.dc.fi.udc.es/~roi/publications/rblanco-phd.pdf>, last accessed date 07th July 2013
- Beng Chin Ooi, Jiawei Han, Hongjun Lu, and Kian Lee Tan. 1996. Index nesting – an efficient approach to indexing in object-oriented databases. The VLDB Journal 5, 3 (August 1996), 215-228.
- Fang Liu, Clement Yu, Weiyi Meng, and Abdur Chowdhury. 2006. Effective keyword search in relational databases. In Proceedings of the 2006 ACM SIGMOD international conference on Management of data (SIGMOD '06). ACM, New York, NY, USA, 563-574.
- Wolfgang Gaul. 2011. Web page importance ranking. Adv. Data Anal. Classif. 5, 2 (July 2011), 113-128.
- Paadal Lyric Portal, <http://www.karky.in/labs/paadal>, last accessed date 08th July 2013.

An Avatar Rendering Hand Gesture For Tamil Words

D.Narashiman, Teaching Fellow (dnarashiman@gmail.com)

Bavatharani Suriyan, Student (bhavi2490@gmail.com)

Dr.T.Mala, Assistant Professor (Sr Gr) (malanehru@annuniv.edu)

Department of Information science and Technology, College of Engineering Guindy, Anna University, Chennai.

ABSTRACT

Advancements in Internet Technology and World Wide Web increase the usage of animated files in web pages. This paved way to overcome usage of video files which have drawbacks such as blocking huge memory storage and high latency time. Therefore, the applications involving the use of multimedia elements tend to increase the usage of animated files rather than video files. Some of the applications that use animated files are story visualization, e-learning applications, mathematical derivation etc. One such interesting application is the usage of avatars for hand gesturing. Avatars are the audiovisual bodies that people use to communicate with each other in Metaverse. Avatar animation has a lot of advantages as they can be designed according to the user requirement, the movements can also be controlled correctly and perfectly as they are given by MEL scripts. The proposed work describes the method of rendering hand gestures of an Avatar for given Tamil pronoun. This system is designed to render hand gestures for given Tamil pronoun which will be useful for Tamil hearing impaired community. The given pronoun undergoes pre-processing using Atcharam, a Morphological analyzer and is classified into four types namely personal pronoun, object pronoun, possessive pronoun and demonstrative pronoun. The hand gestures are already predefined in MEL script for a specific list of pronouns in Tamil. This MEL script helps in matching the pronoun to the defined hand gesture and this is rendered as action on an avatar. This acts as a translator system for hearing impaired people and also helps them in understanding the Tamil word quickly since it uses Tamil variant sign language.

1. INTRODUCTION

The state-of-art technology is towards virtual visualisation of things. All the natural scenes, machine design, building model are animated realistically, human figures are also rendered naturally. These human visual bodies are called the Avatars. There are lots of application where avatars are used in animation world one such application is breaking the communication barriers between Hearing impaired community and speaking community. Even though online tools and software are made available for this process, all are pre-ordered sentences and videos for specific domain like weather forecast, post office automation system. The hearing impaired community avail the service by entering or selecting any of the listed query text. Upon the query selection, corresponding recorded videos will be played. These systems are designed for specific domain, queries are standard and as the video files are in size the loading time is more. More over they cannot be changed dynamically as they are pre-recorded and very expensive. These videos also depend on individuals mannerisms of the person who sign. The camera position and lightings also plays an important role in recording the motion of gestures to capture the minute details otherwise the sign interpretation may go wrong. To overcome the above mentioned drawbacks, video files are replaced by animated avatar. The avatars are created by Maya Embedded Language (MEL) scripts. The scripts are easier to write and the rendering of sign gestures are also faster. The system that has been developed acts as a tool which renders the hand gestures for the Tamil pronouns. This tool can be used in translating Tamil pronoun to sign language gestures and also in teaching and learning sign language. This paper is organized as follows. Section 2 discusses the literature involved in avatar generation and hand gesturing. Section 3 talks about system architecture and various modules. Section 4 gives the performance of the developed system and section 5 gives the conclusion and future work.

2. LITERATURE SURVEY

There is lots of work carried out in translating text to sign language gestures but most of the work deals with generating video files as output and they are domain specific. Almost all the research paper that talk about can be classified as given in figure 1.

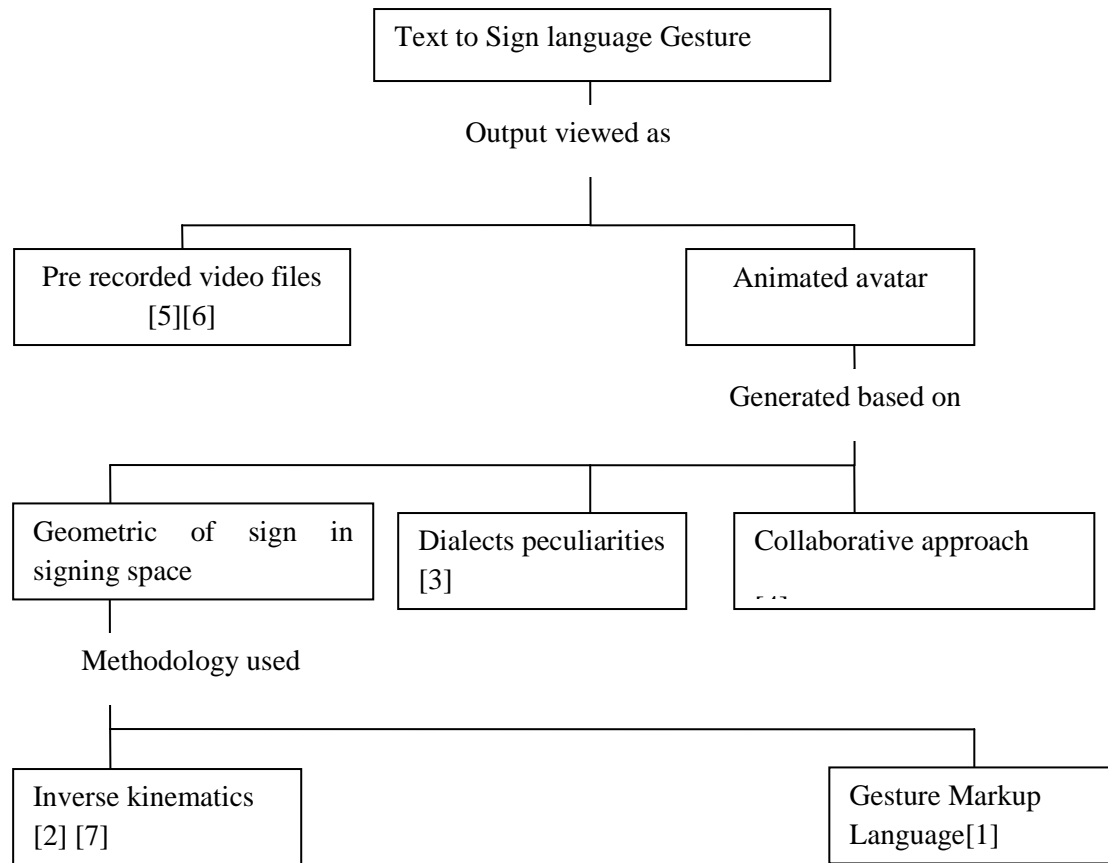


Figure 1: Methods of translating text to sign gestures

J.A.Bangham et.al [1] proposed a system that translates the text to signing avatar using Gesture markup language.

M.Delorme et.al [2] proposed a new approach for generating animated sign sequence based on geometric description of signs. In this approach first the key postures and transition resolutions are found. Secondly the skeleton postures are computed using inverse kinematics and both are merged to render the final output. The main defect in this approach the computation is done extensively for every sign gesture separately.

Mikhail G. Grif et.al [3] proposed a system for translating Russian language to Russian sign language. The work considers peculiarities of the Russian sign language. The system concentrates more on the Russian sign language than on the animation.

Mohamed JEMNI et al. [4] proposed a collaborative approach system in which the user can construct and add the sign to the dictionary through an interface. Sliders are provided in the interface which is used to create the sign and able to store in the database. The system follows client server architecture.

D.Narashiman et al. [5] [6] proposed a system for translating English and Tamil text to sign language videos. The systems identify the words which convey the meaning of the sentence correctly and categorise them as rule based, spatial and finger spelled. These are stored in a knowledge repository. Based on this the final output sequence of text and video ordering is done.

Sam YEATES et al. [7] proposed a Real – time 3D graphics for human modelling and for teaching sign language. The system develops a static pose by kinematic joint angles and renders the animation by interpolating methods.

The proposed system generates the avatar animated sign gesture sequences for the given Tamil text. The avatar creation is specially done using MEL scripts which will be easier to render on web. The next section gives detailed view of the developed system.

3. SYSTEM DESCRIPTION

The system implemented renders hand gesture of an avatar for Tamil pronoun word. The basic components of hand gestures are shape, location, orientation and movement of the hand and palm. These basic components are stored in a database which is dynamically sequenced depending on the pronoun word and hand gestures of an avatar are rendered. Figure 2 shows system architecture of the proposed system.

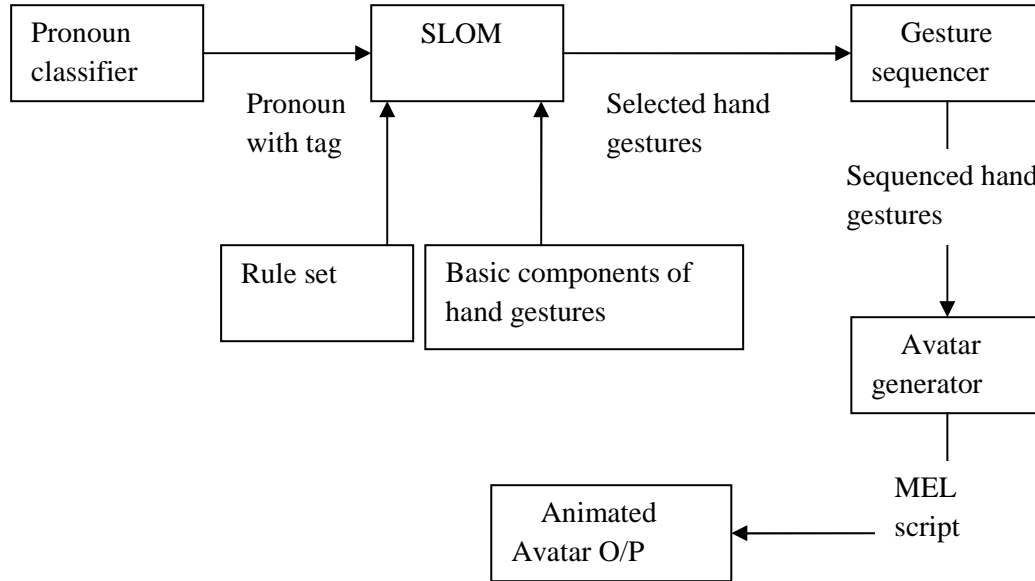


Figure 2: System Architecture

The classifier schemes the words based on different types of pronoun. Depending on this classification, the words are rendered dynamically with their manual gestures defined by the parameter like hand shape, hand location, orientation and movement. The system is designed to generate the gestures for the following pronouns possessive, reflexive, emphasizing, demonstrative and personal.

The PRONOUN CLASSIFIER identifies the pronoun type and also tags the pronoun word with corresponding pronoun tags. The tagged pronoun words are then processed in SLOM module to generate the action sequence based on the rule set. The rules are defined as follows

Rule 1: If the tagged word is possessive pronoun then two actions are performed one for prefix and another for suffix. The suffix sign will be common for all the possessive pronouns. The wrist is closed and moved front. The location of the sign will be varied either in abdomen level or in the chest level.

Rule 2: If the pronoun type is reflexive or emphasizing pronoun then two different signs are done the prefix and suffix. In this case the suffix sign is done by extending the index finger and performing a curve motion from abdomen to chest along with the prefix sign.

Rule 3: For singular demonstrative pronouns the location of the pointing signs is done in the abdomen level.

Rule 4: For plural demonstrative pronouns movements are circular motion and done at the abdomen level.

Rule 5: For singular personal pronoun the action is separated into two, first to mention the gender next the action.

Basic gesture components are stored in a database from where the basic hand gestures are selected. The basic components deal with the shape, orientation, location and movements of hand and finger. Table 1 lists the details of the basic components.

TABLE 1: Basic components of gestures

HAND	RIGHT, LEFT
FINGER	INDEX FINGER, MIDDLE FINGER, RING FINGER, THUMB FINGER, PINKY FINGER
SHAPE	CLOSED, EXTENDED
ORIENTATION	INWARD, OUTWARD, DIAGONAL
LOCATION	FACE, CHEST, ABDOMEN
MOVEMENT	STATIC, RIGHT SHOULDER TO LEFT SHOULDER, CIRCULAR MOTION – CLOCK WISE AND ANTI CLOCK WISE DIRECTION.

TABLE 2: Finger representation

IF	INDEX FINGER
MF	MIDDLE FINGER
RF	RING FINGER
PF	PINKY FINGER
TF	THUMB FINGER
E	EXTENDED
I	INWARD
O	OUTWARD
CH	CHEST
S	STATIC
CM	CURVED MOTION

The selected gesture is then sequenced in the GESTURE SEQUENCER module. The time, location and movement of the gestures arranged in this module and passed to the avatar generator. Pronouns are gestured with right hand. In the sequence the first parameter is shape of the finger, location, orientation and motion. The finger representation is shown in table 2. For the personal pronoun *என்* the sequence is IF_E_I_CH_S, MF_C_I_CH_S, RF_C_I_CH_S, PF_C_I_CH_S, TF_C_I_CH_S. Each parameter corresponds to one finger position. The index finger is extended other fingers are closed and pointing towards chest. Similarly for possessive pronoun

உங்களுடைய is shown by two gestures the order is IF_E_O_CH_S, MF_C_O_CH_S, RF_C_O_CH_S, PF_C_I_CH_S, TF_C_O_CH_S followed by IF_C_O_CH_CM, MF_C_O_CH_CM, RF_C_O_CH_CM, PF_C_I_CH_CM, TF_C_O_CH_CM. The first gesture index finger point outward and all other fingers are closed. The second gesture the index finger is moved in a curved fashion. Similarly the sequences for other pronouns are generated based on the rules. After this the avatar are generated.

The final module is AVATAR GENERATOR. The first part is rendering avatar and the next part is rendering the hand gestures. The first step in creating avatar is modeling a human figure. Then skin, hairs and texture are applied to the avatar. Using Inverse kinematics techniques movements are given to joint chains such as arms and legs can be created. The Inverse Kinematics handle lets to pose and animate an entire joint chain by moving a single manipulator and it automatically rotates all the joints in the joint chain. For example, if you move a hand to a out ward to the body, the other joints in the arm rotate to accommodate the hand's new positioning. Maya is used to create an avatar and render the hand gesture for the given element. The avatar can be created using toolset for each and every action. Improvements to rigging tools, the ability to match Tracks clips, enhancements to the Graph Editor and a new Heat map skinning method offer an enhanced the animation. After finishing the design of the avatar, scripting has to be implemented for the action performance. Maya Embedded Language (MEL) script, it is a powerful scripting language is used to automate many tasks within the Autodesk Maya. The term MEL script refers to the lines of commands typed and executed in the script editor or a series of commands saved in a text file that uses the .mel extension. Expressions can be used to control the behavior of individual nParticles within an nParticle object. Depending on the condition and the gesture sequences, sign gestures are rendered. The following figures 3, figure 4 and figure 5 shows the hand gestures for the word *என் உங்களுடைய, அது* respectively.



Figure 3 shows the hand gestures for the word என்



Figure 4 shows the hand gestures for the word உங்களுடைய



Figure 5 shows the hand gesture for word அது

4. PERFORMANCE ANALYSIS

The analysis is carried out by checking the accuracy of animated character playing the signs. Output generated for the given word are analyzed separately and comparison is made between the correct sign rendered and the incorrect sign rendered for the given Tamil pronoun word. Ten students are asked to identify the sign gesture shown by the avatar. Then the ration is calculated between the correctly identified words to total no of words by each student. From the statistics an average of them is calculated. It is found that about 60 percent of the word are correctly recognized by the student.

5. CONCLUSION

In this paper we have designed a tool that enriches the communication between Tamil speaking community and Tamil Hearing impaired community. The system is designed to generate animated avatar that render the gestures for the Tamil pronouns words. The classifier identifies the pronouns and tags them. The Slom generates the hand gestures for the pronouns and it is sequenced by the gesture sequencer from where the final animation is generated. This work can be made more perfect by enhancing the animation process and by taking into account other parts of speech. This also can be extended to generate a animation for a given story which will be useful for children.

Reference

- J A Bangham, S J Cox, R Elliott, JRW Glauert, and I Marshall, “ Virtual Signing: Capture, Animation, Storage and Transmission – an overview of the ViSiCAST Project” in the proceeding of “Speech and language processing for disabled and elderly people” April 2000
- [2] M.Delorme, M.Filhol and A.Braffort, “Animation Generation Process for Sign Language Synthesis”, in the proceeding of Second International Conferences on Advances in Computer-Human Interactions, 2009.
- [3] Mikhail G.Grif, Yelena A.Demyanenko, Olga O.Korolkavo and Evgeniy B.Tsoy, “Development of Computer Sign Language Translation Technology for Deaf people”, in the proceeding of IEEE, 2011.
- [4] Mohamed Jemi and Oussama Elghoul, ”An Avatar based approach for automatic interpretation of text to sign language”, in the proceeding of 9th European conference for the Advancement of the Assistive technologies in Europe, October 2007.
- [5] D.Narashiman and Dr.T.Mala „An interactive system for conversion of text to sign language, ” in proceeding of International Conference on Information Systems & Software Engineering, 2009, PP.164-167.
- [6] D.Narashiman and Dr.T.Mala “A machine translation system for converting Tamil text to sign-language” in proceeding of 9th Tamil Internet Conference, 2010, PP.572-578.
- [7] Sam Yeast et.al, Eun-Jung Holden and Robyn Owens, “Real-Time 3D graphics for human modelling and teaching Sign Language” in the proceeding of International conference on computer vision and graphics, September 2002.

E-Content To Animation Conversion System (Etacs)

Chakkaravarthy.E chakra.ceg@gmail.com

Narashiman.D dnarashiman@gmail.com

Abstract- This paper describes about the generation of automatic animations for the given e-content information for a particular domain. E-Content to animation conversion system is an innovative idea in the field of Natural Language Processing. It helps children to understand the e-contents more easily. It works by transforming e-content into an easily understandable animations. In this model, we are going to take particular domain e-content as input and we generate simple animations as output for the given e-content information. It aims on developing an animation for the given input text file.

Keywords: Text document classification, Linguistic Analysis, Information Extraction, XML Script Generation, Animation Creation.

1 INTRODUCTION

E-Content to animation conversion system is an innovative idea in the field of Natural Language Processing. It aims on developing an animation for the given input text file. This project focuses on developing animation for the given moral instructions. This project is divided in to two phases. The first phase involves the content generation and creation of template. Template is created using object identifier, spatial inference, temporal inference and environmental inference. In E-Content generation the given text files are trained according to their domain and their respective keywords are displayed with the frequency count. The second phase of the project involves XML script generation from the created template. Then the XML script is passed to openGL code . OpenGL code generates 3D object generation, 3D spatial environment generation and object pose generation. Using OpenGL simple animations can be created. Then, using flash more sample animation objects are created. Flash animations are stored in a particular database. Using java particular flash animations will be executed for the given input. OpenGL code generates 3D object generation, 3D spatial environment generation and object pose generation. Using OpenGL simple animations can be created.

2 CONTENT GENERATION

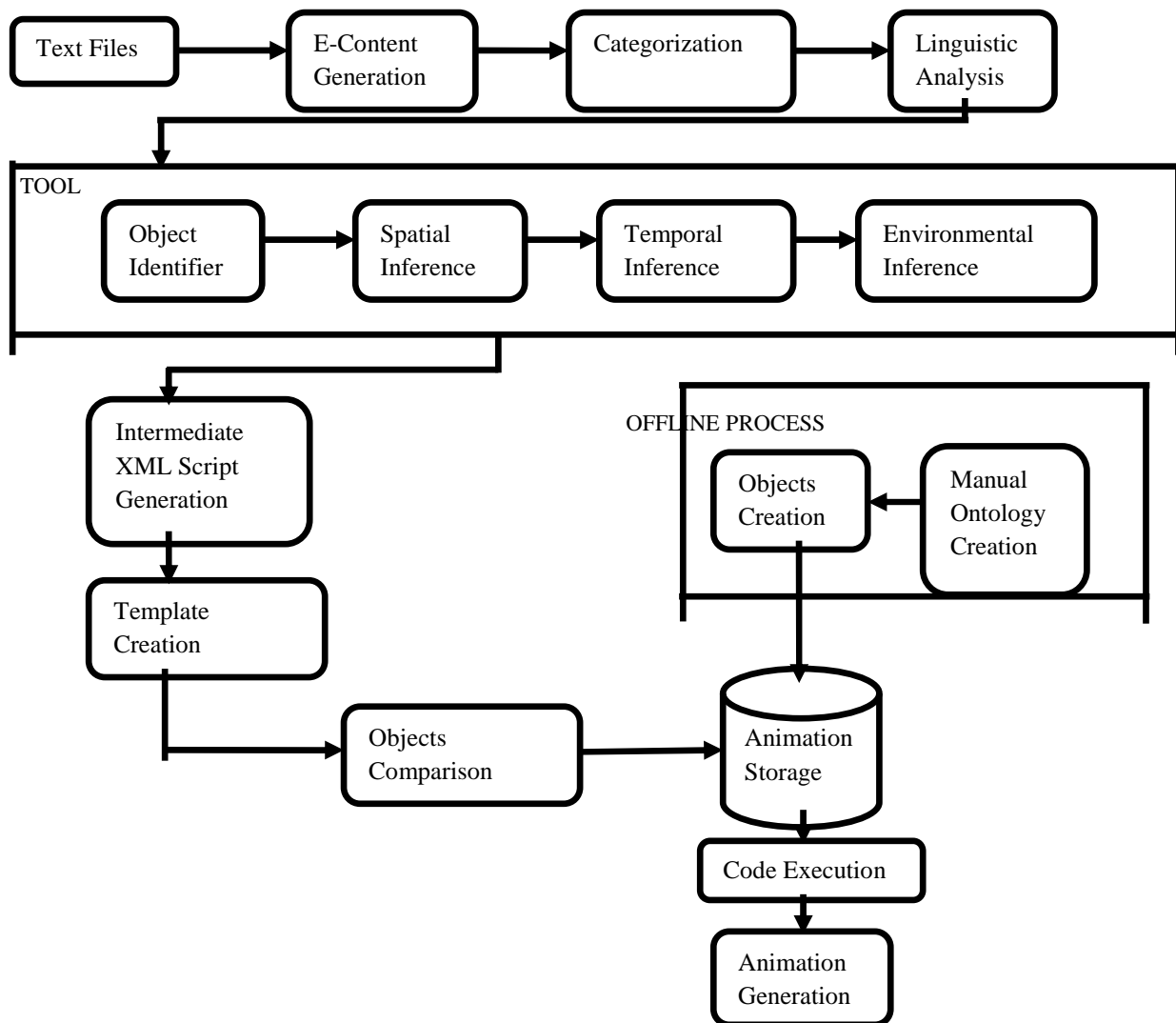
Content generation has been performed to generate the given moral text files for the particular domain. The objective of content generation is to assign entries from a set of prespecified categories to a document. Traditionally this categorization task is performed manually by domain experts. Each incoming document is read and comprehended by the expert and then it is assigned a number of categories chosen from the set of prespecified categories. It is inevitable that a large amount of manual effort is required. A promising way to deal with this problem is to learn a categorization scheme automatically from training examples. The content generation task can be defined as assigning category labels to new documents based on the knowledge gained in a classification system at the training stage. In the training phase we are given a set of documents, each document keywords is compared with the trained domain keywords . Once the Content generation scheme is learned, it can be used for classifying future documents. The content generation task can be defined as assigning category labels to new documents based on the knowledge gained in a classification system at the training stage. In the training phase we are given a set of documents, each document keywords is compared with the trained domain keywords .

3 LINGUISTIC ANALYSIS

In Linguistic analysis, the part-of-speech tagger is used. POS tagger is a grammatical category or word group in a language to which words may be assigned on the basis of how they are used in sentences. The traditional main parts of speech in tamil are noun, verb, adjective, adverb, pronoun, preposition, conjunction, and interjection.

4 ARCHITECTURE DIAGRAM

The overall system architecture explains the various functionalities required to perform the E-content to animation conversion System. E-Content generation involves preprocessing and classification process.



The moral instruction text documents are given as input to the preprocessing phase. In this phase the stop words are removed from each and every text documents in that particular domain. These documents were used as the training set to find the unknown text document belongs to which category . After the stop words are removed, stemming operations is performed then the keywords of that particular domain with the frequency counts were displayed. Then the keywords are stored in a separate file. Similarly three to four domains were trained. Then perform the linguistic analysis for the given input document. After the content generation process, information should be extracted from those documents. It is a type of information retrieval whose goal is to automatically extract structured or semi structured information from unstructured machine-readable documents. The extracted information should fill the template of each document of a particular domain. Thus the template was created using text categorization and information extraction. Template was created for the particular domain. Gate tool was used to extract the information from the text document. After the stop words are removed, stemming operations is performed then the keywords of that particular domain with the frequency counts were displayed. Then the keywords are stored in a separate file. Similarly three to four domains were trained. Then perform the linguistic analysis for the given input document. GATE is an architecture suggests that the elements of software systems that process natural language can usefully be broken down into various types of component, known as resources. Components are reusable software chunks with well-defined interfaces, and are a popular architectural form, used in Sun's Java Beans and Microsoft's .Net, for example.

GATE components are specialized types of Java Bean, and come in three flavours:

- Language Resources (LRs) represent entities such as lexicons, corpora or ontologies;
- Processing Resources (PRs) represent entities that are primarily algorithmic, such as parsers, generators or ngram modellers;
- Visual Resources (VRs) represent visualisation and editing components that participate in GUIs

When using GATE to develop language processing functionality for an application, the developer uses GATE Developer and GATE Embedded to construct resources of the three types. This may involve programming, or the development of Language Resources such as grammars that are used by existing Processing Resources, or a mixture of both. GATE includes resources for common LE data structures and algorithms, including documents, corpora and various annotation types, a set of language analysis components for Information Extraction and a range of data visualisation and editing components. GATE supports documents in a variety of formats including XML, RTF, email, HTML, SGML and plain text.

OpenGL (Open Graphics Library) is a software interface to graphics hardware. The interface consists of a set of several hundred procedures and functions that allow a programmer to specify the objects and operations involved in producing high-quality graphical images, specifically color images of three-dimensional objects. Most of OpenGL requires that the graphics hardware contain a frame buffer. Many OpenGL calls pertain to drawing objects such as points, lines, polygons, and bitmaps, but the way that some of this drawing occurs (such as when antialiasing or texturing is enabled) relies on the existence of a frame buffer. Further, some of OpenGL is specifically concerned with frame buffer manipulation.

5 CONCLUSION AND FUTURE WORK

The unknown documents were categorized and the information was extracted from those documents to create the template. Before categorizing the text documents, it was preprocessed in order to reduce the process time. The content generation was done by using frequency counts. The content generation identifies the particular domain of the document. After the domain was found, it undergoes for linguistic analysis process. In Linguistic analysis it uses POS tagger it produces the tagged output for the given unknown text document. Then template is created by extracting the information from the text documents. Future work focuses on XML script generation for the given template and generate animations for the given text document using OpenGL code.

REFERENCES

- Dhiraj Joshi, James Z. Wang and Jia Li, "The Story Picturing Engine – A System for Automatic Text Illustration", ACM Transactions on Multimedia Computing, Communications and Applications, vol. 2, no. 1, pp. 68-89, 2006.
- Bob Coyne and Richard Sproat, "WordsEye: An Automatic Text-to-Scene Conversion System", SIGGRAPH 2001, Los Angeles, CA, 2001.
- Eva Hanser and Paul Mc Kevitt, "Text-to-Animation: Affective, Intelligent and Multimodal Visualisation of Natural Language Scripts". In: ISEA 2009, International Symposium on Electronic Art. Belfast, Northern Ireland, 2009.
- Diana Inkpen and Darren Kipp "A Prototype Natural Language Interface for Animation Systems", IEEE, 2004.
- Kevin Glass and Shaun Bangay, "Automating the Creation of 3D Animation from Annotated Fiction Text". IADIS International Conference Computer Graphics and Visualization, 2008.
- Johansson, R. et al., "Automatic text-to-scene conversion in the traffic accident domain". In IJCAI-05: Proceedings of the 19th International Joint Conference on Artificial Intelligence. Edinburgh, Scotland, pp. 1073–1078, 2005.
- Kaoru Sumi, Katsumi Tanaka, "Transforming E-contents into a Storybook World with Animations and dialogues using Semantic Tags". International Conference on Advances in Computer Entertainment Technology, 2005.
- Fellbaum, C. editor, "WordNet: An Electronic Lexical Database" MIT Press, Cambridge, USA, 1998.

A Three-level Genre Classification for Tamil Lyrics

Karthikeyan, Nandini Karky, Elanchezhiyan, Rajapandian & Madhan Karky
Karky Research Foundation, Chennai, India.

Abstract

A song is a merging point of two art forms, music and lyrics. Thousands of songs are created every year in Tamil in various forms. Presently, a song's genre is determined by the genre of the music. To the best of our knowledge, we do not have a genre classification for lyrics. In this paper, we propose a three-level genre classification system for Tamil lyrics. The first level, which we call the base classification, classifies Tamil lyrics into 10 base categories that would abstract the theme of the song. The second level classifies the song based on the mood of the song. The third level classifies the song based on the language style. Based on this three-level classification, we have identified 240 buckets in which a Tamil lyric can be placed. In this paper, we describe each level in detail and present genre classification statistics for a collected sample of 1200 Tamil lyrics.

Keyword

Lyric Genre, classification, analysis

1. Introduction

A song is a culmination of two art forms, music and lyrics. From lullabies to rhymes to jingles to original sound tracks, songs travel with us from birth. Tamil, a classical language with rich literature, produces thousands of songs every year in the form of jingles, private albums and original sound tracks of movies. All over the world, music has been well studied and classified in various contexts. In today's world, a genre of a song would denote the genre of the music and not that pertaining to lyrics. Jazz, blues, heavy metals, classical are a few examples of music genres. To the best of our knowledge, a classification of songs based on lyric content does not exist. In this paper, we propose a three-level genre classification for Tamil lyrics, which we believe can be extended to other languages.

In our attempt to classify Tamil lyrics, we identified that a single level classification was insufficient. A lyric has multiple dimensions such as the base concept of the song, the mood conveyed by lyrics and the language style. In this paper, we propose ten base genres, six mood genres and four style genres. We have applied this classification to a collection of 1200 songs in Paadal[1], a lyric portal that facilitates search and ranking of lyrics. We also present a statistical analysis of genre distributions.

This paper is organized into five sections. The following section presents a brief survey of literature relevant to genre classification in general. The third section describes the three levels of genre classification in detail. We present a statistical analysis of genre distribution in section four. The final section summarizes the paper and discusses the work in progress and further extension of this work.

2. Background

In this section, we review literature relevant to this paper. Most of the genre classification research happens in the area of music.

In [1] the authors present a system to extract more than hundred features from music files. The features are used to classify the songs. Features such as instrumentation, texture, rhythm, etc are used to classify the music. The ambiguities in genres are discussed in [2].

In [3] a hybrid classification system is proposed for hierarchical, flat and round robin classification. This again is an automated classification system that mimics human classification. In [4] authors take POS information from lyrics to classify

Lyric Visualization

Rajapandian, Nandini Karky, Elanchezhian, Karthikeyan & Madhan Karky
Karky Research Foundation, Chennai, India.

Abstract

Thousands of Tamil lyrics are being created every year as part of private albums, movie projects, jingles and many more forms. Lyrics play a vital role today in spreading language literacy. Lyrics have exclusive features such as rhymes, similes, metaphors, pleasantness, freshness, genre, mood etc. that differentiate it from the prose form. In this paper, we propose a method to extract these features and visualize a lyric as a flower. The extracted features from a lyric are used to generate a flower. We propose a mapping between the features of a lyric to the features of a flower. We have tested the proposed method over one thousand Tamil lyrics on Paadal, a lyric portal. The visualized flower is used in search results to enhance the search experience of the user. Finding similar lyrics by identifying similar flowers is the core idea behind this user experience feature.

1. Introduction

Tamil is rich in literature. One of the popular literary forms in modern times is lyrics. We have over a million lyrics in the form of original soundtracks of motion pictures, private albums and jingles in Tamil. Quoting or posting context relevant lyrics and favorite lyrics on short messages, blogs and micro blogs is one of the favorite activities of teenagers. Lyrics happen to be their favorite form of literature. Searching for lyrics from a huge database in the form of text or number values is a tedious task. Apart from keywords, lyrics have so many other features that a person might be interested in, such as rhyme style, genre, fresh words, similes and many more. Searching for these features in the form of plain text is a daunting task.

Visualization is a popular technique used to represent complex data as simple readable visual representations. Starting from visualizing simple 2 dimensional data into line graphs, we have complex multi-dimensional data visualization techniques in data warehousing applications.

In this paper, we propose to visualize a lyric as a flower by mapping the statistical features of the lyric to the features of the parts of a flower. Just by looking at the flower for a few seconds, one should be able to determine the features of the lyric. This is the primary objective of this paper.

This paper is organized into six sections. The second section reviews literature related to visualization and lyric features. The third section describes the lyric features and explains the features with examples in detail. The fourth section presents the visualization scheme by describing the mapping. The fifth section talks about the implementation regarding feature extraction. The final section summarizes the paper and concludes with ongoing research and future directions.

2. Background

Visualization by itself is a huge area in Computer Science. We restrict our literature to music and art visualizations only. In [1], authors propose to visualize text readability using a few readability metrics. The technique provides means for readers to see which sections of text are hard to read with complicated words. Lyricon [2], an interesting project, presents a method to automatically select icons for different sections of tunes. Apart from considering the musical features, lyricon considers text features to select icon based on the concept.

‘See the music’ [3] is a meaningful project which tries to visualize music features and emotions for the hearing impaired. Authors of [4] present a visualization of lyrics. Here they use lyric features to extract an event and compose an image from the extracted concepts. The work in [5] presents a visualization technique with lyrics using musical features such as pitch to assist and instruct karaoke singers.

In this paper, we propose visualization of a lyric purely based on ten features of lyrics. The ten features are mapped to features of a flower. The following section explains the ten features in detail.

3. Lyric Features

Independent of the language, a lyric is rich in features not found in prose. Lyric is usually a small piece of work compared to blogs or news articles. The following are the ten features that we identify lyrics and propose to use for the visualization in the next section. All statistical details mentioned in this section were obtained from a collected set of 1000 Tamil lyrics, which were created over a period of 70 years.

3.1 Song Length

Song lyrics do not vary much in length. Average length of a Tamil lyric is 38.56 lines, which we consider as the medium length. In addition, 4 windows of 5 lines on both sides classify songs as very short, short, long and very long.

3.2 Base Genre

Songs are usually classified into genres based on music. In this case, the base genre classifies the songs into 10 classes based on lyrical theme. Character Description, Romance, Philosophy, Festival, Occasion, Relationship, Nature, Patriotic, Spiritual or Miscellaneous are the 10 base genres.

3.3 Mood Genre

Irrespective of the base genre, a lyric can have one of the six moods from Happy, Excited, Tender, Scared, Angry or Sad. The mood genre classifies the lyric based on the words pertaining to the different emotions.

3.4 Style Genre

Tamil has various dialects. The style genre classifies a lyric based on the language style. Traditional, Folk, Contemporary or Mixed are the classes in the style genre.

3.5 Simile Count

Similes are one of the popular features of lyrics. Similes are used to convey complex concepts using simple or known examples or the vice versa. The number of similes in a lyric can be extracted using simile keywords such as போல (poala) போன்ற (poandRa).

3.6 Freshness

A lyric's freshness can be computed in ways such as usage of words, similes, rhymes and many more ways. We consider just the usage of words over a period of ten-year time windows to compute the freshness of lyrics.

3.7 Pleasantness

Ignoring the music and meaning of lyrics, the sounds used in the words would determine the pleasantness of lyrics. Usage of nasals and long vowels increases the pleasantness of a word and thereby a lyric. Pleasantness is computed by modeling the place of articulation of phonemes and the consonant family variations in Tamil.

3.8 Unrecognizable words

Words that do not exist in a dictionary and cannot be analyzed by a morphological analyser are classified as unrecognizable words. We use atchayam [ref] morphological analyser to analyse Tamil words. Non-Tamil words also fall under this category.

3.9 Rhyme Strength

Rhyme is a key feature of lyrics. Lyrics which are rich in rhymes are popular among kids. Tamil has a very unique rhyme schemes and rhyme patterns in lyrics. First letter match (monai), Second letter match (edhugai) and End letter match (Iyaibu) are considered to compute a rhyme score in a lyric.

3.10 Dominant Rhyme

In some lyrics, edhugai or monai may be dominant. Most lyrics have iyaibu as their dominant rhyme. The dominant rhyme is identified by simple count of the rhyme schemes described in section 3.9.

4. Visualization and Mapping

Figure 1 presents a flower structure and the mapping of lyric features explained in the previous section to visual features of a lyric.

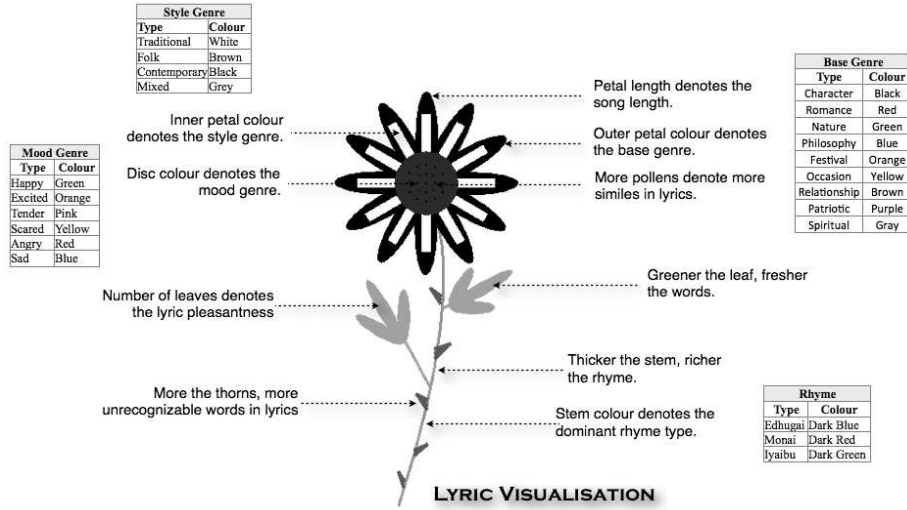


Figure 1. Visualization Mapping

5. Implementation and Results

The visualization presented in this paper was implemented in Java for Paadal[ref] a lyric portal project. 1000+ lyrics were analyzed for the 10 features and visualized. The project was hosted on Linux server. The upper part of the flower was used in search results to visually indicate the genre type as shown in the snapshot in figure 2. The full flower visualization is presented in the lyric page of the portal as shown in figure 3.



Figure 2. Search Results from Paadal [ref]

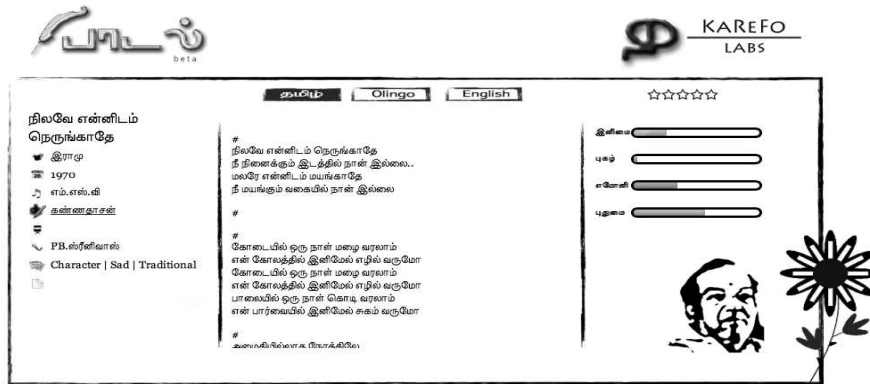


Figure 3. Flower visualization in Paadal portal [ref]

6. Conclusion and Future Work

In this paper, we proposed a visualization technique that extracts ten features from a lyric and maps it to the visual features of a flower. We implemented the visualization in a lyric portal, Paadal. With this technique, users will be able to find their favorite type of song just by looking at the flower in the search results, rather than reading the textual descriptions.

In the future, we plan to find similarities and distance between two flowers in order to group similar flowers. Suggestions to users can be made using these distance metrics. This work can be extended to other languages also by changing the feature extractors to suit the language. Extracting a few more features such as metaphors and abstractions and then compacting the same into similar flower visualization would be challenging research problems for the future.

7. Reference

- Karmakar.S; visualizing text readability; Advanced Information Management and Service (IMS) 6th International Conference, 2010.
- Machida, W; Lyricon: A Visual Music Selection Interface Featuring Multiple Icons; Information Visualization (IV) 15th International Conference, 2011.
- Mori.J; seeing the music; Science and Technology for Humanity (TIC-STH), IEEE Toronto International Conference, 2009.
- Kikuchi.G; Lyrics-based automatic music image generation using scene knowledge for music browsing; Consumer Electronics (ICCE) IEEE International Conference, 2012.
- Rob Diaz-Marino, Sheelagh Carpendale, Saul Greenberg; University of Calgary, 2005.

Scoring Models for Tamil Lyrics

Elanchezhian, TamilSelvi E, Karthikeyan, Rajapandian, Madhan Karky
Karky Research Foundation, Chennai, India.

Abstract

Lyrics are rich in features such as rhyme, pleasantness, similes, metaphors and more. Many of these features are exclusive to lyrics. We have estimated that more than two thousand Tamil lyrics are being created every year in various forms. Modeling the lyric-specific features becomes an essential task in organizing the lyrics for retrieval and analysis. In this paper, we propose three scoring models for analyzing the rhyme, pleasantness and freshness of a lyric. The paper also explains in detail on how the weights were selected for individual models. Scores were computed for 1200 Tamil lyrics based on these models. Discussing the score distribution in detail, the paper concludes with open questions for further research.

Keywords

Rhyme, Pleasantness, Freshness, Lyrics, Scoring models.

1. Introduction

Can art be quantified? We like a painting, we dislike a piece of music, and we find a certain section of a story, good. Is it possible to quantify an art form for comparing one piece of work with another? On the whole, the answer is no. But we can certainly take a few quantifiable features from the art form and compare those features. In a painting, the number of colors used is quantifiable. In music, the pitch and number of instruments are quantifiable features.

Tamil language is rich in literature. We have numerous poetic forms in Tamil. Lyrics, composed with music, are one of the most popular forms of literature. Lyrics from popular songs are being searched and quoted on many blogs, social networking platforms. A few thousands of lyrics are being created every year in form of popular music, motion picture soundtracks, private albums and more. Lyrics have unique features in them such as rhyme, similes, metaphors and many more that are not present in a prose or a news article.

In this paper, we aim to identify those specific features that can be quantified. We try to model the rhyme, pleasantness and freshness of a lyric in order to compute a score for each of these features. Algorithms for computing these scores are presented. The scores are analyzed by comparing them against human judgment for correlation.

This paper is organized into five sections. The following section presents a brief survey of literature relevant to relevant lyric feature extraction and genre classification based on lyric feature. The third section describes the scoring model for Tamil lyrics. We analyze this scoring model by determining the correlation coefficient between scoring model and human judgment in the fourth section. The final section summarizes the paper and discusses the future extension of this work.

2. Background

In Tamil, rhyme refers to the way in which successive lines are strung in a poem or lyric. The occurrence of similar syllables and words in successive lines correspond to alliteration (monai), rhyme (edhugai) and end-rhyme (iyaibu).

Lyric Freshness in this paper refers to the word freshness in lyrics. A song with words that were never used before are considered fresher.

Pleasantness of a lyric is a combined pleasantness of words used in the lyric. Pleasantness of a word does not take into account the meaning of word or the music in the song. The syllables in a word, the presence or absence of long vowels, consonant family the characters belong to and place of articulation determines the pleasantness of a word.

“Poetic Features for Poem Recognition: A Comparative Study” [1] describes the poetic features and how it helps to distinguish and extract poetic feature with which poem can be classified from other type of text.

“Rhyme and style features for musical genre classification by song lyric” [2] proposes genre classification using features like rhyme and part-of-speech.

“LaaLaLaa - A Tamil Lyric Analysis and Generation Framework” [3] generates meaningful Tamil lyrics to a given tune and domain based on the lyric features such as rhyme, meaning and flow. The Rhyme finder is used to choose words that match one or more of the three Rhyme properties (edhugai, monai, iyaibu).

“Special indices for LaaLaLaa Lyric Analysis and Generation Framework” [4] proposes a framework to construct the indices for efficient lyric generation, based on Tamil rhyme scheme such as Rhyme, Meter match and Parts of Speech for fast retrieval.

Hirjee and Brown [5] present a sophisticated tool for extracting rhymes from lyrics based on phoneme frequencies, with a focus on hip-hop styles. This work uses the basic rhyme to study the quality of lyric. A large corpus of lyrics is used for studying rhyme.

In this paper, we propose a scoring model by which a lyric is scored based on rhyming patterns, rhyme schemes (edhugai, monai, iyaibu, Ati EMI, Murru EMI, Inai EMI, Polippu EMI, Oruu EMI, Kulai EMI, Kilkkatuvay EMI, Merkatuvay EMI), Freshness (Timeline and word uniqueness) and Pleasantness (Vowel, consonants classifications & place of articulation) of lyric that are specific to Tamil language to evaluate lyrics.

3. Scoring Models for Tamil Lyric

Modeling the lyric-specific features is an essential task in organizing the lyrics for retrieval and analysis. In Rhyme scoring model based rhyme schemes and rhyme patterns in Tamil. Rhymes schemes are further classified into Ati EMI, Murru EMI, Inai EMI, Polippu EMI, Oruu EMI, Kulai EMI, Kilkkatuvay EMI, Merkatuvay EMI and Rhyme patterns are further classified into Equal Family, Consonant Family, Vowel Family, Consonant Hardness Family, and Meter Match. The following sub section explains the rhyme schemes in Tamil.

3.1 Rhyme Schemes and Rhyme Patterns in Tamil

In Tamil, the grapheme and phoneme are bound stronger than in English. There are 3 characteristic rhyme schemes in Tamil – Monai, Edhugai and Iyaibu .

Two words are said to rhyme in monai if their first letters are the same, in edhugai if their second letters are the same and in iyaibu if their last letters are the same.

Examples:

புறவை (*paRavai*) and பச்சை (*pachai*) rhyme in monai as they start with the same letter. அருவி (*Aruvi*) and விருப்பு (*viruppu*) rhyme in edhugai as they share the same second letter. யாக்கை (*yaakkai*) and வாழ்க்கை (*vaazhkkai*) rhyme in iyaibu as they share the same last letter.

3.1.1 Rhyme Schemes

Further, rhyme can be sub-categorized according to the position of the letters which are involved in the process. Monai between words across lines have different names, if the first letter of a word matches with the first letter of a word in the subsequent line, then it is classified as *ati-monai*.

Example:

உலகத்தை மறந்தேன் உறக்கத்தை மறந்தேன்
உன்னுடன் நான் ஒன்று கலந்தேன்
ulagathai maRandhaen uRakkathai maRandhaen
unnudan naan ondRu kalandhaen

Monai between different words within a line also has different names. The one existing between the first and second words within a line is *inai monai*.

Examples:

நிலவுக்கு நிலவு சுகம் பெற நினைத்தேன்
nilavukku nilavu sugam peRa ninaithaen

Monai between the first and third word is *polippu monai*.

Examples:

உலகத்தை மறந்தேன் உறக்கத்தை மறந்தேன்
ulagathai maRandhaen uRakkathai maRandhaen

OtRu monai is the one which exists between the first and fourth word.

Examples:

நிலவுக்கு நிலவு சுகம்பெற நினைத்தேன்
nilavukku nilavu sugampeRa ninaithaen

The *kulai monai* is the one between the first, second and third words.

Examples:

உயிரே உயிரே உடம்பில் சிறந்தது எது வென்று தவித்திருந்தேன்
uyirae uyirae udambil siRandhadhu edhu vendRu thavithirundhaen

Monai between the first, second and fourth words is *kilkkatuvay monai*.

Examples:

நிலவுக்கு நிலவு சுகம்பெற நினைத்தேன்
nilavukku nilavu sugampeRa ninaithaen

A *merkativay monai* is that which occurs between the first, third and fourth words.

Examples:

அரும்பும் இனிமையும் அதன்வழி ஆணிகொளும்
arumpum inimaiyum adhanvazhi aNikoLum

If all the four words within a line match, then it is a *mutRu monai*.

Examples:

மன்னவா மன்னவா மன்னாதி மன்னன் அல்லவா
mannavaa mannavaa mannaadhi mannan allavaa

The positional categorization discussed above applies to edhugai and iyaibu also by considering the second and last letter of words in each line accordingly. Here, we also refer this kind of categorization as internal rhyme.

The rhyme can be further sub-categorised based on the nature of the letters. If the monai is due to only the consonants and not the vowels, then it is a *varukka monai*.

Examples:

புதைக்கின்ற விதையும் முயற்சிகொண்டால் தான் மியும் கூட தாழ் றக்கும்
puudhaikkindRa vidhaiyum muyaRsigoNdaal thaan
poomiyum kooda thaazh thiRakkum

If the monai exists only due to vowels, not the consonant in the initial position, then it is a *uyir monai*.

Examples:

பரிவான நண்பன் தந்த
தனிவான தோள்கள் கண்டேன்
parivaana naNban thandha
kanivaana thoaLgaL kaNdaen

Monai between long vowels is called as *netil monai*.

வாலிபத்தில் மன்மதன்
லீலைகளில் மன்னவன்
vaalipaththil manmadhan
leelaiikaLil mannavan

The monai due to the three classes of consonants is called an *ina monai*. Since there are hard class (vall-inam), soft class (mell-inam) and middle class (itai-inam), there exist *valina monai*, *Melina monai* and *itai-ina monai*.

Example for valina monai

திங்கள் என்பது பெண்ணாக
செவ்வாய் கோவை பழமாக
thinggaL enbadhu peNNaaga
seyvvaay koavai pazhamaaga

Example for Melina monai

யாருக்கு ஜானகி காத்திருந்தாள்
வண்ண விழியின் வாசலில் என் தேவன் தோன்றினான்
yaarukku jaanagi kaathirundhaaL
vaNNa vizhiyin vaasalil en thaevan thoandRinaan
நெஞ்சில் ஏதோ ஏதோ நினைத்தேன்
மதுக் கிண்ணத்தை இனி நான் தொட மாட்டேன்
nenjjil aedhoa aedhoa ninaitaen
madhuk kiNNathai ini naan thoda maattaen

The nature of letter sub-categorization discussed above applies to edhugai and iyaibu also by considering the second and last letter of words in each line accordingly. Here, we also refer this kind of categorization as an imperfect rhyme.

3.1.2 Rhyme Patterns

When a lyric is input to the system, rhyme scoring begins by extracting rhyme patterns. Rhyme pattern for each scheme is a string, which reflects the letter occurrences in that rhyming spot of each line in the lyric.

Separate patterns for monai, edhugai and iyaibu are extracted from the input lyric by considering the rhyming scheme as shown below.

Pattern	Rhyme Scheme
Rhyme-AA	Ati EMI
Rhyme- AAAA	Murru EMI
Rhyme-AABB	Inai EMI
Rhyme-ABAB	Polippu EMI
Rhyme-ABBA	Oruu EMI
Rhyme-AAAB	Kulai EMI
Rhyme-AABA	Kilkkatuvay EMI
Rhyme-ABAA	Merkatuvay EMI

Table 1: Rhyme patterns for lyric analysis

Here EMI represents edhugai or monai or iyaibu scheme. It is categorized according to the position of the letter involved and we try to extract all high-level rhyme features for scoring the lyric.

In order to extract the rhyme patterns for each scheme, we need to identify the occurrence of similar syllables and words in successive lines. It need not be identical alone but also based on the nature of letters (where syllable sounds belongs to same category).

Based on the nature of letters, we classify the rhyming patterns as

Rhyme Pattern	Description
Equal Family	Identical letters
Consonant Family	Varukka EMI
Vowel Family	Uyir EMI
Consonant Hardness Family	Valina, Melina, Idai-ina EMI
Meter Match	Netil EMI

Table 2: Classification of imperfect rhyme patterns

Given a line or stanza (of any size), we match each subsequent letter to pattern based on the category as specified in table (ii). The rhyme patterns are extracted for the three rhyming schemes (monai, edhugai, iyaibu) separately for each line or stanza.

Example:

Consider a segment of a sample lyric given below.

நினைவே நினைவே நெஞ்சோடு கலந்துவிடு
நிலவே நிலவே விண்ணொடு கலந்துவிடு
ninaivae ninaivae nenjjoadu kalandhuvidu
nilavae nilavae viNNodu ka

Extract every first, second and last letter of each word. The first letters of first line are *ni, ni, ne, ka* mapping to string under:

Equal Family – AABC
Consonant Family – AAAB
Vowel Family – AABC
Consonant Hardness Family – AABC
Meter Match – AABC

Here we can observe that first two letter belongs to equal family category, while when comparing together with third letter there exists a match in the consonant family.

The second letters of first line are *nai, nai, nj, la* mapping to string under:

Equal Family – AABC
Consonant Family – AABC
Vowel Family – AABC
Consonant Hardness Family – AAAB
Meter Match – AABC

The last letters of letters of first line are *vE, vE, tu, tu* mapping to string under:

Equal Family – AABB
Consonant Family – AABB
Vowel Family – AABB
Consonant Hardness Family – AABB
Meter Match – AABB

Here, we extract the rhyme patterns (EMI) for each family within a line, likewise we have to extract the patterns across the lines (ati) also in the same manner.

3.1.3 Rhyme Scoring Model

The pattern string found out as described above is to be scored by the rhyme scorer. We propose a scoring model which handles high level rhyme features such as imperfect and internal rhyme.

The boundary conditions for our model are listed below:

Let,
 t^c be the Number of letters to be compared;
 t^m be the Number of letters matched in equal family;
 t^m_c be the Number of letters matched in consonant family;
 t^m_v be the Number of letters matched in vowel family;
 t^m_h be the Number of letters matched in consonant hardness family;
 t^m_m be the Number of letters matched in meter match;

w be the Weight assigned for each category (family);
 $EScore$ be the Score of the Equal family category;
 $CScore$ be the Score of the Consonant family category;
 $VScore$ be the Score of the Vowel family category;
 $HScore$ be the Score of the Consonant hardness family category;
 $MScore$ be the Score of the Meter match category;

The number of letters matched represents the character found to be equal in the pattern string. We assign weight to each category, thus the letters matched under equal family will have a higher score than the letters under consonant family and so on.

Compute score based on the pattern string

$$\begin{aligned}
 EScore &= (t^m / t^c) * w^e \\
 CScore &= ((t^m_c - t^m / t^c) -) * w^c \\
 VScore &= ((t^m_v - t^m / t^c) * w^v \\
 HScore &= ((t^m_h - t^m / t^c) * w^h \\
 MScore &= ((t^m_m - t^m / t^c) * w^m \\
 Score &= \sum (EScore, CScore, VScore, HScore, MScore)
 \end{aligned}$$

Where,

$$w^e = 0.5, w^c = 0.2, w^v = 0.2, w^h = 0.05, w^m = 0.05.$$

Let w^e , w^c , w^v , w^h , w^m be the weight assigned for equal, consonant, vowel, consonant hardness and meter match family respectively.

Let's see how it works for the previous given example,

ninaivE ninaivE nenjOtu kalandhuvitu

We consider first letter of each word (*ni, ni, ne, ka*) to calculate the score for monai as below:

$$\begin{aligned}
 EScore (AABB) &= 0.5 \\
 CScore (AAAB) &= 0.225 \\
 VScore (AABB) &= 0 \\
 HScore (AABB) &= 0 \\
 MScore (AABB) &= 0 \\
 Score &= 0.725
 \end{aligned}$$

Edhugai and Iyaibu score will also be computed in the same manner as shown above by considering the second and last letter of a word respectively. The average score of all the three rhyming schemes is the overall rhyme score.

3.2 Pleasantness Scoring Model

The classification of consonant family into hard, soft & mid consonants along with the place of articulation is considered for pleasantness scoring. Pleasantness is computed by applying weights to the different classes of characters. Table 1, presents the weights for pleasantness computing.

Vowels		
	Characters	Score
Short Vowels	அ/a, இ/i, உ/u, எ/e, ஒ/o	0.25
Long Vowels	ஆ/aa, ஈ/ee, ஊ/oo, ஏ/ea, ஐ/ai, ஓ/oa, ஔ/au	0.50
Consonants		
Soft Consonants	ங/ng, ஞ/nj, ண/N, ந/n, ம/m, ன/n	0.5
Medium Consonants	ய/y, ர/r, ல/l, வ/v, ழ/zh, ள/L	0.4
Transformations	க/ga, த/dha, ட/da, ப/ba	0.3

Rough	$\acute{k}/k, \acute{c}/ch, \acute{t}/th, \acute{p}/p$	0.2
Greater Rough	$\acute{t}/t, \acute{r}/R$	0.1

Table 1: Weights for Pleasantness Score

As we can see, long vowels and nasals have higher scores where as hard consonants have a lower score.

Pleasantness of every word is computed using the values provided in the table. The average pleasantness of all words in a song would determine the pleasantness of the lyric.

3.3 Freshness Scoring Model

The word fresh has a temporal aspect associated to it. It is not possible to reason about the freshness of a lyric without knowing its time of creation. We divide the entire timeline of lyrics into small windows of 10 years. Let E_1, E_2, \dots, E_n denote the time windows.

For each era, we find the words from lyrics that were used for the first time in that era. We associate those words to the corresponding era. E_1 words are words that were used in a lyric for the first time in E_1 . Stopwords are removed from all the lists.

The freshness score for a new lyric is determined by the following algorithm.

Let,

l be the input lyric.

n be the number words in l

W be the set of words in l

E_i be the i^{th} time window.

$E(w)$ be the time window of the word w .

$E(l)$ denote the time window of the song.

F_{wi}^S be Freshness score of a word i in the lyric.

$F_s(l)$ be the Freshness score of a lyric l

Freshness Score :

begin

initialize ScoreSum := 0;

for each w_i in W

$F_{wi}^S := E(w_i) / E(l)$

ScoreSum := ScoreSum + F_{wi}^S

end for

$F_s = \text{ScoreSum} / n;$

end

The basic idea behind the algorithm is that if a word is created in that era and used in the same era, it has the maximum score of 1. If it is used in the next era, the score becomes half.

4. Results and Evaluation

The Rhyme, Pleasantness and freshness models were presented to a group of linguists, lyricists and language experts who were asked to judge the rhyming, pleasantness and freshness style of a list of lyrics. It is observed that the rhyme score gets a positive correlation between the manual judgment and machine judgment.

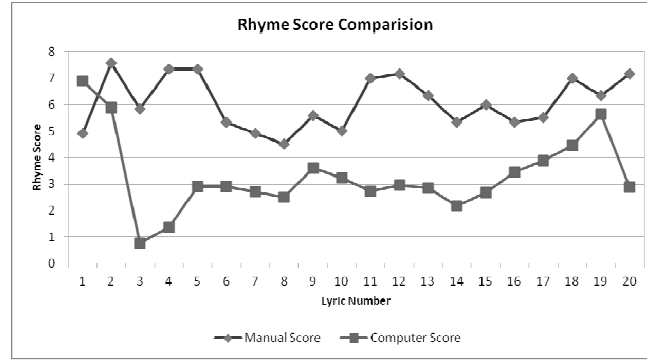


Figure 1: Rhyme Score Comparison

In the Figure 1 the x axis represents the songs and the y axis the Rhyme score yielded by the scoring model and human judgment. The correlation coefficient between scoring model and human judgment is 0.00914.

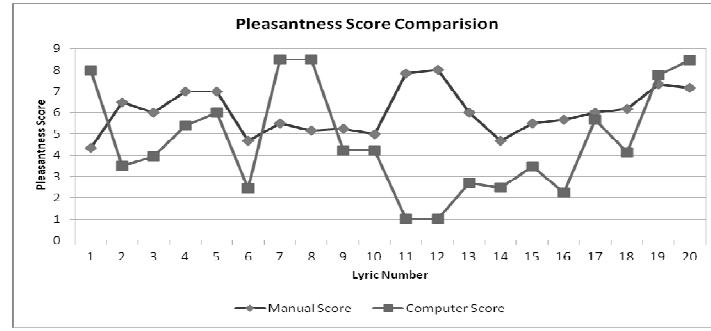


Figure 2: Pleasantness Score Comparison

In the Figure 2 the x axis represents the songs and the y axis the Pleasantness score yielded by the scoring model and that by human judgment. The correlation coefficient between scoring model and human judgment is -0.1317. Pleasantness does not have a strong correlation between human judgment and computer score. The reason could be that the human judges consider songs they are familiar with to be more pleasant, without regarding the pleasantness of individual words.

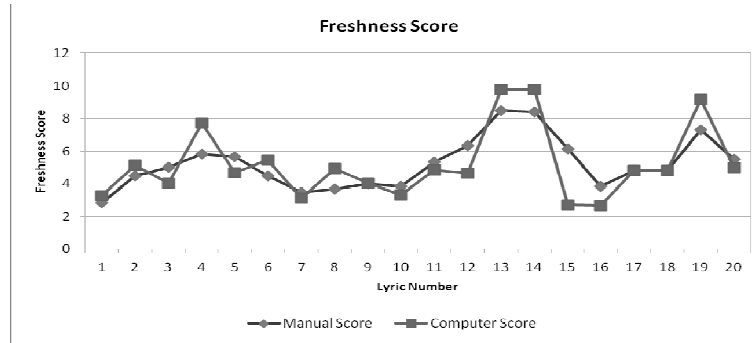


Figure 3: Freshness Score Comparison

In the Figure 3 the x axis represents the songs and the y axis the freshness score yielded by the scoring model and human judgment. The correlation coefficient between scoring model and human judgment is 0.99374. Freshness has a strong correlation between human and machine scores.

5. Conclusions and Future Work

In this paper, we proposed three scoring models for rhyme, pleasantness and freshness of a lyric. The computed scores are being used by Paadal [6] a lyric portal, for visualizing the lyrics and in advanced searches. Studying the pleasantness of words from human perspective is one area we plan to investigate. Though the proposed rhyme scoring model handles the case of internal and imperfect rhyme features, assigning different weights to different categories of rhyme styles may vary the

score. An analysis of what could be a better weight allocation would make an interesting study in the future. Adjusting the weights of rhyme score components and reducing the time window size in freshness scores will open ways for further research in this area.

6. References

- Hamid R. Tizhoosh, Farhang Sahba, Rozita Dara, Poetic Features for Poem Recognition: A Comparative Study.
- R. Mayer, R. Neumayer, A. Rauber: “Rhyme and Style Features for Musical Genre Classification by Song Lyrics, ” Proceedings of the International Conference on Music Information Retrieval, pp. 337–342, 2008
- Sowmiya Dharmalingam., Madhan Karky. “LaaLaLaa - A Tamil Lyric Analysis and Generation Framework” in World Classical Tamil Conference – June 2010, Coimbatore.
- Suriyah M, Madhan Karky, T V Geetha, & Ranjani Parthasarathi “Special Indices for LaaLaLaa Lyric Analysis & Generation Framework” in Tamil Internet Conference – August 2011, Pennsylvania.
- H. Hirjee and D.G. Brown: “Using Automated Rhyme Detection to Characterize Rhyming Style in Rap Music, ” Proceedings ISMIR 2009, 2009.
- HYPERLINK Paadal – Lyric Portal, <http://www.karky.in/labs/paadal>, last accessed 8th July 2013

Modeling basic emotions for Tamil speech synthesis

A G Ramakrishnan and Lakshmi Chithambaran[#]

Department of Electrical Engineering, Indian Institute of Science, Bangalore

[#]Department of Electrical Engineering, M S Ramaiah Institute of Technology, Bangalor

Abstract: This paper explores the modeling of prosody parameters for improving naturalness of Tamil speech synthesis, by studying recorded utterances of speech with and without explicit emotions. To begin with, we look at interrogative and exclamatory Tamil sentences. Prosody parameters namely, pitch contour, energy and duration of each word in the sentences were observed, analyzed and generalized from the un-intonated and intonated, interrogative and exclamatory human speech. Differences in energy level were also analyzed between the two sets of utterances in three different frequency bands. Pitch is by modified in the LP residual domain using DCT. Energy is modified by multiplying the signal by the hypothesized factor and duration is modified as per the duration model by duplicating or removing integer number of pitch periods as necessary. The model was implemented on speech synthesized from Thirukkural TTS, developed by MILE LAB, and the results were found to be satisfactory.

Keywords: TTS, interrogative, exclamatory, pitch, energy, durations, LP, DCT, pitch synchronous, pitch modification, pitch contour, Gaussian modeling.

Introduction

It is known that a listener gets fatigued by listening to synthesized speech for a length of time [2]. This is fundamentally because text to speech conversion systems use normative speech from a native speaker to obtain their basic units for concatenation. Thus, it lacks variations in pitch, duration and energy level, which together are known as prosody [2, 3, and 4]. Further, a human being, even when she/he repeats the same sentence twice, the utterances are not identical, and there are always minute variations in stress levels and local amplitudes and duration. This lack of natural variations in speech attributes makes synthesized speech monotonous [3].

This triggers the need for modeling prosody in synthesized speech. Modeling prosody in languages like French, Spanish and English has a wide research literature. Modeling prosody in Indian languages like Tamil is not a trodden path. Here, interrogative and exclamatory intonations are considered. Prosody parameters are generalized for each of the above and appropriately modeled and implemented to speech synthesized by Thirukkural TTS, developed by MILE LAB.

Modeling interrogative intonation

Pitch contour: Figure 1 shows sample utterances of an interrogative sentence, both with and without intonation along with their pitch contours. It is noticed that the pitch contours of specific interrogation loaded words form a curve, with a rise and fall of pitch relative to the basal value. The mean ratio of maximum to basal pitch is 1.4. The question indicative word may fall in any position of the sentence: starting, somewhere in the middle or last. When the interrogative word (IW) is in the start or middle, it forms the rising part of the curve and the next word forms the falling part. Whereas, if the IW is the last word of the sentence, then it forms the falling part and the previous word forms the rising part of the curve. There are also interrogative sentences in Tamil, which do not have a specific IW; rather they end with the long vowel /aa/. In this case, the pitch significantly rises during the /aa/ sound.

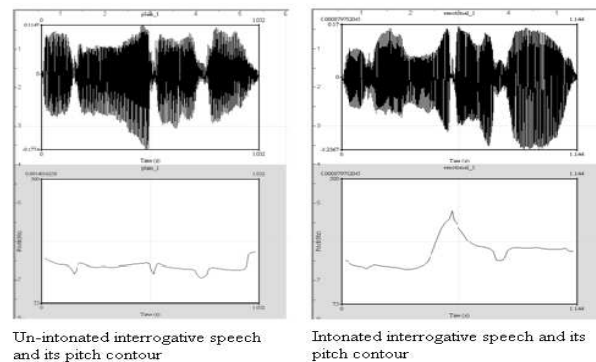


Figure 1: Un-intonated and intonated interrogative speech and their respective pitch contours.

Energy: The energy and duration of the intonated interrogative utterance are higher by an average of 17.5% and 19.7%, respectively, than those of the un-intonated one.

Duration: The durations of the words that fall within the rise and fall region of the pitch contour also follow a distinct pattern. On the average, the duration of the first word in the region is 31.5% less than the un-intonated word, whereas that of the second word is 32% more.

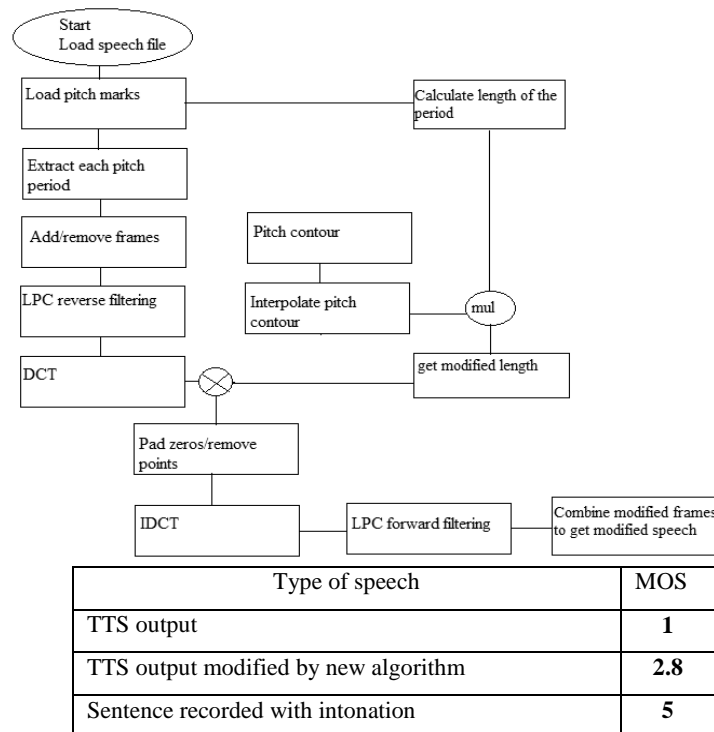
Pitch modification is carried out pitch synchronously by interpolation of LPC residues of each pitch period in the DCT domain. Energy modification is accomplished by multiplying the signal by the hypothesized factor. Durational is modified by duplicating or removing integral number of periods. Tables 1 and 2 compare the mean opinion scores of five natives on un-intonated human speech and TTS speech, respectively before and after the prosodic modifications.

Figure 2: Block diagram implemented for pitch modification as per hypothesis.

Table 1: Comparison of MOS (on a scale of 5) of un-intonated human speech after modification.

Type of speech	MOS
Sentence recorded without intonation	2
Un-intonated recording modified by our algorithm	3.8
Sentence recorded from a human with intonation	5

Table 2: Comparison of MOS of TTS speech after prosody modification.



Contour Analysis: It is noticed that noise in the modified speech is due to the random pitch contour present in the TTS speech. To nullify this effect, pitch contour of the TTS speech is analyzed and if the contour by itself satisfies the hypothesized values within an error of $\pm 5\%$, then the TTS is not processed at all; else, the TTS speech is processed as discussed.

Table 3: Comparison of MOS of TTS speeches as per contour analysis

Type of speech	MOS
TTS output	1
TTS output modified by our new algorithm	3.5
Sentence uttered with intonation	5

Energy Distributions: The fractions of energy in the bands, 0-500 Hz (low frequency band), 500-2000 Hz (mid band) and 2000-8000 Hz (high band) of un-intonated and intonated speech are analyzed. For the first word that falls within the curve, the fractions of energy in low and mid bands of intonated speech are higher on an average by a factor of 3.2 than their un-intonated counterparts. For the second word, the energy in mid frequency band of intonated speech is higher by a factor of 5.43 than their un-intonated counterpart. There is no significant change in the high frequency band for either word. Therefore, by scaling the energies in the respective frequency bands only, we obtain better intonation and clarity of words in the modified speech.

Table 4: Comparison of MOS of TTS speeches as per energy distribution analysis.

Type of speech	MOS (emotion)	MOS (clarity)
TTS output only with pitch modification (TOPM)	3	3.3
TOPM and increase of energy of the signal	3.6	3.6
TOPM and scaling of energy only in specific bands	3.5	3.8

Modeling exclamatory intonation:

It is noticed that the pitch contour of the last two words in the sentence takes a rise and fall pattern. The mean ratio of maximum to basal pitch is 1.8.

Figure 3: Un-intonated and intonated exclamatory speech and their respective pitch contours.

The energy and duration of the intonated exclamatory sentence are higher by an average of 19.3% and 16.4% than those of un-intonated one, respectively. The pitch, energy and duration are modified in the same way as the interrogative case. Table 5 shows the evaluation of the results of the modifications.

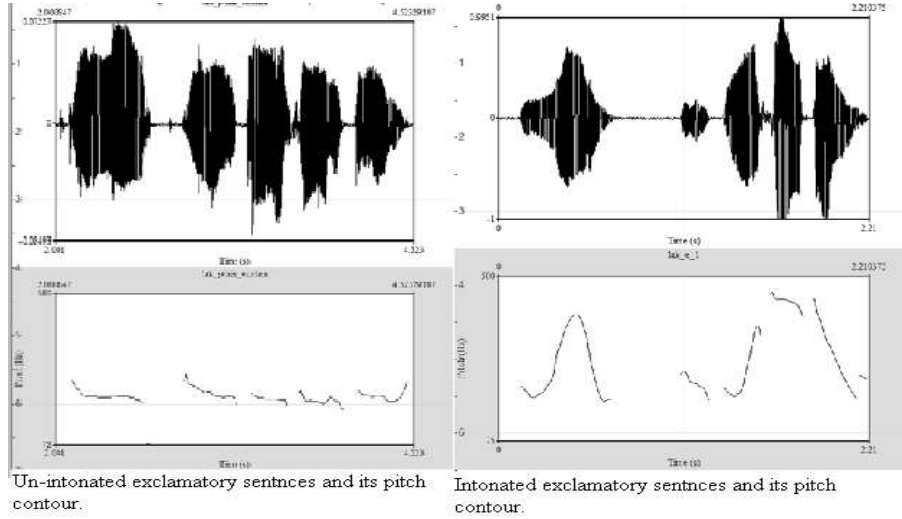


Table 5: Comparison of MOS of un-intonated exclamatory speeches.

Type of speech	MOS
Sentence recorded without intonation	1
Un-intonated recording modified by our algorithm	4
Sentence recorded with intonation	5

Energy in specific bands: As discussed for the interrogative case, energy in different frequency bands are analyzed for exclamatory intonation as well. For the first and second words within the pitch curve, the energies in mid band of intonated speech are 6.25 and 5.4 times those of un-intonated speech, respectively. Change in high frequency band is not significant for both words. Improvement in the quality of processed speech is evaluated and listed in Table 6.

Table 6: Comparison of MOS of TTS speeches as per energy distribution analysis.

Type of speech	MOS (emotion)	MOS (clarity)
TTS output only with pitch modification (TOPM)	3.75	3.5
TOPM and increase of energy of the signal	4.25	3.75
TOPM and scaling of energy in specific bands	4.2	4

Results and Discussion

Prosody of interrogative and exclamatory sentences have been modeled. Modification of the pith, energy and duration of the TTS generated speech gives interrogative or exclamatory intonation, as evaluated by the native listeners. Energy modification in certain frequency bands gives better quality and intended intonation closer to expected natural intonation.

Acknowledgment: The authors thank Mr. Shiva Kumar HR for creating the MILE TTS web demo, used for the experiments reported in this paper.

References

- Fabio Tamburini, "Prosodic prominence detection in speech," Proc. VII ISSPA 2003.
- G.L.Jayavardhana Rama, A.G.Ramakrishnan, M.Vijay Venkatesh, and R.Muralishankar, "Thirukkural - a text-to-speech synthesis system," Proc. Tamil Internet 2001, Kuala Lumpur, August 26-28, 2001, pp. 92-97.
- R. Muralishankar and A. G. Ramakrishnan, "Human touch to the Tamil Synthesizer," Proc. Tamil Internet 2001, Kuala Lumpur, August 26-28, 2001, pp. 103-109.
- R Muralishankar, A.G.Ramakrishnan and P Prathibha, "Modification of pitch using DCT in the source domain," Speech Communication, 2004, Vol. 42/2, pp. 143-154.
- R. Muralishankar and A. G. Ramakrishnan, "Synthesis of speech with emotions," Proc. Intern. Conf. Commn., Computers and Devices, Kharagpur, Dec. 14-16, 2000, pp. 767-770.
- G. L. Jayavardhana Rama, A. G. Ramakrishnan, R. Muralishankar and P. Prathibha, "A Complete Text-to-Speech Synthesis System in Tamil," Proc. IEEE 2002 Workshop Speech Synthesis, Santa Monica, CA USA, Sep. 11-13, 2002, pp. 191-194.
- A. G. Ramakrishnan, Lakshmish N Kaushik and LaxmiNarayana M, "Natural Language Processing for Tamil TTS," Proc. 3rd Language and Technology Conference, Poznan, Poland, Oct 5-7, 2007, pp. 192-196.
- L. R Rabiner and R. W Schafer "Digital Processing of Speech Signals", PHI, 1978.
- Web demo of MILE TTS: <http://mile.ee.iisc.ernet.in/tts>

Conceptual Based Search Engine (Cbse) System For Tamil And English

Saviya shree K V, Project Associate, Department of IST
Umamaheswari E, Research Scholar, Department of CSE
Balaji J, Research Scholar, Department of CSE
Dr. T V Geetha, Professor, Department of CSE
Dr. Ranjani Parthasarathi, Professor, Department of IST
TaCoLa Lab, CEG, Anna University, Chennai

ABSTRACT

Currently Internet provides information in several languages. In order to process information across language and retrieve that information in their own language becomes essential. In order to retrieve information available in different languages cross language information retrieval (CLIR) is the need of the hour, and there is a need to develop systems that can access information from different languages. This paper discusses the design of a Conceptual based search engine (CBSE) for Tamil and English languages. Our system focus on dictionary based approaches and some strategies to handle ambiguities at the different levels of our system. Initially our search system was designed as a Concept based search system for Tamil language and in this paper we have adopted the same approach for English language in which it retrieves documents in both Tamil and English based on the query entered. Our proposed model involves query expansion, enconversion and concept based indexing and searching and ranking. In this work we have also analyzed the pros and cons involved in adopting a uniform approaches for both a morphologically rich languages like Tamil and also fixed order structured languages like English. We have tested our results for Tourism domain with 50, 690 document corpus (25, 690 Tamil documents+ 25, 000 English documents) and achieved 0.51 of precision for both Tamil and English queries (20 queries in Tamil and 20 queries in English) .

Keywords: Concept-Based search engine, Concept-Based Indexing, Enconversion, Query expansion, Universal Networking Language.

1. INTRODUCTION

In recent years, the growth of Web contents in different languages increases drastically. User prefers to search in their native languages and different languages require different translation mechanism to provide multilingual search facility. Though there exist many machine translation systems, they do not consider the semantic of the document and it is suitable only for specific languages. Hence there is a need for a search engine which can tackle both multilingual and conceptual. Conceptual search engine can tackle both semantics and context of the word in the document which provides meaningful results to the user. Keyword based search engine[10][11][12] retrieves results based on the keywords present in the document. Similarly ontology based search engine[13][18] uses hypernymy/hyponymy based ontological relation to know the surrounding context of the word. But considering only hierarchical relation may not be suitable for application such as tourism, news mining system, e-learning, question-Answering etc., In order to overcome the restrictions present in the Keyword based search engine and other semantic based search engines, this Conceptual search engine plays an important role in handling large amount of unstructured data and this search engine deals with the conceptual meaning of a query with the help of semantic representation called Universal Networking Language (UNL).

The UNL representation provides common semantic representation for any languages and does not annihilate the meaning of the document. It can provide both word level semantics and sentence level and also document level semantics. Hence compared to other machine transliteration based bilingual search and rank[16], our approach provides better results without degrading the efficiency of the ranked documents.

Previous multilingual search engine called agroExplorer [2] uses concept only match and it does not consider the link between the concepts. For example for a given query “Mahabalipuram Temple”

it will consider only the concepts “Mahabalipuram” or “Temple”. Our search system will provide result based on the CRC “Mahabalipuram Temple” and also for the concepts as “Mahabalipuram” or “Temple”. Moreover they are not utilizing query expansion which can retrieve conceptually similar documents. Hence ranking approach uses concept match and weight the documents based on document specific properties (Term frequency, Position weight, Inlink and outlink). It is applied for English, Hindi, Marathi and Spanish.

The UNL semantic representation helps us to adopt the same search and ranking mechanism [7] for morphologically rich languages like Tamil and structured languages like English. The resources required for our bilingual search are UW (Universal Word) dictionary, Index based query expansions with semantic representation, Machine Translation (Enconversion), Conceptual search and Rank and finally summary generation for both Tamil and English documents. Here it uses a rule based method for UNL based machine translation system for both Tamil and English, However the dictionary representation is same for Tamil and English documents. We have changed the Morpho-Semantic rules [1] according to the parse tree structure built for the English documents.

The following sections talk about the related work and its methodology in detail. Section 2 compares previous approaches with the proposed approach and Section 3 elaborates each modules of our bilingual search system and in Section 4 we have shown the accuracy of the search results for both Tamil and English documents.

2. RELATED WORK

AgroExplorer[2] recommends to use parser because it will provide lexical information and useful to reduce noisy content. The SQL database[2] is not suitable to process graph based structures like UNL compared to lucene index and We have also tested to store in SQL database but it requires more space and Time complexity is high. When the size of the document or corpus gets increased it is very difficult to perform the process of the indexing[5]. Our system combine both conceptual features and document specific features to know the importance of the concepts in the document, the efficiency of the ranked documents is improved compared to previous UNL concept based search and Rank[7]. We have shown this in result evaluation section 4. The UNL representation provides common semantic representation for any languages and does not annihilate the meaning of the document. It can provide both word level semantics and sentence level and document level semantics. Hence compared to transliteration based bilingual search and rank[16], uses transliteration for bilingual documents and identifies overlapping between two languages. Since transliteration does not cover all the words and the number of overlapping between two documents will be only named entities hence our approach provides better results without degrading the efficiency of the ranked documents . Query expansion[20] uses only dictionary based query expansion and uses transliteration to provide multilingual feature which will lead disambiguation problem. Our approach uses index based query expansion which is retrieved based on the important domain specific concepts which occurs on the top of the documents and more frequently and it uses various levels of filters in order to filter the required documents.

3. METHODOLOGY

As discussed in section 1, our system composes UNL enconversion document processing, UW dictionary[15] as a language independent representation, Index based query expansion[9] to aid conceptual search and rank[7], domain specific template based summary generation[8] for output generation and Conceptual indexing[14] to aid search and rank.

3.1 System Architecture

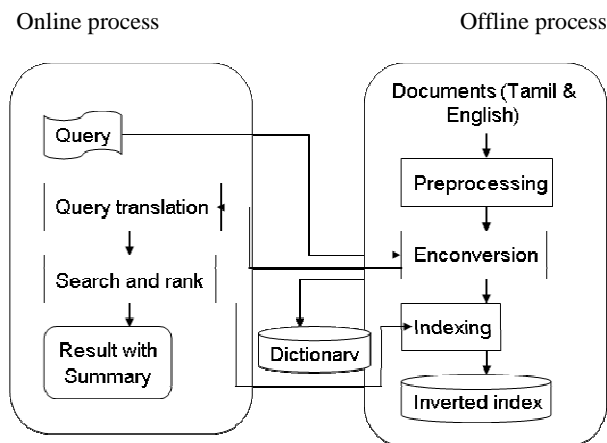


Figure 1: Architecture of Conceptual based search systems for Tamil and English

Our proposed system consists of two segment namely online process and offline process. The offline part consists of preprocessing, enconversion, indexing which is a background process and the online part consists of Searching and Ranking, Query translation and query expansion and this is an active process because it is occurring currently.

The natural language text document which consists of sentences is passed to the enconversion process which is transformed as a UNL graph. The information present in our UNL graph is stored in the Lucene index. When the query is given to the search process, the query gets enconverted using the enconversion process and it searches with the inverted index created. This inverted index consists of two indices CRC indices and C indices which are used for searching process.

The tourism specific documents are finely tuned and preprocessed and this outcome is given to the enconversion process.

3.2 Document Processing

Manning and Schutze[17] discusses four approaches in the Machine translation Technology. They are Word-for-word approach, Syntactic transfer, Semantic transfer approach and Interlingua approach. In our work we are using Semantic transfer approach and Interlingua approach. Interlingua approach is known as a Knowledge representation model in which several techniques available to represent this knowledge such as lists, trees, Semantic networks, schema etc., In this our proposed model follows the Semantic networks which can handle any languages irrespective of its structure.

UNL is a Universal networking language, follows the Interlingua approach which is modeled to represent semantic data extracted from natural language and plays a vital role in the information retrieval applications which requires both semantic and context.

This proposed method can handle both structured languages like English and morphologically rich languages like Tamil. Structured language follows the order of “subject + verb + object” and in morphologically rich language like Tamil the Subject and Object can occur in any order. For instance one example sentence in English as Saranya went to chennai and the same sentence in Tamil can be written as charaNyaa chennaikku chenRaaL (சரண்யா சென்னைக்கு சென்றாள்), chennaikku charaNyaa chenRaaL (சென்னைக்கு சரண்யா சென்றாள்). Hence in Tamil the order of Subject+Verb+Object does not follow any order and gives same meaning for all this sentences. But the semantic representation for all Tamil sentences and English sentences are same which is given in example 1.

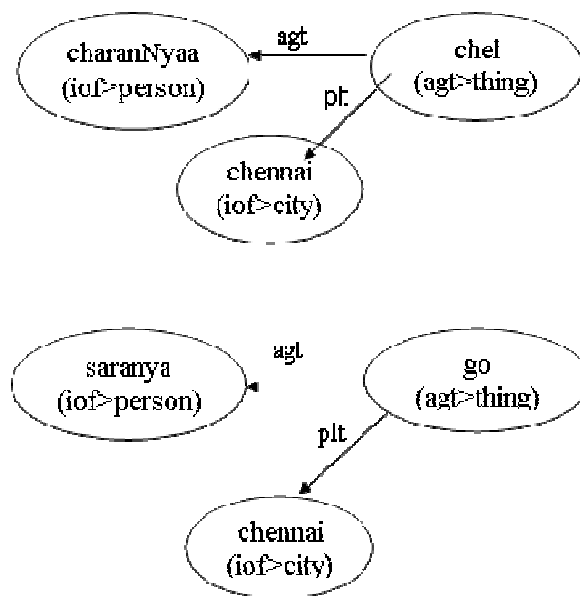


Figure 2: UNL graph representation for Tamil and English sentences

The above example depicts Tamil and English sentences which have same meaning. Universal words represent the concepts and links represents the relation between the concepts.

In the above example saranya, go, chennai are the concepts. agt, plt are the links associated with this concepts. This UNL representation is our background knowledge based system to identify the semantic context of the words present in the document.

There is a separate dictionary for both Tamil and English with same format. Universal Word dictionary consists of the word, headword, constraint, concept id, term id. Multi word Dictionary consists of following fields such as word, First word, Constraint, number of the term present in the multiword, term id, concept id.

Word sense disambiguation problem arises due to the disambiguity in the sense of meaning if a particular word has multiple meaning. For examples in tamil “aaru” represents river and number based on the context it is identified as either it represents river or number. Rule based word sense disambiguation has been introduced to tackle this type of problem for both Tamil [20] and English. Pronoun resolution are also handled for Tamil language.

There are 58 rules for Tamil languages and 43 rules for English languages.

Initially rule based enconversion process for free order language like Tamil was developed by Balaji et al., [1] were taken into contemplation and new rules have been developed for fixed structured language like English. Based on the rules the UNL graph has been constructed.

The following parameters are present in the UNL graph.

Parameters	Descriptions
Term word	The word present in a document.
Head word	Root word of the Term word present in a document.
Constraints	UW constraint of a term word associated with the head word.
Part of speech	Lexical information of a Term word.
Term id	ta_1, ta_2....ta_n represents Tamil term word. en_1, en_2...en_n represents English term word.
Concept id	1, 2... n represent Concept which is used in for searching process.
Term number	This number is based on the sentences.

Table 1: Parameter of the Enconversion

The above set of parameters helps in identifying the importance of concepts in the document inorder to aid in ranking process. Here term id and concept id used to differentiate concept and term based results. Part of speech is used to provide domain level importance for the concepts. Constraints are used to find the context of a word. Term number is used to identify the direction of the concepts associated with other concepts.

3.3 Indexing

This concept based index [14], is an inverted index which maps from concepts to the documents. Usually inverted index contains terms and documents, In our search system, it contains two separate indices such as Concept-Relation-Concept indices (CRC Indices) and Concept indices (C Indices). CRC indices have entire relation between the concepts. C indices represent only the concepts.

For example Chennai maRRum madurai. In this chennai (C2) and madurai (C1) are the concepts and relation between the concepts is “and”. Chennai maRRum madurai will present in the CRC indices.

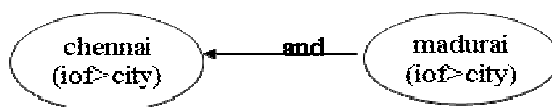


Figure 3: Example for Concept-Relation-Concept (CRC) present in the CRC index

The inverted CRC indices will contain relation (R), Concept1 (C1), Concept2 (C2), term id, concept id, Part of speech tag. chennai and madurai were present in C indices. The inverted C indices consist of concept, Part of speech Tag, concept id, term id.

Each UNL index in the UNL indexer maintains a table of properties such as Language identifier, document identifier, Sentence identifier, Concept Name, Concept ID, Relation Name, term, term frequency, concept frequency and position weight of the concepts and terms. All the above said properties and its associated concepts and terms are stored in Lucene index structure.

The field description is given in Table 2.

Fields	Descriptions
Concept-relation-concept	The relationship between the Concept (C1)-Relation (R)-Concept (C2) belongs to the document set D.
Concept	The list of concepts (Cs) occurred in the document set D.
Language identifier	It uses “ta”, “en” which is used to identify Tamil and English languages of the document set D.
Document Identifier	d1, d2, d3,dn corresponds to list of documents which contains both CRC indices and C indices
Sentence Identifier	S1, S2, S3,Sn corresponds to list of sentence identifier which contains both CRC indices and C indices. This is used to know the position of the term/concept in the document.
Position weight	Concept Position weight
Frequency count	Frequency of occurrence of the concepts in the document
Part Of Speech Tagging	It is used to know the importance of the concepts with respect to the domain of the interest. For example for tourism domain “Named Entities” are more important than other nouns.
Term words	It represents term words present in the document

Table 2: Fields of the indexing

Lucene index is used to capture the above said information present in the document represented in UNL graph. This UNL graph consists of information such as Concepts (C) and Concepts-Relation-Concepts (CRC). UNL Indexer parses the UNL graphs and builds an inverted list on the indices and later this is used in the Search process. These indices are categorized into three different types such as C-R-C (Concept -Relation- Concept) Indices and C (Concept Only) indices, .

In order to provide competent searching and storing of information, the lucene index is separated in to C-R-C, and C index. Inside this we also separate the index with respect to concept id which helps us to reduce search time. These indices are used by both query expansion and search and ranking module which is described in section 3.3 and in 3.4.

3.4 Query Processing

This module consists of both query expansion and translation/enconversion.

Query Translation/ Enconversion

For a given Tamil or English query UNL based semantic representation is given by using light weight UNL enconversion and it also performs light weight Word Sense Disambiguation. The representation of query UNL graph is given below.

For example if the user entered the query “Guindy Park” the query gets enconverted using the enconversion process. Based on the results generated it was given to the query expansion process.

Query Expansion

Our approach uses index based query expansion [9] which is retrieved from the CRC indices and C Indices. Here we are not searching all the expanded concepts, instead important domain specific concepts which occur frequently and also on the top of the documents are given more importance compared to the other. We additionally weight the expanded terms based on its POS tags such as both the concepts in the CRC have Named Entity as Part of speech tag, Number of links, mod/pos relations and frequency of occurrence of the concept/concept-relation-concept with and also across the document. According to the above said features, this proposed approach takes only the top 10 expanded terms.

For example the query gets expanded based on the CRC present in the inverted index.

3.5 Search and Rank

Previous approach on bilingual search and rank [16] uses transliteration for bilingual documents and identifies overlapping between two languages. Since transliteration does not cover all the words and the number of overlapping between two documents will be only named entities.

In this approach we are ranking the documents at three levels, such as Degree of match between query graphs with index graph, concept association based match and Index based match [14]. Previous approach [7] has been extended according to the language preference of the user. If the user enters a query in Tamil, Tamil documents will be given more preference than English. Hence Tamil documents is ranked first if they enter Tamil query else English document is ranked. The snapshot for both Tamil and English query is given below.

Language can be selected based on the user interests.

3.6 Summary

The summary[8] is generated using the tourism specific templates. Subalalitha et.al., (2011)[8] gives the details of summary generation framework. Summary is generated for both Tamil and English. For a given Tamil document it will generate both Tamil summary and English summary and same for English documents.

4. ULT EVALUATION AND CONCLUSION

This search system has been evaluated for tourism domain with 50, 690 documents for 20 Tamil queries and for 20 English queries. We have achieved 0.51 of precision for both the languages. For some of the queries we have identified semantic constraint mismatch between the Tamil and English UW dictionary. As discussed earlier we maintain separate UW dictionary for both Tamil and English. The concepts which are similar in both the dictionary should have same semantic constraint and head word, in order to identify conceptually similar results for both Tamil and English. But due to ambiguous entries in any one of the UW dictionary (Both in Tamil and English) may lead to retrieval of different documents for the same concept. This needs to be rectified by an automation tool to identify the similar concepts with different constraints. Spelling variations needs to be tackled to get conceptually similar documents for both Tamil and English.

5. REFERENCES

- Morpho-Semantic Features for Rule-based Tamil Enconversion:: J Balaji, T V Geetha, Ranjani Parthasarathi, Madhan Karky
- AgroExplorer: a Meaning Based Multilingual Search Engine:: Mrugank Surve, Sarvjeet Singh, Satish Kagathara, Venkatasivaramasastry K, Sunil Dubey, Gajanan Rane, Jaya Saraswati, Salil Badodekar, Akshay, Iyer, Ashish Almeida, Roopali Nikam, Carolina Gallardo Perez, Pushpak Bhattacharyya, AgroExplorer Group Media Lab Asia, IIT Bombay. {mrugank, dubey}, pb@cse.iitb.ac.in
- Using Cross-Language Information Retrieval Methods for Bilingual Search of the Web:: Sung J. Shim Seton Hall University shimsung@shu.edu
- Cross-Language Information Retrieval Via Dictionary-based and Statistics-based Methods:: Fatiha SADAT t, Akira MAEDA tt, Masatoshi YOSHIKAWA * and Shunsuke UEMURA + Graduate School of Information Science, Nara Institute of Science and Technology (NAIST) 8916-5 Takayama, Ikoma, Nara 630-0101. Japan * National Institute of Informatics (NII) CREST, Japan Science and Technology Corporation (JST) E-mail: (fatia-s, aki-mae, yosikawa, uemura)@is.aist-nara.ac.jp
- Multilingual Single Document Keyword Extraction for Information Retrieval :: David B. Bracewell, Fuji REN, and Shingo Kuriowa Department of Information Science and Intelligent Systems, Faculty of Engineering, The University of Tokushima Tokushima, 770-0861 Email: davidbgis.tokushima-u.acj
- Bilingual Search Engine Augmented with Query Expansion:: Nuttanart Facundes and Ratchaneekorn Theva-aksorn
- A multilevel UNL concept based searching and Ranking:: Umamaheswari E, T V Geetha, Ranjani Parthasarathi, Madhan Karky
- Template based Multilingual Summary Generation:: Subalalitha C.N, E.Umaheswari, T V Geetha, Ranjani Parthasarathi & Madhan Karky
- CoRe – Concept Based Query Expansion:: Elanchezhian K, T V Geetha, Ranjani Parthasarathi & Madhan Karky, Tamil Internet Conference, Coimbatore, 2010.

- DBXplorer: A System for Keyword-Based Search over Relational Databases:: Sanjay Agrawal, Surajit Chaudhuri, Gautam Das
- KEYRY: A Keyword-Based Search Engine over Relational Databases Based on a Hidden Markov Model:: Sonia Bergamaschi, Francesco Guerra, Silvia Rota, and Yannis Velegrakis Universit di Modena e Reggio Emilia, Italy University of Trento, Italy velgias@disi.unitn.eu
- Minersoft: A Keyword-based Search Engine for Software Resources in Large-scale Grid Infrastructures:: Marios D. Dikaiakos, Asterios Katsifodimos, George Pallis Computer Science Department, University of Cyprus, Nicosia, CY 1678, Cyprus
- Ontology-based Interpretation of Keywords for Semantic Search:: Thanh Tran, Philipp Cimiano, Sebastian Rudolph and Rudi Studer Institute AIFB, Universität Karlsruhe, Germany {dtr, pci, sru, rst}@aifb.uni-kalsruhe.de
- A Concept Based Semantic Indexing Technique. SWM -08 CoReX:: Subalalitha, T.V.Geetha, Parthasarathi, R., and Karky, M .2008.
- UNDL. 2009. Universal networking digital language. <http://www.undl.org/> Online; accessed 28 September 2009 .
- A Minimally Supervised Approach for Detecting and Ranking Document Translation Pairs:: Kriste Krstovski and David A. Smith.
- Foundations of statistical natural language processing. MIT Press, 1999:: C.D.Manning, H. Schütze, and MIT Press.
- SEWISE : An Ontology-based Web Information Search Engine:: Georges Gardarin, Huaizhong KOU, Karine Zetourni, Xiaofeng Meng, Haiyan Wang PRiSM Laboratory, University of Versailles Saint-Quentin 45 Etats-Unis Road, 78035 Versailles, France and Departement of computer science, Remin University of China, Beijing.
- Query Expansion Techniques for the CLEF Bilingual Track:: Fatiha SADAT, Akira MAEDA, Masatoshi YOSHIKAWA and Shunsuke UEMURA Graduate School of Information Science, Nara Institute of Science and Technology (NAIST)
- On Tamil Word Sense Disambiguation using UNL representation:: Balaji Jagan, Geetha T V, Ranjani Parthasarathi and Madhan Karky, Anna University

Morphological Analyzer for Classical Tamil Texts - A Rule based approach: special Reference to Plural Markers

1. *R.Akilan, Programmer, Central Institute of Classical Tamil, Chennai.*

E-mail: akilan.rp@gmail.com

2. *Prof E.R.Naganathan, Head of Department, Department of Computer science and Engineering, Hindustan University, Chennai. E-mail: ernindia@gmail.com*

3. *Dr. G.Palanirajan, Associate, Central Institute of Classical Tamil, Chennai.*

E-mail: matrixpalani@gmail.com

Abstract

This paper describes the works to build a Morphological Analyzer for Classical Tamil texts using Rule-based approach. The rule-based approach has successfully been used in developing many natural language processing systems. Systems that use rule-based transformations are based on a core of solid linguistic knowledge. The linguistic knowledge acquired for one natural language processing system may be reused to build knowledge required for a similar task in another system. The advantage of the rule-based approach over the corpus-based approach is clear for: 1) less-resourced languages, for which large corpora, possibly parallel or bilingual, with representative structures and entities are neither available nor easily affordable, and 2) for morphologically rich languages, which even with the availability of corpora suffer from data sparseness. Morphology is the study of internal structure of the word. Morphological analysis is a process of segmenting words into morphemes and a process of analyzing the word formation. Morphological analyzer is a tool for any type of Natural Language Processing work. It is a computer program which takes words as input and produces its grammatical structure as output. It identifies and segments the words and assigns the grammatical information. Capturing the agglutinative structure of Tamil words by an automatic system is a challenging job. This paper is going to reveal a rule-based approach for classical Tamil texts

Introduction

Natural Language Processing (NLP) is a computerized approach to analyze the text based on a set of theories and set of technologies. And, being a very active area of research and development, the basic objective of Natural Language Processing is to facilitate human-machine interaction through the means of natural human language.

Morphological analysis of a word is the process of segmenting the word into component morphemes and assigning the correct morphosyntactic information to the morpheme. For a given word, a morphological analyzer (MA) will analyze and return the word and the word class along with the other grammatical information depending upon its word class. MA returns all possible parse for a given word, without considering the context. MA is a very essential tool for languages having rich inflectional and derivational morphology such as morphologically rich languages like Dravidian languages.

Morphological Analyzer is a vital tool in NLP applications. In morphological rich languages, as there are multiple affixation, the finer grammatical information which helps in building efficient NLP applications, can be obtained only from Morphological Analyzer. Morphological Analyzer is required in most of the applications such as information extraction, QA system, machine translation and spell checker. There are several approaches attempted for Morphology for Tamil. We present a methodology for morphological analysis of Tamil, a morphologically rich language, in this paper. We present a rule-based method for Morphology for Classical Tamil, particularly plural marker.

Tamil morphology

Tamil belongs to the Dravidian family of languages. It is one of the Classical Languages. It is a verb-final language and has a relatively free word order; it is an inflectional language. Agglutination is another feature of the language. Tamil morphology is characterized as agglutinative or concatenative, i.e., Words are formed by successively adding suffixes to the root word in series. When suffixes attach to the root several morphophonemic changes take place. The orders in which suffixes attach to a root form determine the morphosyntax of the language and the various changes that take place when a suffix attaches to the root word are called the morphophonemics.

Rule based approach

The rule-based approach has successfully been used in developing many natural language processing systems. Systems that use rule-based transformations are based on a core of solid linguistic knowledge. The linguistic knowledge acquired for one natural language processing system may be reused to build knowledge required for a similar task in another system. The advantage of the rule-based approach over the corpus-based approach is clear for: 1) less-resourced languages, for which large corpora, possibly parallel or bilingual, with representative structures and entities are neither available nor easily affordable, and 2) for morphologically rich languages, which even with the availability of corpora suffer from data sparseness. These have motivated many researchers to fully or partially follow the rule-based approach in developing their natural language processing tools and systems. In this paper we address our successful efforts that involved rule-based approach for Morphological Analyzer for Classical Tamil texts.

Challenges in Morphological Analyzer for Classical Tamil

Tamil is a classical language which belongs to the Dravidian language family. Tamil literature has existed for over two-thousand years. The morphological structure of Classical Tamil is quite complex since it inflects to person, gender, and number markings and also combines with inflections that indicate aspect, mood, causation, attitude etc in verbs. A single verb root can inflect for more than two-thousand word forms. Noun root inflects with plural, oblique, case, postpositions and clitics. A single noun root can inflect for more than five hundred word forms including postpositions. The root and morphemes have to be identified and tagged for further language processing at word level. The structure of verbal complex is unique and capturing this complexity in a machine analyzable and generatable format is a challenging job. The formation of the verbal complex involves arrangement of the verbal units and the interpretation of their combinatory meaning. Phonology also plays its part in the formation of verbal complex in terms of morphophonemic or *sandhi* rules which account for the morphophonemic changes due to inflection.

Methodology

Finite State Automata (FSA)

FSA is a model of behavior composed of a finite number of states and transitions between these states. FSA is an abstract device used for recognizing simple syntactic structures or patterns. An automata is normally depicted by directed graph, called State Diagram and it is also represented in a tabular form as State Table. An FSA, as a string processing device, accepts strings as input and decides if the structure is correct, that is, it either accepts or rejects the string. From a mathematical perspective it is regarded as a function, mapping a set of string to the set {Accept, Reject}.

Plural markers

Considering the limitation of space available, This paper deals mainly with nouns with plural marker suffixes in Classical Tamil texts.

The form of a noun that typically denotes more than one person, thing, or instance. Contrast with singular.

Plural markers

The different forms of plural markers are *kaḷ* and *mār*

The *kaḷ* forms may be summarized as following

- 1) The *kaḷ* forms should not hibernate in the following cases because some of the case *kaḷ* may be in the root words
Makaḷ, Tiṇkaḷ, Makkaḷ
In such cases, the MA will return a result based not on the Rules of the FSA but on the corpus of root words provided in the M.A. in such cases the corpus-based approach comes into play and has to be integrated into the M.A.
Ex.
mīṇ eṇi paratavar maḥa makaḷ (naḥ 101:8)
peṇmai yaḥutta makaḷeṇ kiḥiviyum (Tol-col 650:2)
maḥavarē naḷ āyar makkaḷ nerunai (kali 102:30)
- 2) To be remove the *kaḷ* forms in the following cases, after removing the plural maker the remaining word gives the meaning full root word. In such cases the M.A. follows the rules- based approach
Kalavaikaḷ, Kiḷaikaḷ
Kalavai + kaḷ, Kiḷai + kaḷ

Ex.

kēḷivatu uṇṇu kiḷaikaḷ tuṇcupa (Nālaḷyār191:2)
kalavaikaḷuṇṇu kaḷḷippar valavaikaḷ (Nālaḷyār268)

- 3) To remove the ṇkaḷ forms in the following cases after removing the plurals add m in the root word and remove ṇ in the suffix word

Iṭaṇkaḷ -> Iṭa + ṇkaḷ -> Iṭam (+m) + (-ṇ) ṇkaḷ -> Iṭam + kaḷ
kuṇaṇkaḷ -> kuṇa + ṇkaḷ -> kuṇam (+m) + (-ṇ) ṇkaḷ -> kuṇam + kaḷ

Ex.

aṇṇa iṭaṇkaḷalvaḷ ellām (Tol 1154:2)
kuṇaṇkaḷaip pāṇāṭṭum tōḷaṇ vantīyāṇ kol (Kali 71:18)

- 4) To remove the ṟkaḷ forms in the following cases after removing the plurals add l in the root word and remove ṟ in the suffix word

coṟkaḷ -> co + ṟkaḷ -> col (+l) + (-ṟ) ṟkaḷ -> col + kaḷ

Ex.

tiruntupu nī kaṟṟa coṟkaḷyām kēḷpa (Kali 81:13)

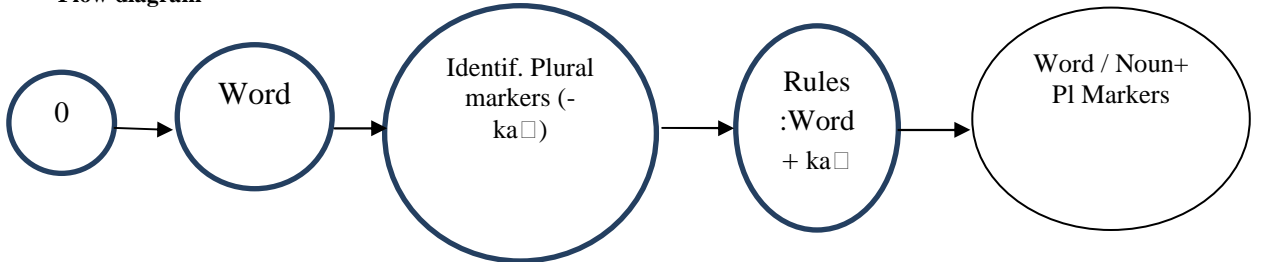
- 5) To remove the ṭkaḷ forms in the following cases after removing the plurals add ! in the root word and remove ṭ in the suffix word

nāṭkaḷ -> nā + ṭkaḷ -> nā! (+!) + (-ṭ) ṭkaḷ -> nā! + kaḷ

Ex.

oṇṇu iṭa iṭṭu varumēlniṇ vāḷnāṭkaḷ (Ainti 70-56:3)
maṇṇāḷvēntē niṇ vāḷnāṭkaḷ (Cilampu 28:125)

Flow diagram



Testing of Morphological Analyzer

The morphological analysis of words identifies root and suffixes of a word. Generally, rule-based approaches are used for morphological analysis which are based on a set of Rules and Dictionary that contains root words and morphemes. For testing the M.A. rules two thousand Classical Tamil root word corpus was chosen. The words were in the nominative case and had only plural markers and no other case markers. It was found that the M.A. methodology was useful in morpheme analysis.

Conclusion

This paper has described the Morphological Analyzer for Classical Tamil rule-based approach; in this paper rule-based approach is applied for Plural marker. These rule-based approaches for Plural markers produce the result with accuracy. In future, using this approach we can develop a rule-based approach for the analyzing not only of Plural markers but also of other markers and grammatical variations.

References

- Vijay Sundar Ram R, Menaka S and Sobha Lalitha Devi, “Tamil Morphological Analyser”, in “Morphological Analysers and Generators”, (ed.) Mona Parakh, LDC-IL, Central Institute of Indian Languages, Mysore, pp. 1 –18. 2010
- Anand Kumar M, Dhanalakshmi V, Soman K.P, “A Sequence Labeling Approach to Morphological Analyzer for Tamil Language”, International Journal on Computer Science and Engineering Vol. 02, No. 06, 2010
- Beesley, Kenneth R. Arabic, “Finite-State Morphological Analysis and generation”, Proceedings of the 16th International Conference on Computational Linguistics, Vol.1.Copenhagen, Denmark. Page No: 89-94, 1996
- Viswanathan, S., Ramesh Kumar, S., Kumara Shanmugam, B., Arulmozi, S. and Vijay Shanker, K. “A Tamil Morphological Analyzer”, Proceedings of the International Conference On Natural language processing ICON 2003, Central Institute of Indian Languages, Mysore, India. pp. 31–39, 2003
- Parameshwari K, “An Implementation of APERTIUM Morphological Analyzer and Generator for Tamil”, an e journal of Language in India (www.languageinindia.com), May 2011 Special Volume: Problems of Parsing in Indian Languages, May 2011
- Khaled Shaalan, "Rule-based Approach in Arabic Natural Language Processing" in International Journal on Information and Communication Technologies, Vol. 3, No. 3, June 2010

Development of Hints in Tamil using the Techniques of Natural Language Generation

Rajeswari Sridhar, V. Janani, G. Rasiga, G. Monica

rajisridhar@gmail.com, kvv.emails@gmail.com, rasiga.g@gmail.com, shruthiksha@gmail.com

Department of Computer Science and Engineering, College of Engineering, Guindy, Anna University, Chennai

Abstract

In this paper we propose a method to generate Tamil sentences yielding to paragraph generation. The proposed method uses Tamil words as the input. Parts of Speech tagging of the words are done by using a Morphological Analyzer. The system is based on the assumption that the input hints are available in Subject-Object-Verb pattern that is followed by most of the sentences in Tamil language. The hints development algorithm will generate Tamil sentences for all sentence patterns that start with a subject. Proper addition of Vetrumai (வேற்றுமை உருபு) suffixes are done to the entities by the rules of Tamil grammar. Gender of the subject is determined by means of a Look-up and based on the Tense preference given by the user, the verb form of the root word is modified to the correct form. This sentence generation uses the lexical resources fed to the system. The efficiency of the system is assessed to be 75%, considering the accuracy of the output, the processing time and the algorithm used.

1. Introduction

Development of hints into apprehensive text using the Techniques of Natural Language Generation is one field which has not been researched upon much [1]. While there are many scholarly articles for the summarization of voluminous text in English and other languages, there seem to be very few articles on the expansion of hints into a complete passage. But such a system, which when developed, will prove beneficial to those who use shorthand and hints to record information and later wish to develop that into a complete passage. Furthermore, we propose to develop the system for one of the natural languages, Tamil. This paper is organized as follows: Section 2 talks about the related work in this field and Section 3 talks about our proposed system which is the Development of Hints in Tamil using the techniques of Natural Language Generation, where we introduce the algorithm used, the rules defined and the complete flow of working of our system, Section 4 discusses on the assumptions and constraints of the existing system, Section 5 talks about the results and Section 6 concludes the paper with possible future extensions.

2. Related work

Many scholarly articles have been published for Morphological Generation and Morphological Analysis in Tamil. We try to focus this work as a sentence generation process which could subsequently lead to paragraph generation. A correct POS will identify the correct suffixes that need to be added to the root word. In order to POS tag the input hints, we used the Morphological Analyzer developed by Anna University's Tamil Computing Laboratory (TACOLA) [2]. In their work [2], Morphological analysis and POS is done for the input words wherein the words are tagged as noun, verb, and entity. In addition, for verbs, the root word and the suffixes that specify the tense and gender of the word are also retrieved by the tool. Now, the words should be put into sentences. For this we referred to the work of S. Lakshmana Pandian and T.V. Geetha [1] in which they have done sentence generation by identifying the role of each word in the sentence as agent, co-agent etc., Using the case suffixes mentioned by P.Anandan, Ranjani Parthasarathy and T.V.Geetha [3] and by referring to Tamil grammar rules for adding vetrumai, we determine the suffixes to be added. We referred to the work of K.Rajan et al [4], to know what modifications need to be done when two words occur together in a sentence in Tamil. In their paper, they discuss about the changes like addition, deletion, alteration etc., when two or more morphemes or words occur together. With these as basis, we modified each work and added more to develop the given hints into a paragraph which is discussed in the following section.

3. Proposed Algorithm

Our Algorithm follows the steps outlined in Figure 1.

The block diagram describes the proposed methodology for generating completed, meaningful and grammatically correct sentences from the given hints (from an input file). The whole methodology is carried out in 2 phases.

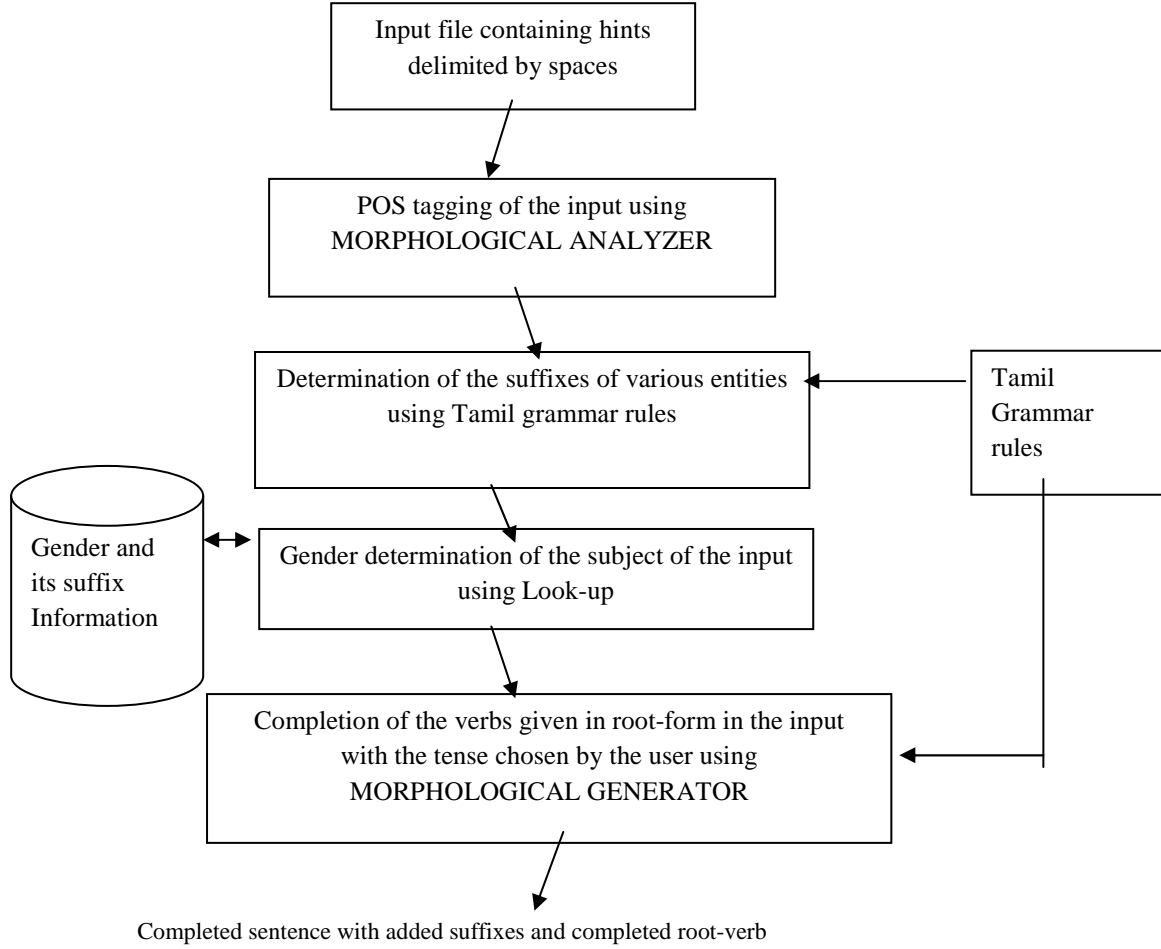


Figure 1 Block diagram of the Sentence generation system

PHASE 1

Input file is a text document containing the hints where the words are in root-form. This file is given as an input to the morphological analyzer tool [2] which does the POS tagging and distinguishes noun, verb, entity, etc. Tagged words which are the output of this phase are given as the input to the next phase.

PHASE 2

Now, using Tamil Grammar rules, suffixes for the entities are determined and added. Gender of the 'subject' is determined using Look-up. Tense information is received from the user which is used to convert the verbs which are in root-form by using a morphological generator. At the end, the completed sentences are obtained.

The algorithm is explained in the following sub-sections.

3.1 Parts Of Speech Tagging

The first step in the hints development work is the POS tagging. Using Morphological Analyzer tool, the input is tagged as either entity, Noun or Verb. For instance the following is the input and output of the Morphological Analyzer.

Input : இராமன் கார்த்திக் கடை செல் புத்தகம் வாங்கு

Output : இராமன் (Raman) +Entity
கார்த்திக் (Karthik) +Entity

கடை (Kadai) +Entity
செல் (Sel) +Verb
புத்தகம் (Puthagam) +Noun
வாங்கு (Vaangu) +Verb

3.2 Partition of Input into Phrases

After POS tagging, the input is partitioned into phrases. The challenge in partitioning is to decide on the phrase boundary. As a constraint, our project assumes the first word of the input as subject and divides the input into smaller phrases of the form Subject-Object-Verb. If there are multiple verbs in the input, the algorithm considers the first verb that occurs after each object for the addition of suffixes. Considering the above example, the phrases are taken in the following order:

- 1) இராமன் கார்த்திக் செல் (Raman Karthik Sel)
- 2) இராமன் கடை செல் (Raman Kadai Sel)
- 3) இராமன் புத்தகம் வாங்கு (Raman Puthagam Vaangu)

3.3 Addition of வேற்றுமை (Vetrumai) Suffixes

After phrasing, now suffixes are to be added by considering the POS. The following are some of the rules that have been formulated by referring to Tamil grammar rules:

- 1) If the type of the verb can be classified as either ஆக்கல், அழித்தல், அடைதல், நீத்தல், ஒத்தல், உடைமை, then the suffix 'ஐ' will be added to the object.
- 2) If the verb cannot be classified as any one of the above, Named Entity Recognition of the object is done. i.e, The object is classified as name, place, animal or thing. And this classification was done using Look-up.
 - i) If the object is a place or a thing and the verb is one of 'உள்ளது', 'இருந்தது' or 'இருக்கின்றது' then the suffix 'இல்' is added.
 - ii) If the object is a place and the verb's root is either 'வா', 'செல்', or 'பேர்', then the suffix 'கு' is added.
 - iii) If the verb is either 'ஏற்றது', 'ஏற்றவன்' or 'ஏற்றவள்', then the suffix 'கு' is added.
 - iv) If none of the above cases are satisfied, the Named-Entity-Recognition of the subject is also performed.
 - v) If the Subject is either a name, animal or a thing and so is the object, then the suffix 'ஓடு' is added with the object.

3.4 Morphological Generation using Punarchi Vithigal (புணர்ச்சி விதிகள்)

After determining the sequence of words and their corresponding suffixes the following Tamil Punarchi Vithigal are used to construct the words based on suffixes.

- 1) “இ, ஈ ஐ வழி யவ்வும் ஏனை

உயிர்வழி வவ்வும் ஏ முன் இவ்விருமையும்

உயிர்வரின் உடம்படுமெய் யென்றாகும்”

(i.e.) if the object ends with a 'உயிர்', character and the suffix starts with 'இ, ஈ, ஐ', then it becomes Object + 'ய்'+suffix, else if the suffix starts with 'உயிர்' other than 'இ, ஈ, ஐ', then it becomes Object + 'வ்'+suffix and if the suffix starts with 'ஏ', then it can either be Object + 'வ்'+suffix or Object + 'ய்'+suffix and joins using rule 2.

Example : சென்னை + இல் = சென்னை + ய் + இல் = சென்னையில்

- 2) “உடல் மேல் உயிர் வந்து ஒன்றுவது இயல்பே”

(i.e.) if the object ends with a 'மெய்,' and the suffix starts with a 'உயிர்' character, then these two fuses together and forms a 'உயிர்மெய்'

Example : இராமன் + ஓடு = இராமனோடு

3) “தனிக்குறில் முன் ஒற்று உயிர் வரின் இரட்டும்”

(i.e.) if the object consists only a ‘குறில்’ and ends with a ‘மெய்,’ and the suffix starts with a ‘உயிர்’, then that ‘மெய்’ doubles and joins using the standard rules mentioned above.

Example : கல் + ஐ = கல் + ல் + ஐ = கல்லை

4) “உயர் வரின் உக்குறள் மெய் விட்டோடும்”

(i.e.) if the object ends with a ‘கு’ and the suffix starts with a ‘உயிர்’, then ‘உ’ gets eliminated from ‘கு’ and joins using rule 2.

Example : விறகு + ஐ = விறக் + ஐ = விறகை

5) “மகர ஈற்று அஃறினைப் பெயர்கள் ‘அத்து’ என்னும் சாரியை பெற்று வேற்றுமை உருபுகளை ஏற்கும்”

(i.e.) if the object is a named entity for non-living things and ends with ‘ம்’ then it becomes Object + ‘அத்து’+suffix and joins using the standard rules mentioned above.

Example : மரம் + கு = மர + த் + த் + உ + க்கு = மரத்துக்கு

6) Adding suffix ‘கு’

* If there is a ‘உயிர்மெய்’ at the end of the object, it becomes Object + ‘க்கு’

Example : சென்னை + கு = சென்னைக்கு

* If the object ends with an ‘ஆ’ it becomes - Object + ‘வுக்கு’

Example : அமெரிக்கா + கு = அமெரிக்காவுக்கு

* If the object ends with a ‘மெய்’, it becomes - Object + ‘உக்கு’ and joins using the standard rules mentioned above.

Example : கோயில் + கு = கோயிலுக்கு

4. Assumptions and Constraints

1) The project will work for select male and female names only since the Morphological Analyzer Tool that we used for POS Tagging is programmed to accept only a finite set of proper nouns.

2) While converting the root verb into complete verb, only the gender and tense are taken into account, the number (singular/plural) is not considered and is by default singular.

3) The algorithm adds one of the four suffixes (வேற்றுமை உருபு) - கு, ஐ, ஓடு, இல் for the generation of complete sentences.

4) The project will work under the assumption that the first word of the input is always the subject.

Sample input:

இராமன் ஆடு ஆசிரியர் கவர் (Tense specified as present tense)

Output of the Morph-Analyzer:

இராமன் +Entity

ஆடு Verb

ஆசிரியர் +Noun

கவர் +Verb

Output of Sentence Generator:

இராமன் ஆடுகிறான் ஆசிரியரை கவர்கிறான்

5. Results and Analysis

Around 50 sentences as hints were given as input to our system and it developed syntactically correct sentences in 36 of the cases which is 72%. The sentences that could not be handled by our system was due to the constraints imposed by the limited words in look-up and the first word not being the subject of a few sentences. In addition the errors in Morphological analyzer percolated and affected the sentence generation.

6. Conclusion and Future work

Presently the system developed by us caters to the development of hints in Tamil, albeit with a few constraints and assumptions. We intend to extend our work in the near future with no constraints imposed, and develop a fully functioning Hints Developer in Tamil, by extending the Morphological analyzer tool to identify words in wide range of domains and also try to handle any sentence form (by identifying the subject of the sentence instead of assuming the first word to be the subject). We also intend to make use of Unsupervised Learning methods for Morphological Analysis of the given input. The context of the given input could be understood to determine the semantics to be used for generating sentences.

References

- 1) S. Lakshmana Pandian, T.V. Geetha, “Semantic Role Labeling for Tamil Documents”, International Journal of Recent trends in Engineering, Vol 1, No. 1, May 2009.
- 2) Anandan, P., Saravanan, K., Parthasarathi, R., & Geetha, T. V., “Morphological analyzer for Tamil”, Proceedings of ICON2002, pp. 3-10, 2002.
- 3) Anandan, P., Geetha, T.V., and Parthasarathi, R., “Morphological Generator for Tamil”, In Proceedings of the Tamil Inayam Conference, Malaysia, 46—54, 2001
- 4) K.Rajan, V.Ramalingam and M. Ganesan, “Machine Learning for Sandhi Rules in Tamil”, Proceedings of the 11th International Conference INFITT, pp. 141 – 146, 2012.

Paeri: Evolving Tamil Name Generation Algorithm

Rajapandian, Karthikeyan, Elanchezhiyan & Madhan Karky

Karky Research Foundation, Chennai, India.

Abstract

Finding new and unique names for newborns as well as new terms in science, technology or any domain is a challenge in any language. The sounds that constitute a name, together primarily identify an object. In this paper, we propose an algorithm, Paeri that uses character-level statistics from Tamil dictionary words and traditional Tamil male and female names to create pronounceable phonetic combinations that can be used to name any object. Paeri algorithm learns the frequencies of beginning and end characters of male and female names from a traditional name list. It also learns the co-occurrence frequency of characters from one lakh Tamil words. The name list and word database are discarded once the learning is completed. Starting letter, name length, uniqueness factor & sex preference are optional inputs for the Paeri algorithm. Paeri can generate infinite number of names if we do not limit the number of letters. If we limit the number of letters in a name to five, Paeri can generate approximately 86 crore names based on Tamil phonetics. In this paper we analyze the Paeri algorithm and how it learns preferred and non-preferred names, based on user feedback.

1. Introduction

Naming a new born or a newly discovered object or concept is unique to every society. Some cultures name newborns after Gods and religious beliefs whereas other cultures name children out of sounds they hear out of random objects. A name serves as an identity to a person. It is not unique. In most cultures names have an associated meaning to them. A lot of research has been conducted in every culture regarding the roots of names. In an alternate view point, a name is a mere combination of sounds in a particular order. We do not research the meaning associated with names such as Buddha, Gandhi, Mandela or Theresa. In a foreign land, a person's name may be nothing but a series of sounds stitched together. Most of us would have an interesting story with the way our name is pronounced in a different culture/country. However if we dig deeper and deeper into the meaning of a name, we would end up with meaningless sounds.

With the growing population and modern day parents seeking unique names for their newborns, we identified the need for an automated name generator based on Tamil phonetics. In this paper, we propose Paeri, an algorithm to generate names based on analysis of how Tamil words have been formed and how proper nouns have been created using combination of sounds. Paeri is an evolving algorithm that learns good and bad names based on user feedback and updates itself to create better names.

This paper is organized into five sections. The following section lists some popular name generation projects relevant to this paper. The third section describes the Paeri algorithm in detail. The implementation and results are discussed in the fourth section. The fifth section concludes with summary and future work.

2. Background

Many existing name generators take first names and last names from a database and generate different combinations of first and last names. Tamil baby name lists are provided by many sites [2], [3], [4], [5] are a few of those sites. In [1], authors analyze the frequencies of characters in Tamil dictionary words. The interesting aspects of this paper are the statistics on most frequent characters and the list of 17 characters that are never used in any of the dictionary words. The work presented in [6] uses simple 2 and 3 level markov chains to form English names from regular words with simple grammar. Similar markov chain based name generation is also presented in [7]. A consonant vowel pattern based algorithm is presented in [8] again for English names.

Paeri algorithm presented in this paper unlike the methods mentioned above, applies statistical knowledge acquired from existing names and dictionary words to generate pronounceable names.

3. Paeri: Algorithm

In this section we present the Paeri algorithm. Paeri algorithm works in two phases. The first phase, which works offline, is where words obtained from Tamil dictionaries and Tamil names databases such as [2], [3] & [4] are analyzed. The offline process is explained in the following steps.

3.1 Paeri: Offline

Step 1: For each character in Tamil, create a record that stores the following statistics:

% of words in dictionary and name lists that have the character in starting position

% of words in names list that have the character as male ending.

% of words in names list that have the character as female ending.

Step 2: For every combination of two characters in Tamil, build a two dimensional matrix to record the following statistics from names lists and word database.

Frequency of the first letter followed by the second letter.

Frequency of the second letter followed by the first letter

Step 3: Discard the name list and word database. Build an index based on Tamil characters that holds the statistics obtained from Step 1 & 2.

3.2 Paeri: Online

Inputs:

Len -> length of the Name

FirstLetter -> is the starting letter of the name (optional)

gender -> gender

u -> uniqueness factor

Initialise i = 0;

Let N be the string that holds the name.

Initialise N= null;

If FirstLetter != NULL

N = FirstLetter;

Else

N = generateRandomFirstLetter (u);

End If

For i = 1 to (Len - 2)

Let pLetter: = generateNextLetter (N[i], u);

N: = N + pLetter;

i := i + 1;

End for

Let endLetter = generateRandomLastLetter (gender, u);

Let penultimateLetter = generatePenultimateLetter (pLetter, endLetter, u);

N: = N + penultimateLetter + endLetter;

Return N;

3.3 Algorithm Analysis

The basic idea behind the offline part is to acquire the knowledge of what letters are used to start words and names, how male and female names end and what letters come before and after which letters. Once learnt, the database and name lists are discarded as they are not needed for the name generation anymore.

The online part has four inputs: length of the name to be generated, starting letter of the name, gender and uniqueness factor. The uniqueness factor is a value between 0 and 1. Values closer to 0 would generate unique names by selecting low frequent first letters, middle letters and penultimate letters. Value closer to 1 would generate names that look very simple and common.

The method generateRandomFirstLetter (u) picks a random first letter from the recorded statistics based on the uniqueness factor u.

The method generateNextLetter (char, u) picks a random letter from a set of letters that are possible after character char based on the uniqueness factor u.

Method generateRandomLastLetter (gender, u) pics a random last letter from the set of characters.

Selecting the penultimate letter is different from selecting other letters. The primary reason behind selecting the last letter before selecting the penultimate letter is to ensure a smooth bridge to the finish.

The method generatePenultimateLetter (pLetter, endLetter, u) works by selecting a letter from the intersection of set of possible next letters of pLetter and set of possible previous letters to the end letter. This intersection would ensure that selecting the penultimate letter would not lead to an unpronounceable sound leading to the end letter. This is ensured as we have only very few options for end letters for both male and female names.

The block depicted in Figure 1 shows online and offline phases. The database mentioned in the offline part is a statistical database of all characters. None of the names or dictionary words are stored in the database.

The algorithm evolves by learning good names and bad names based on user selection. The offline process step 1 - 3 is repeated with the user selected names as the name list. The frequency information is updated to the existing statistic in the database.

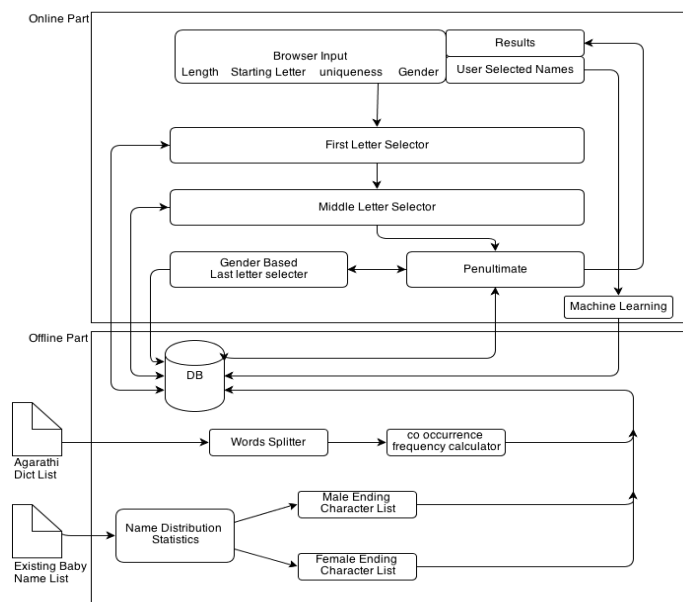


Figure 1 : Paeri Tool - Snapshot

5. Implementation and Results

Paeri - online name generator tool [ref] was developed based on the Paeri algorithm described in section 3.1 and 3.2. The tool was developed in Java and hosted on Linux server. JQuery was used for the interface. A snapshot of the working tool is given in figure 2. The tool provides means to select starting letter and to fix the name length between 2 and 5. The uniqueness factor can be given a value between 1 and 10. The gender can be selected by clicking on the gender icon.

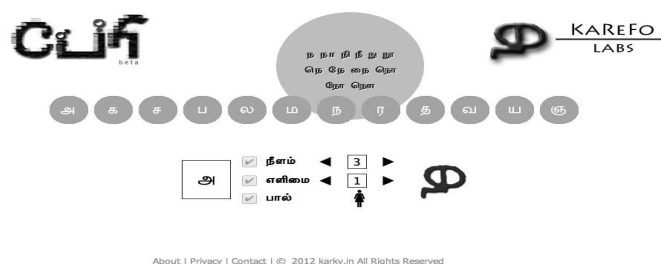


Figure 2 : Paeri Tool – Snapshot

If we do not limit the length of the name, Paeri can generate infinite number of names. By limiting the name length to 5, we can generate a maximum of 86 crore unique pronounceable Tamil names.

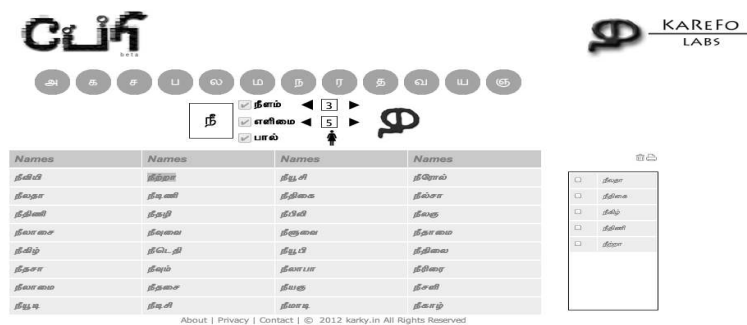


Figure 2 : Paeri Tool Results Snapshot

For a name starting with நீ, with length 3, uniqueness factor 5 and gender as female some of the names generated were நீலதா, நீதிகை, நீகிழ், நீதினி & நீற்றா.

The tool is now being used widely by parents to name their children and entrepreneurs to name their organizations. The name Paeri was generated using the very same algorithm. சமன் (saman) & மலிங்கை (malingai) were the first boy & girl names selected using Paeri online tool for naming newborns.

6. Conclusion and Future work

In this paper, we discussed the need for a name generator based on Tamil phonetics. We presented Paeri, an algorithm to generate names based on statistics obtained from dictionary words and traditional name lists. The evolving nature of the algorithm was also discussed. Implementation of the algorithm as the Paeri online tool was discussed and the results analyzed. Around 86 crore names can be generated using Paeri tool by limiting the maximum length of a name to 5.

Generating an associated meaning from the sounds that created the name will be an interesting topic for further research. Similarly generating a name from a given meaning or explanation is a possible area for research.

7. Reference

- Karthika Ranganathan, Elanchezhian.K, T.V.Geetha, Ranjani Parthasarathy, Madhan Karky, Frequency Analysis of Tamil Alphabets, National Seminar on Computational Linguistics and Language Technology, March 2011.
- HYPERLINK "<http://www.koodal.com/tamilnames/>" <http://www.koodal.com/tamilnames/> last accessed date 30th June 2013.
- HYPERLINK "<http://babynames.looktamil.com/show/pure-tamil-names>" <http://babynames.looktamil.com/show/pure-tamil-names>, last accessed date 30th June 2013.
- HYPERLINK "<http://tamilcube.com/babynames/tamil-baby-names.aspx>" <http://tamilcube.com/babynames/tamil-baby-names.aspx>, last accessed date 30th June 2013.
- HYPERLINK "http://www.yourastrology.co.in/english/page/baby-names_96" http://www.yourastrology.co.in/english/page/baby-names_96, last accessed date 30th June 2013.
- Name Generators for NPCs, HYPERLINK "<http://www.godpatterns.com/2005/09/name-generators-for-npcs-to-populate.html>" <http://www.godpatterns.com/2005/09/name-generators-for-npcs-to-populate.html>, last accessed 8th July 2013.
- Name Generation, HYPERLINK "<http://pcg.wikidot.com/pcg-algorithm:name-generation>" <http://pcg.wikidot.com/pcg-algorithm:name-generation>, last accessed 8th July 2013.
- Rndom Name Generation Algorithm, HYPERLINK "<http://www.tinyfrogsoftware.com/random-name-generator-algorithm/>" <http://www.tinyfrogsoftware.com/random-name-generator-algorithm/>, last accessed 8th July 2013.

Semi-Automatic Query Testing Tool

Ganesh J, Project Associate, Department of IST

UmaMaheswari E, Research Scholar, Department of CSE

Dr. T V Geetha, Professor, Department of CSE

Dr. Ranjani Parthasarathi, Professor, Department of IST

TaCoLa Lab, CEG, Anna University, Chennai

ABSTRACT

Measuring the effectiveness of information retrieval from the World Wide Web search engines is challenging and expensive because of human relevance decisions involved. However, both for developers and users, it is important to know the effectiveness of web search engines. Developers of search engines need to evaluate the search results for every change that is carried out in the query processing process or the search and rank process or when new filtering methods are applied. Typically this evaluation has to be done for a significant number of queries. Users of search engines may also want to compare the results of different search engines for their queries. Hence, there is a need to automate the process of relevance judgment. This paper proposes a semi-automatic tool for relevance judgment, which requires a human evaluator to evaluate the relevance of a document to a query only once. In the development framework, many versions of the search engine would be developed, each of which needs to be evaluated with the help of this tool, this process is simplified. The tester or the human evaluator rates the URL's depending on their relevance to the given query the first time that it is encountered. This information is stored in a database. The subsequent time when the different version of the search engine is tested, the retrieved URL's are first compared with the results stored in the database. If there is a match, that URL need not be judged again. Only the new set of URL's retrieved have to be judged thus saving considerable human effort and time. An overall precision score such as MAP or P@N is automatically calculated by the tool. This tool has been used in the development of a concept-based search engine, and it has helped to reduce the testing time by about 50% to 75% depending on the novelty of the results obtained. A generic user can also use this tool, to evaluate different search engines, by firing the same set of queries to the different search engines, and save time in evaluating the results obtained.

1. Introduction

Search engines are designed to help users to rapidly find useful information on the Web. The growth of the World Wide Web is an exotic phenomenon. In 2008 the number of web pages at about one trillion unique URLs, announced that Google Search. The performance of search engines can be evaluated using various measurements, such as precision, coverage, response time, recall, and interface. In this paper we focus on automatically calculating the precision of search engine [10], using previous evaluation reports. Precision is commonly defined as the ratio of relevant documents from the retrieved documents. Precision of Web search engines is important because [1] users are attracted by getting relevant documents, and [2] the precision results are useful in identifying which ranking and indexing method is more useful for information retrieval. In previous studies, precision of search engines has been fully manually evaluated based on general subject queries and some specific query domains. The major benefit of manual precision evaluation is its accuracy. The drawback is that it is biased, time-consuming, and cannot adjust well to the ever changing search engines and other concept based semantic search engines [6][10]. Here, we present an automatic precision evaluation method for concept based search engine [10]. In general, Search engines are evaluated based on a small set of sample queries by computing relevance scores based on 'more relevant', 'less relevant', 'irrelevant', 'partial relevant' and 'site cannot be accessed'. Since our search system [10] is based on the concepts rather than terms, the relevance judgment is assigned with the linguistic experts. Concept based search system cannot be tested as normal search engines. In Tamil, the term water falls (English) can be written in different ways like "Aruvi" "Neervzhchi". If the search term is "Aruvi", if the documents containing only "Neervzhchi", the normal term search engine [6] cannot be able to retrieve results for both "Aruvi" and "Neervzhchi". But our search engine can retrieve both the results which is more relevant. Hence automatic relevance judgment requires term level match between the query and document which is not possible for conceptual search engine [10]. In order to evaluate search engines conceptually manual efforts are required. Hence we use linguistic experts to assign relevance rating for each retrieved documents. The precision score is automatically calculated with the help of relevance rating.

This paper is organized as follows. Section 2 provides comparative study of related work for this proposed methods. In Section 3, describes the methodology used to automate the relevance metric for a given query. In Section 4 discusses accuracy of query automation tools for precision. Section 5, discusses conclusions and future work.

2. Related Work

The study of Abdur Chowdhury and Soboroff (2002) [8] work, which measures the performance of search engines automatically. In this study, initial query–document pairs are constructed randomly. Then, for each search engine, mean reciprocal rank is computed for over-all query–document pairs. If query–document pairs are reasonable in such a way that it should match with the query term, then this method could be useful. Fazli Can, Rabia Nuray and Ayisigi B. Sevdik (2003) [1] proposed a method to measure the performance of eight search engines through its precision and recall, and binary user relevance judgments is needed for finding the correct documents. There is also the Yi Shang _and Longzhuang Li (2002)[4] study, which involves a general approach for statistically evaluating precision of search engines on the Web, using three relevance scoring methods from popular IR (Information Retrieval) similarity measures and one common mimic manual evaluation approaches are used. Another approach proposed by Liwen Vaughan (2003) [2] evaluates the measurements of Web search engine. In this approach [2], a general approach is used to evaluate the precision and recall. Here, an experiment was conducted to test the performance comparison of three common search engines.

3. Methodology

Conceptual search engines [10] are used for evaluating the search capabilities in terms of concepts and relations, rather than terms. Before giving input to this automatic evaluation system, different relevance rating tag is assigned. This is explained below

- If the content of the web page document closely matched the subject matter of the search query, then it was categorized as ‘more relevant’ and it was assigned as a parameter ‘Y’ value is 1.
- If the content of the web page document is not closely matched the subject matter of the search query, then it was categorized as ‘partial relevant’ and it was assigned as a parameter ‘P’ value is 0.5.
- If the content of the web page document contains some links related to the subject matter of the search query, then it was categorized as ‘less relevant’ and it was assigned as a parameter ‘L’ value is 0.5.
- If the content of the web page document is not related to the subject matter of the search query, then it was categorized as ‘not relevant’ and it was assigned as a parameter ‘N’ value is 0.
- If the site can’t be accessed due to the following problems: page cannot be find, page under construction, and some other technical faults, then it was categorized as ‘site cannot be accessed’ and it was assigned as a parameter ‘X’ value is 0.

Using the above relevance judgments, we can calculate the precision score [7].

This proposed method calculates precision score at two levels. In the first level it assigns rating values 1, 0.5, 0 based on its relevance and computes precision. In the second level when different version of the same search engine requires to be tested, it compares the previous test report and assigns the previous score for the same set of resulted URLs for the same set of queries. It is not required to assign human judgment if the results are unique in both the versions. This considerably reduces human effort in assigning the relevance score.

4. Evaluation

As described earlier, we evaluate search engines precision using five relevance score rating methods. The experiments were based on short queries containing 2, 3 or 4 terms. The queries were derived from tourism domain. We have used existing conceptual based search engine [10] which is developed for Tamil language in our experiments. This automatic evaluation tool can also be adopted for any term/concept search engine which requires testing for different versions. Here top 10 results are considered for relevance judgment.

In the case of human-based evaluation, the linguistic experts decide the relevance of the top 10 pages of the search engine in a concept relevance rating method. The linguistic expertes were given sufficient guidance to assign any degree of relevance among the five rating method.

The average precision is computed in the following way. First the results are retrieved from the conceptual search engine, and then the average of the precisions for each query (P@5 and p@10) is computed by this method after assigning relevance tags. This report is stored for future reference. If a new version requires to be tested, query-document pairs are retrieved and compared with the previous test reports. If it finds similar queries and documents, it assigns the previous relevance score, instead of checking again. Hence the relevance judgment is calculated only once. Fig. 1 and Fig. 2 show the snapshot of the automated test system.



Fig: 1 Automatic

Query Number	Query Word	Date	URL	Doc Id	Relevancy Tag
43	infotamil.ch	3/3/2013	http://www.lahbaikudikadu.com/pages/ladies_section_detail.aspx?id=40	d19216881129242211145.txt	N
43	infotamil.ch	3/3/2013	http://www.lahbaikudikadu.com/pages/ladies_section_detail.aspx?id=40	d19216881129242211145.txt	N
43	infotamil.ch	3/3/2013	http://tamil.webdunia.com/miscellaneous/literature/magazine/0705/11/1070511017_1.htm	d19216881129268491769.txt	N
43	infotamil.ch	3/3/2013	http://tamil.webdunia.com/miscellaneous/literature/magazine/0705/11/1070511017_1.htm	d19216881129268491769.txt	N
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2ME24eUYd302gOodd3OgE20ndY2e4YMMca2oAA2	d19216881322523394626.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2ME24eUYd302gOodd3OgE20ndY2e4YMMca2oAA2	d19216881322523394626.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2ME24eUYd302gOodd3OgE20ndY2e48MCCA2oAU2	d19216881129242123454.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2ME24eUYd302gOodd3OgE20ndY2e48MCCA2oAU2	d19216881129242123454.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac3ME34eUYd302gOodd3OgM20ddY2e4YMMca2oAS2	d19216881129263714274.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac3ME34eUYd302gOodd3OgM20ddY2e4YMMca2oAS2	d19216881129263714274.txt	R
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2CE34eUYd302gOodd2OgM20jdY2e4iMCCA2oAK2	d19216881129262244760.txt	
43	infotamil.ch	3/3/2013	http://www.infotamil.ch/ta/view.php?22AoCac2CE34eUYd302gOodd2OgM20jdY2e4iMCCA2oAK2	d19216881129262244760.txt	

Fig: 2 Semi-automatic

5. Conclusions

In this paper, we present a semi-automatic method for evaluating precision of a concept based search engine by applying conceptual relevance feedback given by the linguistic experts. The five scale rating scheme is used to judge relevant, less relevant, partial relevant and irrelevant results which is obtained from the search engine. In this dynamically changing environment method, there is a need to change the approaches according to the user preferences and the human relevance judgment is inconsistent and it may vary person to person. Hence automating the human judgments for semantic search engines require machine learning methods to know the neighboring context of the query. It is a challenging task to adopt fully automatic test tool for semantic search engines which is our future work.

6. References

- [1] Fazli Can, Rabia Nuray, Ayisigi B. Sevdik. Automatic performance evaluation of Web search engines. Information Processing and Management: an International Journal, 2004.
- [2] Liwen Vaughan. New measurements for search engine evaluation proposed and tested. Information Processing and Management: an International Journal, 2004.

- [3] H. Vernon Leighton and Jaideep Srivastava. First 20 Precision among World Wide Web Search Services (Search Engines). Journal of the American Society for Information Science, 1999.
- [4] Yi Shang and Longzhuang Li. Precision Evaluation of Search Engines. World Wide Web, 2002.
- [5] Chu Heting and Rosenthal Marilyn. Search Engines for theWorldWide Web: A Comparative Study and Evaluation Methodology. Proceedings of the ASIS Annual Meeting, 1996.
- [6] Girit, Hasan and Eberhard, Robert and Michelberger, Bernd and Mutschler, Bela. *On the Precision of Search Engines: Results from a Controlled Experiment*. In: 15th Int'l Conference on Business Information Systems (BIS 2012), 2012.
- [7] Kumar, B.T. Sampath; Pavithra, S.M. Evaluating the searching capabilities of search engines and metasearch engines: comparative study. Annals of Library and Information Studies (ALIS), 2010.
- [8] Abdur Chowdhury and Ian Soboroff. Automatic Evaluation of World Wide Web Search Services. In Proceedings of the Acm Sigir Conference, 2002.
- [9] Jansen, B. J., & Pooch, U. A review of Web searching and a framework for future research. Journal of the American Society for Information Science and Technology, 2001.
- [10] A multilevel UNL concept based searching and Ranking:: Umamaheswari E, T V Geetha, Ranjani Parthasarathi, Madhan Karky.

Patti Vaithiyam — An Information Extraction System for Traditional Tamil Medicines

J. Betina Antony^{#1} and G.S.Mahalakshmi^{#2}

[#]Department of Computer Science and Engineering, Anna University
Chennai, Tamil Nadu, India

¹betinaantony@gmail.com

²gsmaha@annauniv.edu

Abstract— This work attempts to extract traditional medicine related information from tamil siddha documents generally published as ‘health tips’ in websites, blogs and magazines. Sometimes a sub-section of it is also popularly known as ‘Paatti Vaithiyam’. It extracts names of the traditional medicines that are mentioned in the unstructured documents. The idea is to record the name of the items as well as the cure in connection with the process of preparing the home remedy. The output is obtained in a structured format that is easily understandable.

Keywords— Information extraction, tamil siddha documents, traditional medicine, Paatti Vaithiyam

I. INTRODUCTION

One of the most ancient medical systems known, Siddha medicines originated in the southern part of India, from Tamil nadu. It is a part of the trio Indian medicines – Ayurveda, Siddha, Unani. ‘Siddhargal’ or ‘Siddhars’ were believed to be the founders of this oldest system of medication. The thousands of texts produced by them laid the foundation for Siddha Medicine. The word ‘Siddha medicine’ means medicine that is perfect. Based on the mode of application, they are classified into 32 categories of *internal medicines* and 32 categories of *external medicines*.

Traditional Knowledge Digital Library is an Indian traditional knowledge repository containing mainly information about medicinal plants and formulations used in the Indian medical system. The main objective of the library is to protect the knowledge about traditional and ancient medical practices from bio-piracy and unethical patents. There are also various text documents written on Siddha medicines based on the information gather from the ancient manuscripts obtained. Few books that are known are ‘*Pogar-7000*’ that deals with almost all subjects of Siddha medicine especially metals and minerals and ‘*A Scientific Journal from national Institute of Siddha*’ containing the scientific research oriented articles on Siddha medicines. Thus there are many sources of traditional medicine information. Hence it is essential to create an Information Extraction (IE) system to efficiently utilise these resources and provide effective usage of the information obtained.

Tamil medical documents are mainly comprised of unstructured text and automatic processing of these texts is still a challenging task in the field of Natural Language Processing (NLP). This is mainly due to the migration of interest of the people from traditional medicines to allopathic or modern medicines. Also high quality studies are essential to compare and evaluate the value of traditional Indian drugs. Our system tries to retrieve these unstructured documents and process them to obtain necessary and valuable information.

This paper is organized as follows: Section II gives an overview on the Information Extraction systems for the Biomedical domain in various other languages. Section III provides details about the approach suggested. Section IV deals with the experiment results and evaluation while Section V gives the conclusion and some discussions for future work.

II. RELATED WORK

Information Extraction is considered a successful language processing technology to obtain information from unstructured text. The basic technique of IE paradigm is to extract entities by shallow analysis, recognize its references, update database and fill templates [6]. The shallow analysis usually involves pattern matching in regular expressions.

Though there are no actual works on traditional tamil medicine, a number of information extraction works have been carried out for Biomedical documents [10][7] and Clinical records[15][8][16] by string matching or rule based methods. One such information extraction system is MedEx [16], which identifies the name of the medicines and also signature information such as strength, route and frequency from discharge summaries. One important challenge in these clinical records is the heterogeneity of data (narrative or coded).

IE finds its application in many prototypes which are used in the extraction process:

A. Concept Extraction

MetaMap[3], a effective mapping procedure, makes use of knowledge intensive approach and computational linguistic techniques to map biomedical text to the UMLS Metathesaurus. The biomedical texts contain classified concepts and hierarchical relationships. The two main problems faced by this system are detection of idiosyncratic text and resolution of ambiguity. Clashing concepts can also be resolved by classification based on semantic types.

B. Named Entity Recognition (NER)

The concept extraction process is supported by a several Named Entity Recognition (NER) systems. Various approaches for NER include Hidden Markovian Model (HMM) [14][4][17] and Support Vector Machine (SVM)[9][5]. POSBIOTM-NER[13] is a Biomedical NER extraction system that uses SVM machine learning approach to build and expand a NER Dictionary by SVM training. This NER system adopts edit-distance measure, an additional input to resolve spelling variant problem.

C. Word Sense Disambiguation (WSD)

Word Sense Disambiguation (WSD) is the task of selecting the appropriate sense for words in the given context. A number of graph-based approaches have been identified to identify the intended meaning of words in a context [12] that investigates the connectivity measures of the graph and identifies how this method performs comparably to the state-of-art. One such method that makes use of UMLS MetaThesaurus and Personalized PageRank, a state-of-the-art algorithm is explained in [1]. Disambiguation here is done by converting the tables from the thesaurus into graphs and applying the algorithm to select best sense for each ambiguous word.

III. MATERIALS AND METHODS

A. Medical Documents

The medical documents are the text documents containing the remedial instructions for a particular disorder. The documents contain the following tags:

- 1) <நோய்> : This tag contains the name of the disorder for which the cure is suggested. The contents of this tag may be more than one.
- 2) <மருந்து>: This tag provides the remedial measures to be taken for the given disorder. The instruction is composite from which various other informations such as ingredients used, special notes, medicine preparation procedure and dosages instructions can be obtained.

B. Methods

a) Document Retrieval:

Document retrieval is the process of retrieving relevant documents for a given query. The query may be a set of keywords and the relevant documents are text documents whose contents contain these keywords or related words [11].

For a given query (say அஜீரணம்), the relevant documents are obtained by matching the elements in the <நோய்> tag. Eg. If <நோய்>அஜீரணம்<நோய்> is present for the query, the document is retrieved.

b) Morphological Analysis:

Morphological analysis is the process by which individual words are analysed into their components and non-token words such as punctuations are separated from the words. The process assigns syntactic categories to all the words in a sentence.

For tamil, morphological analysis can be done using a tool called 'Atcharam' [2]. Eg. For a word <தண்ணீரில்>, the analysed words will be

தண்ணீர் < Noun & 100 >

இல் < Locative Case & 504 >

Here தண்ணீர் is the morpheme obtained by morphological analysis.

c) Category Tagging:

Categorizing a set of text into predefined classes is called text categorization. Text categorization can be done by many machine learning techniques. However we employ manual mechanism to tag keywords into name of disease, ingredient or other.

Eg. In the sentence 'வாய்துர்நாற்றம் போக்கவும் ஏலக்காய் சிறந்த மருந்து', வாய்துர்நாற்றம் falls into the category of disease and ஏலக்காய் falls into category of ingredient. Rest are not keywords hence are tagged as other.

C. Assumptions

Few assumptions to be considered in this system are:

- i. Biomedical documents are manually tagged and stored based on the name of the disease and name of the core medicinal ingredient.
- ii. NE Dictionary contains list of named entities that include name of the disorders and also names of a variety of medicinal elements.
- iii. Anaphoric pronouns are avoided in the instructions to avoid co-reference ambiguities.

D. Algorithm:

The algorithm for the proposed system is as follows

STEP 1: Retrieve the relevant tagged documents on receiving a query.

STEP 2: For each document retrieved,

2a: Do Morphological Analysis using an Analyser to obtain morphemes.

2b: Do Category Tagging using the NE Dictionary.

STEP 3: From the category tagged document, Assign

3a: the element in tag <DD> to name of the disorder

3b: the elements in tag <MM> to the list of ingredients required

3c: the contents of tag <மருந்து> from original document to the procedure.

STEP 4: Feed the Information to the template. Output is obtained in a structured format.

Fig.1 gives the overall working procedure for the system.

Fig. 1 Traditional Medicine Information Extraction

IV. RESULTS AND DISCUSSION

A. Dataset

The information extraction routine was carried out for nearly 160 documents covering over 30 different common disorders ranging from simple disorders like hiccups (விக்கல்) or dry cough (வறட்டு இருமல்) to diabetes (சர்க்கரை or நீரிழிவு நோய்).

B. Sample Query:

The information extracted for a given query is depicted in the following snapshots.

For a given query- அஜீரணம், the relevant tagged document is retrieved (Fig.2).

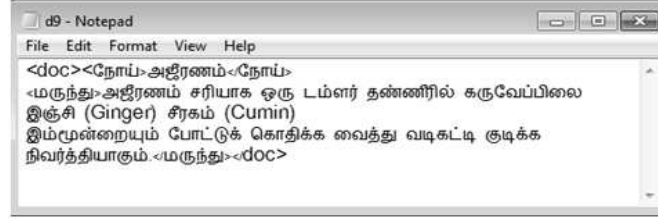


Fig. 2 Retrieve relevant document for the given query

The instructions in the document are set as input to the Morphological Analyser. Fig.3 shows the analysis output and the morphemes obtained.

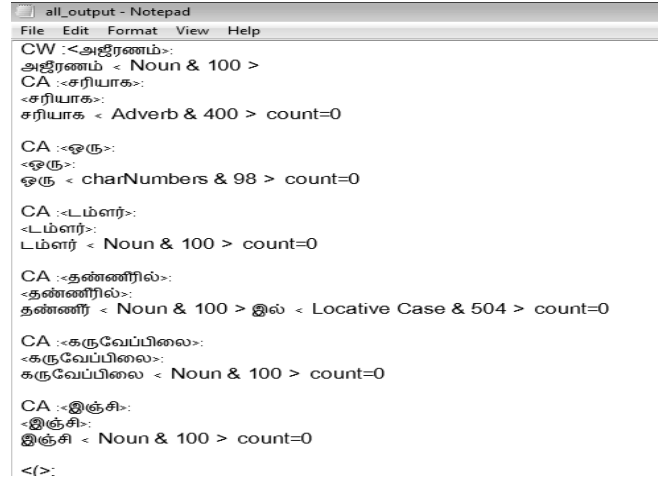


Fig 3. Morphological Analysis

The keywords are now extracted and tagged into their categories (Fig 4 & Fig.5).

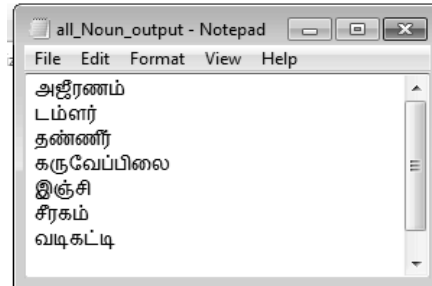


Fig 4. List of Morphemes

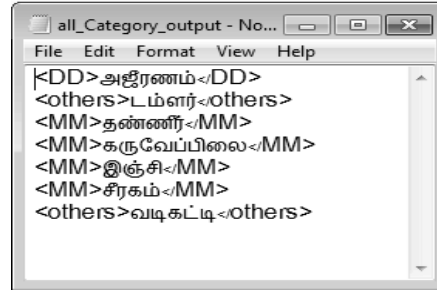


Fig 5. Category Tagging

The remedial instruction from the original document from the tag <மருந்து> is obtained (Fig.6).

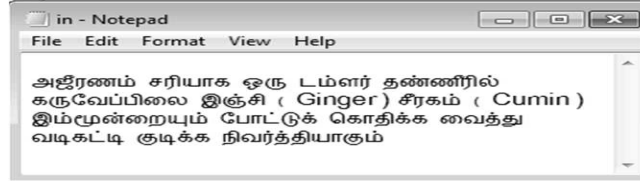


Fig 6. Remedial Instructions

The information about the ingredients used and the remedy preparation instruction are fed to the template. The output is displayed in a structured format (Fig.7).

நோயின் பெயர் :	அஜீரணம்
மருந்து :	
தேவையான பொருள்கள் :	கருவேப்பிலை இஞ்சி சீரகம் தண்ணீர்
மருந்து செய்யும் முறை :	அஜீரணம் சரியாக ஒரு டம்ளர் தண்ணீரில் கருவேப்பிலை இஞ்சி (Ginger) சீரகம் (Cumin) இம்மூன்றையும் போட்டுக் கொதிக்க வைத்து வடிகட்டி குடிக்க நிவர்த்தியாகும்

[Prev](#) [Next](#)

Fig 7. Output in structured form

C. Evaluation of Results and discussion

The arduous part in the given framework is tagging of the keywords into their respective categories. For experimental purpose, the whole process was run for 40 files, each containing a remedial instruction for a particular disorder. The f-score was found to be 0.782. The fall in the value of f-score is mainly due to the lower recall (0.72). The precision however was found to be quite high as manual category tagging in involved (0.858).

TABLE 1

Incorrectly tagged words by the Morphological Analyser

Type of word	Example	Tagged format
Compound noun [noun + noun]	கோபுரம்தாங்கி	<கோபுரம்தாங்கி> கோபுரம் < Noun & 100 > தாங்கி < Noun & 100 > <கோபுரம்+தாங்கி> count=2
Partly tagged words	எலுமிச்சம்	<எலுமிச்சம்>: மிச்சம் < Noun & 100 >
Noun tagged as blended word	கண்டங்கத்திரி	<கண்டங்கத்திரி>: கண்ட < Adjective & 300 > ங் <sandhi> கத்திரி < Noun & 100 >

Compound words [proper noun + common noun]	அருகம்புல் சாறு	<அருகம்புல்>: <அருகம்புல்>: அருகம்புல் < Noun & 100 > count=0 CA :<சாறு>: <சாறு>: சாறு < Noun & 100 > count=0
Noun tagged as particle	பால்	<பால்>: பால் < Particle & 305 > count=0
Ambiguous noun	பாலை	<பாலை>: பாலை < Noun & 100 > count=0
Untagged word	சந்தனக்கட்டை	Not tagged
Entities	சர்க்கரை	<சர்க்கரை>: சர்க்கரை < Entity & 50 > count=0
Compound noun [adjective + noun]	வெள்ளை பூண்டை	
Noun tagged as pronoun	நீர்	<நீர்>: நீர் < Pronoun & 104 > count=0
Unknown words	ஓமவள்ளி	<ஓமவள்ளி>: <unknown>
Element tagged as different element	மாதுளம்	<மாதுளம்>: மா < Noun & 100 > த் < Sandhi & 1401 > உளம் < Noun & 100 > <மா+த்+உளம்> count=2

The fall in recall value is due to various reasons. Table 1 shows the types of words that were tagged incorrectly by the Analyser. The main reason for fall in f-score is found to be the absence of the medicinal terms in the analyzer dictionary. For eg. வெற்றிலை (Betel leaves), a common word is tagged <unknown> by the morphological analyser.

Another reason for incorrect extraction is due to the splitting of a sing word into two words. For eg. அதிமதுரம் (Liquorice) is split into two separate nouns, அதி and மதுரம் by the analyser. Hence a single ingredient is split into two, thereby preventing it from being tagged. These kinds of errors can be rectified by adding the unknown words and the compound nouns to the analyser dictionary.

Since the keyword extraction involved extracting only proper nouns, certain ingredient elements that were tagged as entity were not retrieved. Eg. சர்க்கரை (sugar) was placed in the category of entities. Hence to overcome this error, words tagged as <Entity> are also considered as a keyword and are subjected to category tagging.

A very common difficulty that was observed in most of the text is two words denoting the same item. Eg. அத்தி பழம் (Fig) has two words but denotes the same fruit. However when analysing, they are tagged as two separate nouns. Though அத்தி denotes the fruit fig, பழம் here is a common noun. Hence it gives an ambiguous output. Similar error can be found in கமலா ஆரஞ்சு and தேங்காய் எண்ணை. This factor can be resolved by considering common nouns and the adjacent proper noun and tagging them as a single word. This however may not be successful in all cases.

Another difficulty that was observed was to identify the correct sense of certain words in the given context. For eg. நீர் (Water) and நெய் (ghee) are proper nouns denoting elements that are used to prepare the remedy. But they are tagged as pronoun and verb respectively.

V. CONCLUSION AND FUTURE WORK

Thus the given system can extract traditional medicine information from tamil health tip documents and display the essential output in a structured format. Traditional medicine terms are extracted and tagged. The system may be extended to support disease based or ingredient based queries or both. To avoid the errors due to missing terms in the category tagged data, the NE dictionary can be built using machine-learning algorithms. Also to avoid the incorrect understanding of terms in the context, the Word Sense Disambiguation (WSD) can be addressed.

ACKNOWLEDGMENT

This project is funded by **Centre for Technology Development and Transfer**, Anna University, Chennai under Sanction No: 2984/CTDT-1/RSS SIP/CEG/2013 Batch-5 for “Student Innovative Project under Research Support Scheme”.

REFERENCES

- Agirre, Eneko, Aitor Soroa, and Mark Stevenson. "Graph-based Word Sense Disambiguation of biomedical documents." *Bioinformatics* 26.22 (2010): 2889-2896.
- Anandan, P., Ranjani Parthasarathy & Geetha, T.V., 2001. “*Morphological Analyser for Tamil*”, ICON 2002, RCILTS-Tamil, Anna University, India.
- Aronson, Alan R. "Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program." In *Proceedings of the AMIA Symposium*, p. 17. American Medical Informatics Association, 2001.
- Chopra, Deepti, and Sudha Morwal "named entity recognition in english using hidden markov model." *International Journal* (2013).
- Ekbal, Asif, and Sivaji Bandyopadhyay. "Named entity recognition using support vector machine: A language independent approach." *International Journal of Electrical and Electronics Engineering* 4.2 (2010): 155-170.
- Grishman, R. Information Extraction: Techniques and Challenges. In M.T. Pazzienza (Ed.), *Information Extraction* (Int. Summer School SCIE-97), Springer Verlag, 1997.
- Haddow, Barry, and Beatrice Alex. "Exploiting multiply annotated corpora in biomedical information extraction tasks." *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco*. 2008.
- Harkema, Henk, et al. "Information extraction from clinical records." *Proceedings of the 4th UK e-Science All Hands Meeting*. 2005.
- Kazama, Jun'ichi, et al. "Tuning support vector machines for biomedical named entity recognition." *Proceedings of the ACL-02 workshop on Natural language processing in the biomedical domain-Volume 3*. Association for Computational Linguistics, 2002.
- Kulick, Seth, et al. "Integrated annotation for biomedical information extraction." *Proc. of the Human Language Technology Conference and the Annual Meeting of the North American Chapter of the Association for Computational Linguistics (HLT/NAACL)*. 2004.
- Monz, Christof. "Document retrieval in the context of question answering." *Advances in Information Retrieval*. Springer Berlin Heidelberg, 2003. 571-579.
- Navigli, Roberto, and Mirella Lapata. "An experimental study of graph connectivity for unsupervised word sense disambiguation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32.4 (2010): 678-692.
- Song, Yu, Eunji Yi, Eunju Kim, Gary Geunbae Lee, and Soo-Jun Park. "POSBIOTM-NER: a machine learning approach for bio-named entity recognition." *Korea* 305 (2004): 350.
- Tjong Kim Sang, Erik F., and Fien De Meulder. "Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition." *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*. Association for Computational Linguistics, 2003.
- Uzuner, Özlem, Imre Solti, and Eithon Cadag. "Extracting medication information from clinical text." *Journal of the American Medical Informatics Association* 17.5 (2010): 514-518.
- Xu, Hua, Shane P. Stenner, Son Doan, Kevin B. Johnson, Lemuel R. Waitman, and Joshua C. Denny. "MedEx: a medication information extraction system for clinical narratives." *Journal of the American Medical Informatics Association* 17, no. 1 (2010): 19-24.
- Zhou, GuoDong, and Jian Su. "Named entity recognition using an HMM-based chunk tagger." *proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2002.

தமிழ் மென்பொருள் உருவாக்கம் : தேவைகளும் திட்டங்களும்

முனைவர் ப.அர.நக்கீரன்

இயக்குநர், தமிழ் இணையக் கல்விக்கழகம்

சென்னை, இந்தியா

சுருக்கம்

தமிழ் மென்பொருள்களை உருவாக்குவதற்கான முயற்சிகள் பல ஆண்டுகளாகச் செய்யப்பட்டு வருகின்றன. ஆனால் முழுமையான மென்பொருள்கள் இன்னும் சந்தைக்கு வரவில்லை. இதற்கு, தமிழில் மென்பொருள்களை உருவாக்குவதற்கான அடிப்படை ஆராய்ச்சியும், ஆராய்ச்சியின் முடிவுகளும், முடிவுகளின் அடிப்படையில் அமைந்த தகவல் களஞ்சியமும், தகவலின் அடிப்படையில் அமைந்த செயல்விதிகளும், விதிகளைச் செயல்படுத்தத் தேவையான திட்டங்களும், நிதியும், இன்னபிற வசதிகளும் இல்லாததே காரணமாகும்.

எனவே தமிழ் மென்பொருள் உருவாக்கத்திற்குத் தேவையான தகவல் தளத்தை உருவாக்கும் முயற்சியில் தமிழ் இணையக் கல்விக்கழகம் முனைந்திருக்கிறது. இதன் முதல் கட்டமாக, இலக்கணக் குறிப்புடன் கூடிய சொல்களஞ்சியத்தையும், சொல் தொடர்க் களஞ்சியத்தையும், மின் அகராதியையும், பேச்சுத் தரவுத் தளத்தையும் உருவாக்கும் திட்டத்தை தமிழ் இணையக் கல்விக்கழகம் செயல்படுத்திக் கொண்டிருக்கிறது.

இத்தகவல் தளத்தைப் பயன்படுத்தித் தமிழில் மென்பொருள்களை உருவாக்கும் சூழலையும் த.இ.க. உருவாக்கி வருகிறது. தமிழ் மென்பொருள் உருவாக்கத் தேவையான கணிப்பொறிகள், மென்பொருள்கள், கருவிகள் மற்றும் அறிவுரைச் சேவைகளை உள்ளடக்கிய தமிழ் மென்பொருள் உருவாக்கும் மையம் (Tamil Software Incubation Centre) ஒன்று தமிழ் இணையக் கல்விக்கழகத்தில் நிறுவப்பட்டுள்ளது.

ஆக்கப்பூர்வ எண்ணங்களைக் கொண்டிருப்போர் தேவையான ஆய்வு மேற்கொள்ளவும், சோதிக்கவும், மேம்படுத்தி, செம்மைப்படுத்தி உலகிற்கு வழங்கவும் இந்த மையம் துணை புரியும்.

இத்திட்டங்களின் முன்னேற்றம் பற்றி எடுத்துரைப்பதே இக்கட்டுரையின் நோக்கமாகும்.

முன்னுரை

இன்று மிக வேகமாக வளர்ந்து கொண்டிருக்கும் ஒரு துறை கணினித் தொழில் நுட்பமாகும். மூன்று மாடிகளில் செயல்பட்டுக் கொண்டிருந்த கணினியின் உருவம் சிறுத்து மேசைக்கு வந்து, மேசையிலிருந்து மடிக்கு வந்து, கையில் வந்து செல்பேசிகளில் உலாவரும் காலம் இது.

இந்தக் கணினி யுகத்தின் வேகத்திற்கு ஈடு கொடுக்காத எந்த மொழியும் வளர முடியாது; ஏன், வாழவே முடியாது. தமிழுக்கும் இது பொருந்தும். எனவே கணினித் தொழில் நுட்பத்துக்கு ஏற்பத் தமிழின் வளர்ச்சியைத் திட்டமிடுதல் இன்றியமையாத தேவையாகும்.

உலகத் தமிழர்கள், அரசுகள், கல்வி நிறுவனங்கள், ஊடகம், வணிகம் என்று பல்துறைகளுக்கும் தேவையான மென்பொருள்களை உருவாக்கி, நிறுவி, பராமரித்துக் கணினிவழித் தமிழ்ப் பயன்பாட்டைப் பெருக்குவதே இதற்கான வழியாகும்.

தமிழ் மென்பொருள்கள் பல உருவாக்கப்பட்டுள்ளன. ஆனால் அவை எதுவும் முழுமை பெற்றுப் பயனாளர்களைச் சென்றடையவில்லை.

தமிழ் மென்பொருள்களின் இன்றைய நிலை

இன்று உருவாக்கப்பட்டு வரும் சில மொழிபெயர்ப்பு மென்பொருள்களைப் பயன்படுத்தினால், 'Tamil Virtual Academy offers academic programmes' – என்பதைத் தமிழ் மெய்நிகர் அகாடமி கல்வித்திட்டங்கள் வழங்குகிறது (கூகுள் மொழிபெயர்ப்பு) என்றோ, தமிழ் இணையக் கல்விக்கழகம் விரும்பிக் கொடு கல்வி நிகழ்ச்சி நிரல் (இந்திய அரசு மொழிபெயர்ப்பு மென்பொருள்) என்றோ மொழிபெயர்க்கும். தமிழ் இணையக் கல்விக்கழகம் கல்வித் திட்டங்களை வழங்குகிறது என்று சரியான மொழிபெயர்ப்பைத் தராது.

Kavin waited for the train, but the train was late – என்பதை, கவின் ரயில் காத்திருந்தார். ஆனால் தாமதமாக இருந்தது (Google) என்று மொழிபெயர்க்கும். கவின் இரயிலுக்காகக் காத்திருந்தான். ஆனால் ரயில் தாமதமாக வந்தது – என்பதே சரியான மொழிபெயர்ப்பாக இருக்கும்.

இதேபோல் ‘அவன் கணக்கில் புலி’ – என்பதை Tiger in his account (Google) என்றோ, He is maths tiger – என்றோ மொழிபெயர்க்கும். He is an expert in maths – என்று சரியாக இருக்காது.

மேலும் இன்றுள்ள சொல்பிழை திருத்திகளைப் பயன்படுத்தி,

அகர முதல வெழுத்தெல்லா மாத்ரி

பகவன் முதற்றே உலகு

என்ற குறளைச் சரிபார்த்தால் வெழுத்தெல்லா முதற்றே என்ற சொற்களைத் தவறு என்று காட்டும்.

அச்சிலிருந்து பேச்சாக மாற்றும் மென்பொருள்கள் அம்மா, அவர், ஐ, கள் என்ற சொற்களைச் சரியாக ஒலிக்கும். ஆனால் அம்மாவை, அவர்களை என்ற சொற்களைச் சரியாக ஒலிக்காது. பேச்சை அச்சாக மாற்றும் மென்பொருள்களின் நிலையும் இதுவேதான். இவற்றுக்கெல்லாம் அடிப்படைக் காரணங்களாகப் பின்வருபவற்றைக் கூறலாம்.

1. தமிழில் தரவுத் தளங்கள் தேவை என்பதை உணர்ந்து, திட்டமிட்டு உருவாக்கப்படாமை.
2. தமிழில் மென்பொருள்களைப் பயன்படுத்துவோர் குறைவு என்பதால், மென்பொருள் உருவாக்குவோர் ஆர்வம் காட்டாமை.
(அல்லது)
மென்பொருள் இல்லாமையால் பயன்படுத்துவோர் குறைவு.
3. மென்பொருள் உருவாக்குவோரிடம் ஒரு ஒருங்கிணைப்பு இல்லாமை.
ஒரே மென்பொருளை பலர் அரைகுறையாகச் செய்து, பாதியிலேயே கைவிட்டு விடுவது; அதுவரை அவர்கள் செய்த பணியை மற்றவர்களிடம் சொல்லாமல், தாங்களே கழுக்கமாக வைத்துக் கொள்வது.
இதனால் ஒன்றிலிருந்து மற்றொன்று என்ற வளர்ச்சி வேகம் தடைபடும். அனைவரும் சக்கரங்களையே உருவாக்கிக் கொண்டிருந்தால், வண்டி எப்பொழுது தாயாராகும்?
யார் என்ன செய்து கொண்டிருக்கிறார்கள் என்ற புரிதலே இல்லாமல் தமிழை வளர்க்க முயற்சி செய்து கொண்டிருக்கிறோம்.
4. அறிவும் திறனும், ஆக்கப் பூர்வ எண்ணங்களும் கொண்டிருப்போருக்குத் தேவையான நிதியும் வாய்ப்புகளும், வசதிகளும் கிடைக்காமை.

இந்தக் காரணங்களை நீக்குவதற்கான முயற்சிகளைத் தமிழ் இணையக் கல்விக்கழகம் முன்னெடுத்திருக்கிறது.

த.இ.க.வின் திட்டங்கள்

1. தமிழில் உள்ள சொற்களுக்கான இலக்கணக் குறிப்பு

தமிழில் உள்ள ஒவ்வொரு சொல்லுக்கும் மொழியியல் அடிப்படையில் இலக்கணக் குறிப்பு கணினி புரிந்து கொள்ளும் வகையில் தரப்படும். இதனால், கணினியால் சொற்களைச் சரிபார்க்கவும், அவற்றை இலக்கணப்படி பிரிக்கவும், கோர்க்கவும் இயலும். மேலும் குறிப்பிட்ட சொல்லுக்கு அடுத்து இன்ன சொல்தான் வர வேண்டும் என்ற விதிகளின்படி கணினியால் ஒரு சொல் தொடரைத் தீர்மானிக்க முடியும்.

எடுத்துக்காட்டாக, ஒரு பெயர்ச்சத்தை அடுத்து ஒரு பெயர்ச்சொல்தான் வரவேண்டும். ஒரு வினையெச்சத்தை அடுத்து, ஒரு வினைதான் வரவேண்டும் என்ற விதியின்படி, சொற்களைச் சரியாகக் கோர்க்கக் கணினியால் முடியும்.

எ.கா:

<u>மலர்ந்த முகம்</u>		<u>பறந்து</u>	<u>வா</u>
பெ.எச்	பெ	வி.எச்	வி
<u>சிவந்த</u>	<u>மேனி</u>	<u>மோதி</u>	<u>பார்</u>
பெ.எச்	பெ	வி.எச்	வி

மேலும் வேற்றுமை உருபைப் பிரித்துக் காட்டுவதால், ஒரு சொற்றொடரில் எந்த வேற்றுமை உருபு வரவேண்டும் என்பதைக் கணினி புரிந்து கொள்கிறது.

எனவே கணினிக்குத் தமிழ்ச் சொற்களின் இலக்கணம் அடிப்படையாகத் தெரிந்திருக்க வேண்டியது அவசியமாகும். இதற்கு இத்திட்டம் துணை செய்யும்.

2. தொடரியல் மற்றும் பொருண்மையியல் விரிதரவு

கணினிக்குச் சொல் இலக்கணம் மட்டுமல்லாது, தொடரியலில் எழுவாய் மற்றும் பயனிலை, பெயர்த்தொடர், சொல்லுருபுத் தொடர் போன்ற தொடர்களையும், பெயரெச்சத் தொடர், குறையெச்சத் தொடர் போன்ற எச்சத் தொடர்களையும், பொருண்மையியலில் உவமை, உருவகம், குறிப்பொருள், உணர்பொருள் ஆகியவற்றையும் தெரிந்திருக்க வேண்டும். ஏனெனில் எழுவாய், பயனிலை எங்கு வரும் என்பதைக் கணினியால் கணிக்கமுடியும். தமிழ்மொழி இடதுபக்கக் கிளை வடிவில் (Left side branch) வளரும் மொழியாகும்.

சான்று

நல்லவன்

அவன் நல்லவன்

நேற்று வந்த அவன் நல்லவன்

வண்டியில் நேற்று வந்த அவன் நல்லவன்

பொருண்மையியலில் உவமை (simile), உருவகம் (metaphor), குறிப்பொருள் (connotation), உணர்பொருள் (denotation) போன்றவைகளைச் சுட்டிக்காட்டுவதால் சரியானதொரு மொழிபெயர்ப்பு மென்பொருளை இதனால் உருவாக்க முடியும்.

3. உச்சரிப்புடன் கூடிய தமிழ் மின் அகராதி

இவ்வகராதி உதவியுடன் தமிழ் மொழிபெயர்ப்புக்கான நல்லதொரு மென்பொருளை உருவாக்க முடியும். ஏனெனில் தமிழ்ச் சொற்களுக்கான ஆங்கிலப் பொருள் மற்றும் இலக்கணக் குறிப்பு, பயன்பாடு போன்றவைகள் இவ்வகராதியில் அமைந்துள்ளன. இதனால் புலி என்ற சொல்லுக்கு ஒரு விலங்கு என்பது மட்டுமல்லாமல், வல்லமை என்ற பொருளும் கிடைக்கும்.

4. பேச்சொலி தரவு

பேச்சொலியை அச்சாகவும், அச்சைப் பேச்சொலியாகவும் மாற்றக் கூடிய மென்பொருள்களை உருவாக்க ஒரு பேச்சொலி தரவு இன்றியமையாத தேவையாகும். ஒவ்வொரு சொல்லும் அது வரும் இடத்திற்கேற்ப ஒலி அளவில் மாறுதல் ஏற்படும். எனவே ஒவ்வொரு சொல்லுக்கும், அதன் வடிவங்கள், பேசுவோர் (ஆண், பெண்), குழல் ஆகிய வேறுபாடுகளையும் புரிந்து கொண்டு செயல்பட, கணிப்பொறிக்கு ஏராளமாகப் பேச்சொலி தரவுகள் தேவைப்படுகின்றன. இதன் மூலம் ஒருவர் பேசும்போதே அவரின் சொற்களை எழுத்துருவாக மாற்றும் வகையில் ஒரு மென்பொருளை உருவாக்க முடியும்.

5. தமிழ் மென்பொருள் உருவாக்கும் மையம்

தமிழ்நாட்டில் தமிழ் மென்பொருள்களை உருவாக்குவோர் மிகக் குறைவாகவே உள்ளதற்கு, அவர்களுக்கு போதிய வசதிகள் இல்லாமையும் ஒரு காரணமாகும். எனவே ஆக்கபூர்வ எண்ணங்களைக் கொண்டிருப்போர், அந்த எண்ணங்களைச் செயல்படுத்தத் தேவையான ஆய்வு மேற்கொள்ளவும், பரிசோதிக்கவும், மேம்படுத்தவும், செம்மைப்படுத்தி உலகிற்கு வழங்கவும் தேவையான கணிப்பொறிகளும், மென்பொருள்களும், மற்ற கருவிகளும், சேவைகளும், இடமும் கொண்ட மையம் ஒன்றைத் தமிழ் இணையக் கல்விக்கழகத்தில் நிறுவும் பணிகள் தொடங்கப்பட்டுள்ளன.

இம்மையத்தில், மாணவர்களும், நிறுவனம் சாரா தனியாளர்களும், சிறு, குறு நிறுவனங்களும் அவர்களுக்குரிய கட்டணங்கள் செலுத்திச் சேர்ந்து கொள்ளலாம்.

த.இ.க. வல்லுநர் குழு, அவர்களின் திட்டங்களை ஆராய்ந்து, அதனடிப்படையில் அவர்கள் சேர அனுமதியளிக்கும். ஒவ்வொருவருக்கும் அச்சப் பொறி, நகல் எந்திரம், மென்பொருள்கள் கொண்ட நூலகம் ஆகியவை பொதுவானதாக இருக்கும். கலந்துரையாடல் கூடமும், கருத்தரங்கக் கூடமும் ஏற்படுத்தப்படும்.

த.இ.க.வின் வல்லுநர் குழு, ஒவ்வொரு திட்டத்தையும் தொடர்ந்து கண்காணித்து, தேவைப்படும் அறிவுரைகளையும் தகவல்களையும் வழங்கும்.

உருவாக்கப்படும் மென்பொருள்களைச் சந்தைப்படுத்தத் தேவையான வணிக நிதி ஆலோசனைகளை, உரிய வல்லுநர்களைக் கொண்டு வழங்கும்.

ஒரு மென்பொருள் உருவாக்கத் தேவையான அனைத்து வசதிகளையும், மென்பொருள்களையும், கருவிகளையும் கொண்ட அமைப்பாக இம் மையம் உருவாக்கப்படும்.

முடிவுரை

இத்திட்டங்களும் த.இ.க. மின் நூலகத்தில் பனுவல் வடிவில் உள்ள நூல்களும், இணைய வகுப்பறையில் உள்ள தமிழ் விரிவுரைகளும் தமிழில் மென்பொருள்கள் உருவாக்குவதற்கு ஒரு தூண்டுகோளாக அமைய வேண்டும்; அதனால் பல மென்பொருள்கள் உருவாக வேண்டும் என்பதே எங்கள் விருப்பம். எனவே இத்தரவுகள் அனைத்தும் பலரும் பயன்பெறும் வண்ணம் திறந்த வெளியில் கிடைக்கும்.

ஒரு மென்பொருள் உருவாவது அதன் தேவையைப் பொறுத்து வேகமெடுக்கும். பயன்படுத்துவோர் மிகுதியானால், மென்பொருள்களின் எண்ணிக்கையும் மிகுதியாகும்.

பயன்படுத்தாத எதுவும் பாழ்பட்டு அழியும் என்பதைத் தமிழர்கள் நினைவில் கொள்ள வேண்டும். கணித்தமிழுக்கு மட்டுமல்ல, தமிழுக்கும் இது பொருந்தும்.

அரிய பொருளெதையும் –

தேடியளிக்கச் சிலர்போதும் – மற்றதனைக்

கூடிச் சுவைக்க குவலயத்தை நாமழைப்போம்.

- குலோத்துங்கன்

பார்வை:

1. அ.ஜேம்ஸ், ப.அர.நக்கீரன், தமிழ்மென்பொருள் உருவாக்கத்திற்கு அடிப்படைத் தேவையான தமிழ்த் தரவுகள் (பக்.100) 11ஆம் உலகத் தமிழ் இணைய மாநாட்டு கட்டுரைகள், 2012.
2. உலகத் தமிழ் இணைய மாநாட்டுக் கட்டுரைகள்
3. Translate.google.co.in
4. <http://tamilvu.org/coresite/html/cwann>

Standardize fonts to catapult Tamil to the digital forefront

A G Ramakrishnan

Department of Electrical Engineering, Indian Institute of Science, Bangalore, India.

Abstract

A strong case is made for the standardization of a few Tamil fonts both for official publications (including school and college books) of the Governments and public display boards (including hoardings, road signs, name boards of establishments) in countries, where Tamil is (one of) the official languages of the State. The basic purpose is to guarantee a great recognition performance for both optical character recognizers and camera based document analysis and recognition systems. In this internet and mobile era, this committed action is bound to tremendously increase access to information for all through the search media, as well as boost the accessibility to all kinds of print information to people with visual disability. With the use of the ubiquitous smart phones, this also can give immediate access to translation of any Tamil material (printed or camera captured) to all the valuable tourist and business visitors. If we are the first in the world to do this, then we can catapult Tamil language to the digital forefront.

Preamble

Standardization is the sine qua non of twenty first century world living, which has made affordable and ultimate technology available to common man. Unless one observes deeply, it may not be immediately obvious to the common man that the many tools and appliances that have become part of our daily lives are attributable to the innumerable standards that have been created by extended discussions and common agreement between the industrial houses, scientists and Governments, as the case may be. For example, in spite of the fact that some of us may have issues with the current Unicode standard, no one can deny the tremendous increase it has provided to access of information on the internet. This is in general true, that even an imperfect standard is a lot better than near perfect, multiple technologies that do not communicate to each other. This is also immediately evident today, where the key inputs, representation, transmission and rendering of Tamil text has not been standardized across the many manufacturers of mobile phones of the world. Partly, this is also due to the lack of adequate pressure on the involved parties by the concerned public and Government machinery. Because of this, we are not able to do the simple thing of texting in Tamil through our mobiles in an era, where one can take a colour photo and send it anywhere in seconds!

Gains due to standardization of communication technology

Without standardization of the modulation/demodulation, transmission technology, as well as the design and manufacturing technology of VLSI processors, we won't be enjoying the amazing benefits of talking to anyone in the world by the click of a button. Standardization of the several layers of the internet protocol has connected the world to an unbelievable level, that experts in any field (say medicine or art or archaeology) are able to instantly share their new knowledge, as well as benefit from the knowledge of other experts across the globe, by posting their questions on the internet discussion forums. Standardization of medical image formats through DICOM has resulted in considerable progress in the area of medical image processing, which has enabled sharing of crucial diagnostic information between different imaging modalities. This has increased possibilities for information fusion from the multiple modalities such as x-ray computed tomography (CT) and magnetic resonance imaging (MRI), thus facilitating improved diagnosis to the benefit of both patients and physicians.

Negative Examples

We will now see some examples as to how non-standardization leads to difficulties for the people. In India, there are several educational boards that have distinct syllabi and method of evaluation at the school level: State Boards, Central Board of Secondary Education (CBSE) and Indian Council of Secondary Education (ISCE). Because of these differences, when these students apply for admission to colleges offering engineering education, they are asked to write separate entrance examinations, based on the performance in which these candidates are given admission in the college. Once again, there is difference of opinion between different deemed Universities and the State Governments with respect to these entrance examinations, because of which one needs to write multiple entrance examinations!

Virtual Education

Standardization of Unicode has made every Tamil website and digital document accessible to everyone and information exchange and dissemination has been highly facilitated. Standardization of image formats through jpeg, tif, etc. has resulted in easy sharing of photographs and other pictures through internet. Standardization of video formats has resulted in the revolutionary phenomenon of youtube, which has facilitated rapid sharing of very valuable information such as educational videos. Many reputed Universities of the world have made available open courseware through the internet, which has extended the reach of great teachers to committed students anywhere in the world. The availability of audiovisual pedagogy on demand, as opposed to one's reading of text from a book, has enhanced the possibility of rapid learning.

e-Shopping

In today's automated world, we have standardized the dimensions of our dresses too, which has made easy and affordable availability of readymade outfit. In fact, one is able to buy clothes, shoes and other bodyware online today, only because of standardized dimensions, image formats, secure online payments and international standardization of plastic money. The introduction of standardized sizes of doors have made replacement of damaged doors easy. In the same way, standardization of fonts for Government documents, forms and school books will go a very long way in enabling easy access to information through optimized OCR technology for the people with visual disability and even the majority population.

Enhanced interaction between human beings

In the internet era, use of mobiles has become ubiquitous for various applications such as sending emails, finding routes, seeing the meaning for a word, listening to music, making railway and bus reservations, taking photographs and short videos. Standardization of one or two good Tamil fonts for road displays, mile stones, billboard and signboards is bound to rapidly lead to reliable technology for mobile based capture, recognition and transliteration or translation of text. While we can happily use Tamil for all personal and official communications, it will not be a barrier for our friends, business associates and visitors, who read and speak other languages of the world.

Effectiveness of postal automation through standardization

In the United States of America, the sizes of the postal envelopes have been standardized to automate sorting efficiently through machines. Non-standard sizes of postal articles are penalized by an enhanced rate of postage.

Camera based document analysis and recognition

In today's world, there is too much of interdependence between the states and even countries in terms of economy, business, pleasure travel, etc. Outsourcing of various services has also necessitated wide spread travel of personnel to new places, where they may not know the local language and/or script. A unique phenomenon in India is that the language spoken changes every time and in every direction one travels around 500 km. This is not true even when one travels 1000 km in USA, Russia or China. Accessibility to information for business and tourist travellers and hence workability will be enhanced considerably, if the traveller can read and/or translate information in the foreign language (in our context, Tamil) to her/his language. Similarly, when a Tamil native is in other territories, she/he can use this technology to quickly translate necessary information in the local language to Tamil. Technology exists today, that can detect, localize [1] and extract [2] text reliably from images captured on the road using digital or mobile camera (scene images). Even if this text is curved or wavy, we can detect the contour and align the text line image to be horizontal [3], before attempting to recognize the text. Currently, since the font size, style and background are not standardized, the algorithm needs to have special capabilities [4] to deal with such images. This has necessitated active research in this area [5] and creation of many standard annotated databases [6] and tools [7]. Many of these will be obviated if we standardize the fonts, font sizes, and also foreground and background colors of the road signs, display boards, etc.

For example, in restaurants, one can (i) capture the menu items on the mobile camera, recognize it and quickly translate it to one's own language; (ii) identify names of shops and establishments; (iii) effortlessly identify the destinations of buses and trains; (iv) follow signs in airports and so on.

Recognition of printed text

For the past decade, we have been engaged in developing technology to digitize school and college books [8, 9] for the benefit of students with visual disability. Using our OCR technology, close to 200 Tamil school, college and other books have been converted to Braille books by Worth Trust, Chennai, which are being regularly used by blind students. Indian law

has made amendments to the copyright law, so that any book can be converted for the exclusive use of blind people, without violating the copyright laws. However, when different books use different fonts and formats, one is not able to always assure the best possible recognition performance by an OCR. However, at least for the school and college books published or authorized by the Government machinery, if the page layout and fonts are standardized, then the OCR can be optimized to give ultimate performance for such standard books. This will facilitate rapid digitization of existing books for such purposes. Similarly, if the court proceedings, real estate registration documents, Government notifications, forms, etc. are in standardized Tamil fonts, then any time the need arises, they can be digitized with great accuracy and speed.

Optical character recognizers can give a very high performance, if they are tuned to one or more specific fonts. While one may prefer to use artistic fonts for certain personal documents such as marriage invitations and award citations, it is highly desirable that we standardize fonts for all displays meant for the public, such as road names, bus and train boards. One can quickly capture such information using a mobile camera, recognize the text and transliterate or translate it to the target language. Since many of these boards contain common and proper nouns, even transliteration will go a long way in giving the essential information needed by a visitor. However, since the text involved is usually small, even effective translation may not present a great technical challenge.

Conclusion

If the merits available are sufficiently understood, then through committed planning of the Governments, the key players can be educated and the entire arena transformed in no time. Let us together cause this happen to Tamil first in the world; then we can catapult Tamil language to the digital forefront in the Universe.

References

1. T. Kasar and A. G. Ramakrishnan, "Multiscript and multioriented text localization from scene images, " Proc. 4th International Workshop on Camera-based Document Analysis and Recognition (CBDAR 2011), pp. 1-14, 2011, Beijing, China.
2. T. Kasar and A. G. Ramakrishnan, "COCOCLUST: Contour-based color clustering for robust text segmentation and binarization, " Proc. 3rd workshop on Camera-based Document Analysis and Recognition (CBDAR 2009), pp. 11-17, 2009, Spain.
3. T. Kasar and A. G. Ramakrishnan, "A method to extract and align text of an image captured and a system thereof, " Indian Patent Office Reference. No: 109/CHE/2011, 2011.
4. Deepak Kumar, M. N. Anil Prasad and A. G. Ramakrishnan, "NESP: Nonlinear enhancement and selection of plane for optimal segmentation and recognition of scenic word images, " Proc. International Conference on Document Recognition and Retrieval (DAR) XX, 5-7 February 2013, San Francisco, CA USA.
5. Deepak Kumar and A. G. Ramakrishnan, "Power-law transformation for enhanced recognition of born-digital word images, " Proc. 9th International Conference on Signal Processing and Communications (SPCOM 2012), 22-25 July 2012, Bangalore, India.
6. Deepak Kumar, M. N. Anil Prasad and A. G. Ramakrishnan, "Benchmarking recognition results on camera captured word images datasets, " Proc. Workshop on Document Analysis and Recognition (DAR 2012), 16 December 2012, IIT Bombay, Mumbai, India.
7. T. Kasar, D. Kumar, M. N. Anil Prasad, D. Girish, A. G. Ramakrishnan, "MAST: Multi-script annotation toolkit for scenic text, " Proc. Joint Workshop on Multilingual OCR and Analytics for Noisy and Unstructured Text Data (J-MOCR-AND 2011), pp. 1-8, 2011, Beijing, China.
8. Peeta Basa Pati and A. G. Ramakrishnan, "Word level multi-script identification, " Pattern Recognition Letters, 2008, Vol. 29, pp. 1218-1229.
9. D. Dhanya and A. G. Ramakrishnan, "Optimal feature extraction for Bilingual OCR, " Proc. Fifth IAPR Workshop on Document Analysis Systems DAS-02, Princeton, NJ, August 19-21, 2002, pp.25-36.

Notes on Tamil Orthography – puLLi, kAl, ai, ja, etc.

By Mani M. Manivannan, Senior Director of Engineering, Symantec Corporation, Chennai, India
And Chair, Working Group 2 on Tamil Unicode, INFITT.

Abstract:

Tamil orthography is of interest to Tamil computing. There have been suggestions to add new diacritics on top of the series of Tamil consonants with inherent a (அகரம் ஏறிய உயிர்மெய் வரிசை) to represent foreign sounds in loan words in Tamil. In this article, we will show that the genius of Tamil artists has created all possible shapes on top of the consonants with inherent a, to represent puLLi (புள்ளி, U+0BCD). This makes it very hard to consider any additional diacritics in that position to represent foreign characters. Also, as Tamil typefaces evolved from inscriptions and palm leaves to print, they became easier to read. Yet, there are still a few characters that present difficulties such as the discriminability of AI and JA (ஐ, U+0B90 and ஜ, U+0B9C), vowel signAA or kAl (஁, U+0BBE) and pure consonant RA (ர, U+0BB0+U+0BCD, when the descender is absent in the typeface) as well as the AU length mark (஁, U+0BD7) and LLA (ள, U+0BB3) confusion. In this article, we review some of these challenges and make some proposals to address them. This is of interest to students as well as internet users as disambiguation becomes important for Tamil domain names.

கட்டுரைச் சுருக்கம்:

சீர்முறையில் தமிழ் வரிவடிவம் பற்றிய குறிப்புகள் - புள்ளி, கால், ஐ, ஜ, முதலானவை பற்றிச் சில கருத்துகள்.

அகரமேறிய தமிழ் மெய்யெழுத்துகளின் மீது பிறைக்குறி போன்ற மீக்குறிகளை இட்டு வடமொழி மற்றும் பிறமொழி ஒலிகளைக் குறிப்பிடலாம் என்று அண்மையில் சிலர் பரிந்துரைத்து வருகிறார்கள். ஆனால், அகரமேறிய உயிர்மெய்யெழுத்தின் மீதும், அதன் உள்ளும் என்ன குறியீடுகளை இட்டாலும், அவற்றைத் தமிழர்கள் புள்ளியாகத்தான் பார்க்கிறார்கள் என்பதை இந்தக் கட்டுரை சான்றுகளுடன் நிறுவுகிறது. எனவே இப்படிப் பட்ட மீக்குறிகளை எழுத்தின் மேலோ, அதற்கு உள்ளேயோ இடுவது குழப்பத்தைத் தான் விளைவிக்கும். தமிழ் வரிவடிவங்கள் கல்வெட்டுகள், சுவடிகளிலிருந்து அச்சுக்கு மாறிய போது அவை பெரும்பாலும் சீராக்கப் பட்டன. இருப்பினும், ஒரு சில எழுத்துகள் இன்னும் மாணவர்களுக்கும், இணையப் பயனர்களுக்கும் குழப்பம் தருபவை. இந்தக் கட்டுரையில் ஐ/ஜ குழப்பங்கள், துணைக்கால், ரகர ஒற்றுக் குழப்பங்கள், ஓளகாரக் குறி, ளகரக் குழப்பங்கள் பற்றி ஆய்ந்து சில பரிந்துரைகள் தந்துள்ளோம். தமிழில் இணையக் கொற்றப் பெயர்கள் பரவலாகும் முன்னர் இவற்றிற்குத் தீர்வு காண்பது நன்மை தரும்.

The Blackwell Encyclopedia of Writing Systems (Coulmas 1996:379–80) defines Orthography as follows: “Correct spelling and that part of grammar that deals with the rules of correct spelling. An orthography is a normative selection of the possibilities of a script for writing a particular language in a uniform and standardized way. All orthographies are language specific.” Besides being concerned about spelling and grammar of a language, an orthography defines all aspects of writing including placement and shapes of characters, diacritics as well as spelling of loan words all in accordance with the rules of the language it codifies (Karan, 1996)

A *script* on the other hand represents the graphic form of the units of a writing system. (Coulmas (2003:35)). While *orthography* refers to the standardized variety of a given language-specific writing system, a script can be used by multiple languages.

Language communities may be expected to be in control of their orthographies with politics and socio-linguistics playing an important role in the design or reform of an orthography. Governments often influence orthographies through their language policies. With the advent of globally connected networks, non-state players such as multi-national corporations and non-governmental standards bodies that serve the needs of such corporations can have a disproportionate influence on orthographic designs. It is important for the language communities to engage all standards bodies early and work with them closely to make sure that their needs are addressed satisfactorily.

We can see the Tamil encoding standards as an example of the divergence between the efforts of the language community and those of the remote standards bodies. The initial design of ISCI (Indian Standard Code for Information Interchange) and

the Tamil block of Unicode had little or no participation from the Tamil community or even the state government of Tamil Nadu. The standards created by the latter eventually had to yield to the former and we are frequently forced to deal with the idiosyncrasies of the standards designed by people who weren't part of the Tamil language community. The best way to avoid surprises is for a language community to be aware that non-state players far beyond their borders can have much more influence on their language than their representative government and to engage these bodies primarily on technical terms in addressing the needs of the language. The good news is that frequently such bodies are more willing to address their concerns than their governments.

Any Tamil speaker who inspects the Tamil Unicode block would be surprised to see that the very first character listed at (<http://www.unicode.org/charts/PDF/U0B80.pdf>) is the character Tamil Sign Anusvara (see Fig. 1).

Various signs

The anusvara should not be confused with the use of a circular glyph for the pulli.

0B82 ூ TAMIL SIGN ANUSVARA

- not used in Tamil

0B83 ௃ TAMIL SIGN VISARGA

= aytham

- in fonts which display the Tamil pulli as a ring shape, the glyph for aytham also uses rings

Figure 1. Anusvara in Tamil Unicode Block

The sample glyph shows a circle above a character that it is to combine with. The annotation says it is not used in Tamil and an additional annotation saying that the “anusvara should not be confused with the use of a circular glyph for the pulli” has been recently added. In orthographic terms, anusvara belongs to Sanskrit language to represent a Sanskrit based sound using a written form that is characteristic of Devanagari or other scripts designed to represent Sanskrit and related languages. Tamil language does not use anusvara nor does it have a written form for the anusvara. As we show in Figure 2, Tamil uses circular glyph for puLLi quite regularly.



Figure 2. Circular glyphs used as puLLi in modern Tamil (2012-2013)

The author collected hundreds of samples over a period of 18 months starting in January 2012 taking pictures of various glyphs that represent puLLi in Tamil. Figure 2 shows a sample of such collection ranging from print articles, wall posters, business signs, wall advertisements, political cartoons etc. It is quite obvious that the circular glyph is read as a puLLi in Tamil and any use of a circular glyph to represent a foreign character in the Tamil milieu is certain to cause confusion.

It is with this in mind that an attempt was made by Shriramana Sharma to request deprecation of this anusvara character. (Sharma L2/12-018, **Request to deprecate 0B82 TAMIL SIGN ANUSVARA**). It was opposed by Naga Ganesan who argued that the anusvara sign is in use in Tamil script to transliterate from neighboring Indian scripts (Ganesan L2/12-051, **Proposal to change the glyph shape of Tamil Sign Anusvara (U+0B82)**). Interestingly, Ganesan had earlier opposed Sharma's extended Tamil proposal which was designed to transliterate Sanskrit and other Indian languages that havearga letters in addition to writing Saurashtri in Tamil script. At best, Ganesan's proposal is inconsistent. While it is to be expected that orthography wars will break out whenever attempts are made to change established norms triggering political and socio-linguistic controversies, here the usual linguistic politics of Tamil Nadu gets really muddled. We have an orthodox Sanskrit scholar associated with the Sankara Mutt of Kanchi fighting to remove an obviously Sanskrit orthographic element from the Tamil block while a member of the Kongu landed gentry is arguing for the Sanskrit orthographic element in Tamil to make it easier for transliterating Sanskrit in Tamil script!

One more example of Ganesan's inconsistent stand is his reference to McAlpin's use of a circular glyph as a diacritic to represent spoken Tamil (McAlpin, 1976:12, <http://nganesan.blogspot.com/2012/04/circular-dot-diacritic-in-tamil.html>). In the Unicode mailing list, Ganesan suggested correcting the notation on Anusvara sign with "used in Spoken Tamil representation for nasalization" even though there is no evidence that the McAlpin notation found any followers since 1976.

Diacritics:

Superposed circle (ஃ). The letter is not to be pronounced, but it may have a predictable effect on the preceding vowel. The letter will have its full pronunciation when followed by another vowel. This diacritic is used only finally and contrasts with the normal pulli (஁).

ஃ, ஃ --nasalize the preceding vowel with:
அஃ = [ɛ̃] or [ʌ̃] and அஃ = [ɔ̃]

ஃ, ஃ -- no vowel change. Used mainly in citation forms.
அஃ = [o:] for dialects with this shift.

Superposed check (ஃ). The consonant has an optional full release which is heard as a weak vowel of indeterminate quality or not, but the surrounding phonology is as if it were followed by a full vowel, i.e., the consonant is not part of a true consonant cluster.

இருக்கறேன் irukkrē = MLT இருக்கிறேன்

Figure 3. McAlpin's diacritic proposals using circular glyphs

Usage of diacritics to render foreign sounds, loan words, academic notation, etc., is not a new idea. There have been other proposals to use diacritics to render Sanskrit (Sharma, L2/10256), Hindi (Manikutti, 2012) or other foreign sounds (Sevakumar, 2010) in Tamil using the Tamil character set. However, such usage in random proposals, printed texts or a few other non-standard sources needs to be weighed against standard practices of language community. A rational system of diacritics limited to specialist user groups has its merits as long as the orthographic principles of Tamil are not impacted.

It is notable that in more than 2000 years of close interaction with Sanskrit, prakrits and neighboring languages, Tamil orthographic tradition has not integrated such orthographic practices as conjunct consonants, independent representation of aspirate and voiced sounds etc. from these languages, into its native repertoire. The only concession that it has made is to borrow an extremely limited number of grantha characters (ஸ, ஷ, ஜ, ஹ, ஸ்ரீ, and ஷ்). Even under the influences of the imperial Mauryas (300 BCE) and later the Pallavas (600 CE) who patronized the Pallava Grantha script that is a cousin of the modern Tamil script, orthographic tradition of Tamil had retained its fierce independence. (Gift Siromoney, 1983:21). It is this independence that has enabled Tamil script to retain its simplicity and the avoidance of conjunct consonants is an important feature of this simplicity.



Figure 4. Conjunct consonant ஷ decomposed in modern Tamil

Even when Tamil borrows a conjunct consonant such as the grantha ஷ (KSHA), it breaks it up into constituent parts in the course of time. (See Figure 4). The only other conjunct consonant in Tamil grantha (ஸ்ரீ, <0BB6, 0BCD, 0BB0, 0BC0>) is treated like a symbol similar to Tamil OM (U+0BD0, ஓ) rather than as a conjunct consonant because most Tamils do not know about the consonant ஸ்ரீ (U+0BB6) and are not aware that ஸ்ரீ is a conjunct consonant (ஸ்ரீ + ரீ).

As for using diacritics to render foreign sounds and loan characters, one needs to consider the Tamil practice of treating ANY diacritic or symbol on top of or inside a Tamil consonant as just another glyphic variant of pulli. The author has an interesting collection of photographic samples numbering into several hundreds that show this practice. In Figure 5, we show a very tiny sample of this practice.

[illegible]

However, since the developer communities outside of Tamil milieu is unfamiliar with such nuances of the script, they build implementations that may end up confusing the Tamil user community. For example, Nokia has implemented Tamil in their mobile offerings and have included “Tamil” anusvara with a circle above a consonant. At first I thought these to be a glyphic variant of puLLi. But I suspect that this is a serious implementation of anusvara. See Figure 6. (Maag, 2012) While implementers may read the Unicode annotation on anusvara cautioning them not to confuse “the circular glyph for puLLi”, users won’t and this will certainly create confusion.

ஸூ ஹூ கூ ஸூ ஸூ க்ங் ச ஜ்ஞ ட் ண்
த்ந் ன் ப் ம் ய் ர் ற் ல் ள் ழ் வ் ஸ் ஷ் ஸ்
ஹ் கூ ஸ் ஸ் க்ங் ச ஜ்ஞ ட் ண் த்ந் ன் ப்
ம் ய் ர் ற் ல் ள் ழ் வ் ஸ் ஷ் ஸ் ஹ் க்ஷ்
ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ்
ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ் ஸ்

ORTHOGRAPHIC CHALLENGES

While modern Tamil typefaces are easy on our eyes, ancient inscriptions and palm manuscripts are harder to read without a lot of training. While lack of spaces between words, punctuation, etc., are part of the reason for this difficulty, one can see that even with the advent of printing Tamil was somewhat difficult to read in the early days. The evolution of typefaces in Tamil starting with mid 19th century has made Tamil script quite modern and easy on the eyes. And yet there are some characters that remain difficult to read both on the printed page and on the computer screens. A study (Narayanan, 2011) reports that poor readers had difficulty in reading Tamil words that contained irregular orthography. They had difficulty in discriminating between ஐ and ஜ. They read ஒள as ஒ ள. The Tamil Nadu government standard on Tamil Unicode implementation refers to homographic confusion between the Tamil vowel sign AA or kaI (◌ஊ, **U+0BBE**) and pure consonant RA (ற, **U+0BB0+ U+0BCD**). Similar homographic confusion exists between AU length mark (◌ஊ, **U+0BD7**) and LLA (**U+0BB3**).

Fortunately, the resolution for solving the latter two problems is straightforward and besides the recommendation from the Tamil Nadu standard on Unicode, the Tamil Nadu state government text books for primary school children have implemented them. The Tamil Development department of the Tamil Nadu state government had clarified to the Tamil Nadu text book society what standard forms for pure consonant RA, RI and RII (ர், ரி, ரீ) they should use (Figure. 7). Nearly all the Tamil Nadu text books follow this recommendation though the author did find a few inconsistencies.

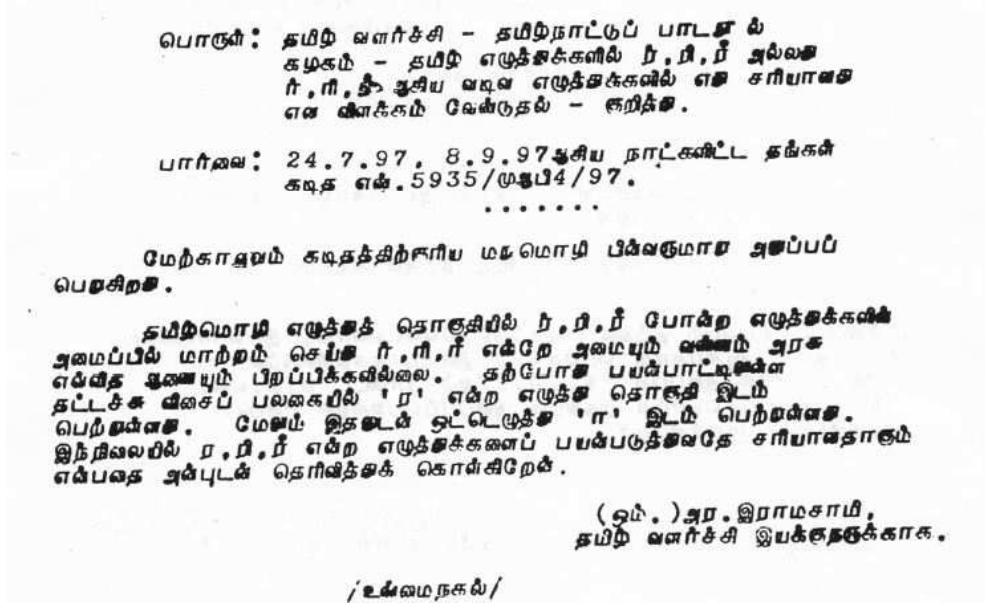


Figure 7. Tamil Nadu government's clarification on the orthography of ர், ரி, ரீ

Most of the Tamil fonts follow this convention though not all. Almost none of the magazines and newspapers in Tamil Nadu do. It is hoped that this research paper would alert the type foundries of supporting these changes and implement this consistently.

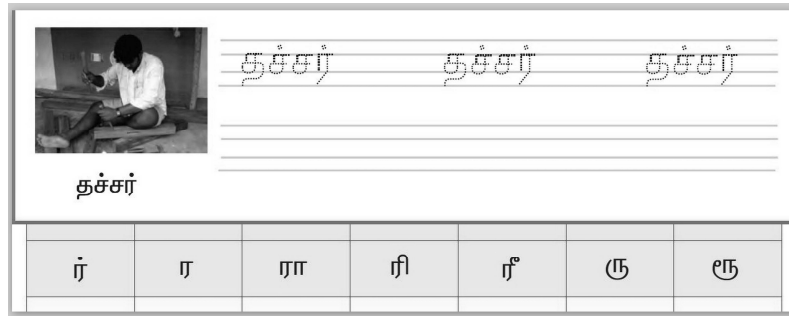


Figure 8. Rakara mey series, Tamil Nadu state text book for Standard 1, Vol III, Page 44

The simple addition of the descender makes the series consistent and avoids confusion. Figure 8 shows a page from the Tamil Nadu government text book (2013) for primary school children implementing this.

Similarly the government text books have implemented a way to disambiguate between the AU length mark and the Tamil Letter LLA. They have made the AU length mark smaller and identical to the glyph in the Tamil letter UU. The following figures show the pages from Tamil Nadu government text book.



Figure 9. The Au series, Tamil Nadu state text book, Standard 1, Vol III, pp 26-27

The Unicode reference glyph in the letters UU and AU is even better (Figure 10). That glyph cannot possibly be mistaken for Tamil letter LLA (௪).



Figure 10. The Au marker glyph from the Unicode Standard in comparison with the letter LLA

The discriminability between AI and JA (ஐ, U+0B90 and ஐ, U+0B9C) is much harder as can be seen from figure 11. It becomes even harder when designers use standard character sizes rather than the “uppercase” character sizes of AI and JA since the space required to disambiguate the lower glyphs is more restricted. The standard size AI and JA are esthetically more pleasing and are preferred by a larger number of designers. The best way to increase the discriminability is through elegant typeface design.



Figure 11. Various versions of JA series and AI in modern Tamil

What is difficult for the street sign designer need not be the same for computer typeface designer. The lower half of both the characters AI, JA, ஐ, ஐ can be designed with lesser weight and the bottom curves in JA need not be symmetric. An example of the weight variation that can be used to address this is similar to the character g in Franklin Gothic and Garamond typefaces as in: g (Franklin Gothic) and g (Garamond). Typeface designers have variety of other choices.

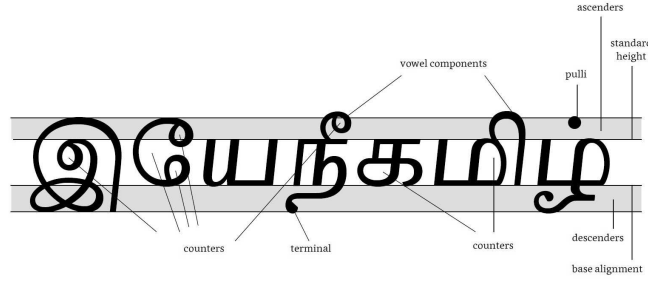


Figure 12. Tamil script alignment (Vargas, 2007)

Evolution of Tamil typefaces is a fascinating subject. Fernando de Mello Vargas has collected rare details about the Tamil typeface evolution. Anyone interested in Tamil typeface design will benefit by reading his Masters thesis (Vargas, 2007). In figure 12 given above Vargas describes the Tamil script alignment. While the letter இ sits below the baseline in his diagram above (and in some modern designs), it was common until a few decades ago to make it sit on the baseline with the ascender extending to the topline. Other alignment details in the figure are consistent with the current practice.

In general, as observed from the lettering in modern Tamil Nadu, the designers appear to have several basic characters which sit on the baseline. These are mostly short vowels (அ, ஈ, உ, ஊ, எ, ஓ) with ஐ, ஒ, ஓ as optional members of the set and the following consonants (க, ங, ச, ட, ண, ப, ம, ய, ல, வ, ள, ன, ஸ, ஜ). These characters don't have ascenders or descenders. The puLLis and vowel modifiers of these characters go into the ascender or descender spaces. The long vowels (ஆ, ஏ) and (ஔ, த, ந, ர, மு, ற, ஷ, ஹ) all have glyphs that rest on the base line with descenders that may not extend to the bottom line. Sometimes letters like ஔ, ந are shortened with no descenders at all. The letter இ used to sit on the baseline with only an ascender glyph completing it until the 70's. The ascender stroke for இ is nowadays seen as a separate stroke with no continuity with the rest of the glyph. The letters ஐ, ஒ, ஓ, and ஜ used to sit on the baseline and their body was completely within the middle space in some old designs. Several modern designers treat these as characters with their bottom half as descenders. I have given below pictures with both designs for comparison.



Figure 13. Characters with and without descenders and shorter au marker



Figure 14. Vowels and AI character with and without descenders and shorter au marker



Figure 15. JA characters with and without descenders

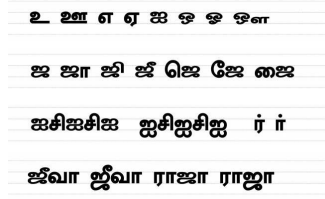


Figure 16. Characters with and without descenders for comparison

Recommendations:

1. Continue to request Unicode Technical Committee to deprecate 0B82 Anusvara Sign and annotate it such that the circular glyphs above the character are not used for that at all.
2. Encourage more technical participation in INFITT WG02 and UTC meetings and lists.
3. Implement your typefaces to follow the Tamil Nadu state text book model for representing AU length mark and RA pure consonant, RI and RII (ர், ரி, ரீ)
4. Use your best typeface designs to disambiguate AI and JA (ஐ, ஜ)
5. Consider making the characters ஐ and ஜ sit on the base line instead of the bottom line and the body of the character fit within the middle.
6. Review all your typefaces to see if readers unfamiliar with the script or those with learning disability may find your characters difficult to disambiguate
7. Consider making the character இ sit on the base line instead of the bottom line.
8. Continue to monitor Tamil Orthography and Typography discussions also
9. Discuss the need to represent loan words in Tamil or render foreign sounds in dictionaries with a rational set of diacritics with Tamil characters.
10. Pay attention to smartphones where the characters are tiny.

Acknowledgements:

Thanks to ‘Thendral’ Jeevamani Balan for the discussions on the characters and his comparison diagrams in figures 13 through 16 above. Thanks to Tamil Virtual University (Prof. Nakkeeran and Mr. Mahesh) for the copy of the Tamil Development Department memo. Thanks to Cartoonist Bala for excerpts from his cartoon. Thanks to Prof. C. R. Selvakumar for a nice discussion on the diacritics. Thanks to Dr. S. Palaniappan for a long discussion on the subject matter and for his valuable input. Thanks to my wife Asha for putting up with me while I was taking pictures of posters, wall signs, name boards for the last couple of years. She is really relieved that this paper is complete and hopes that I will stop being crazy!

References:

The Au series, Tamil Nadu state text book, Standard 1, Vol III, pp 26-27

<http://www.textbooksonline.tn.nic.in/Books/Std01/Std01-III-TamEng.pdf>

Baker, Philip. 1997. Developing Ways of Writing Vernaculars: Problems and Solutions in a Historical Perspective. In: Andrée Tabouret-Keller, Robert B. Le Page, Penelope Gardner-Chloros, and Gabrielle Varro (eds.), *Vernacular Literacy: A Re-Evaluation*. Oxford: Clarendon Press, 93–141.

Coulmas, Florian. 1996. *The Blackwell Encyclopedia of Writing Systems*. Oxford: Blackwell.
2003. *Writing Systems: An Introduction to their Linguistic Analysis*.
Cambridge: Cambridge University Press.

Ganesan, Naga., L2/12-051, Proposal to change the glyph shape of Tamil Sign Anusvara (U+0B82)).

<http://www.unicode.org/L2/L2012/12051-tamil-anusvara.pdf>

Karan, Elke., 1996. WRITING SYSTEM DEVELOPMENT AND REFORM: A PROCESS, B.A. thesis, University of North Dakota.

Maag, Dalton., 2012. “Type Design meets Tamil”, 14 June 2012,

[Nokia implementation of “Tamil” anusvara]

<http://nokiapureblog.daltonmaag.com/2012/06/type-design-meets-tamil.html>

Manikutty, Anand, A Modest Proposal towards Modernizing the Tamil Writing System (July 7, 2012). Available at SSRN: <http://ssrn.com/abstract=2184552> or <http://dx.doi.org/10.2139/ssrn.2184552>

McAlpin, David W., 1976. Core Vocabulary for Tamil, Institute of South Asia Studies, University of Pennsylvania, Philadelphia, PA 19174, USA. page 12 as quoted at <http://nganesan.blogspot.com/2012/04/circular-dot-diacritic-in-tamil.html>

Narayanan, B. (2011). Role of knowledge of orthography and automaticity in differentiating good and poor readers in Tamil. Poster presented at the Language Literacy and Cognitive Development (LLCd) Symposium, 16th – 17th December held at Bangalore.

Rakara mey series, Tamil Nadu state text book for Standard 1, Vol III, Page 44

<http://www.textbooksonline.tn.nic.in/Books/Std01/Std01-III-TamEng.pdf>

Sevakumar, C. R., 2010. தமிழ் எழுத்துமுறையின் தனித்தன்மையும் தமிழெழுத்துச் சீர்திருத்தமும், செ. இரா. செல்வக்குமார், தமிழ்ச் செம்மொழி மாநாடு, கோவை.

Sharma, Shriramana., L2/12-018, Request to deprecate 0B82 TAMIL SIGN ANUSVARA

<http://www.unicode.org/L2/L2012/12018-tamil-anusvara-depr.pdf>

L2/10256, Proposal to encode characters for Extended Tamil

<http://www.unicode.org/L2/L2010/10256-extended-tamil.pdf>

Siromoney, Gift., 1983. Origin of the Tamil-Brahmi script, Seminar on "ORIGIN EVOLUTION AND REFORM OF THE TAMIL SCRIPT", pp.21-29, The Institute of Traditional Cultures, University Buildings, Madras-600005. http://www.cmi.ac.in/gift/Epigraphy/epig_origin.htm

Tamil Nadu government's clarification on the orthography of ற், றி, றீ (Private communication with Tamil Virtual University)

T.N.G.O. on Unicode, 2010. G.O. (Ms) No.29 Dt: June 23, 2010

Information Technology – Committee constituted for declaring Unicode as current standard of Tamil Coding – Report submitted - Standards Prescribed for 16-bit coding of Tamil Script – Orders – http://www.tn.gov.in/go_view/dept/17?page=1

Unicode Tamil block (<http://www.unicode.org/charts/PDF/U0B80.pdf>)

Vargas, Fernando de Mello., 2007. Dissertation for M.A. in Typeface Design, Dept. of Typography and Graphic Communication, The University of Reading, United Kingdom

பல்வகை உணர்விகளில் மொழி மாதிரியங்களின் பயன்பாடு :

ஒரு புதிய கண்ணோட்டம்

ஆ. க. ராமகிருஷ்ணன், ப. அருள்மொழி

மறுமொழி ஆய்வகம், மின் பொறியியல் துறை, இந்திய அறிவியல் பயிலகம், பெங்களூரு.

கட்டுரைச் சுருக்கம்

இந்தக் கட்டுரையானது தமிழில் வாய்மொழி மற்றும் ஒளிவழி எழுத்துரு உணர்தல் (speech and optical character recognition) முறைமைகளில் மொழி மாதிரியங்களின் (language models) பயன்பாட்டில் எதிர்கொள்ளப்படும் சில குறிப்பிட்ட இடர்பாடுகளை விவாதிக்கிறது. இத்தகைய மொழி மாதிரியங்கள் ஐரோப்பிய மொழிகளில் மிகுந்த இலாபகரமாகப் பயன்படுத்தப்பட்டாலும், அவை அவ்வாறே தமிழில் பயன்படுத்தப்படும்பொழுது, தமிழின் மொழியியல் செறிவு காரணமாக அம்மொழிகளில் காணப்படாத சில புதிய இடர்பாடுகளை எதிர்கொள்ள நேரிடுகிறது. இது தமிழின் புணர்ச்சி விதிகள் (sandhi rules) மற்றும் உருபனியல் மாற்றச் செறிவு (morphological richness) காரணமாகவும் சொற்களோடு விசுவாசம் இணைந்து பொருள்தரும் பண்பு (agglutination) மற்றும் பகுதி கட்டில்லாச் சொல்வரிசை முறை (Partially free word order) காரணமாகவும் ஏற்படுகிறது. ஒரு மீப்பெரு பனுவலை (large corpus) அல்லது பனுவலில் இருந்து பெறப்பட்ட பெரும் சொல்வங்கியை n-கிராம் மொழி மாதிரியம் கொண்டு பகுப்பாய்வு செய்ததன் மூலமாகப் பெறப்பட்ட பல்வேறு புள்ளி விவரங்களை முன் வைத்து, அதன் மூலம் தமிழில் எத்தகைய மொழி மாதிரியங்களை எவ்வாறு பயன்படுத்தலாம் என்பது குறித்த எங்கள் கருத்துக்களையும் இங்கு சமர்ப்பிக்கிறோம்.

முன்னுரை

தற்போதைய கணினிமுறை பெருஞ் சொற்களஞ்சிய தொடர் வாய்மொழி உணர்தல் (large vocabulary continuous speech recognition - LVCSR) முறைமைகள் 2, 00, 000-க்கும் மேற்பட்ட எண்ணிக்கையிலான பெரும் சொல்லகராதியைப் பயன்படுத்துகின்றன. சொற்களஞ்சியத்தின் பருமன் அதிகரிக்கும்போது அதையொத்து விட்டர்பி சொல் அணிக்கோவை (word lattice), இன்ன பிற தேடுதல்களில் கணிப்பு சிக்கற்பாடு அதிகரிக்கின்றது. மொழியியல் செறிவு காரணமாகத் தமிழ் சொல்வங்கியானது பனுவலின் அதிகரிப்பு விகிதத்துக்கு ஏற்ப (மடக்கை விகிதத்தில்) அதிகரிக்கிறது. ஒரு வினை வேர்ச்சொல்லின் (root verb) வருவித்த படிமங்கள் (derived forms) மிகச் சிலவே உள்ள ஆங்கிலம் மற்றும் ஹிந்தி போன்ற மொழிகளைப் போலல்லாமல் தமிழானது ஒரு வேர்ச்சொல்லுக்குச் சில ஆயிரங்களுக்கு மேற்பட்ட வருவித்த படிமங்களைக் கொண்டுள்ளது [1]. வினைச்சொற்களின் ஒவ்வொரு உருபனியல் கூட்டப்பட்ட வடிவமும் காலம், இடம், பொருள், எவல், மறுப்பு, அழுத்தம், கேள்வி போன்ற பல்வேறு வகையான பொருளை வெளிப்படுத்தும் இயல்பில் புதிய எல்லைகளைக் கொண்டுள்ளது. சொல்லிலக்கண விதிகளில் அருகாமை சொல்லிலக்கணக் கூறுகளுடன் இணையும் தன்மையானது ஒருவிதமான நெகிழ்வுத் தன்மையைக் கொண்டுள்ளது. இவ்விதமான நெகிழ்வுத் தன்மையானது சொல்லிலக்கணக் கூறுகள் வெவ்வேறு இடங்களில் இணைய அனுமதிக்கிறது. இதனால் தமிழில் உருபனியல் மாற்றச் செறிவானது வினைச்சொற்களின் பல்வேறு வடிவங்களின் எண்ணிக்கையை ஆச்சரியப்படுத்தக்க வகையில் அதிகரிக்கிறது [2]. உதாரணமாக நாங்கள் பயன்படுத்தும் வரையறுக்கப்பட்ட சொல்வங்கியில் உள்ள 'வா' என்கிற வேர்ச்சொல்லின் வருவித்த படிவங்களின் தனிப்பட்ட எண்ணிக்கை 4567 ஆகும். இதன் மூலமாகத் தமிழ் போன்ற செறிந்த மொழிகளுக்கான உணர்தல் மூலோபாயங்கள் (recognition strategies) ஆங்கிலம் மற்றும் ஹிந்தி போன்ற மொழிகளுக்கான மூலோபாயங்களை ஒத்திருக்கவியலாது என்பது புலனாகிறது. எனவே தமிழுக்கு ஏற்றார்போல மாற்று மூலோபாயங்கள் வகுக்கப்பட வேண்டியதின் அவசியம் இங்கு வலியுறுத்தப்படுகிறது.

வாய்மொழி உணர்வியில் மொழி மாதிரியங்கள்

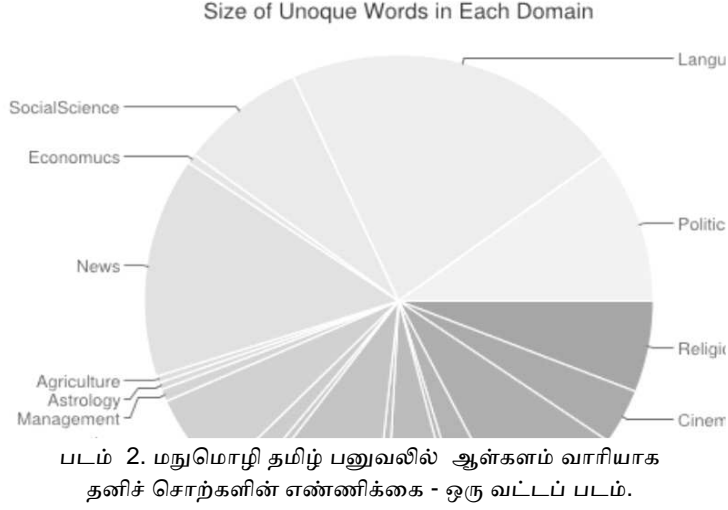
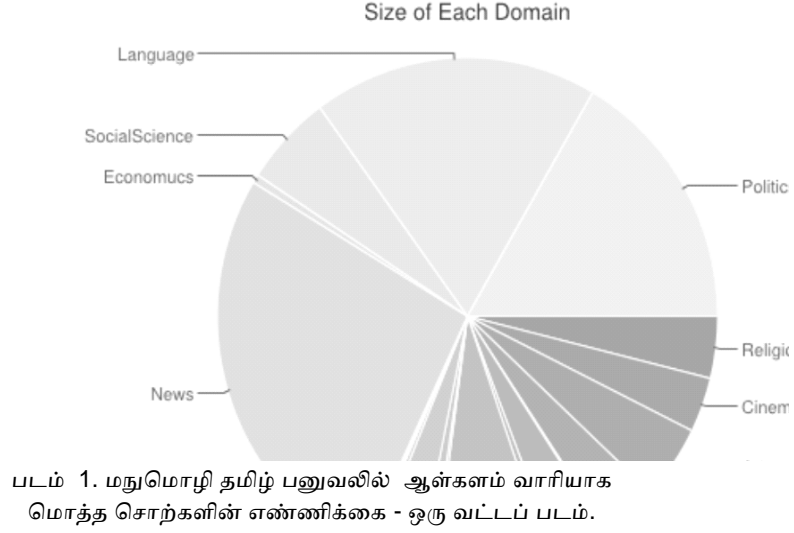
LVCSR முறைமைகள் இரண்டு லட்சம் சொற்களுக்கும் மேற்பட்ட எண்ணிக்கையிலான பெரும் சொல்லகராதியைப் பயன்படுத்துகின்றன. ஆனால் ஒரு சராசரி மனிதனின் மூளையில் பதிந்திருக்கும் சொற்களின் எண்ணிக்கை அதாவது மனிதனின் கற்றுக்கொண்ட சொல்லகராதி இருபதாயிரத்திலிருந்து [3] ஐம்பதாயிரம் [4] வரை தான். இதனால், கணினிகளுக்கு மட்டுமே சாத்தியமாகக்கூடிய ஒரு மீப்பெரு கட்டற்ற சொல்லகராதியானது, உணர்தல் முறைமைகளில் மேலும் குழப்பங்களைத் தருவிக்கவே வாய்ப்பு அதிகம் என்பது அறியப்படுகிறது. நாம் உரையாடும்போது, ஆள்களம் நமக்கு பேசுவதற்கு முன்பே பிடிபட்டுவிடுகிறது. அதன்பின், பல்பொருள் ஒருமொழிச் சொற்கள் பயன்படுத்தப்படினும், மனித மூளையின் தேடுதலானது, அச்சமயத்தில் விவாதிக்கப்படும்

ஆள்களத்துக்கு உட்பட்டே அமைகிறது. மேலும் கட்டயடுதல் செய்பணி (dictation tasks) முறைகள் பொதுவாக ஒரே அல்லது ஒருசில தொடர்புடைய ஆள்களங்களுக்கு உட்பட்டே அமைகின்றன.

அட்டவணை 1: தமிழ்ச் சொல்வங்கியின் அளவு, தனிச்சொற்களின் எண்ணிக்கை, ஆள்களத்தைத் தேர்வு செய்வதால் பெறப்படும் பேரகராதி மற்றும் சிக்கற்பாட்டு குறுக்கம், வரையறுக்கப்படாத ஒற்றைப் பேரகராதியின் தனிச்சொற்களின் எண்ணிக்கை : ஓர் ஒப்பீடு.

ஆள்களம்	சொற்களின் எண்ணிக்கை	தனிச் சொற்களின் எண்ணிக்கை	சொல்லகராதியின் குறுக்க விகிதம் (%)	ஒப்பீட்டு விகிதம் (ஆள்களத்துக்கு உட்பட்டது - %)
மொழியியல்	4179938	614184	35.8	85.3
கலை	1055330	222333	12.9	78.9
அறிவியல்	769531	136931	7.9	82.2
விளையாட்டு	831656	82384	4.8	90.0
புவியியல்	59124	17051	0.99	71.2
சட்டம்	189321	45738	2.7	75.8
சமயம்	878975	169277	9.8	80.7
செய்தி	6056139	408833	23.8	93.2
திரைப்படம்	779432	103163	6.0	86.7
ஜோதிடம்	37757	13194	0.7	65.1
மேலாண்மை	85183	23116	1.3	72.8
பொருளாதாரம்	120157	17873	1.0	85.1
சமூகவியல்	1309273	227952	13.3	82.6
அரசியல்	3779525	285040	16.6	92.5
சரித்திரம்	1681028	241082	14.0	85.6
வேளாண்மை	42754	12396	0.7	71.0
கல்வி	126546	23749	1.4	81.2
கவிதை	30105	17065	0.99	56.7
மற்ற பிற	699317	166283	9.7	76.2
மொத்தம்	22707568	1716256	--	92.5

இந்நிலையில் தகுந்த ஆள்களத்தை முன்னறிவதன் மூலமாக அறியப்படும் அறிவுப்புலத்தை வாய்மொழி மற்றும் எழுத்துரு கண்டறிதலில் பயன்படுத்துவது இத்தகைய குழப்பங்களைக் குறைப்பது மட்டுமில்லாது, கணிப்பு சிக்கற்பாட்டையும் குறைத்து துல்லிய அளவையின் படித்தரத்தை உயர்த்தும் என்ற கோட்பாட்டை இங்கு முன்மொழிகிறோம். இந்நோக்கில் இங்கு 2.2 கோடிக்கும் மேற்பட்ட சொற்களைக் கொண்ட பனுவல் பகுப்பாய்வுக்கு உட்படுத்தப்பட்டுள்ளது. இப்பனுவலானது 14 தனிப்பட்ட ஆள்களங்களாக (domains) வகைப்படுத்தப்பட்டுள்ளது. அவை முறையே மொழியியல், கலை, விளையாட்டு, புவியியல், சட்டம், சமயம், அறிவியல், செய்தி, அரசியல், திரைப்படம், ஜோதிடம், மேலாண்மை, பொருளாதாரம் மற்றும் சமூகவியல் ஆகியவை. இவற்றிலிருந்து பல்வகைப்பட்ட ஆள்களங்களில் உள்ள தனிச்சொற்கள், சொல் மற்றும் எழுத்து மட்ட n-கிராம் புள்ளிவிவரங்கள் போன்ற பலவகைப்பட்ட பகுப்பாய்வு முடிவுகள் தருவிக்கப்பட்டுள்ளன. அட்டவணை 1, சில பொதுவான ஆள்களங்களையும் அவை ஒவ்வொன்றிலும் உள்ள தனிச்சொற்களின் எண்ணிக்கையையும் அவற்றின் குறுக்க மற்றும் ஒப்பீட்டு விகிதங்களையும் காட்டுகிறது. உதாரணமாக, மொத்தச் சொல்வங்கியில் உள்ள தனிச்சொற்களின் எண்ணிக்கை 17.16 லட்சமாக இருந்தபோதிலும் 'அறிவியல்' ஆள்களத்தில் உள்ள தனிச்சொற்களின் எண்ணிக்கை 136931 ஆகும். ஆக, நாம் அறிவியல் துறை சார்ந்த ஒரு சொல்வதெழுதல் பணியில் ஈடுபடும்போது, 136931 சொற்களாலான சொல்வங்கியைப் பயன்படுத்தினால் போதுமானது. இது மென்பொருளின் வேகத்தையும் துல்லியத்தையும் நன்கு அதிகரிக்கும் என்பது திண்ணம். வட்டப் படங்கள் 1 மற்றும் 2-ம் மேற்கூறிய கருத்துக்களை வலியுறுத்துகின்றன.



எவ்விரு ஆள்களங்களுக்குள்ளும் பொதுவான சொற்கள் குறைவே.

இரண்டாம் அட்டவணை, ஒரு உதாரணமாக, "சமயம்" என்ற ஒரு ஆள்களத்துடன் வெவ்வேறு ஆள்களங்கள் தனித்தனியே சேரும்போது மொத்தச்சொற்கள் எத்தனை, அவை இரண்டிற்கும் பொதுவான சொற்கள் எத்தனை என்பதைக் காட்டுகிறது. "சமயத்"தில் தனிச்சொற்கள் 169277 உள்ளன . மொத்த சொற்களுடன் ஒப்பிடும்போது பொதுச் சொற்கள் 13 அல்லது அதற்கும் குறைவான விழுக்காடுகளே உள்ளன. ஆகையால் ஆள்களம் பொருத்தே நாம் சொல் வங்கியைப் பயன்படுத்த வேண்டும் . அட்டவணை 2 ஒரு மாதிரியே.. உண்மையில் 19 ஆள்களங்களுக்கிடையே, 171 இருகன பிணைப்புகளைச் சோதிக்க முடியும்.

அட்டவணை 2: இரு ஆள்களங்களுக்கிடையே உள்ள பொதுவான தனிச்சொற்கள். எல்லா வரிசைகளுக்கும் பொதுவான ஆள்களமான "சமயம்" 169277 தனிச்சொற்கள் கொண்டுள்ளது.

ஆள்களம்	தனிச் சொற்கள்	மொத்தம் (கூட்டு)	பொதுவான சொற்கள்
திரைப்படம்	103163	260853	11587
கலை	222333	345710	45900
விளையாட்டு	82384	240884	10777

பல ஆள்களங்களுக்குப் பொதுவான தனிச்சொற்களின் எண்ணிக்கை

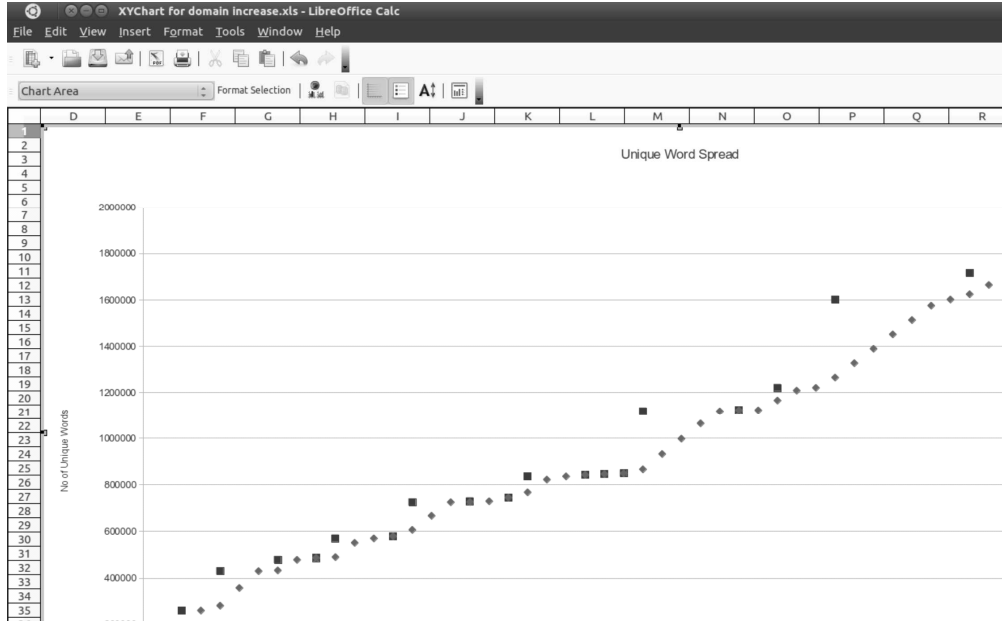
அட்டவணை 3-ஆனது ஆள்களங்களுக்குப் பொதுவான தனிச்சொற்களின் எண்ணிக்கை, ஆள்களங்கள் சேர்ச் சேர வெகு விரைவில் குறைந்து கொண்டே வருவதைக் காட்டுகிறது. இது, தனிச்சொற்களும் ஒவ்வொரு துறைக்கும் ப்ரத்யேகமானவையே என்பதை உணர்த்துகிறது. ஆக எல்லாத்துறைகளையும் சேர்த்த மொத்தச்சொல் வங்கியை உபயோகிப்பதில் எவ்வித லாபமும் இல்லை என்பது புலனாகிறது.

அட்டவணை 3: இரண்டு, மூன்று, நான்கு எனப் பல ஆள்களங்களுக்குப் பொதுவான தனிச்சொற்கள். ஆள்களங்களின் எண்ணிக்கை அதிகமாக, அதிகமாக பொதுச் சொற்கள் குறைந்து கொண்டே வருகின்றன.

ஆள்களம்	தனிச் சொற்கள்	முழுவதற்கும் பொதுவான சொற்கள்
சமயம்	169277	169277
திரைப்படம்	103163	11587
கலை	222333	9960
விளையாட்டு	82384	5887
கவிதை	17065	1854
அறிவியல்	136931	1705
கல்வி	23749	1046
சரித்திரம்	241082	1044
புவியியல்	17051	711
சட்டம்	45738	680
மற்ற பிற	166283	680
மேலாண்மை	23116	589
ஜோதிடம்	13194	503
வேளாண்மை	12396	421
செய்தி	408833	420
பொருளாதாரம்	17873	392
அறிவியல்	227952	392
மொழியியல்	614184	392
அரசியல்	285040	392

கையெழுத்து உணர்வியல் மொழி மாதிரியங்கள்

கையெழுத்து மற்றும் அச்சிடப்பட்ட பனுவல்களைக் கையாளும் ஐரோப்பிய மொழிகளுக்கான ஆவணக் கண்டறிகை முறைமைகள் கண்டறிதலிலும், உணர்விழைகளைத் திருத்துவதிலும் மீப்பெரும் சொல்லகராதிகளைப் பயன்படுத்துகின்றன. படம் 3, தமிழ் மொழியின் ஒரு முக்கியமான இயல்பைத் தெளிவாகக் காட்டுகிறது. ஒவ்வொரு ஆள்களமாகச் சேர்த்துக் கொண்டு செல்கையில், தனிச்சொற்களின் மொத்த எண்ணிக்கை, தவிட்டாமல், மென்மேலும் வளர்ந்து கொண்டே இருப்பதுதான் அது. ஆகவே, தமிழ் மொழியானது, ஒரு வரையறுக்க முடியாத தனிச்சொல் வங்கியை உடையது. அதனால், சொல் மட்டப் பேரகராதி (word level lexicon) கொண்டு பிழை திருத்துவது என்பது தமிழ் மொழிக்குச் சரிவராது. எங்கள் ஆய்வில், எமில்லெ (Emille) பனுவலிலிருந்து தேர்ந்தெடுக்கப்பட்ட 2.4 லட்சம் தனிச்சொற்கள் கொண்ட பேரகராதியில் நாங்கள் கையெழுத்து உணர்வியைச் சோதனை செய்ய உபயோகித்த சாதாரணமான 2000 வார்த்தைகளில் 1530 வார்த்தைகளே இருந்தன [5] என்பது குறிப்பிடக்கூது. ஈரெழுத்துப் (bigram) படிமங்களால் மட்டுமே கையெழுத்து உணர்வியின் பிழைகளைத் திருத்த இயலும் [6]. அப்படியும் பல ஜோடி வார்த்தைகள் (உ: அவன், அவள்; வருவான், வருவாள்) கடைசி எழுத்தில் மட்டும் வேறுபடும் என்பதால், அவற்றிற்கு ஈரெழுத்துப் படிமங்களும் தவறான சொற்களைக் கொடுக்க வாய்ப்பு உள்ளது. ஆக, அத்தகைய சூழ்நிலையில் நாம் உணர்வியின் எழுதப்பட்ட சொல்லிலிருந்து கீற்றக்குறியீடுகளுக்கு (symbols) துண்டுபடுத்தல் (segmentation) [7] மற்றும் உணர்தல் (recognition) ஆகிய பணிகளின் துல்லியத்தையே முயன்று அதிகரிக்க வேண்டியிருக்கிறது [8].



படம் 3. தமிழ் மொழியின் செறிவு : பனுவலில் ஒவ்வொரு ஆள்களமாகச் சேர்க்கும்போது (நீல நிறம்) தனிச் சொற்களின் எண்ணிக்கை தங்கு தடையின்றி வளர்தல். ஒவ்வொரு தடவையும் மற்றுமொரு லட்சம் பனுவல் சேர்க்கப்படும்பொழுது தனிச்சொற்களின் வளர்ச்சி (சிவப்பு நிறம்)

தமிழில் நீளமான சொற்கள் அதிகம்

தமிழில் ஒவ்வொரு வினைச்சொல்லுக்கும் குறைந்த பட்சமாக என்பது உருபனியல் அமைப்புகள் (paradigms) உள்ளன. கூட்டு வினைச்சொற்களையும் கணக்கில் கொண்டால் இது இன்னும் பல மடங்காக உயரும் [RVR2003]. தமிழ் உருபனியல் மாற்றங்கள் எழுவாய், பயனிலை, செயப்படுபொருளைக் கண்டறிவதன் மூலம் பிரதிபெயர்களின் ஆதியைக் கண்டறிதலில் (pronominal resolution) பயன்படுமளவுக்கு பொருள் செறிந்தது [9]. இத்தகைய செறிவு தெலுங்கு போன்ற மற்ற திராவிட மொழிகளிலும் உண்டு [10]. இச்செறிவே சொந்திருத்தி போன்ற முறைமைகளை உருவாக்குவதைக் கடினமாக்குகின்றன [11]. இவ்வித உருபனியல் செறிவானது ஒரு ஆங்கிலச் சொற்றொடரை மொழிபெயர்க்கையில் ஒரு வார்த்தையில் பொருள் தரும் வகையில் அமைந்துள்ளது [12]. உருபனியல் மாற்றச் செறிவினால் தமிழ் வார்த்தைகளின் சராசரி நீளம் அதிகமாக இருக்க வாய்ப்பு உள்ளது. இதனைச் சோதிப்பதற்காக வெவ்வேறு நீளமுள்ள தமிழ் தனி வார்த்தைகளின் புள்ளி விவரங்களைச் சேகரித்தோம். இவை அட்டவணை 4-ல் கொடுக்கப்பட்டுள்ளன. இங்கு நீளம் என்பது தமிழ்ச் சொல்லிலுள்ள உயிர்மெய் எழுத்துக்களின் எண்ணிக்கையே. உதாரணமாக, "கல்கி" என்ற சொல்லில் மூன்று எழுத்துக்கள் உள்ளன. ஆயின், ஒருங்குறி எண்ணிக்கை ஐந்தாகும்.

அட்டவணை 4: வெவ்வேறு நீளமுள்ள தமிழ் தனி வார்த்தைகளின் புள்ளி விவரங்கள். மொத்த எண்ணிக்கையை ஒட்டி, இறங்கு வரிசைப்படி கொடுக்கப்பட்டுள்ளன.

எழுத்துக்கள் / சொல்	மொத்த எண்ணிக்கை	எழுத்துக்கள் / சொல்	மொத்த எண்ணிக்கை
10	135486	5	41230
11	130425	18	30471
9	128475	4	23679
12	122741	19	21826
8	115506	20	14660
13	107806	21	10195
7	90559	3	8465
14	90164	22	6818
15	72450	23	4512
6	64978	24	3035
16	55613	25	1924
17	41860	2	1833

முடிவுரை மற்றும் வருங்கால ஆய்வுப் பணிகள்

பை-கிராம் இடுகுறி அல்லது எழுத்து சார்ந்த புள்ளி விவரங்கள் வெவ்வேறு ஆள்களங்களுக்குக் குறிப்பிடத்தக்க அளவில் வேறுபடுகிறதா என்பதை உற்றறிதல் மூலம் பல பயனுறு தகவல்களைப் பெறலாம். உதாரணமாக, இப்புள்ளிவிவரங்கள் சங்கத்தமிழ்ப் பாடல்களுக்கும் நவீன தமிழ் தரவுகளுக்கும் குறிப்பிடத்தக்க விதத்தில் வேறுபட்ட தனிச்சிறப்பியல்புகளைக் கொண்டிருக்கும். மேலும் இலக்கண, சொற்பொருளியல் மற்றும் புணர்ச்சி விதிகளை வரையறுப்பதன் மூலம் உணர்தல் நெறிமுறைகளின் செயலாக்கம் வலுவடையலாம்.

மேற்கோள்கள்

1. Rajendran, S., Viswanathan, S. Ramesh Kumar, "Computational morphology of Tamil verbal complex, " Language in India, Vol. 3 : 4 April 2003.
2. Annamalai, E. Dynamics of Verbal Extension in Tamil. Thiruvananthapuram: Dravidian Linguistics Association, 1985.
3. Richard Lederer. A celebration of English, good grammar, and wordplay. Marion Street Press, 2012.
4. David Crystal, "The Stories of English", Overlook TP Publishers, 2004. ISBN:978-1-58567-719-1
5. Suresh Sundaram, Bhargava Urala and A. G. Ramakrishnan, "Language models for online handwritten Tamil word recognition, " Proc. Workshop on Document Analysis and Recognition (DAR 2012), 16 December 2012, IIT Bombay, Mumbai, India.
6. Suresh Sundaram and A. G. Ramakrishnan, "Bigram language models and reevaluation strategy for improved recognition of online handwritten Tamil words, " revised manuscript under review, ACM Transactions on Asian Language Information Processing (TALIP), 2013.
7. Suresh Sundaram and A. G. Ramakrishnan, "Attention-feedback based robust segmentation of online handwritten isolated Tamil words, " ACM Transactions on Asian Language Information Processing (TALIP), Vol. 12 (1), March 2013, Article No. 4.
8. Suresh Sundaram and A. G. Ramakrishnan, "Performance enhancement of online handwritten Tamil symbol recognition with reevaluation techniques, " revised manuscript under review, Pattern Analysis and Applications, 2013.
9. Murthy, Kavi Narayana, L. Sobha, and B. Muthukumari, "Pronominal resolution in Tamil using machine learning, " Proc. First Intern. Workshop on Anaphora Resolution (WAR-I), pp. 39-50. 2007.
10. K. Narayana Murthy, "Parsing Telugu in the UCSG Formalism, " Proc. Indian Congress on Knowledge and Language, vol 2, Jan. 1996, pp 1-16, Central Institute of Indian Languages, Mysore.
11. Ranjani Parthasarathy, Geetha T.V., Morphological Analyzer for Tamil, TI 2001, Chennai.
12. Vu, Ngoc Thang, and Tanja Schultz, "Initial experiments with Tamil LVCSR, " Proc. Intern. Conf. Asian Language Processing (IALP), Hanoi, Vietnam. pp. 81-84, IEEE, 2012.

A tool that converted 200 Tamil books for use by blind students

Shiva Kumar H R and A G Ramakrishnan

Department of Electrical Engineering, Indian Institute of Science, Bangalore.

Abstract

A versatile tool has been created, with user-friendly interface, for the rapid and efficient conversion of printed Tamil books to Braille books for the use of persons with visual disability. The tool has been developed in Java using Eclipse SWT and runs on Linux, Windows and Mac operating systems. This tool has been developed as an open source project and is available under the Apache 2.0 license from code.google.com. An individual scanned page or all the pages of a whole book can be recognized by this tool. The average time taken for digitizing a Tamil page is two seconds. The output can be saved in RTF, XML or BRF (Braille) format directly, by the click of a button. There is a provision for manually selecting the individual columns of a two-column printed page or even marking the individual rectangular text blocks of a page with a more complex Manhattan layout. The user can modify the reading order of the so-selected text blocks. This information of ordered text blocks is passed on to the Tamil OCR integrated at the backend of the tool and hence the recognized Tamil text in Unicode is put together in the same reading order. In the case of books with identical or very similar text layouts across its pages, such an user-defined layout can be saved as a custom layout and automatically applied to segment the other pages of the same book or a different book also.

Keywords: Tamil OCR, Braille conversion, digitization, column layout marking, automated error detection, error correction, text block separation, line segmentation correction, XML format.

Introduction

In the past two years, when our Tamil OCR was being used by Worth Trust, Chennai on a regular basis, we conducted a study on the work flow and the time taken for the various steps involved in converting the printed book to a Braille book. We realized that while it takes only a few minutes to OCR the whole book, it takes two to three days to correct the OCR errors, convert the Unicode to Braille text and reformat it to the Braille printer and page requirements. This incited us to take upon ourselves the task of developing a comprehensive tool that handles all the issues involved in the work flow and thus reduce the overall processing time from a week to a single day.

On account of this, the tool offers a convenient graphical user interface, as shown in Fig. 1, for effecting corrections of the recognized text. The input image and the output text are displayed side by side in two different windows, which can be simultaneously zoomed in and out. This results in alignment of text lines between the source image and the OCR output text. Our Indic Keyboard interface [5], again available under Apache 2.0 licence from code.google.com, can be used to edit the Tamil text before saving.

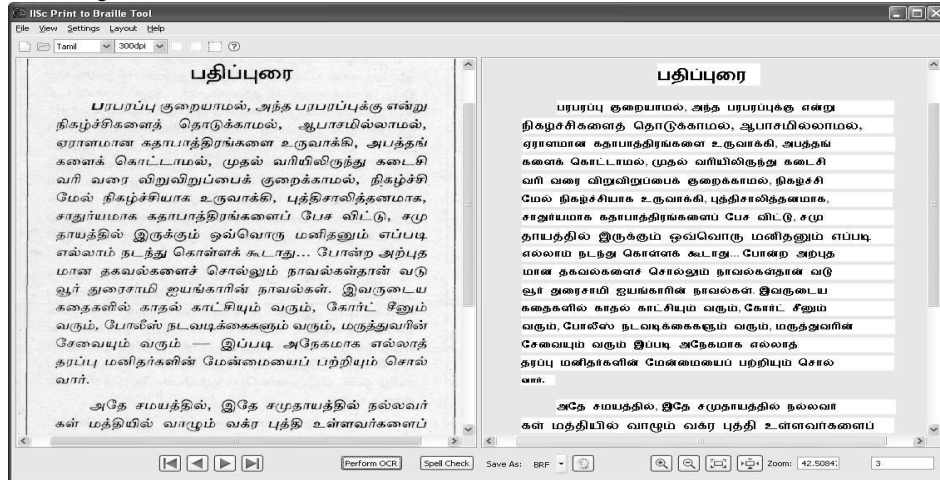


Fig. 1. Aligned display of input image and output text for easy verification and editing.

Output Edit Facility

Using the OCR that has been interfaced, the text check and edit facility tries to identify the wrongly recognized words and suggests a list of alternatives to select from, as well as provisions to edit the recognized word or type a completely new word. The suggested new word

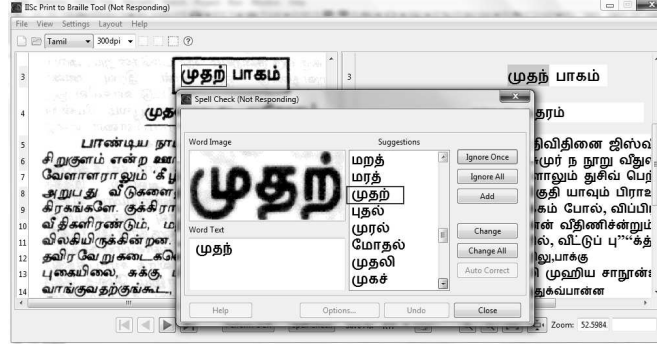


Fig. 2. GUI to correct recognition errors with suggested list of alternatives, etc.

can be used to replace the current instance of the wrongly recognized word or all instances occurring in the entire book.

Correction of logical layout detection

Figure 3 shows the GUI for selecting one of the standard logical layouts for the printed page to be recognized. The user can edit the wrong layout segmentation of the OCR or the standard layout applied by the tool itself. This facility is illustrated in Fig. 4.

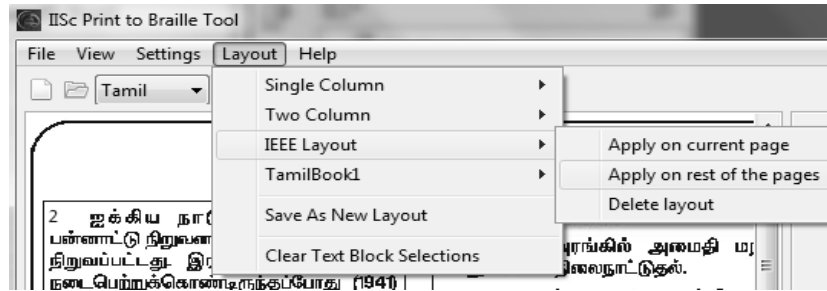


Fig. 3. Tool supports user correction of layout segmentation by the OCR.

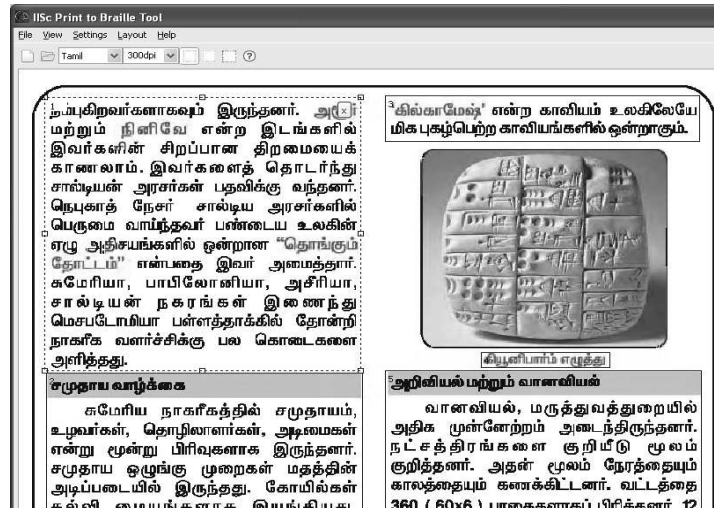


Fig. 4. Facility to mark and edit text blocks and their logical order.

Correcting Line Segmentation Issues

Similarly, if the OCR wrongly segments the text lines, the user can rectify the under or over segmentation of the text lines, using the convenient interface available, shown in Fig. 5.

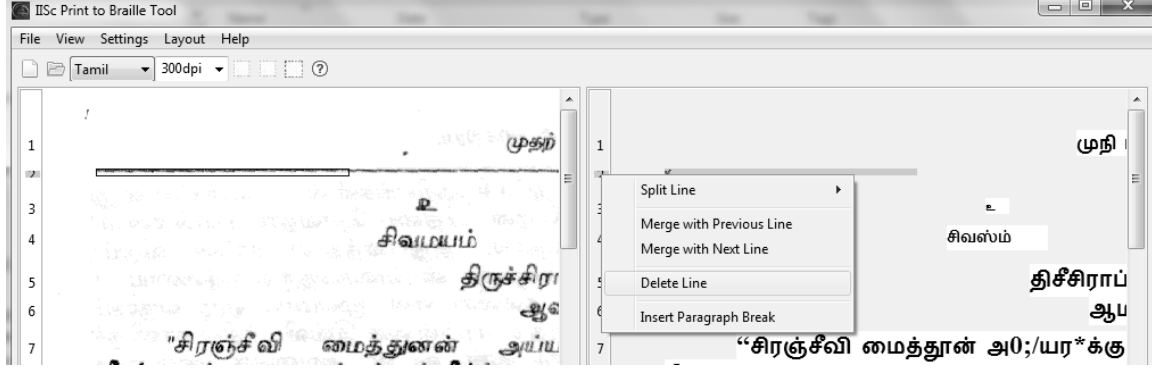


Fig. 5. Facility provided by the tool to correct wrong line segmentation by the OCR.

Standard XML

Figure 6 shows a part of the XML file automatically saved by the tool for a page of text. The text block, text line and word numbers and their coordinates are saved in a hierarchical fashion. This can be used to reconstruct the original document with its logical structure, if required. Also, OCR error analysis can be performed, if the corresponding ground truth is available at any of these levels.

```

1 <?xml version="1.0" encoding="UTF-8" standalone="yes"?>
2 <page xmlns="http://mile.ee.iisc.ernet.in/schemas/ocr_output" skew="0.0" ImageLoc="">
3   <block type="Text" BlockNumber="1" rowStart="362" rowEnd="488" colStart="559" colEnd="1833">
4     <line LineNumber="1" rowStart="392" rowEnd="460" colStart="559" colEnd="1833">
5       <word unicode="4." WordNumber="1" colStart="607" colEnd="0"/>
6       <word unicode="இத்தமியில்" WordNumber="2" colStart="0" colEnd="0"/>
7       <word unicode="பாசிம்" WordNumber="3" colStart="0" colEnd="0"/>
8       <word unicode="கிமி" WordNumber="4" colStart="0" colEnd="0"/>
9       <word unicode="1922-கிமி" WordNumber="5" colStart="0" colEnd="0"/>
10      <word unicode="1945" WordNumber="6" colStart="0" colEnd="1801"/>
11    </line>
12  </block>
13  <block type="Text" BlockNumber="2" rowStart="498" rowEnd="2930" colStart="297" colEnd="1153">
14    <line LineNumber="2" rowStart="522" rowEnd="573" colStart="297" colEnd="1153">
15      <word unicode="பாசிம்" WordNumber="1" rowStart="522" rowEnd="573" colStart="402" colEnd="557"/>
16      <word unicode="என்னும்" WordNumber="2" rowStart="522" rowEnd="573" colStart="600" colEnd="765"/>
17      <word unicode="சொல்" WordNumber="3" rowStart="522" rowEnd="573" colStart="808" colEnd="934"/>
18      <word unicode="உசில்" WordNumber="4" rowStart="522" rowEnd="573" colStart="978" colEnd="1142"/>
19    </line>
20    <line LineNumber="3" rowStart="580" rowEnd="633" colStart="297" colEnd="1153">
21      <word unicode="என்னும்" WordNumber="1" rowStart="580" rowEnd="633" colStart="308" colEnd="464"/>
22      <word unicode="இலத்தின்" WordNumber="2" rowStart="580" rowEnd="633" colStart="486" colEnd="676"/>

```

Fig. 6. A part of a sample XML file from the tool, illustrating the hierarchical structure and the details of the information recorded.

Current users

Worth Trust Chennai has already converted 172 school and college Tamil printed books using this tool and converted them to Braille books, which are already being used by over a hundred blind students. Sri Ramakrishna Math, Chennai is using our tool to digitize out-of-print books published by them on the occasion of the 150-th anniversary of Swami Vivekananda. AuroLok Trust, Gujarat is using it to convert Tamil books on philosophy related to Sri. Aurobindo. Parankushachar Institute of Vedic Studies, Bangalore is using it to convert a voluminous translation of *Naalaayira Divya Prapandham*. Samskriti Foundation, Mysore is making use of it for similar objectives. Also, the library of Pondicherry University is using this tool for the benefit of the visually challenged users.

Licensing

Currently, the tool is available together with the Tamil OCR for free use by any non-profit organization for digitizing books for the visually challenged or any other social objectives, after a signed agreement. However, people interested in using it for other purposes can contact for a separate licensing agreement.

Some of the enhancements planned are to provide automated segmentation of text from graphics and layout analysis for multicolumn documents and complex layouts. It is also planned to extend the facility for manual selection of text blocks from rectangular to arbitrary shaped regions employing user-defined polygonal windows. Also on the cards is the facility to automate the manual insertion of information such as “image removed” at locations, where the graphics or photos have been removed. Further, it is intended to make the automatic correction of multiple occurrences of the same recognition error across a book more robust by carrying out sophisticate verifications.

Acknowledgments

The authors thank Technology Development for Indian Languages (TDIL), Department of Information Technology, Government of India for funding a national level research consortium project on OCRs in Indic scripts and thus facilitating the development of this socially useful tool. Ours thanks are also due to the undergraduate interns from RVCE and through the summer fellowships of Indian Academy of Science. and MILE lab project assistants for contributing to the development of some of the sub-modules of the PrintToBraille tool and to Worth Trust, Chennai for their useful inputs and feedback.

References

- [1] IISc-MILE Print to Braille Tool - an open-source project on [code.google.com](http://code.google.com/p/ocr-gui-frontend).
<http://code.google.com/p/ocr-gui-frontend>.
- [2] XSD for OCR Output XML. http://ocr-performance-evaluator.googlecode.com/svn/trunk/OCR_Model/src/main/resources/ocr_output.xsd.
- [3] JAXB classes for OCR Output XML. http://ocr-performance-evaluator.googlecode.com/svn/trunk/OCR_Model/src/main/java/org/iisc/mile/ocr/model.
- [4] IISc MILE TTS Web Demo. http://mile.ee.iisc.ernet.in:8080/tts_demo/.
- [5] Open source, downloadable Indic Keyboard Interface, under Apache 2.0 licence. <https://code.google.com/p/indic-keyboards/>

A Multimodal Framework for the Recognition of Ancient Tamil Handwritten Characters in Palm Manuscript Using Boolean Bitmap Pattern of Image Zoning

E.K.Vellingiraj, Asst. Professor and Dr.P.Balasubramanie, Professor

Department of Computer Science & Engineering, Kongu Engineering College, Perundurai, girirajek@rediffmail.com

Abstract:

Tamil is one of the oldest languages in the world with rich literature. In the ancient days, the writers, especially in Tamilnadu, used palm leaves to encrypt their writing. A very good example of the usage of Palm leaf manuscripts to store the history is Tamil grammar book named Tolkappiyam which was written during 4th B.C. The ancient literature includes many palm leaf manuscripts that contain Sangam works, classics, Saiva, Vaishnava and Jain works, medical works, food, astronomy & astrology, vaastu & Kaama shastra, jewellery, music, dance & drama, medicine, Siddha and so on. Over the 3, 500 Tamil manuscripts are available in Saraswathi Mahal Library located in Thanjavur, Taminadu, India. In this library, only a few palm leaf manuscripts are digitalized and many are to be digitalized so as to enable quick reference in the future. The objective of the proposed research is to develop the model that can recognize Tamil characters from palm manuscripts and convert them into text format. In the field of handwritten character recognition, image zoning is a widespread technique for feature extraction since it is rightly considered to be able to cope with handwritten pattern variability.

1. Introduction:

Tamil Handwritten character recognition is one of the most difficult tasks in the pattern recognition system. There are lots of difficult things which can be solved through image processing technique: separating each character, recognizing character fonts and written styles used in different centuries. Many researchers try to apply many techniques for breaking through the complex problems of Tamil handwritten character recognition. The optical character recognition (OCR) is one of the techniques which can be defined as the process of recognizing either printed or handwritten text from document images and converting them into text format. There are many algorithms being used in the literature to perform this conversion task for specific language. In Tamil language, there are so many researchers who research using Tamil OCR but they are unable to obtain not more than 90% of accuracy. The problem is to go one step up to recognize the ancient Tamil characters. The objective of the current research is to recognize the ancient Tamil characters from old palm manuscripts by converting them into text format.

2. Related Work:

According to [1], the image of palm manuscripts is used to extract a text line. This conventional text line extraction can be roughly classified into four categories:

- i. Projection based methods
- ii. Hough transform based methods
- iii. Bottom up grouping methods
- iv. Image segmentation based methods

Most of works based on text line segmentation can be roughly categorized as bottom-up or top-down approaches. In the top-down methodology, a document page is first segmented into zones, and a zone is then segmented into lines, and so on. Projection based methods is one of the most successful top-down algorithms for printed documents and it can be applied on handwritings only if gaps between two neighboring handwritten lines are sufficient. Projection based methods allow the efficient extraction of text lines [2], [3]. We use this method for palm manuscripts.

3. Methodology:

Here, all details of the proposed system design are given. First, the overall framework of the ancient Tamil handwritten character recognition system is given.

3.1 System Architecture Overview:

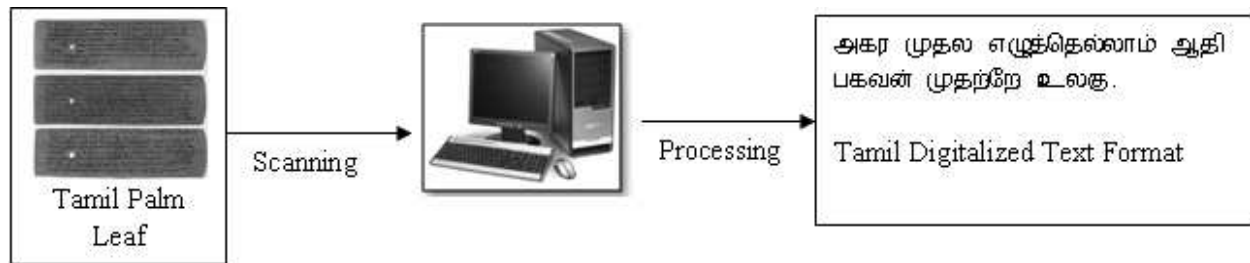


Fig. 1: Framework of the Ancient Tamil Handwritten character recognition in Palm manuscripts

3.2 System Structure Chart:

Based on the system framework in the previous section, the Tamil palm leaf image is converted into Tamil text format. This framework includes, i) Image scanning ii) Image preprocessing iii) Feature extraction iv) Character recognition v) Text conversion.

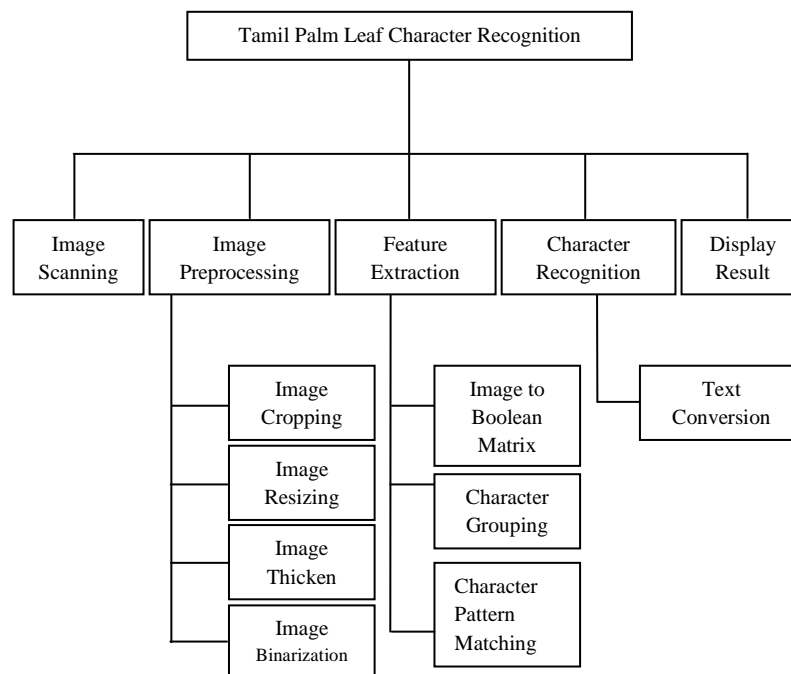


Fig 2: Structure chart of Ancient Tamil handwritten characters recognition by image zoning using the Boolean matrix

- i. **Image Scanning:** In the first stage, the Tamil palm leaf manuscripts belonging to different centuries would be collected from various places in Tamilnadu. These manuscripts are scanned by 4800 dpi scanner and stored in Jpeg format.
- ii. **Image Preprocessing:** In the image preprocessing module, the proposed system would prepare a palm manuscript handwritten character image for the feature extraction module. This stage consists of five sub-processes: a) image cropping b) segmentation c) image re-sizing d) image thickening and e) Image binarization. Each of these sub-processes are given below:

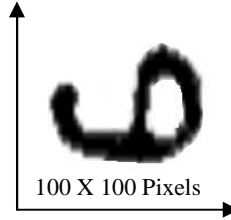
- a) *Image cropping*: This process involves the cropping of each word. The scanned palm leaf image would have white space. Using this white space, words are cropped.



- b) *Segmentation*: There are three types of segmentation. Line segmentation, word segmentation and character segmentation. Researchers used various techniques for segmentation like threshold techniques, region based method, edge based method, graph based methods, clustering methods, compression based methods, histogram methods, watershed transformation and model based methods.

Here we will take the edge detection method to segment the characters. The edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally has discontinuities. Marr-Hildreth algorithm is one of the edge detection, which has a method of detecting edges in digital images.

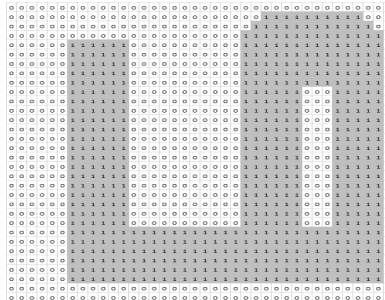
- c) *Image re-sizing*: Each segmented character is in different size. So, it is necessary to change all the characters into equal size. In the proposed method, the character image is re-sized into 100X100 pixels.



- d) *Image thickening*: Each darkened pixel of the re-sized character is thickened through darkening the nearest pixels. Using the nearest algorithm, a thin character is changed into thicker character through darkening the color of the nearest pixel for a target range.



- e) *Image binarization*: Each character is stored in Boolean matrix in either 0's or 1's. Using the image zoning technique [5], all the dark pixels are stored in 1's and light pixel in 0's.



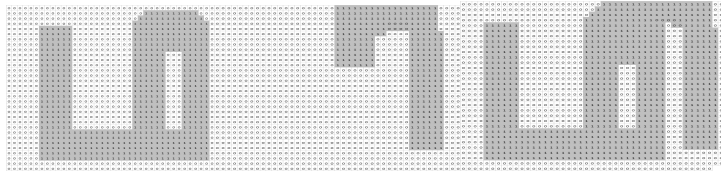
- iii. *Feature extraction*: This feature extraction module extracts the basic components of Tamil characters. There are three sub modules a) Image conversion into Boolean Matrix b) Image Grouping and c) Character Pattern Matching the details of which are given below:

- a) *Image conversion into Boolean Matrix*: Each character from palm leaf is converted into Boolean matrix. Similarly, all the actual Tamil character sets (Fig 4) are also converted into the Boolean matrix because palm

1	2	3	4	5	6	7	8	9	10
அ	ஆ	இ	ஈ	உ	ஊ	ஏ	ஐ	ஒ	
11	12	13	14	15	16	17	18	19	20
ஐ	ஊ	க	ங	ச	ரு	ட	ண	த	ந
21	22	23	24	25	26	27	28	29	30
ப	ம	ய	ர	ல	வ	ழ	ள	ற	ங
31	32	33	34	35	36	37	38	39	40
கு	ந	க	ங்	ப	ண்	து	நு	பு	(மு
41	42	43	44	45	46	47	48	49	50
ய	ரு	வ்	மு	ந்	று	ஐ	ம	சு	
51	52	53	54	55	56	57	58	59	60
மு	டு	கு	மு	க	ரி	ஃ	ஓ	ஔ	
61	62	63	64	65	66	67			
ஈ		ஊ	ஊ	ஊ	ஊ	ஊ			

[illegible]

142



A simple algorithm that is can be used for character matching is as follows:

M is the Boolean matrix one by one in all character sets in Tamil scripts (Character set 1 to 67). N is the Boolean matrix for current handwritten character in palm manuscripts.

1. Predefine M in all Boolean matrix
2. Check the current character Boolean matrix N with M
3. If the range of both Boolean matrix is set to 1, the characters are identified (Character set no 1 to 55)
4. If not, check the part of the Boolean matrix with the entire predefined matrix (Character set no 1 to 30)
If Boolean matrix is matches, then check all the other parts of the Boolean matrix (Character set 56, 57, 58, 63, 64, 65)

5. Conclusion:

In this paper, we have proposed a simple method for converting ancient Tamil handwritten scripts into text format. There are thousands of Tamil palm manuscripts that are yet to be digitalized. The aim of this paper is to convert the palm manuscript image into digitized text format. However, our method has some difficulties in handling cases such as cursive Tamil script, merging of two Boolean matrixes, and a hole in palm manuscript image. These are only some basic issues which can be overcome through future extension of character recognition.

References:

- Hyung Il Koo and Nam Ik Cho, "Text Line Extraction Chinese Documents Based on an Energy Minimization Framework", IEEE Trans. On Image Processing, Vol.21, no.3, pp 1169-1175, Mar 2012.
- G. Nagy, S. Seth, and M. Viswanathan, "A Prototype Document Image Analysis System for Technical Journals, " Computer, vol. 25, no. 7, pp. 10–22, Jul. 1992.
- F. Shafait, D. Keysers, and T. M. Breuel, "Performance Evaluation and Benchmarking of Six-Page Segmentation Algorithms, " IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 6, pp. 941–954, Jun. 2008.
- Y.Liang, M.C.Fairhurst & R.M.Guest, "A Synthesisd Word Approach to Word Retrieval in Handwritten Documents", Elsevier Pattern Recognition, Vol.45, PP 4225-4236, June 2012.
- Giuseppe Pirlo, Donato Impedovo, "Adaptive Membership Functions for Handwritten Character Recognition by Voronoi-Based Image Zoning", IEEE Trans on Image Processing, Vol 21, No 9, PP 3827-3836, Sep 2012.
- Chomtip Pornpanomchai, Verachag Wongsawangtham, Satheanpong Jeungudomporn, and Nannaphat Chatsumpun, "Thai Handwritten Character Recognition by Genetic Algorithm (THCRGA)", IACSIT Journal of Engineering and Technology, Vol 3, No 2, Apr 2011.
- Qiu-Fend Wang, Fei Yin, and Cheng-Lin Liu, "Handwritten Chinese Text Recognition by Integrating Multiple Contexts, IEEE Trans on Pattern Analysis and Machine Intelligence, Vol 34, No 8, Aug 2012.
- Tiji M Jose and Amitabh Wahi, "Recognition of Tamil Handwritten Characters using Daubechies Wavelet Transforms and Feed-forward Back Propagation Network", IJCA, Vol 64, No 8, PP 0975-8887, Feb 2013.
- Jin Chen, Daniel Lopresti, " Model Based Ruling Line Detection in Noisy Handwritten Documents", Pattern Recognition Letters, Elsevier, 2012.
- A Bharath and Sriganesh Madhvanath, "HMM-Based Lexicon-Driven and Lexicon-Free Word Recognition for Online Handwritten Indic Scripts", IEEE Trans on Pattern Analysis and Machine Intelligence, Vol 34, No 4, Apr 2012.

Analysis of Statistical Feature Extraction Approaches Used in Tamil Handwritten OCR

Antony Robert Raj.M¹, Abirami.S², Murugappan.S³
^{1,2} Department of Information and Science and Technology
Anna university, Chennai – 600 025
Email: antorobert@gmail.com, abirami_mr@yahoo.com
³ Department of Computer Science and Engineering
Annamalai University, Chennai – 608 002

Abstract. Tamil Handwritten Character Recognition System (THCRS) is used to recognise the characters from Tamil handwritten document images. THCRS is one of the upcoming and challenging tasks in Optical Character Recognition Environment. Various phases get involved in the process of Character Recognition such as Pre-processing, Segmentation, Feature Extraction and Classification. Out of which, extraction of features from digitized documents plays a crucial role. This paper is intended to analyze and present a survey on the feature extraction mechanisms available for THCRS based on statistical approaches. This survey also attempts to penetrate and utters the strength and weaknesses of many researches that catered solution in THCRS. In addition, this paper tries to propose the limitations addressed by various statistical algorithms, which may help the researchers to employ best statistical algorithms according to their requirements.

I. INTRODUCTION

Optical Character Recognition (OCR) is a process which recognizes the exact character from digitized printed and handwritten documents. It also helps to extract the proper character object from glassy and damaged medium such as historical handwritten documents, palm scripts and so on. THCRS is one among the sub areas of character recognition.

THCRS applications are confronting many disputes due to high character sets of Tamil language and inconsiderable variation of Tamil handwritten characters. Tamil language consists of 247 letters which includes consonant, vowels and special characters[17]. Additionally different styles of various characters are used to denote Tamil numerals, days, years and so on. Recognition of Tamil handwritten characters is quite complex due to different styles and dissimilar shapes. This sort of variation occurs due to the writing style of individuals, mood, age factor and so on. Figure 1 shows the various writing styles of few Tamil characters.

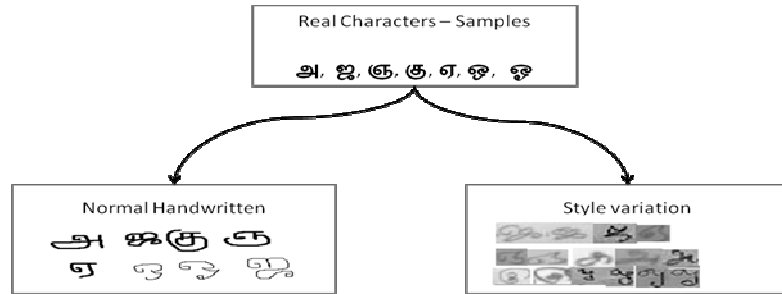


Figure. 1. Character samples at various levels

In THCRS, various phases has been performed to recognize the characters, they are Pre-processing (Binarization, Thinning, Noise removal, Skew and Slant correction and Normalization), Segmentation, Feature Extraction and Classification[18]. Among these phases feature extraction plays crucial role. Feature extraction has been performed in literature using three different types of extraction techniques called structural, statistical and hybrid. The structural features are based on geometrical features that are taken from the structural properties of the character. The statistical features are quantization properties which include zoning, projection, profile and crossing of characters. The Hybrid technique is obtained from the combination of statistical and structural.

2. STATISTICAL APPROACHES

Statistical feature extraction method is one of the pixel based quantization approach. In most of the Tamil recognition system, the character identification based on the statistics has been used to achieve the proper result. The pixel values have been analysed in statistical approach to obtain the spatial distribution. The area and the size of the character region have been measured using average value of the pixels which are present in the character region of the image. Various methods are available in the literature to get the statistical values out of them the initial approaches are calculating the mean and the variance from the pixel values, standard deviation, maximum and minimum levels of intensity, averaging pixels, grouping pixels, relationships between the pixels and so on. As a next level statistical approaches tries to find the Gradient values by calculating the distance between the pixels such as dimension and edge distribution. In other statistical approaches Fourier, Gabor, Wavelet and Gabor channels are used to calculate multiscale and time frequency features. The normalization is one of the most essential pre-processing steps before statistical feature extraction which have been employed in most of the research works. The following figure 2 sketches the techniques used in the literature to extract the features which lying under statistical approach of Tamil OCR.

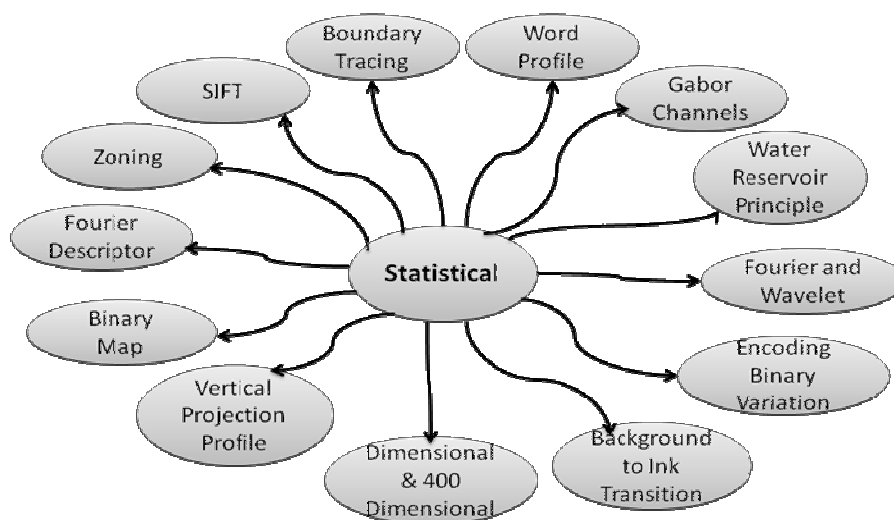


Figure 2. Statistical feature extraction techniques.

A. ZONING

Zoning is a pixel distribution based approach, which is done by various procedures. The character region has been divided into many zones to calculate the pixel values. In character identification environment, considerable significant works have been reported using zoning methodology. Vertical projection based pixel distance [2] is one of the zone based approach where Tamil numerals (0 to 9) are taken into account. Here the substantial samples of the Tamil numerals have been collected from various people and they accomplished 90% accuracy. Normal samples with limited complexity have been exhibited. As an initial attempt vertical projection based pixel distribution has adapted for zone based approach. Normalization has been considered as a pre-processing step to standardize (50X50) the image and select the features. As numerous features are present in the Tamil characters, calculation of pixel distance using vertical projection is not sufficient to find all the features. The primary characters (247 characters) have not been considered in this process. The highly challengeable factors such as different styles of characters (lot of variation, abnormal writing and continuous writing) and discontinuities present in any of the characteristic features lead this procedure to obtain negative results.

Another zoning procedure employed in research works [4] [7], have been deliberated on the document quality, noise and variability of the character documents. The reasonable pre-processing steps involved in this process achieve the beneficial accuracy but consumes more time. Very normal styles of character samples of 34 characters have been experimented in this process. This algorithm does not work well if the characters contain more curve features and similar shapes.

Pixel densities calculated from the image centroid and zone centroid [6] was another effective process in zoning procedure. Here the Tamil numerals (0, 1, ..., 9) have been chosen for recognition. The data samples of Tamil numerals have been

collected and normalized into 50X50 sizes. 96% accuracy has been accomplished by this process. Limited pre-processing steps may save time but the pixel density calculation from image and zone centroid consumes more time. The image and zone centroid concepts helped to identify the curve related features. The limited complexities have been considered in the collected samples. This process has failed to recognize the characters in different style and distortion.

The pixel density calculated from each zone employed in [10], where 26 characters have been debated in the recognition process and accomplished 96.9% accuracy. Considerable pre-processing steps have been used to obtain the perfect image, but it reduces the speed of the process. This process failed to consider the various abnormal handwritten and variations in character style.

B Boundary Tracing

Here, the image was converted into two tone images (foreground 1 background 0) then the boundary tracing approach [3] has been applied to extract the contour of the image. 4 connectivity or 8 connectivity neighbouring algorithm has been used to trace the inner and outer boundary or outer boundary from the region of the image. 97% accuracy has been furnished in the recognition of 30 different characters. No characters have been exhibited in this process, also failed to recognize the characters with different shape and style.

C. Encoding Binary Variation

In the binary variation approach [5], the binary pixel values (1's) have been counted based on row, column and diagonal wise to discover the features. Tamil numerals (1 to 10) have been employed for recognition purpose. Very few pre-processing steps have been applied to get proper region. Limitations are character complexity, style variations and missing selection of primary characters (247 characters).

D. Determination of Small Groups

The statistical feature extraction system, based on number of pixel transitions (1 to 0 and 0 to 1) has been employed in [8]. It has afforded good accuracy to numerous character sets. Limited sample sets (2 to 10 people) have been considered, but that has been written by people of various age groups. The limitation of handwritten character samples has not been exhibited properly. Character selections have not been clear in this process and it provides negative results when similar shaped characters are taken into account. The noise removal is one of the essential pre-processing steps in this appendage because this procedure is based on pixel transition, but it has not been employed in this process.

E. Binary Mapping

In the binary mapping approach [9], the character regions of image have been divided into small blocks to find the maximum 1's. This have contributed good accuracy rate and also used to extract the feature from the formal character sets. This process might not be desirable in case of character samples containing different styles and shapes. Limited samples of 11 characters out of 247 have been employed on this process.

F. Dimensional Method

The capable method to produce an intended result is Dimensional method [11] [14], where Tamil numerals and eight letters from 247 characters have preferred for recognition. Here the character region of the image has been divided into many blocks. Then direction chain code (Structural method) and gradient of the image (Statistical method) have been calculated as features from those blocks. Numerous characters with challengeable samples have not been considered in this procedure. Limited pre-processing steps may lead this process very fast but it may not provide good achievement in recognition.

G. Water Reservoir

The water reservoir was another approach [12], where the water have been poured from different direction of the character sample to find the reservoir features such as top reservoir, bottom reservoir, left reservoir, right reservoir, water overflow, normal distance between reservoir baseline and water level of reservoir and the profile information. This has been a very efficient procedure to extract the feature effectively from printed and very formal handwritten Tamil characters. Few characters in Tamil don't have provision to allow the water filling method (water reservoir) to identify the characteristic feature. In this circumstance, the character features are split into many pieces and apply the water to identify the curves or

lines. Finally those features can be merged to apply the character identification techniques. In [12], a number of pre-processing steps have been used to obtain better clarity feature. Very formal limited characters have been considered in this time consuming process

H. Fourier and Wavelet Transform

The Fourier and wavelet transform have provided lot of contribution in Tamil OCR [13]. No proper pre-processing and feature selection steps have been found to point out the features like curves and curvature, as the Tamil characters contain many curves and loops. Ancient characters were informal which contains more twists. Hence the vertical and horizontal line identification has not been enough to point out the character features. Better recognition results could be achieved if focussed more on pre-processing steps, mainly in the Ancient character recognition.

I. Projection Profile and Moment Based.

In projection profile and word profile [15][16] scenario, sum of all foreground intensity value, upper and lower word profile from upper and lower boundaries have been taken as features. Proper results have been achieved for a good sample set which contains 10 Tamil Handwritten characters.

J. Scale Invariant Feature Transform (SIFT)

SIFT Feature is one of the renowned techniques on the Pattern recognition environment. In Tamil Handwritten OCR the SIFT [1] have been provided enormous results but failed to provide good result if character samples were abnormal (Informal) and similar shaped. Out of 247, selected 20 characters have been chosen for this process. This would be one of the time consuming process.

III. PROBLEM DEFINITION

Data collection must be maximized and amended to face the real time challenges in Tamil Hand written character recognition environment. Overall, many research works have faced problem to identify the character with abnormal handwritten and similar shaped. Statistical feature extraction methods have not desired many pre-processing steps (it depends on the extraction methods), but pre-processing steps are essential to reduce the error rate and remove the unwanted signals from the image. There are 247 characters to identify, but 90% characters are yet to be selected. Classification algorithms can't provide 100% results directly, but better result could be obtained only by selecting and extracting the best feature sets. Table 1 and 2 points out the various issues left out by many researches in Tamil handwritten recognition.

TABLE 1. ANALYSIS OF FEATURE EXTRACTION TECHNIQUES USED IN TAMIL HANDWRITTEN OCR

S. No	Algorithm Used	Positives	Vital Analysis
1	Zoning	<ul style="list-style-type: none"> Tamil Numerals Beneficial results Good Sample sets Not a time consuming process 	<ul style="list-style-type: none"> Less character sets Considered only formal characters Need more pre-processing steps. Failed to remove unwanted pixels
2	Boundary Tracing	<ul style="list-style-type: none"> 30 Characters Not a time consuming process Good accuracy 	<ul style="list-style-type: none"> More characters are yet to be considered Feature selection is not proper Character specification is not clear Negative results when variation in character shape and style
3	Encoding binary Variation	<ul style="list-style-type: none"> Success in formal characters Good accuracy Tamil Numerals 	<ul style="list-style-type: none"> Negative results will obtain from the abnormal and similar shaped characters Feature selection is not sufficient
4	Determination of small Groups	<ul style="list-style-type: none"> Not a time Consuming process 	<ul style="list-style-type: none"> Handwritten characters have not been selected Character selection is not clear Negative results if the characters are informal or similar shaped Noise removal process is needed in this phase
5	Binary Mapping	<ul style="list-style-type: none"> 11 characters selected Good feature selection 	<ul style="list-style-type: none"> Limited number of samples Need more Pre-processing steps
6	Dimension	<ul style="list-style-type: none"> Capable to provide good results 	<ul style="list-style-type: none"> Feature selection is not enough

		Variation [5]	- க, உ, ட, ச, று, கூ, எ, அ, கூ, ம)			above problems
		Dimensional method [14]	8 (௭ ௭ ௭ ௭ ௭ ௭ ௭ ௭)	2690	96 %	
5	2011	SIFT [1]	20 (அ, ஆ, இ, ஈ, உ, ஊ, எ, ஏ, ஐ, ஒ, ஓ, க, ங, ம, ல, ட, ண, த, ப, ய)	2000	79 %	No improvement from above problems
		Vertical projection profile, word profile, background to ink transition [16]	10 (40 from 4 document),	-	80.75%	
6	2012	Fourier and wavelet Transform [13]	-	-	98 %	Only improvement: Ancient characters
		Vertical projection profile, word profile, background to ink transition [15]	10 (40 from 4 document)	-	79 %	

IV. CONCLUSION

This paper focuses mainly on the feature extraction issues available in various statistical THCRS researches. Feature Extraction heavily depends on the algorithmic complexity and pre-processing steps used. Only limited and similar characters have been considered for recognition. Most of the procedures are time consuming, writer dependent and character dependent. Supporting character combinations are less frequently used. Most of the process depends on the writers and characters. Researchers who wish to work in Tamil handwriting character recognition can address the above open challenges in future.

References

- [1] A. Subashini and N.D Kodikara, "A Novel SIFE-based Codebook Generation for Handwritten Tamil character Recognition", 6th IEEE International Conference on Industrial and information systems, Page (s): 261 – 264, 2011
- [2] S.V Rajashekararadhya, Vanaja Ranjan P, Manhunath Aradhya V N "Isolated Handwritten Kannada and Tamil Numeral Recognition: A Novel Approach", First IEEE International Conference on Emerging Trends in engineering and Technology, Page (s): 1192-1195, 2008.
- [3] J. Sutha, N. Rama Raj, "Neural network based offline Tamil handwritten character recognition System", IEEE International Conference on Computational Intelligence and Multimedia Vol : 2, page (s): 446 – 450, 2007.
- [4] N. Shanthi, K. Duraiswami, "A Novel SVM -based Handwritten Tamil character recognition system", Springer, Pattern Analysis & Application, Vol –13, No.2, 173-180, 2010.
- [5] Bremananth R and Prakash A, "Tamil Numerals Identification", IEEE International Conference on Advances in Recent Technologies in Communication and Computing, page (s): 620 – 622, 2009.
- [6] S.V Rajashekararadhya, Dr. Vanaja Ranjan P, "Efficient Zone based Feature Extraction Algorithm for Handwritten Numeral Recognition of Four Popular south Indian Scripts", International Journal of Theoretical and Applied Information Technology, pages: 1171 – 1181, 2008.
- [7] N. Shanthi, K. Duraiswami, "Performance Comparison of Different Image size for Recognizing unconstrained Handwritten Tamil character using SVM", Journal of Computer Science Vol-3 (9): Page (3) 760-764, 2007.
- [8] U.Bhattacharya, S.K Ghosh and S.K Parui, "A Two Stage Recognition Scheme for Handwritten Tamil Characters", Ninth International Conference on Document Analysis and Recognition Vol :1 page (s): 511 – 515, 2007
- [9] K. Sarveswaran and Ratnaweera, "An Adaptive Technique for Handwritten Tamil Character Recognition", International Conference on Intelligent and advanced Systems, page (s): 151 – 156, 2007

- [10] S. Hewavitharana, H.C Fernando, "A Two Stage Classification Approach to Tamil Handwritten Recognition", Tamil Internet, California, USA, 2002
- [11] U.Pal, T.Wakabayashi, N.Sharma, and F.Kimura, "Handwritten numerals Recognition of six popular Indian Scripts", IEEE, ninth international conference on Document Analysis and Recognition, Vol 2: Page (s). 749-753, 2007
- [12] Sukalpa chanda, Srikantha Pal and Umapada Pal, "Word-wise Sinhala Tamil and English Script Identification Using Gaussian Kernal SVM", IEEE, 2008
- [13] S. Raja Kumar, V. Subbiah Bharathi, "An Offline Ancient Tamil Script Recognition from Temple Wall Inscription using Fourier and Wavelet Features", European Journal of Scientific Research, ISSN 1450-216X, Vol.80 No, 4, 2012
- [14] T. Wakabayashi, U. Pal, F. Kimura and Y. Miyake, "F-Ratio Based Weighted Feature Extraction for Similar Shape Character Recognition", IEEE, Tenth International Conference on Document Analysis and Recognition, Page (s): 196-200, 2009
- [15] A.N. Sigappi, S. Palanivel "Query Word Image Retrieval Scheme for Handwritten Tamil Documents", International Journal of Applied Information System, Vol.1, No-1, Page (s): 1-5, 2012
- [16] A.N. Sigappi, S. Palanivel and V. Ramalingam, "Handwritten Document Retrieval System for Tamil Language", International Journal of Computer Application, ISSN: 0975-8887, Vol.31, No.4, Page (s)-42-47, 2011
- [17] Abirami.S and Manjula.D, "Feature String based intelligent information retrieval from Tamil document images", International Journal of Computer Applications in Technology, Vol.35, No.2, pp.150-164.
- [18] Antony Robert Raj and Abirami.S, "A Survey on Tamil Handwritten Character Recognition", CS&IT, CSCP, 2012.

Offline Tamil Handwritten Character Recognition

¹Jegadheeswaran J, ²Abirami S, ³Vijayalakshmi M

^{1, 2, 3} Department of Information Science and Technology, College of Engineering, Anna University, Chennai

¹jjegadhees@gmail.com ²abirami_mr@yahoo.com ³vijim@annauniv.edu

Abstract— Handwriting recognition is the ability of a computer to recognize characters and other symbols that have been written by hand in natural handwriting. It may be either online or offline. Off-line handwriting recognition involves the automatic conversion of text handwritten in an image into letter codes which are usable within computer and text-processing applications. On-line handwriting recognition involves the automatic conversion of text the moment as it is written on a digitizer, where a sensor picks up the pen-tip movements as well as pen-up/pen-down switching. This paper describes a system for recognizing offline Tamil handwritten characters into machine printed Tamil characters format. Data samples are collected from different writers on A4 sized documents. They are scanned at a resolution of 300 dpi and stored as jpeg images. Various pre-processing operations have been performed on the digitized image to enhance the quality of the image. Then the individual characters are extracted from the given document image and the important features of the character images are derived and classified using the classifier. Zone based feature extraction has been adopted here which calculates pixel densities for Nine different zones in the character image. Later, features like number of black and white pixels, number of transitions, and the direction of contour movement in each zone has been calculated. Totally, these features are considered as feature set. Later, characters are identified by classifying these featuresets using SVM.

Keywords: Handwritten recognition, pre-processing, noise removal, segmentation, feature extraction, classification, SVM.

1. INTRODUCTION

In document image processing, the paper documents are initially scanned and stored in the hard disk. It is easy to define document image processing as scanning-storing-retrieving. The final outcome of document image processing will be in compatible electronic format, which makes documents to access easier. Document image processing comprises of a set of simple techniques and procedures, which are used to work upon the images of documents and convert them from pixel information into a format that can be read by a computer.

The document image may be printed or handwritten. The printed document means that the documents are written by some electronic devices such as digital pen etc. handwritten document means that the documents are written by ink pen. Handwritten character recognition (HCR) is a growing area in academic and production fields. The recognition system has two types. They are online and offline. Off-line handwriting recognition is the subfield of Optical Character Recognition (OCR). Optical character recognition is the mechanical or electronic conversion of scanned images of handwritten or printed text into machine-encoded text. There are several works already done in several languages like Telugu, Hindi, English, Tamil etc. Here we discuss about Tamil Character Recognition (TCR). Recognition of any handwritten characters with respect to any language is difficult, because, the handwritten characters differ in written format, intensity, scale and orientation, not only from person to person but also according to the mood of the same person[22]. The applications of OCR are conversion of handwritten document to an editable soft format, Recognition of postal addresses for automated postal system, data acquisition in bank checks, Data and word processing.

Tamil belongs to the southern branch of the Dravidian languages. Tamil is a most popular language in south side of India and also the people lives in srilanka, Malaysia, Singapore are speaking in Tamil. The Tamil language has totally 247 characters they are 12 vowels, 18 consonants, 1 special character named aytam and an additional 216 combinant letters representing a total of 247 combinations of a consonant and a vowel, a mute consonant, or a vowel alone[21]. Therefore, in this paper we proposed a new technique for Tamil handwritten character recognition. The remainder of the paper is organized as follows: Section 2 deals with literature survey. Section 3 deals about the proposed system architecture. Section 4 describes about results and performance analysis, and conclusions are mentioned in section 5.

II. RELATED WORK

S.V. Rajashekaradhyya et. al has been proposed the projection distance metric and zoning based scheme for numeral recognition from Kannada and Tamil numerals [4]. A nearest neighbor classifier is used for subsequent purpose.

Vanaja Ranjan et. al has been proposed a zone based hybrid feature extraction algorithm scheme towards the recognition of off-line handwritten numerals of two popular south-Indian scripts [6]. The zone based feature extraction method used here for feature extraction.

Ramanathan et. al has been proposed a new technique for Optical character Recognition using Gabor filters and Support Vector machines (SVM) [10]. This model recognizes the characters of six different fonts in English and Twelve different fonts in Tamil.

Shanthi and K. Duraiswamy describes a system for recognizing offline handwritten Tamil characters using support vector machine (SVM) [12]. Pixel densities are calculated for 64 different zones of the image and these values are used as the features of a character.

Ponmathavan et. al are done a Tamil font recognition based on global texture analysis [16]. The main objective of this proposal is to employ support vector machines (SVM) in identifying various fonts in Tamil.

In this paper we proposed a recognition system for handwritten Tamil characters. The scanned document image is preprocessed to ensure that the characters are in a suitable format. Then the line, word and characters are segmented and features are extracted. Finally Support Vector Machine has been used for training of extracted features and recognition of characters.

III. PROPOSED ARCHITECTURE

Basically off-line handwritten character recognition system includes three stages. They are image preprocessing, feature extraction, and classification.

The process of handwriting recognition involves extraction of some defined characteristics from the character image called features to classify an unknown character into one of the known classes. Preprocessing is primarily used to improve the quality of the input image. A feature extractor is must need for efficient data representation and extracting meaningful features for later processing. A classifier assigns the characters to one of the several classes. The architecture of the proposed system is shown in Figure 1.

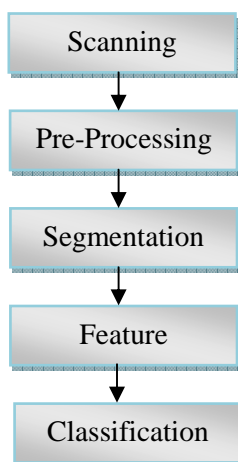


Figure 1 Offline Handwriting Recognition Architecture

A. Preprocessing

There is a need of preprocessing steps in character recognition to improve the quality of the digitized scanned document image to achieve for better recognition results. The output of this module is noise free image. The following steps are performed during preprocessing.

- 1) Binarization
- 2) Noise Removal
- 3) Thinning

B. Binarization

Separating the foreground (ink) from background (paper) is called binarization. This means, the binary image can have 2 possible values. 1 and 0 representing white and black respectively. Here a threshold value has been fixed for a given scanned document image. By using this value, the image can be binarized. This method known as global thresholding. The binarization steps are given below.

Step1: Select an initial estimate for the variable T. (Initial estimate is the midpoint between the minimum and maximum intensity values in the image.)

Step2: Segment the image using the threshold value T. This will produce two groups. They are G1, consisting of all pixels with intensity values $>T$, and G2, consisting of all pixels with intensity value $< T$.

Step 3: Compute the average intensity values μ_1 and μ_2 for the pixels in regions G1 and G2.

Step 4: Compute the new threshold value: $T = \frac{1}{2} (\mu_1 + \mu_2)$

Step 5: Repeat steps 2 to 4 until the difference in T in successive iterations is smaller than a predefined parameter T. After binarization, noise removal has been performed.

C. Noise Removal

Due to poor quality of the scanner, some noises exist with the scanned document image. This will affect the results of the recognition to obtain better results, noises need to be removed from the scanned document image.

D. Thinning

Next to that, thinning has been done to obtain the skeleton of the image by peeling off the image pixels without changing the originality of the image. We have used hit and miss transform for thinning. These are explained in following steps.

Step1: Sweep the image with one of the structuring elements shown in Figure 2.

Step2: If the 3x3 image pattern matches the structuring element (blanks denote don't-care points), then put one on the corresponding location of the result image, otherwise put a zero.

Step3: Invert the resulting image and perform binary AND of it with the initial preprocessed image. This removes the points produced by the first structuring element.

Step4: Repeat processes 1-3 until all structuring elements have passed over the image (each element takes the output of the previous one as input).

Step5: Repeat processes 1-4 until the image doesn't change anymore.

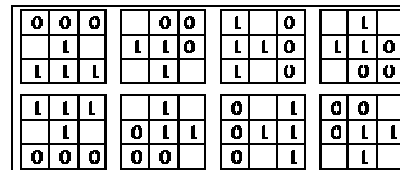


Figure 2 Structuring elements

E. Segmentation

Segmentation is a process of separating lines, words, and even characters of a hand written or machine printed document. It is a crucial step as it extracts the meaningful regions for analysis. There are several approaches for segmenting the region of interest. For handwritten document, this is quite difficult. The details of line, word and character segmentation has been discussed as follows.

F. Line Segmentation

The horizontal projection value for each row has been calculated for line segmentation. Horizontal projection value is nothing but number of on pixels presents in the row. Lines are segmented using horizontal projection profile[21].

G. Word and Character Recognition

Vertical projection values for each column has been calculated to aid word and character segmentation. Vertical projection value is nothing but number of on pixels present in that the column. Words and Character are segmented through Vertical projection profiles[21].

H . Feature Extraction

Transforming the input data into a set of feature is called feature extraction. If the features are carefully chosen it is expected that the feature set will extract the relevant information from the input data. Extracted features are used for classifying the character images. Here the features derived are zone based features, number of on pixels present in the image (column and row wise) known as ‘encoding binary variation’, number of transitions and directional features from the contour of the image. The output of this module is the set of features that are derived from the character image.

Zone Based Feature Extraction

Here the character image is divided into 9 zones. This is done by irrespective of the size of the image. Here a combination of icz and zcz algorithm has been used and they are explained in the following steps[6].

- Step 1:** Compute the Centroid of the input image.
- Step 2:** Divide the input image into 9 zones.
- Step 3:** Compute the pixel distance between the image centroid to each pixel present in the zone.
- Step 4:** Repeat step 3 for the entire pixels present in the zone.
- Step 5:** Compute the average pixel distance for all pixels present in the zone.
- Step 6:** Compute the centroid of the zone.
- Step 7:** Compute the pixel distance between the zone centroid to each pixel present in the zone.
- Step 8:** Repeat step 7 for the entire pixel present in the zone.
- Step 9:** Compute the average distance between the points are calculated in zone.
- Step 10:** Repeat the steps 3-9 sequentially for the entire zone and it is repeated for all the 9 zones.

Totally zone centroid & image centroid features are obtained for classification and recognition.

Encoding Binary Variation

In this method, number of 0's present in the image has been calculated[5]. This is done independently on all the 9 zones present in the image. The steps in this method are given below.

Algorithm for Number of On Pixels in Row and Column Pixels:

- Step 1:** Set count1 = 0 and count2=0.
- Step 2:** Increment the variable count1 to 1 if the image pixel value is 0 during row wise processing and increment the variable count2 to 1 if the image pixel value is 0 during column wise processing and store the variable count1 and count2 into an array.
- Step 3:** Repeat the process until row >= height of the zone and column count >= width of the zone.
- Step 4:** This is repeatedly done for all the 9 zones present in the image.

Number of Transitions:

In this method, number of transitions (i.e. 1's to 0's or 0's to 1's) for each zone has been calculated[23]. The steps are explained in the following steps.

- Step 1:** Set count=0
- Step 2:** Count the number of transitions in a zone until it reaches the size of the zone. If the transition is occur then increment the count value.
- Step 3:** This process is repeated for all the 9 zones present in the image.

Contour Formation:

Here the boundary of the pixels from the image has been traced. Here 8-neighborhood pixel algorithm has been used here boundary tracing. The direction of traversing will be from either clockwise or anti-clockwise. Then the direction of the contour moving can be detected by using chain code. The algorithm is explained below.

Begin

Set B to be empty.

From left to right scan the cells until a black pixel, s, of M (input image) is found

Insert s in B

Set the current boundary pixel p to s i.e., $p=s$

Backtrack i.e., (move to the pixel from which has been entered)

Set c to be the next clockwise point in M (p)

While c not equal to s do

 If c is black

 Insert c in B

 Set $p=c$

 Backtrack (move the current pixel c to pixel from which p was entered)

Else

 Advance the current point c to the next clockwise point in M (p)

End While

End

Where

B - Array to store the contour.

S - Starting black pixel position

P - Current boundary point

After contours function, chain codes of the image has been traced for finding the direction of movement of the contour. The table 1 given below shows the positions of the next pixels might be occur. It shows the codes 0 to 7.

Table 1. Chain Codes

3	2	1
4	Current pixel (I, J)	0
5	6	7

The pixel positions for 0 to 7 chain codes are shown in terms of I and J coordinates in the table2.

Table 2. Next pixel positions corresponding to chain codes

Current pixel at coordinate (i, j)		
Code	Next row	Next column
0	I	J+1
1	I-1	J+1
2	I-1	J
3	I-1	J-1
4	I	J-1
5	I+1	J-1
6	I+1	J
7	I+1	J+1

finally all the features have been extracted and given to the classifier for recognition of characters.

I. Classification

The support vector machine is a new classifier that is extensively used in many pattern recognition applications nowadays. The SVM uses the principle of structural risk minimization. Regarding the pattern classification problem, the SVM demonstrates a very good generalization performance. SVM is a binary classifier that separates any two classes by finding a hyper plane of maximum margin between the two classes. The margin means that the minimal distance from the separating hyper plane to the closest data points. The SVM gives better results, when the margin is maximal. The outcome of the SVM is based on the data points that are at the margin and are called support vectors. There are two approaches to extend the SVM to make suitable for multiclass classification. The first one is (ONO) and the other one is one against all (ONA). Here we have used ONA approach where N SVM classifiers are performed to separate one of N mutually exclusive classes from all other classes. The output of classification is the recognized character class.

IV. EXPERIMENTAL RESULTS

Feature set consists of zone based features such as number of on pixels in row and column in the zone, number of transitions in the zone, and directional features. With the combination of Feature set and support vector machine, a high accuracy recognition has been realized. A database of 20 users has been divided into two categories. The training set consists of the writing samples of 10 users selected at random from the 20, and the test set, of the remaining 10 users. A portion of the training data was also used to test the system. In the training set, a recognition rate of 100% has been achieved and the test set results has been shown in the following table. From these results we obtain that, the training set produced much higher recognition rate than the test set. Support vector machine provides better recognition results which are shown in Table 3.

Table 3. Performance analysis

CHARACTER	TOTALLY TRAINED	TEST SAMPLE	CORRECTLY CLASSIFIED	MIS-CLASSIFICATION	ACCURACY
௮	20	12	9	3	85%
௮	20	10	8	2	80%
௮	20	8	6	2	75%
௮	20	6	4	2	67%
௮	20	4	3	1	75%
௮	20	11	8	3	72%
௮	20	6	4	2	67%
௮	15	7	5	2	71%
௮	15	6	4	2	66%
௮	15	9	6	3	66%
௮	15	5	4	1	80%
௮	15	6	4	2	66%
௮	15	4	3	1	75%
௮	15	8	7	1	87%
௮	15	7	5	2	71%
௮	15	9	7	2	77%
௮	15	6	4	2	66%
௮	15	6	4	2	66%
௮	15	5	3	2	60%
௮	15	7	6	1	85%
௮	15	7	5	2	71%

V. CONCLUSION

This paper provides a system to recognize offline handwritten Tamil characters using support vector machine. Here features are extracted from the character image using techniques such as zone centroid, image centroid, encoding binary variation in each zone, number of transition in each zone and the movement of the contour. Based on the features, characters are classified into a resultant class. We are able to classify 90-100 characters. Performance analysis clearly shows that the recognition accuracy is acceptable for different Tamil characters. The recognition errors were mainly due to abnormal writing of writers which arise among and ambiguity similar shaped characters.

Future work can include extracting more and more robust features to achieve better discrimination power. The recognition accuracy of the individual characters can be further improved by using powerful classifiers.

References

- R.Jagadeesh Kannan, R.Prabhakar and R.M. Suresh “Offline cursive tamil handwritten recognition”, IEEE International Conference on Security Technology, VOL.4, pp. 159-164, 2008.
- Akshay apte, harshad gado “Tamil character recognition using structural features”, IEEE Transactions On pattern recognition, VOL.11, pp. 174-178, November 2005.
- J.sutha, N.ramaraj “Neural Network Based offline Tamil Handwritten character recognition system”- International Conference on Computational Intelligence and Multimedia Applications, pp. 446-450, 2007.
- S.V. Rajashekararadhya, P. Vanaja Ranjan and V.N. Manjunath Aradhya, "Isolated handwritten Kannada and Tamil numeral recognition: A novel approach", First International Conference on Emerging Trends in Engineering and Technology ICETET 08, pp.1192-1195, 2008.
- Bremananth R and Prakash S, “Tamil numerals identification”, IEEE International Conference on Advances in Recent Technologies in Communication and Computing, pp. 620-622, 2009.
- S.V. Rajashekararadhya, P. Vanaja Ranjan, “Zone-Based Hybrid Feature Extraction Algorithm for Handwritten Numeral Recognition of Two Popular Indian Scripts”, IEEE conference on World Congress on Nature & Biologically Inspired Computing, pp. 526-530, 2009.
- P. Banumathi and Dr. G. M. Nasira “Handwritten Tamil Character Recognition using Artificial Neural Networks”- IEEE transactions on pattern recognition, pp. 1028-1032 April 2011.
- U. Bhattacharya, S. K. Ghosh and S. K. Parui “A Two Stage Recognition Scheme for Handwritten Tamil Characters”, Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), pp. 732-736, 2007.
- Shashikiran K, Kolli Sai Prasad, Rituraj Kunwar, A. G. Ramakrishnan, “Comparison of HMM and SDTW for Tamil Handwritten Character Recognition”, IEEE conference on pattern recognition, pp. 253-256, 2010.
- R.Ramanathan, S.Ponmathavan, N.Valliappan, L.Thaneshwaran, Arun.S.Nair and Dr.K.P.Soman, “Optical Character Recognition for English and Tamil Using Support Vector Machines”, International Conference on Advances in Computing, Control, and Telecommunication Technologies, pp. 610-612, 2009.
- A. Subashini and N. D. Kodikara “Novel SIFT-based Codebook Generation for Handwritten Tamil Character Recognition” 6thInternational Conference on Industrial and Information Systems ICIIS 2011, pp. 261-264, Aug. 16-19, 2011.
- N. Shanthi, K. Duraiswamy “A novel SVM-based handwritten Tamil character recognition system”, International conference on pattern anal application, vol.13, pp. 173-180, 2009.
- N. Shanthi and K. Duraiswamy “Performance Comparison of Different Image Sizes for Recognizing Unconstrained Handwritten Tamil Characters using SVM”, Journal of Computer Science 3 (9), pp. 760-764, 2007.
- R.M.Suresh, L.Ganesan “Recognition of Printed and Handwritten Tamil Characters Using Fuzzy Approach”, Sixth International Conference on Computational Intelligence and Multimedia Applications, pp. 543-548, 2005.
- R.Indra Gandhi and Dr.K.Iyakutti “An Attempt to Recognize Handwritten Tamil Character Using Kohonen SOM”, International Journal of Advance Networking and Applications, volume: 01, pp. 188-192, 2009.
- R.Ramanathan, S.Ponmathavan, N.Valliappan, L.Thaneshwaran, Arun.S.Nair and Dr.K.P.Soman, “Tamil font Recognition using gabor filters Support Vector Machines”, International Conference on Advances in Computing, Control, and Telecommunication Technologies, pp. 613-615, 2009.
- S. Hewavitharana and H. C. Fernando, “A Two Stage Classification Approach to Tamil Handwriting Recognition”, International conference on tamil internet california, pp. 423-429, 2002.
- A. Subashini and N. D. Kodikara, “Bag-of-keypoints Approach for Tamil Handwritten Character Recognition using SVMs”, International Conference on Advances in ICT for Emerging Regions - ICTer2011, pp. 102-107, 2011.
- U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts", Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2007.
- B.V.Dhandra, R.G.Benne and Mallikarjun Hangarge, “Handwritten Kannada Numeral Recognition Based on Structural Features”, International Conference on Computational Intelligence and Multimedia Applications, pp. 224-228, 2007.
- Abirami. S and Manjula. D, ‘Feature String based intelligent information retrieval from Tamil document Images’, international Journal of Computer Applications in Technology, Vol. 35, No. 2, pp. 150-164.
- Antony Robert Raj. M and Abirami. S, “A survey on Tamil Handwritten Character Recognition”, CS&IT, CSCP, 2012.
- Toni. M and Rat. R Manmatha (2003), “Features for Word spotting in historical manuscripts, ICDAR, pp. 218-222.

Tamil Script Recognition using Smart Phone

Rex Aantonny. P^{*}, Iniya Nehru .E⁺

^{*}Software Executive, Cadgraf Digital Pvt Ltd, Chennai, India, ¹joyemperor@gmail.com

⁺National Informatics Centre, Chennai, India, ²nehru@nic.in

Abstract— Having an archive of important documents is always a good idea; this could be done with a filing system which involves a lot of papers or it could be done digitally with the use of computers. However, a filing system could be unpleasant to navigate though and time consuming. With the aid of Optical Character Recognition (OCR), this is a technology enables conversion of different types of documents into searchable and editable data. This can be performed more efficiently. Documents that are normally take hours to digitize can be performed in a matter of minutes and with the resurgence of mobile technology, it would be fitting to have such an application on a mobile phone. The world of OCR is having a lot of conflicts in identification and also in isolation of the characters, particularly identification of the ancient Tamil scripts differ from document style, which is the collection of text characters of various style, form, trend and situation along with complicated texture backgrounds. Hence, the extraction of texts in is difficult as well as challenging task. Using our proposed method it is seen that the extraction of Tamil script is easier compared to other existing methods.

Keywords— Character recognition, Translation, Language Identification.

INTRODUCTION

Current OCR (optical character reader) technology [1-6] is widely applied to Zip identification, product inspection and classification, document identification, vehicle number recognition, picture recognition, slips and checks. OCR technology in the United States of America was finished in one-stage from starting in the 1950s to early 1970s. Since 1966, Pattern Information Processing System project became the instrument of development by participating many companies. On-lain OCR research to recognize at the same time with handwriting was the first attempt in 1959. There are many character recognition systems such as auto-inspection system of Japan's NEC, mail sorter of Germany's Siemens, face, fingerprint identification, document processing system, the field of artificial intelligence, document pattern recognition / analysis, automatic processing of checks, number / word / string recognition of Canada Concordia University.

Most of the character identification program will be recognized through the input image with a scanner or a digital camera and computer application. The computer and scanner provide a spatial size problem. To rectify this problem we propose a methodology that provides a character recognition system using smart phone. Character identification software developed for smart phones with an emphasis on speed, spatial, hardware, financial limitations can be solved. But performances of smart phone and computer are different; the speed of massive character identification is varied. Hardware speeds up the development of smart phones; this issue seems to be resolved as soon as possible. In this paper, the character identification method is presented by using OCR technology and smart phone. This paper as follows. Section II provides the proposed character recognition method. Section III shows the important of the database. In Section IV says the Neural network formation. In section V simulation of the ocr and the section VI we summarize the main results.

PROPOSED CHARACTER RECOGNITION METHOD

Here we creating a language recognition technique using, Tesseract-OCR [2] and Mezzofanti. Tesseract-OCR 2.03 version supports some languages such as English. From the 3.00 version, the Tamil language (current script only) is supported. It employs the image processing library called as leptonica. Tesseract-OCR is used in AOSP and eyes-free project. Mezzofanti is an open-source Android Application. It recognizes the characters in the image taken by the camera using the Tesseract library. The app in version 1.0.3, uses Tesseract 2.03 version, so it has the drawback that does not support many language identification. For that reason, we implement Mezzofanti as Tesseract 3.0 version. Tesseract 3.0 is build to NDK and then the source code of the packages associated with the Tesseract is downloaded. Mezzofanti source should be modified.

Eclipse or Ant is used to build the Mezzofanti. Mezzofanti application and dictionary and pre-learning data files have to be installed. Mezzofanti for Eclipse or Ant is to build, and installed in Android smart phone. Mezzofanti and Tesseract can add simply a language that is not supported by default; the database only for that language can be easily applied. Proposed system can support full mode and line mode. Full mode can recognize the entire document, and lines mode can be recognize one line of image. The mismatched characters and errors are deals to increase the efficiency that prefer at the result area.

Fig.1 shows the screen shots for the proposed recognition system on smart phone.

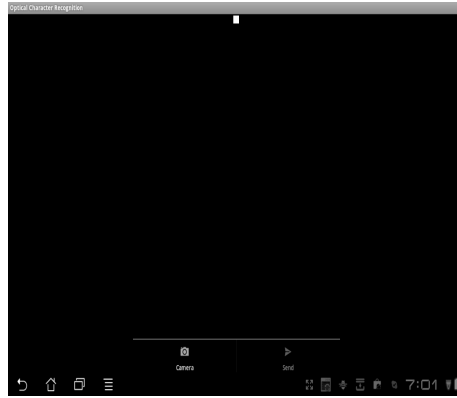


Figure1. Camera Activity

There is an option to change the line mode and full mode . At the line mode only identify the letters inside the white area. Identify the character on the white area, and then transfer upon the result area recognize.

The recognition result is shown as follows. Fig 2 shows the digitalized image of the Ancient Tamil. After being captured by smart phone camera, the data is processed by binarization for object conversion that shown in figure3. Then we can see the recognition results on the screen of smart phone as shown in figure 10 (b).

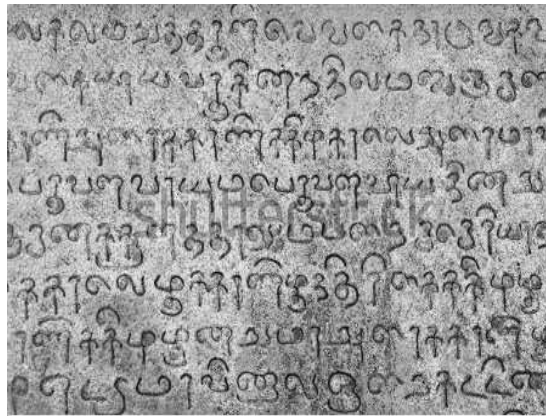


Figure2. Example of natural scene of Ancient Tamil

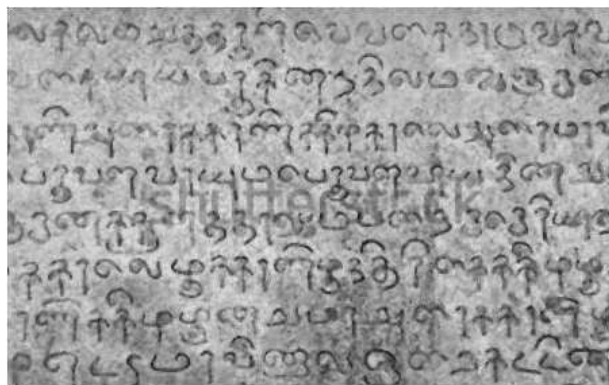


Figure3. Binarization of captured data

DATABASE QUALITY

The proposed mechanism is fully depending upon the dataset where we declare in the database. In Ancient Tamil language the written script was getting cultivated at each century (depending upon the area).The major thing is we have to train the data set for each century and make them as a cluster for make the searching process, this will help to increase the efficiency of the process.

The efficiency of the database used for matching the character, basically the identification of Tamil characters are really tough compare to the English letters to overcome this here we used neural networks and matrix formation of the characters are used as inputs and it makes more sense to solve all the problems.



Figure4.Example of 6th Century Tamil Letter«

The figure5 shows the 6th century Tamil character «, the character « is converted into binarized that is 0 or 1 for the efficiency of the process. Here the character is transferred to matrix form 63*64*3 by using Matlab algorithms after that its finalized data used to train the Neural network.

In the conversion of the character is performed at each pixel of the image its check the noise level and produce the matrix. Figure 6 shows the matrix form of the Tamil character « .

185	184	186	182	162	167	160	187	182	183	191	188	187	171	172	172	197	185	188	197	197	198	168	
124	121	194	192	197	200	196	204	183	187	184	103	107	108	27	24	26	115	121	121	196	200	195	173
114	124	192	198	199	204	196	184	183	186	100	109	108	24	29	27	121	122	121	196	198	204	167	
121	116	195	192	192	197	194	197	182	181	183	105	99	101	29	25	26	122	116	122	200	196	197	171
0	0	98	88	97	189	185	191	205	203	204	100	94	97	0	0	0	99	94	95	187	190	192	183
0	0	91	96	88	184	186	207	204	205	200	96	93	0	0	2	93	98	96	192	190	191	174	
0	0	94	88	93	184	189	185	201	204	203	94	94	93	1	0	0	96	93	91	190	189	191	183
19	13	15	14	15	135	130	133	200	199	200	112	118	114	0	2	0	116	121	121	192	191	192	175
20	14	12	13	12	130	133	131	197	196	198	112	113	113	2	1	1	117	120	118	192	193	193	178
8	15	12	13	14	135	125	134	197	201	200	115	113	113	0	0	0	116	121	119	193	194	194	176
169	168	0	0	0	129	140	133	239	239	239	100	104	103	0	0	1	119	123	122	198	200	199	178
164	164	0	0	0	134	140	134	245	238	245	99	98	102	0	0	0	119	120	121	199	200	199	175
171	166	0	0	0	139	128	136	241	238	244	98	104	98	0	0	0	119	119	120	198	199	199	170
94	93	8	6	6	100	103	100	150	156	151	91	95	94	0	0	1	126	125	124	199	199	199	169
99	98	3	4	5	98	102	104	152	156	150	89	93	91	0	0	0	126	125	124	199	199	199	169
98	94	7	6	5	100	103	99	150	150	156	92	94	0	0	0	0	126	126	125	199	198	198	167
22	24	0	0	1	12	8	10	2	1	2	38	36	37	3	6	7	126	127	129	192	192	194	162
29	20	1	4	0	10	12	10	1	1	3	31	39	41	2	6	7	126	126	128	192	193	194	163
23	26	0	1	3	14	10	8	0	5	0	37	38	32	7	8	7	128	127	129	193	193	194	162
127	133	13	15	11	43	51	49	85	77	82	20	23	24	0	0	4	126	125	127	195	195	195	165
133	131	12	9	9	46	51	43	77	84	84	20	22	29	9	4	127	125	127	196	196	196	170	
131	125	15	9	16	50	47	50	84	79	85	20	23	21	5	0	5	129	127	128	197	198	197	169
130	136	5	16	10	47	45	46	85	82	78	24	26	24	5	4	5	129	129	130	199	199	198	171
203	205	173	172	176	220	218	220	241	241	244	86	87	87	0	0	0	124	125	129	201	203	203	177
206	200	174	176	170	220	217	220	244	242	244	84	89	86	0	0	0	126	126	127	205	205	205	178
204	208	177	169	177	222	223	215	245	240	244	87	83	93	0	0	1	125	125	126	203	201	205	177
186	182	203	202	195	189	188	191	194	190	195	83	78	76	5	0	0	123	122	123	200	199	201	177
154	154	199	207	200	190	185	188	151	154	150	78	81	81	0	0	0	118	119	121	197	196	199	175
157	151	205	199	200	192	195	187	153	152	153	77	73	76	0	0	1	118	119	121	193	191	192	172
196	201	163	165	167	133	135	138	39	31	33	14	10	12	11	7	9	120	121	119	188	186	187	169
195	195	168	162	166	133	135	131	29	35	33	12	14	9	8	10	11	119	120	118	185	182	184	167
195	190	163	165	159	135	132	132	34	31	34	9	9	12	8	11	10	119	122	121	187	185	186	168
50	47	14	14	13	0	0	0	6	3	5	40	44	39	7	9	8	102	104	99	208	202	209	154
52	47	14	16	11	0	0	0	3	5	7	41	38	46	12	9	9	105	101	107	207	211	203	159
44	44	13	13	7	0	0	0	7	0	6	42	42	42	6	10	10	103	111	104	205	210	211	157
18	25	3	4	7	14	20	19	159	161	160	89	90	97	0	0	0	118	121	124	216	212	204	174
15	18	8	7	9	21	23	17	159	157	162	89	93	96	0	0	0	120	123	124	206	212	215	169
23	12	15	8	6	20	23	24	162	161	162	93	95	95	0	1	0	119	125	123	209	212	214	173
125	123	163	168	173	188	188	186	207	205	201	154	150	154	106	105	106	185	175	179	188	182	186	178
123	124	173	169	167	189	190	189	208	209	211	147	151	149	106	107	107	178	182	182	183	186	189	178
122	129	169	169	170	190	190	192	208	207	210	154	149	154	105	108	107	177	186	180	189	187	186	183
199	192	174	175	169	207	211	207	159	158	153	179	182	179	191	190	192	205	197	195	159	157	161	175
195	198	169	177	169	210	208	203	156	166	159	175	180	181	192	193	194	192	200	208	153	156	158	177
199	197	174	170	170	210	208	211	159	152	160	181	181	179	192	191	196	198	203	195	157	158	158	178

Figure5.example of 6th Century Tamil « matrix form

NEURAL NETWORK FORMATION

This need for accuracy is so demanding. Artificial Neural Network is so strong that it has even caused millions of people to learn an entirely new way to write. It has also made way that was easier for computers to detect. Unistroke recognition algorithms, like the popular raphiti used on Palm devices require the user to adapt instead of the device, essentially the antithesis of natural interface design. The technology has proven to be very accurate: Each character is written with a single stroke. This solves the character level segmentation problem that previously plagued handwriting recognition. The curve drawn between pen down and pen up events can be recognized in isolation.

Unistroke recognition algorithms can be relatively simple because there is no need to decide which parts of the curve belong to which character or to wait for more strokes that belong to the same character as is often

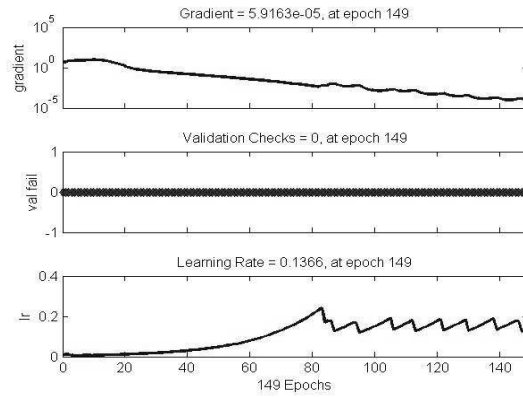


Figure6: Training state of the ANN for the Tamil character «

The figure 6 shows the graph of training state of the neurons.

The case when we try to recognize conventional handwriting. There have been many studies in the past using spiking neuron models to solve different problems. They provided a biologically plausible learning algorithm for realizing RBFs (Radial Basis Functions), which themselves are quite powerful in function approximation, pattern classification etc. In this study, spiking neurons were used to compute RBFs by storing information in their delays. The time difference between the pre and the post synaptic spikes was used to learn these delays. Neural networks use a set of processing elements (ornodes) analogous to neurons in the brain. These processing elements are interconnected in a network that can then identify patterns in data once it is exposed to the data, i.e the network learns from experience just as people do. This distinguishes neural networks from traditional computing programs, which simply follow instructions in a fixed sequential order.

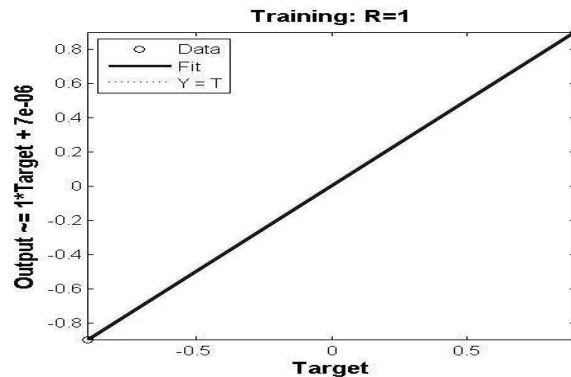


FIGURE7: REGRESSION STATE OF THE ANN FOR THE TAMIL CHARACTER «

The regression state of neurons is shown in figure7 . It's shows the regression rate of the neurons in the training of neural network.

There are many types of neural networks. One famous concept isotones neural network, a two-level network. In this concept, all images are down sampled before being used, which prevent the neural network from being confused by size and position. To develop a handwriting recognition system that is both as reliable as Unistroke, and natural enough to be comfortable, the system must be highly adaptable. Creating software that is as adaptable as its users are unique is a very challenging problem for conventional computer algorithms. This is why many people in the field of handwriting recognition are turning to neural networks to perform the recognition processing. Adaptable by their very nature, neural networks can bring to the computing world software that molds and conforms in ways algorithms like Unistroke never could. To line up the neurons with their recognized letters, each letter image in the network will be trained and fed into the network and the winning neuron is determined. The output neuron with the largest output value is considered the winner.

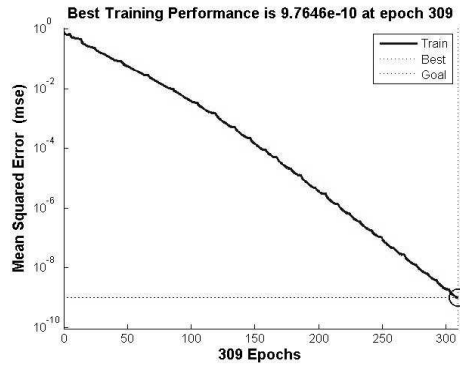


FIGURE8: PERFORMANCE STATE OF THE ANN FOR THE TAMIL CHARACTER «

the training state of the neurons shown in the figure8. There the best performance occurs in the 9.7646e-10 at epoch 309.

SIMULATION

To identify the conflicts in the natural scene we first implemented the ocr in English and evaluated its performance, those results are calculate as following

$$\text{Performance} = \frac{\text{Number of correctly recognized Character}}{\text{Total number of characters}}$$

For Tamil

Document type	Number of characters	Identification rate
Dark lighting	25	67.5%
Small fonts	75	80.1%
Wide letters	35	75.1%

Table 1: Performance comparison

Performance table for Tamil language in different styles of documents rate is shown in the Table1.

CONCLUSIONS

In this paper, character recognition system was implemented by using the Android smart phone. The implementation process of the system was described to recognize the characters in the document using the camera screen. Photo data taken by a smart phone can be compared with the database of the system, and then the characters can be identified; the recognized character can be created to a text file to take advantage. And also this method didn't need any network provider. This is more cost effective.



Figure 9 (a): capturing image by user

In figure 9 (a) shows the user capturing Tamil letters, that is taken as a digital document and the conversion takes place. The output of the taken image is shown in figure 9 (b).



Figure 9 (b): Extracted data from the database or server

WORK IN PROGRESS

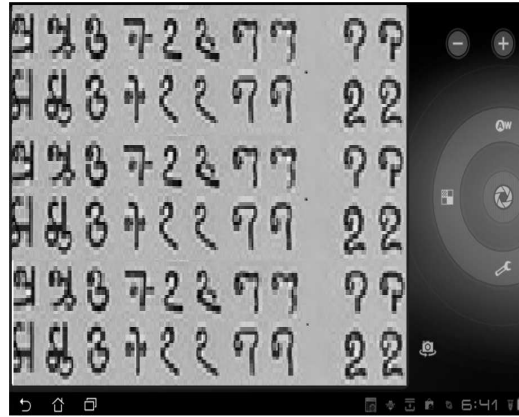


Figure 10 (b):Input image taken by the user

Shown figure 10 (a) is the image of ancient Tamil script captured by user and digitalized output after processing is shown in figure 10 (b).

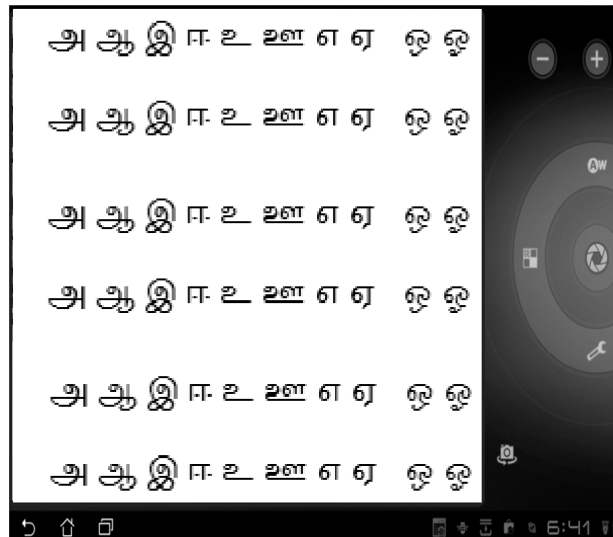


Figure 10 (b):Translated image

REFERENCES

- J.L. Blue, G.T. Candela, P.J. Grother, R. Chellappa, C.L. Wilson, Evaluation of pattern classifiers for fingerprint and OCR applications, *Pattern Recognition*, 27 (4), 1994, 485–501.
- R. Smith, An overview of the Tesseract OCR engine, *Proc. Int. Conf. Document Analysis and Recognition (ICDAR 2007)*, 2007, 629-633.
- M. Seeger, Binarising camera images for OCR, *Proc. Int. Conf. Document Analysis and Recognition*, 2001, 54-58.
- K. Wang, J. A. kangas, Character location in scene images from digital camera, *Pattern Recognition.*, 36 (10), 2003, 2287-2299.
- C. C. Chang, S. M. Hwang, D. J. Buehrer, A shape recognition scheme based on relative distances of feature points from the centroid. *Pattern Recognition*, 24 (11), 1991.1053-1063.
- T. Bernier, J.-A. Landry, A new method for representing and matching shapes of natural objects. *Pattern Recognition*, 36 (8), 2003. 1711-1723.
- Pattern recognition using multilayer neural-genetic algorithm-Yas Abbas Alsultanny□, Musbah M. Aqel- PII: S0925-2312 (02)00619-7

Use of Tamil Grammar Rules for Correcting Errors in Optical Character Recognised Document

Rajeswari Sridhar, Remitha Rathi L, Rithya P, Nivrutha P

rajisridhar@gmail.com, remitharathi@gmail.com, rithya3@gmail.com, nivmeen@gmail.com

Department of Computer Science and Engineering, Anna University, Chennai

ABSTRACT

This work attempts to correct the errors that are typically encountered in an OCR system. This work aims at correcting the typical errors that are produced by Tamil OCR which includes errors in 'ottru', 'sandhi', and misinterpretation of characters, not recognition of characters and failure of marking the end of the sentence. The system uses Morphological Analyser to find the incorrect words in the document. Tamil grammar rules are used to correct Ottru and Sandhi errors. A bigram Probabilistic model at the word level is constructed to validate the correct sequence of text. The tagger is used to tag the words which aids in the process of identifying the end of the sentence. After correcting the errors the modified OCR system had a word accuracy of 91.75%.

Keywords

Automatic error correction in OCR, Post Processing of Tamil OCR, Tamil grammar rules, Optical Character recognition (OCR), Tamil Documents.

1. Introduction

In this modern computer era, there is a need for information to be available in digitized format [1]. But most of the business data and historical literature are available only on paper which is a hurdle for manipulating and extracting information from this data [2]. Business data involves numbers and in the current scenario, business analytics tools are available to help in manipulating and analysing this data. Data from Literature is one of the critical areas where it still remains on paper. This poses a great challenge for the Scholars to gain knowledge about a particular literature [3].

Literature in languages like English, Hindi, Sanskrit, Tamil, etc is varied and rich. In this paper, we discuss about the attempt that we have made to convert Tamil literature by means of OCR and make it web searchable. The reason behind choosing Tamil is due to the rich literature and being one of the oldest languages [4]. In addition, Tamil is also a morphologically rich language and is parent to many new languages. Therefore one simple way to process and make Tamil documents available on the Internet is by means of performing optical character recognition (OCR). This process is however very difficult to achieve high accuracy. The difficulty arises in multiple scenarios, namely a wide set of characters and close resemblance of some characters. In addition to the character level difficulty, the partial free word order nature of the language [5] makes it difficult at the sentence level to generate documents from OCR error free. The converted document typically contains a lot of errors arising from the above said challenges of the language. Manual correction was the only way to correct the errors but it is a cumbersome task for huge documents. Therefore, we developed a system which automatically corrects the errors in the OCR'd document.

1.1 Literature Survey

OCR error correction using Morphological parsing [6] was performed for Bangla, an Indian language, to separate the word as root and suffix. The technique is based on morphological parsing where two separate lexicons of root words and suffixes are used. The algorithm detects the candidate root-suffix pairs of each input string, validates their grammatical agreement and the root/suffix part in which the error has occurred. The correction is made to the error part of the input string by means of a fast dictionary access technique. To do so, the information about the error patterns generated by the OCR system is examined, and some alternative strings are generated to fix the erroneous word. Among the alternative strings, those satisfying grammatical agreement in root and suffix are finally chosen as suggested words. The authors claim that their error correction module for Bangla language can correct 84.22% of single character errors. This algorithm however was not able to correct multiple character errors in a single work.

In a work for post processing in OCR [7] for English, the algorithm focuses on three types of errors. Non- word error, isolated word error and error related to context. Non word error is due to the similar feature misinterpretation of characters. This is corrected by maintaining N-Gram table for letters and also by dictionary lookup technique. Isolated word errors are

due to typographical error which is corrected by analysing the common typographical errors by maintaining confusion matrix and by minimum edit distance technique. Errors related to context are corrected by statistical language model based on word bigram probabilities. This work reported an average correction rate between 74% and 88%.

Another work [8] focuses on correcting the errors in English based on the context. The POS Tagger is used to tag the word which is used to correct the real word error. The assumption the authors have made is, if two or more words have the same POS tag then there could be ambiguity. The authors have developed a bi-gram at the word level to correct the errors encountered by the OCR. The context information found by the system using bi-gram model was not sufficient to correct the errors related to context since it is based only on POS information while other information about the words have not been used. The post processing work for English thus focussed on the fixed word order of the language and cannot be adapted directly for Tamil.

We were motivated by the work done by the authors Karen Kukich [6], Xiang Tong and David A Evans [7] and designed a bigram model to correct errors at the word level to correct root, their suffixes or both. But due to the inherent nature of the Tamil language other errors like 'Otru' (Phonetic stress given to a particular consonant), 'Sandhi' (Phonetic stress based on context for a consonant between words) are also part of the OCR'd document. These errors are corrected by designing a rule based technique specifically for Tamil by designing algorithms using Tamil grammar rules. This was motivated from the work done for Bangla [6], however, the work done for Bangla cannot be used directly as they have designed algorithm to correct single character errors. In addition, the language characteristics of Bangla and Tamil are not the same. Hence we incorporated Tamil grammar rules and designed algorithms to correct multiple character error in a word. We designed algorithms that convert the grammar rules of Tamil to their corresponding representations for correcting 'Otru' and 'Sandhi' errors. Sentence delimiter was also identified and corrected by using Tamil grammar rules.

In section 2 architecture of the system is explained. In section 3 the results are analysed. In section 4 performance measures are discussed. In the final section conclusion and future works are discussed.

2. Materials and Methods

The overall system architecture is shown in Figure 1. The system aims at correcting the errors produced by Tamil OCR given the scanned Tamil document which is first of a kind in this area. The first step is to correct the ottru error in the document using the ottru rules in Tamil Grammar [9]. Then the Ottru corrected input is given to the Morphological Analyser. Analyser [10] separates the root from the word and looks up the dictionary for validation. If the root word is incorrect, using modified form of Levenshtein algorithm [11] the root word matcher, lists the nearest matching words. Then the words are validated using bigram probabilistic model. On the other hand if the analyser is not able to separate the suffix it represents that the error is in the suffix. So the word is passed to the Prefix Matcher. It starts finding the root in the prefix part of the word. The remaining part is given to the Suffix Corrector. It generates the correct suffix using the automata of formation of suffix. Morphological Generator [12] forms the whole word with its root and suffixes using inflectional rules. The final step involves removal of the sandhi error using rules [13] and marking the sentence delimiter for the corrected document.

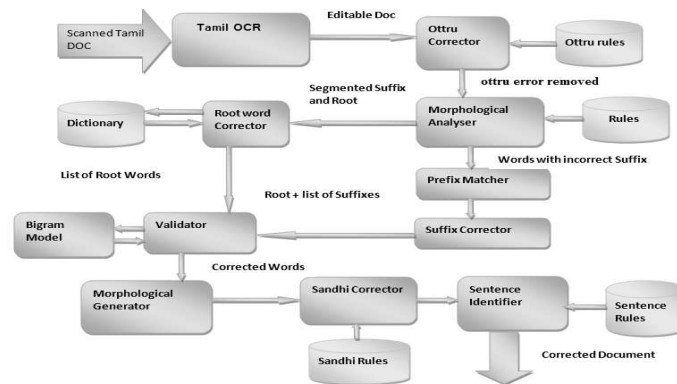


Figure 1: Architecture of the Error Correction algorithm

2.1 Contributions in the proposed system

Ottru Corrector – Corrects Ottru error based on the formulated rules.

Morphological Analyser- Designed to split the suffix and root from the word.

Root Word Corrector - It suggests the nearest matching root word for a given incorrect word

Prefix Matcher- It lists the word that matches with the prefix of a given word.

Suffix Corrector - It corrects incorrect suffixes by validating against automata.

Bi-gram model – A Word level model is constructed using the words from “Dhinamani” news paper.

Sandhi Corrector- Corrects Sandhi error based on the formulated rules.

Each of the modules is described below.

2.2 Ottru Corrector

The first step in the error correction algorithm is the ottru error corrector. This is performed by analyzing the occurrences of consonant in a Tamil word and formulating rules accordingly to correct the Ottru error in a given document. The three types of consonant clustering in Tamil language are given below.

- **Iirottru mayakkam**

A consonant (Meiezhuththu) can be followed by different consonant not by vowel-consonant combination (Uyir Meiezhuththu). After these three consonants \dot{y} (y), \dot{r} (r), \dot{z} (zh) the following consonants \dot{k} (kk), \dot{ng} (ng), \dot{ch} (ch), \dot{nga} (nga), \dot{th} (th), \dot{nn} (nn), \dot{pp} (pp) \dot{m} (m) can occur.

Example: வாம்ப்பு (vaaypu), தீர்ப்பு (theerppu), வாழ்க்கை (vazhkkai)

- **Udannilai meimayakkam**

A consonant can be followed by its corresponding vowel-consonant combination (Uyir Meiezhuththu). Except \dot{r} (rr), \dot{z} (zh) all the consonants falls under udannilai meimayakkam

Example: மக்கள் (makkal)

- **Vettrunilai meymakkam**

A consonant can be followed by a possible vowel-consonant.

Example: தங்கம் (thangam)

After \dot{ng} (ng) the vowel-consonant combination of \dot{k} (kk) only occurs. No other vowel-consonant combination occurs after \dot{ng} (ng)

Based on the above specified types we have formulated the generalised Ottru rules. Some of the rules are

Rule 1: In Tamil, letters ர-ந (ra-na) do not occur consecutively

Rule 2: In Tamil, letters ந-து (na-thu) do not occur consecutively

Rule 3: The same consonants cannot repeat consecutively.

Rule 4: In Tamil, the ங-க (nga - ka) cannot occur consecutively.

These rules were represented in the database for use by the correction algorithm.

2.3 Morphological Analyser

Table 1: Inflectional rules for obtaining root from whole word.

Input String	Root + suffix	Combine Character	Inflectional Rules
மரங்கள் (marangal)	மரம்+கள் (maram + kal)	ங் (ng)	Rule 5: ங் (ng) is to be removed and ம் is to be added to identify the root word after discarding the suffix கள் (kal)
கையால் (kaiyaal)	கை+ஆல் (kai + aal)	ய் (yy)	Rule 6: After removing ஆல் (aal), ய் (yy) should be deleted
படிக்கிறான் (padikkiraan)	படி+கிறு+ஆன் (padi + kiru + aan)	க் (kk)	Rule 7: After ஆன் (aan) is removed உ் (vu) is added. Then க் (kk) is deleted after discarding கிறு (kiru)
வருகிறான் (varukiraan)	வா+கிறு+ஆன் (vaa + kiru + aan)	ரு (ru)	Rule 8: After ஆன் (aan), கிறு (kiru) are removed. வரு (varu) should be modified as வா (vaa)

The ‘Ottru’ corrected input is given to the Morphological Analyser and is an editable document. We have designed the root word identification by our new algorithm which is based on discarding suffixes. In our approach, the word is parsed from right to left to delete suffixes, by initially identifying the smallest suffix which is appended to identify the longest suffix. The

process of incrementing the suffix and discarding them is continued to identify the proper root. This poses a problem due to the inherent nature of Tamil language wherein, the root +suffix, introduces an additional character or characters to combine them. Therefore this combining character is required to identify the root. So, rules have been modified in our algorithm, such that these combining characters are retained and using these characters along with the prefix, the root words are identified. The process of identifying the root words are based on the inflectional rules with the assumption that the suffixes are identified correctly. These rules are obtained by referring to Tamil Grammar. Table 1 shows some probable combinations of how the root word is obtained from the whole word after removing the suffix/suffixes using these inflectional rules. In the process of determining the root, the suffix that was discarded, is stored separately to be used at a later stage for generating the document after correcting errors. In addition to identifying the root word, we use a look up and tag the root words as noun, verb, adjective and adverb. Thus the output of this module is the tagged root word and the suffixes in that word.

2.4 Root Word Corrector

After determining the root, there may be errors that occur in the root word. This error correction is handled by the root word corrector module by assuming the suffixes in the word are correct. In Tamil grammar, there is a set of suffixes that are possible depending upon the POS information of the root.

From Table 2, we can conclude that depending on the suffix the correction at the root word level can be limited to a smaller domain. To carry out this we used the nearest matching strategy using Levenshtein algorithm. Levenshtein algorithms list the probable words having the same POS that matches an incorrect root. Using our constructed word level Bi-Gram Probabilistic model the suggested word is validated and the correct word is chosen based on context.

Table 2: Noun and Verb Suffixes

Noun Suffix	KaL, ukku, aana, aaga, il, aal, Utan, athu, mel, kiizh, poola, paRRi, kuRittu, paarttu, taaNTi, tavirttu, tavira, ozhiya, oTTi, koNTu, vaittu, viTTu, maatiri, vita, patilaaka
Verb Suffix	Padu, AthaRku, mal, th, t, R, in, kir, kinRu, p, v, um, ull, Een, oom, aay, aan, aaL, aar, aarlaL, varkaL, van, var, vaL, illa

2.5 Prefix Matcher

A typical word can have error at the prefix or the suffix or both. In this work, we assume that if the root word is not being identified by the Morphological Analyser then there is error in the suffix. A Prefix Matcher algorithm is used to identify the root word in the event of error in suffixes. It starts parsing the string from the beginning and finds the longest matching root word by a look up and separates the word into root and suffix. The root is processed by the root word corrector as explained earlier and the suffix other than the root is given to the Suffix Corrector.

2.6 Suffix Corrector

The suffixes of the words that are processed by the prefix correction are given as input to the module. We constructed automata which represent suffixes from right to left based on Tamil grammar rules and used this automata to validate and correct the suffix. The idea behind constructing automata for suffix correction is based on the principle that Tamil words support multiple suffixes for a given prefix.

Example 1:

மரங்கள் (marangal) → மரம்+கள் (maram + kal)

மரத்தினுள்ளே (marathinulle) → மரம்+இன்+உள்ளே (maram + in + ulle)

மரங்களுக்கு (marangaluku) → மரம்+கள்+உக்கு (maram + kal + ukku)

In example 1 for a prefix மரம் (maram), suffixes like கள் (kal), இன் (in), உள்ளே (ulle), உக்கு (ukku) are possible that can be appended with root to form a valid noun.

Example 2:

செய்து (seidhu) → செய்+த்+உ (sei + th + vu)

செய்தான் (seithaan) → செய்+த்+ஆன் (sei + th + aan)

செய்யாமல் (seyyaamal) → செய்+ஆ+மல் (sei + aa + mal)

In example 2 for a prefix செய், suffixes like உ (vu), ஆன் (aan), ஆமல் (aamal) are possible that can be appended with root to form a valid verb. Figure 2 and Figure 3 shows the possible path of suffixes that can combine with the root word to form a valid word of noun and verb respectively which we have derived by referring to Tamil grammar rules.

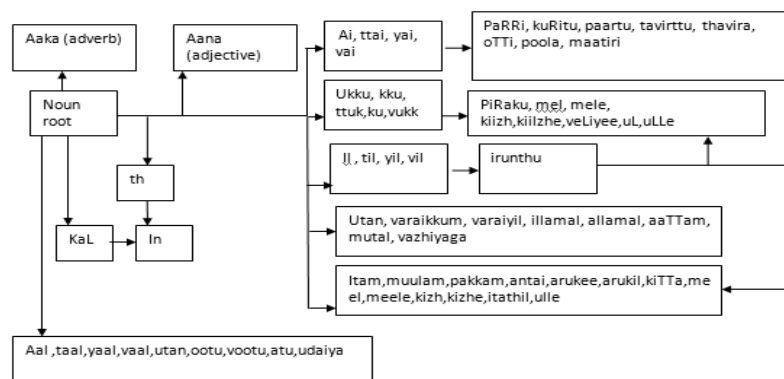


Figure 2: The flow of suffixes that can be appended with the root word of Noun

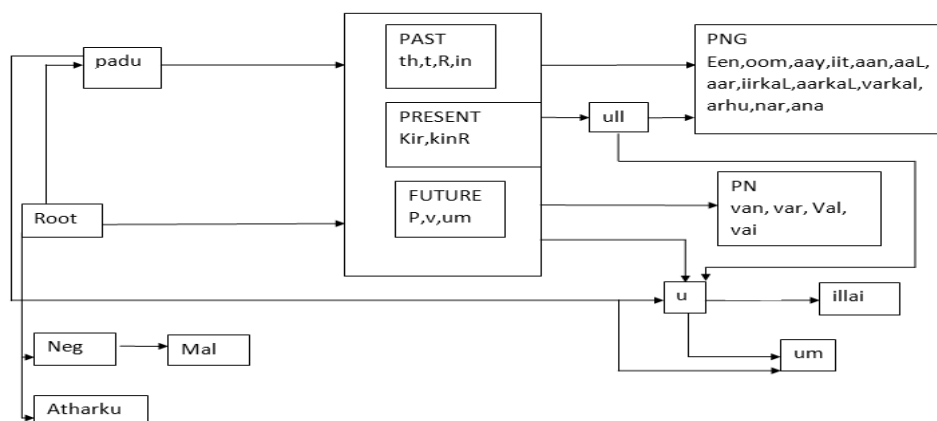


Figure 3: The flow of suffixes that can be appended with the root word of Verb

This flow of suffixes with root word is used to construct the automata for validating the suffixes that are suggested by the Suffix Corrector. Since we are correcting suffixes the automata is constructed from the last suffix to the first suffix. The suffix represents states of the automata and transition represents the presence of a previous suffix given a valid current suffix. We start analysing the suffix from the last and when the suffix nearly matches to the state in the automata we proceeded to the next state in the path and this process goes on. If the path is terminated with final state then we conclude that the list of suffixes found through the path are valid and is used to form the complete word after correcting the root.

2.7 Bi-gram Model

The algorithm for root corrector gives a list of probable words and the Suffix Corrector has corrected a given suffix. To identify the correct root, from the list of available root word, context information is essential. This context information is created by constructing a Bigram model for the root words. So, from the sequence of list of probable words, the bi-gram probability of the Bi-gram model is used to choose the adjacent words.

2.8 Morphological Generator

After determining valid root words and their correct or corrected suffixes, the Morphological Generator is designed to combine the root with the suffix to form the actual word. Inflectional rules of the Tamil language are used to generate the words. During the process of combination, as explained in Table 1, combine character or characters is used for generating the words.

2.9 Sandhi Corrector

Using sandhi rules in Tamil, sandhi error is corrected. Some of the rules that we have incorporated in this work for error correction are listed below.

Rule 9: வன்தொடர் குற்றியலுகரத்திற்குப் பின் வலி மிகும் (van thodar kutriyalugarathirku pin vali migum).

Rule 10: இரண்டாம் வேற்றுமை விரி பின் வலி மிகும் (irandaam vetrumai viri pin vali migum).

Rule 11: நான்காம் வேற்றுமை விரி பின் வலி மிகும் (naangaam vetrumai viri pin vali migum).

2.10 Sentence Identifier

In Tamil, the sentences usually end with pattern Verb+...+PNG (Person Number Gender). Sentence Identifier identifies Verb+...+PNG tag in the word and finds the end of the sentences.

3. Result Analysis

To test our system for error correction we attempted the output of Ponvizhi Tamil OCR [14]. This requires training of fonts in the input documents. The font file with the trained characters is loaded for recognising. We have tested 200 documents from Dhinamani newspaper with an average of 70 words for each document. The output of each module is described below.

3.1 Ottru error Correction

Table 3 shows the words corrected by the Ottru corrector using the rules in section 2.2.

Table 3: Result of Ottru corrector

Input String	Output String	Description
ஊரந்து (oorandhu)	ஊர்ந்து (oorndhu)	Corrected using Rule1 and Rule 2 of section 2.2
மககள் (makakal)	மக்கள் (makkal)	Corrected using Rule 3 of section 2.2
சிங்கம் (singagam)	சிங்கம் (singam)	Corrected using rule 4 of section 2.2

3.2 Morphological Analyser

Morphological Analysis is done for nouns and verbs. There are 28 paradigms for nouns and 32 paradigms for verbs. We have about 69 suffixes for nouns and 48 suffixes for verbs. Each suffix has different inflectional rule to join with different paradigms. The results of Morphological Analyser are showed in Table 4.

Table 4: Results of Morphological Analyser

Input String	Root + Suffix	Inflectional rules
பழங்கள் (pazhangal)	பழம்+கள் (pazham + kal)	Rule 5 of section 2.3
கணினியால் (kaniniyaal)	கணினி+ஆல் (kanini + aal)	Rule 6 of section 2.3
படிக்கிறான் (padikkiraan)	படி+கிறு+ஆன் (padi + kiru + aan)	Rule 7 of section 2.3
வருகிறான் (varukiraan)	வா+கிறு+ஆன் (vaa + kiru + aan)	Rule 8 of section 2.3

3.3 Prefix Matcher and Suffix Corrector

Prefix Matching and Suffix Correction is analysed as follows

Example 1:

ஆளுகைக்கன (aalugaikkana) → ஆளுகைக்கான (aalugaikkaAna)

The word ஆளுகைக்கன contains error in suffix part so it is passed to Prefix Matcher by Morphological Analyser. ஆளுகை is found by Prefix Matching and க்கன is given to Suffix Corrector which identifies ஆன as valid suffix by reaching the final state of the automata.

Example 2:

அலுவலககளின (aluvalakakalina) → அலுவலகங்களின் (aluvalkangalin)

Similarly அலுவலகங்களின் (aluvakangalin) is passed to the Prefix Matcher. அலுவலகம் is found by Prefix Matcher and - களின் (kalian) is passed to the Suffix Corrector.

Suffix Corrector takes the word and processed from right to left of the word. Since it follows the automata in figure 4 இன் (in), கள் (kal) are found as suffixes and these suffixes and root word are given to the generator for constructing the whole word.

Example: களின்



Figure 4: Automata for validating suffix

3.4 Root Word Corrector

Our system also corrects the errors at word level in the categories specified in Table 5 using Root Word Corrector.

Types of Error	Incorrect words	Correct words	Description
1. Similar Feature character errors	அரக (araga)	அரசு (arasu)	The incorrect characters are corrected by similar feature analysis at root word corrector
	லிண்ணப்பம் (linnappam)	விண்ணப்பம் (vinnappam)	
2. Error in whole words	இ—சம் (I – sam)	இலவசம் (ilavasam)	This is corrected by root word corrector
	வ—ங்கு (va – ngu)	வழங்கு (vazhangu)	Root word corrector suggests possible words like வழங்கு, வணங்கு and bigram validates as வழங்கு by seeing the context
3. Partially Recognised Character Errors	-தாடர்ச்சி (-thadarchi)	தொடர்ச்சி (thodarchi)	All these words are corrected by root word corrector
	பயன்ப-டு (payanpa – du)	பயன்பாடு (payanpaadu)	
4. Missing character	ங்களிப்புடன் (ingalippudan)	பங்களிப்புடன் (pangalippudan)	
5. Interpretation of additional character for single one	வினை—ட்டு (vilai – ittu)	வினையாட்டு (vilaiyattu)	

Table 5: Types of Error occurred in Root word with examples

3.5 Sandhi error correction

Example 1: மக்களுக்குத் தேவைபடும் (makkalukkutha thevaipadum)

Using Rule 9 in section 2.10, it is corrected as மக்களுக்குத் தேவைபடும் (makkalukkuth thevaipadum)

Example 2: அவனைக் காப்பாற்று (avanaika kaappaatru)

Using Rule 10 in section 2.10, it is corrected as அவனைக் காப்பாற்று (avanaikk kaappaatru)

3.6 Sentence Identifier

Example 1: உயரிய விருது வழங்க முதல்வர் உத்தரவிட்டுள்ளார்.

Here verb+ull+PNG (aar) pattern occur in உத்தரவிட்டுள்ளார் (uththaravittullaar) so it is marked as end of the sentence.

Example 2: நிலக்கரியை நீர்வழிப் போக்குவரத்து மூலம் எடுத்துச் செல்ல

ஒப்புக் கொண்டுள்ளது.

In Example 2 கொண்டுள்ளது (kondullathu) has verb+tense+ull+PNG (athu) so it is marked as the end of the sentence.

4. Performance Measures

This section details on testing procedures done over the completed modules in order to ensure the correctness of the implementation. We tested 200 documents from Dhinamani newspaper and observed the following measures.

The Average Error Rate in 200 documents is calculated as

Average error rate in the documents = Total No. of errors in documents / Total no. of documents (1)

Similarly the average correction rate is calculated as

Average Correction Rate = Total No. of Errors Corrected / Total No. of Errors occurred (2)

Finally a graph is drawn with error and correction rate of various types of error using the value calculated by the relationship given in Equation (1) and (2).

The manually computed correction rate for Sentence Identification is found to be 95.7% while the Average Correction Rate of a document was computed to be 91.75%.

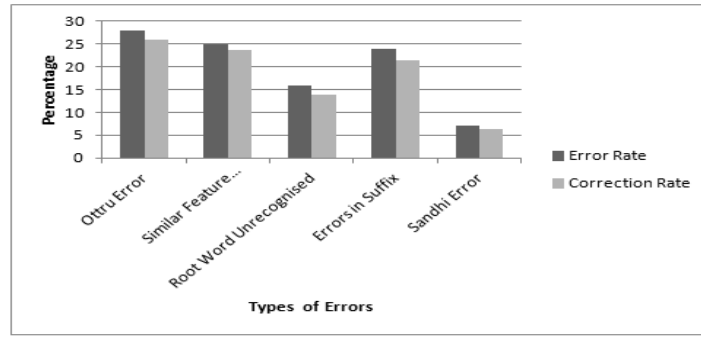


Figure 5: Graph showing types of error with its error and correction rate.

5. Conclusion and Future Work

Thus our system was successful in reducing the maximum possible errors in the document. But still there are some errors which are ambiguous. This can be improved by using Trigram Approach instead of Bigram which incorporated more contextual information. Suffix Correction can be enhanced by rules or by learning which helps in finding the appropriate suffix related to the context. Instead of our approach of look-up to tag root words, learning algorithms could be incorporated for the same.

References

- [1] Karen Coyle, "Mass Digitization of Books", *the Journal of Academic Librarianship*, Vol. 32, Issue 6, pp. 641-645, 2008.
- [2] Umesh Kumar and Jagdish Lal Raheja, "Document Presentation Engine for Indian OCR: A Document Layout Analysis technique", *International Journal of Recent trends in Engineering and technology*, Vol.3, no.3, pp.182-186, 2010.
- [3] John Burns and kevins Hawkins, "Providing next Generation Tools For Scholars: Jstor's Advanced Technology Research Group", *International Journal of Humanities and Arts Computing*, vol.4, pp.141-149, 2010.
- [4] Tholkappiyar, Tholkappiyam, Project Madurai, 2001.
- [5] T.Dhanabalan and T.V.Geetha, "Unl Deconverter for Tamil" *International Conference on the Convergence of Knowledge, Culture, Language and Information Technologies*, pp. 1-6, 2003.

- [6] U. Pal, P. K. Kundu AND B. B. Chaduri “OCR Error Correction of an Inflectional Indian Language Using Morphological Parsing” *Journal of information science and engineering*, vol. 16, pp.903-922, 2000
- [7] Karen Kukich “Techniques for automatically correcting words in Text”, *ACM Computer Surveys*, vol.24 no.4, pp.377-439, 1992.
- [8] Xiang Tong and David A Evans, ”Statistical Approach to Automatic OCR Error Correction in Context”, *Fourth workshop on very Large Corpora*, pp.88–100, 1996.
- [9]<http://www.tamilvu.org/otru/c0211603.htm>
- [10] Vijay Sundar Ram R, Menaka S and Sobha Lalitha Devi, ”Tamil Morphological Analyser”, *3rd National Conference on Recent Advances and Future Trends in IT* Punjabi University, Patiala, Punjab, pp.1-18, 2009.
- [11]<http://en.wikipedia.org/wiki/Levenshteindistance>.
- [12] Anand Kumar M, Dhanalakshmi V V, Rekha R U, Soman K.P, Rajendran S, ” A Novel Data Driven Algorithm for Tamil Morphological Generator”, *International Journal of Computer Applications*, vol. 6, no.12, pp.52-56, 2010.
- [13]<http://www.tnpsctamil.in/2012/01/1.html>
- [14] Dr V. Krishnamoorthy, “OCR Software for Printed Tamil Text”, *Tamil Internet*, pp.99-101, 2002.

Identification of Tamizh script on Tablet PC

Bhargava Urala and A G Ramakrishnan

Department of Electrical Engineering, Indian Institute of Science, Bangalore, India.

Abstract

Research on Tamil script has mostly dealt with the identification of isolated letters. Different classifiers such as, hidden Markov models (HMM), neural networks and support vector machines (SVM) have been used along with a selection of features such as Fourier, wavelet, angular and directional features. The challenge of Tamil word recognition that has been addressed in the literature so far can be divided broadly into two approaches: (i) Recognition of individual strokes and then concatenating them using appropriate models to detect a word or a compound character. (ii) Grouping of strokes to form symbols or stroke groups, which may or may not constitute a compound character and recognition of these symbols to form words. The second method needs a framework for efficient segmentation of strokes into stroke groups and post processing techniques based on statistical language models. Use of word lists have also been investigated to enhance Tamil word recognition rates.

We portray experiments with preprocessed (x, y) coordinates, Fourier and derivative features for recognizing individual Tamil characters and the integration of segmentation, feature extraction, classification and post-processing steps into a dynamic link library for use on a Tablet PC.

Preprocessing

Each stroke group is smoothed with a Gaussian kernel, normalised to a 1 X 1 box and resampled to contain a fixed number (64) of points. This preprocessing removes jitter and accounts for variations in size and speed. An example of a preprocessed stroke group is shown in Figure 1 (a).

Deriving Features

We have studied the use of local and global features extracted from the preprocessed data for training the classifier. As the names suggest, local features capture minor variations occurring in the data in small neighbourhoods and global features capture the overall shape of the character. We use x, y coordinates of normalized symbols and their first derivative as the local features and truncated Fourier coefficients as the global features.

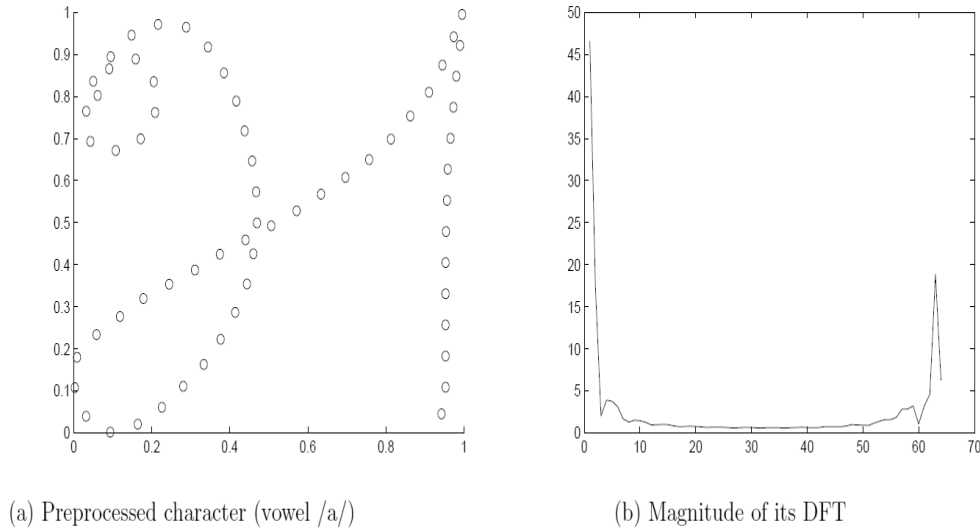


Figure 1: A sample preprocessed character and the magnitude of its DFT

It is well known that in the case of discrete Fourier transform (DFT), most of the energy in the signal is often contained in very few coefficients. Therefore, we also experimented with truncating the Fourier transform to less number of coefficients. The features were extracted from the samples of the training set of IWFHR isolated Tamil symbol database [1] and validated

on the test set of the same database. We list below the features (local, global and combined local and global features) that we experimented with.

1. Preprocessed (x, y) coordinates: Each normalized symbol consists of 64 (x, y) points. Therefore, the length of this feature vector is $64 \times 2 = 128$.
2. Fourier descriptor: In this case, the preprocessed coordinate vector is treated as a vector of 64 complex points $z = x + jy$ [2] and its DFT is taken, which results in a 64-point complex valued vector. Therefore, the length of feature vector is $64 \times 2 = 128$.
3. Truncated DFT: In this case, the Fourier coefficients obtained from DFT are truncated to 32 complex points. We experimented with lengths of 8, 16 and 32. The reconstructions from 8, 16 and 32 complex points can be seen in Figs. 2 (a), 2 (b) and 2 (c), respectively. It can be seen that 32 complex points give an acceptable reconstruction. Hence, the chosen feature vector length is $32 \times 2 = 64$.
4. Combination of preprocessed (x, y) coordinates and truncated DFT features - Feature vector length = $128 + 64 = 192$.
5. Concatenation of the sequence of preprocessed (x, y) coordinates, truncated DFT coefficients and the sequence of first derivative features, Accordingly, the feature vector length is $128 + 64 + 128 = 320$.

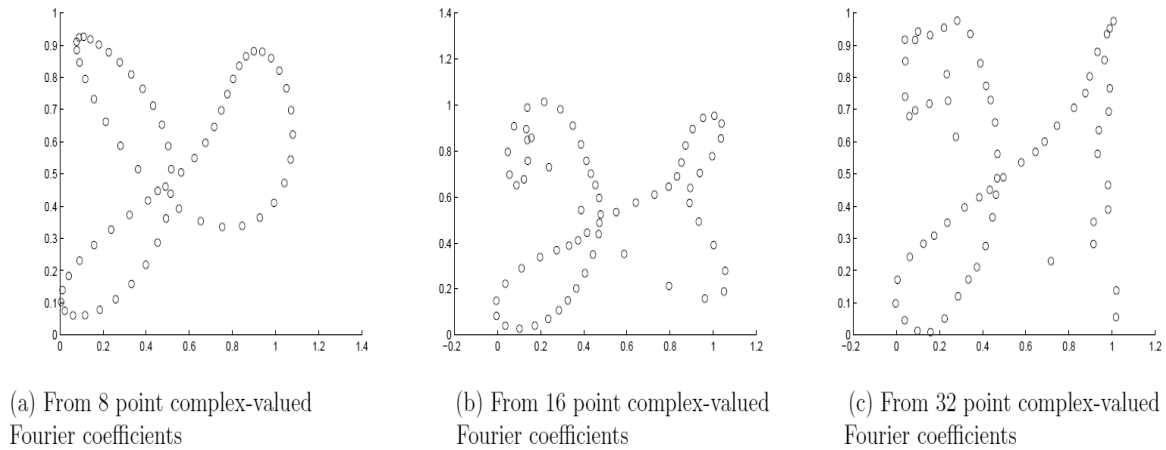


Figure 2: Reconstruction of the character from the truncated Fourier transform coefficients

Segmentation framework

Segmentation of handwritten strokes into Tamil symbols consists of two steps. The initial segmentation is based on the horizontal overlap between bounding boxes of consecutive strokes. The second step is called attention feedback [3] and involves class labels and SVM confidence values obtained from the classifier stage as well as certain pen and stroke displacement statistics measured from a large Tamil handwritten character database to improve the results of initial segmentation.

SVM Classifier

We use a support vector machine (SVM) with radial basis function (RBF) as kernel trained on the features extracted from the training database, which consists of a large number of samples of 155 unique Tamil symbols. The SVM gives a class label and an associated confidence level (between 0 and 1), both of which are used for improved word recognition.

Post processing

Statistics of co-occurrence of Tamil symbols estimated from the Emille corpus of Tamil text along with SVM confidence levels are used to generate N-best choices of symbol strings for a given handwritten word [4], using a Viterbi lattice. The symbol string is then converted to a Unicode string using a finite state transducer.

Application and usage on Windows Tablet PC

The Windows Tablet PC has a touch screen and a pen interface, which can be used to write and record handwritten data. An application developed by GIST, CDAC Pune serves as the front end, whereas dynamic link libraries (.dll) developed by us in C++ with functions and blocks to achieve the tasks described in the previous sections, act as the recognition engine. Along with the recognition of Tamil words, provision for recognition of Indo-Arabic numerals has been made with the same engine as well. A snapshot of the application in use is shown in Fig. 3. The engine has a high recognition rate exceeding ninety percent and a speedy performance with the average recognition time for a single symbol or stroke group being about 45 msec. Since the application is meant for practical use, a heuristics-based algorithm has also been developed and implemented with the engine, which detects and corrects the delayed strokes and also detects and removes the overwritten strokes. The online handwritten word recognition engine has been integrated with multiple form-filling applications.

Acknowledgments

The sample application we used to illustrate the recognition performance of our Tamil engine has been developed by CDAC, Pune as part of a consortium project funded by Technology Development for Indian Languages (TDIL), Department of Information Technology, Government of India. We thank CDAC, Pune and TDIL for the same. We also thank Dr. Suresh Sundaram, Assistant Professor, IIT Guwahati for his suggestions on this work.

References

1. Madhvanath, S and ucas, S M. IWFHR 2006 online Tamil handwritten character recognition competition. Proc. Intl Conf. Frontiers in Handwriting Recognition.
2. J. Rajkumar, K. Mariraja, K. Kanakapriya, S. Nishanthini, and V. S. Chakravarthy. Two schemas for online character recognition of Telugu script based on Support Vector Machines. In Proc. ICFHR, pp. 565–570, 2012.
3. Suresh Sundaram and A. G. Ramakrishnan, “Attention feedback based robust segmentation of online handwritten words,” Indian Patent Office Reference. No: 03974/CHE/2010.
4. Suresh Sundaram, Bhargava Urala and A. G. Ramakrishnan, “Language models for online handwritten Tamil word recognition,” Proc. Workshop on Document Analysis and Recognition (DAR 2012), 16 December 2012, IIT Bombay, Mumbai, India

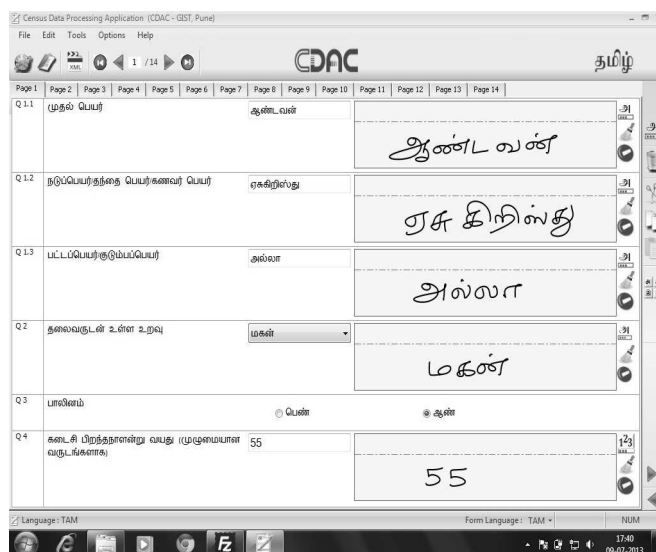


Fig. 3. A snapshot showing the recognized outputs of our Tamil handwriting recognition engine.

Recognition System for Tamil Sign Language Using Hand Gestures

Vijay Jeyakumar^[1], Dr.K.Bommanna Raja^[2]

Assistant Professor^[1], Professor & Head^[2]

Department of Biomedical Engineering^[1]^[2],

PSNA College of Engineering and Technology, Dindigul, India

vijay@psnacet.edu.in, dr.k.bommnnaraja@gmail.com

ABSTRACT

Hand gesture recognition system can be used for interfacing between computer and human using hand gesture. This work presents a technique for a human computer interface through hand gesture recognition that is able to recognize static gestures from the Tamil Sign Language hand alphabet. The first step of the method is to read all stored templates. Then get the real video from the webcam and convert it into number of frames. Convert the RGB frame to HSC frame and set the threshold at 0.5 by Otsu thresholding method. Compare each pixel value in each frame with this threshold and segment the sign from the frame. Compare it with the stored template and the corresponding text will be displayed. This recognition system aids the deaf and dumb people to communicate with the others who does not know about their sign languages.

Keywords: Gestures, Sign language, Tamil

1. INTRODUCTION

The idea is to make computers understand human language and develop a user friendly Human Computer Interfaces (HCI). Making a computer understand speech, facial expressions and human gestures are some steps towards it. Gestures are the non-verbally exchanged information. A person can perform innumerable gestures at a time. Since human gestures are perceived through vision, it is a subject of great interest for computer vision researchers. This paper aims to determine human gestures by creating an HCI. Coding of these gestures into machine language demands a complex programming algorithm. An overview of gesture recognition system is given to gain knowledge.

2. RECOGNITION METHODS OF FOREIGN LANGUAGES

Research has been limited to small scale systems able of recognizing a minimal subset of a full sign language. Andrew [1] developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power glove [3]. The most successful commercially available glove is by far the VPL Data Glove.

It was developed by Chan Wah[4] during the year 2002. It is based upon patented optical fiber sensors along the back of the fingers. Star-ner and Pentland developed a glove-environment system capable of recognizing 40 signs from the American Sign Language (ASL) with a rate of 5Hz. Hyeon-Kyu Lee and Jin H. Kim presented work on real-time hand-gesture recognition using HMM (Hidden Markov Model). Kjeldsen and Kendersi devised a technique for doing skin-tone segmentation in HSV human interfaces, in this method, the hand regions are extracted from multiple images obtained by a multiviewpoint camera system, and constructing the “voxel Model.” Hand pose is estimated. Chan Wah Ng, Surendra Ranganath presented a hand gesture recognition system, they used image furrier descriptor as their prime feature and classified with the help of RBF network. Their system’s overall performance was 90.9%. Claudia Nölker and Helge Ritter [16] presented a hand gesture recognition modal based on recognition of finger tips, in their approach they find full identification of all finger joint angles and based on that a 3D modal of hand is prepared and using neural network.

The latest research work in this paper is undergoing in the guidance of Riza Atiq. The method used was Canny edge detection method, but the drawback was both hands cannot be used. The signer can use only right hand. Xiaole bai, focused on Pakistan sign language by using the method Boltay Haath system but the limitation is it could not recognize abduction between fingers.

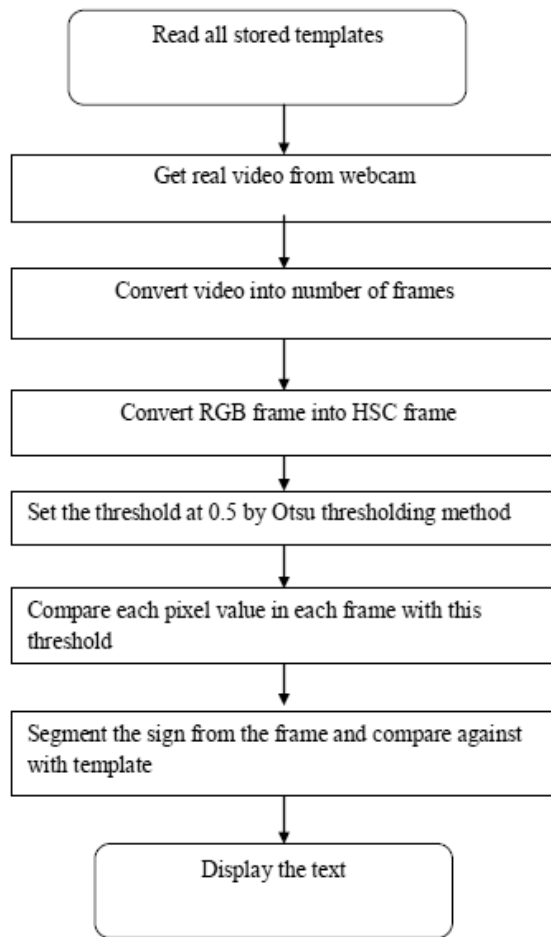


Fig 1 Overview of System Architecture

3. MATERIALS AND METHODS

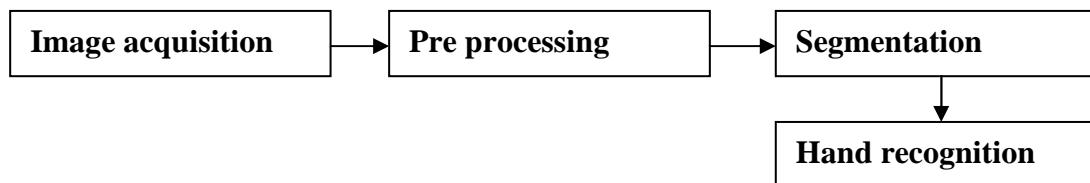


Fig 2 Proposed methodology

3.1 Image Acquisition

For the purpose of training the palm images are captured. The run time images are captured using web camera. The images are captured in a high intensity environment directed to illuminate the image source which is held at black background so as to avoid shadow effects. The images are captured at a specified distance (typically 1.5 to 2 ft) between camera and signer. The distance is adjusted by the signer to get the required image clarity. In order to avoid the images of arm extension from palm, it is better to use a black wrist band on the signers arm.

3.2 Pre Processing

Preprocessing is very much required task to be done in hand gesture recognition system. Preprocessing is applied to images before segmentation. Preprocessing consists of Morphological filtering, which includes dilation, erosion, opening and closing. This step also includes the conversion of videos into number of frames.

3.3 Segmentation

A very good segmentation is needed to select a adequate threshold to extract hand from background .i.e. there is no part of hand should have background and background also shouldn't have any part of hand. In general, the selection of an appropriate segmentation algorithm depends largely on the type of images and the application areas. The Otsu segmentation algorithm was tested and found to give good segmentation results for the hand gestures and was, therefore, selected .Otsu algorithm is nonparametric and unsupervised method.

3.4 Hand Recognition

In this step we should compare each pixel value in each frame with its threshold the sign from each frame is compared with the template and corresponding text will display.

4. SYSTEM METHODOLOGY

This method had been developed by T.Shanableh for recognizing isolated Arabic sign language gestures in a user independent mode. We modified it for Tamil sign language. For our work, we have taken the srilankan sign language to build up the system. In this method the signers wore gloves to simplify the process of segmenting out the hands of the signer via color segmentation. Many Researchers utilized special devices to recognize sign language. Hand shape and motion are extracted easily and accurately using these devices. However they are very expensive and will reduce the naturalness of sign language communication. The most important step is the conversion of RGB image to HSC image. An RGB image, sometimes referred to as a "true color " image, is stored in MATLAB as an m-by-n-by-3 data array that defines red, green and blue color components for each individual pixel. RGB images do not use a palette. If we take close look to the image after converting to HSC image, we find that background may have some 1s which is known as background noise and hand gestures have some 0s that is known is gesture noise. These errors can lead to a problem in segmentation of hand gesture so we need to remove these errors. A morphological filtering [4] approach has been applied using sequence of dilation and erosion to obtain a smooth, closed, and complete contour of a gesture. Segmentation in our proposed hand gesture recognition system is done by Otsu algorithm. The algorithm treats the segmentation of a RGB image into a HSC image as a classification problem in which the two classes (in this case, hand and background) are generated from the set of pixels within the HSC scale image. There are total L levels in image (0-255).Using a threshold T, the image is segmented into 2 classes. The optimum threshold K^* is determined as the value of k which maximizes the ratio of between class variance to the total class variance.

In image processing finding edge is fundamental problem because edge defines the boundaries of different objects. Edge can be defined as sudden or strong change in the intercity or we can say sudden jump in intensity from one pixel to other pixel. By finding the edge in any image we are just reducing some amount of data but we are preserving the shape. Hysteresis uses two thresholds and if the magnitude is below the first threshold, it is set to zero (made a no edge). If the magnitude is above the high threshold, it is made an edge. And if the magnitude is between the 2 thresholds, then it is set to zero unless there is a path from this pixel to a pixel with a gradient above T_2 .

Step 1: In order to implement the canny edge detector algorithm, a series of steps must be followed. The first step is to filter out any noise in the original image before trying to locate and detect any edges. And because the Gaussian filter can be computed using a simple mask, it is used exclusively in the Canny algorithm. Once a suitable mask has been calculated, the Gaussian smoothing can be performed using standard convolution methods . A convolution mask is usually much smaller than the actual image. As a result, the mask is slid over the image, manipulating a square of pixels at a time. The larger the width of the Gaussian mask, the lower is the detector's sensitivity to noise. The localization error in the detected edges also increases slightly as the Gaussian width is increased.

Step 2: After smoothing the image and eliminating the noise, the next step is to find the edge strength by taking the gradient of the image. The Sobel operator performs a 2-D spatial gradient measurement on an image. Then, the approximate absolute gradient magnitude (edge strength) at each point can be found. The Sobel operator uses a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows).

Step 3: Once the edge direction is known, the next step is to relate the edge direction to a direction that can be traced in an image. Then, it can be seen by looking at pixel whose value is "1", there are only four possible directions when describing the surrounding pixels - 0 degrees (in the horizontal direction), 45 degrees (along the positive diagonal), 90 degrees (in the vertical direction), or 135 degrees (along the negative diagonal). So now the edge orientation has to be resolved into one of these four directions.

Step 4: After the edge directions are known, non maximum suppression now has to be applied. Non maximum suppression is used to trace along the edge in the edge direction and suppress any pixel value (sets it equal to 0) that is not considered to be an edge. This will give a thin line in the output image.

4.1 Localized Counter Sequence

After edge detection we get a boundary of hand in image that is our contour of hand image. Now an algorithm is applied on the contour to track it in clockwise direction and the contour pixel are numbered sequentially. Some of the screen shots are shown below.



Fig 3 Gesture of letter “अ”

Fig 4 Gesture of letter “ए”

Fig 5 Gesture of letter “इ”

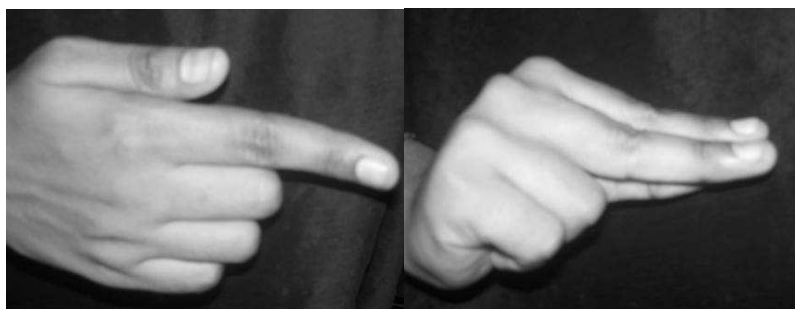


Fig 6 Gesture of letter “उ”

Fig 7 Gesture of letter “ऋ”

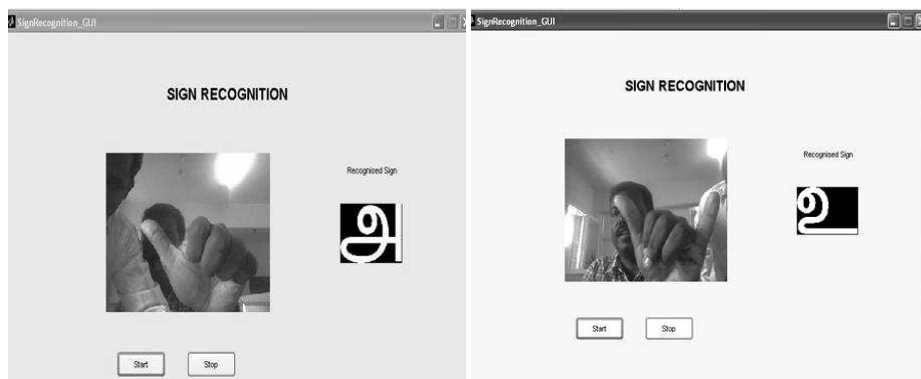


Fig 8 Screenshot of output “अ”

Fig 9 Screenshot of output “ए”

5. CONCLUSION:

In our work, the preprocessing of gesture recognition system which includes image acquisition, segmentation and morphological filtering method. RGB images is converted into HSC image consisting hand or background. Morphological filtering techniques are used to remove noises from images so that we can get a smooth contour. Otsu thresholding algorithm is used for segmentation. Canny edge detection technique is used to detect the border of hand in image. A contour tracking is applied to find the contour and pixel in contour is numbered sequentially. Local contour sequence for any arbitrary pixel is calculated as perpendicular distance from the chord connecting end points of window size w . Hand recognition system can be useful in many fields like robotics, computer human interaction and so make this offline system for real time will be future work.

6. REFERENCES:

- Andrew Wilson and Aaron Bobick, "Learning visual behavior for gesture analysis, " In Proceedings of the IEEE Symposium on Computer Vision, Coral Gables, Florida, pp. 19-21, November 1995.
- Arun Katkere, Edward Hunter, Don Kuramura, Jennifer Schlenzig, Saied Moezzi, and Ramesh Jain, "Robogest: Telepresence using hand gestures", Technical report, University of California, San Diego, Visual Computing Laboratory, Technical Report No. VCL-94-104, December 1994.
- Bowden and Sarhadi, "Building temporal models for gesture recognition" British Machine Vision Conference, pages 32-41, 2000.
- Chan Wah Ng, Surendra Ranganath, "Real-time gesture recognition system and application", Image Vision Comput, 20 (13-14): 993-1007, 2002
- Christopher Lee and Yangsheng Xu, "Online, interactive learning of gestures for human robot interfaces" Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996.
- Claudia Nölker and Helge Ritter, "Visual Recognition of Continuous Hand Postures", IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 31, no. 1, February 2001.
- Etsuko Ueda, Yoshio Matsumoto, Masakazu Imai, Tsukasa Ogasawara. "Hand Pose Estimation for Vision-based Human Interface", IEEE Transactions on Industrial Electronics, Vol. 50, No. 4, pp. 676–684.
- Hank Grant, Chuen-Ki Lai, "simulation modeling with artificial reality technology (smart): an integration of virtual reality simulation modeling", Proceedings of the Winter Simulation Conference, 1998.
- Henrik Birk and Thomas Baltzer Moeslund, "Recognizing Gestures From the Hand Alphabet Using Principal Component Analysis", Master's Thesis, Laboratory of Image Analysis, Aalborg University, Denmark, 1996.
- Hyeon-Kyu Lee and Jin H. Kim, " An HMM-Based Threshold Model Approach for Gesture Recognition" IEEE transactions on pattern analysis and machine intelligence, vol. 21, no. 10, october 1999.
- <http://www.happinesspages.com/baby-sign-language-FAQ.html>
- <http://www.prima.inrialpes.fr/FGnet/data/MoeslundGesture/database>
- J. Edward Jackson, "A Users Guide to Principal Components", Wiley Series in Probability and Mathematical Statistics, A Wiley-Interscience Publication, 1st edition, 1991.
- Jennifer Schlenzig, Edward Hunter, and Ramesh Jain, "Recursive spatio-temporal analysis: Understanding Gestures", Technical report, Visual Computing Laboratory, University of San Diego, California, 1995.
- Matthew A. Turk and Alex P. Pentland, " Face recognition using eigenfaces", IEEE Society Conference on Computer Vision and Pattern Recognition, pages 586–591, Lahaina, Maui, Hawaii, June 3-6 1991.
- Richard Watson, "Gesture recognition techniques", Technical report, Trinity College, Department of Computer Science, Dublin, July, Technical Report No. TCD-CS-93-11, 1993.
- Rick Kjeldsen and John Kender, "Finding skin in colour images", In Proc. IEEE Int. Conf. on autom. Face and Gesture Recognition, pages 3 12-3 17, 1996.
- Thad Starner and Alex Pentland, "Real-time American sign language recognition from video using hidden markov models", Technical Report No. 375, M.I.T Media Laboratory Perceptual Computing Section, 1995.
- Theodore Brun. "Teckensprks Lexikon". Bokforlaget Spektra AB, Halmstad, 1974.
- Thomas G. Zimmerman, Jaron Lanier, Chuck Blanchard, Steve Bryson, Young Harvill, "A hand gesture interface device", SIGCHI/GI Proceedings, conference on Human factors in computing systems and graphics interface, p.189-192, April 05- 09, , Toronto, Ontario, Canada, 1987.

Usage of Tamil in Internationalized Domain Names

Balaji Vasudevan

balaji.vasudevan20@gmail.com

As the Internet becomes a daily part of our lives, we become increasingly aware of the many areas of its influence. One such area is Internet domain names. Internet domain names are important in that they are the principal identifiers of a website. Traditionally, ASCII character sets have been used to represent these domain names. However, as the number of non-English and non-Latin users of the Internet increase globally, the need to address the multilingual global population.

1. What is IDN?

An **Internationalized domain name** (IDN) is an Internet domain name that contains at least one label that is displayed in software applications, in whole or in part, in a language-specific script or alphabet, such as Arabic, Chinese, Russian, **Tamil** or the Latin alphabet-based characters with diacritics, such as French.

Ex: www.sample.இலங்கை

www.sample. சிங்கப்பூர்

There are two types of IDN domains: partial (IDN.xx) or full (IDN.IDN). A **Partial IDN** has the extension still in Latin-based characters, while **Full IDNs** are completely represented in that language's native character. Internationalized Domain Names will increase Internet accessibility worldwide, particularly for people using non-Latin alphabets such as Tamil, Arabic, Chinese, Cyrillic, Greek, Hebrew, Japanese, and Korean. Many countries have started registrations for IDN.IDN domains. Until these extensions are approved by ICANN and added to the root routers, they cannot be used globally. Pre-orders are already being offered through **101domain.com**.

2. Vision of IDN

- Solving the final barrier to widespread adoption of Internet in non-English speaking communities
- Reducing the digital divide specifically in non-English speaking communities

3. Need for IDN

Although users of languages based on Latin characters, either natively (e.g. English) or in a transliterated form (e.g. Malay), do not have linguistic problems with the current domain name system, native speakers of Arabic, Chinese, Japanese, Korean, Tamil, Thai and others who use non-ASCII scripts remain at a considerable disadvantage. In an attempt to solve this problem, as well as generally provide for improved multilingual and multiscript support, a process of "internationalization" of the Internet's Domain Name System (DNS) has been under way.

4. **Role played by ICANN and Language Communities:** The **Internet Corporation for Assigned Names and Numbers** is a non-profit private organization headquartered in United States. It maintains registries of Internet protocol identifiers, and for the management of the top-level domain name space (DNS root zone), which includes the operation of root name servers. Until recently, the Root Zone was limited to a set of characters conforming to US-ASCII (American Standard Code for Information Interchange) or "Latin" alphabets. This changed with the introduction of Internationalized Domain Names (IDNs), which introduced top-level domains (TLDs) in different scripts and enabled Internet users to access domain names in their own language.

5. Study of Domains and the Technical Aspects of the multilingualization:

Domain name is

- Human readable identifier of an entity within the Internet
- Substitute of an IP address

These domain name space is divided into a) Country Code top-level domains (ccTLD, eg- co.uk) and Generic top-level domains (gTLD, eg- .gov, .edu).

Sample Link name: <http://உதாரணம்.பரிட்சை>

DNS name: xn--zkc6cc5bi7f6e.xn--hlcj6aya9esc7a

The DNS domain name space has a hierarchical structure (see Figure 1 **Error! Reference source not found.**) and is used to identify entities in the Internet. Each node in the structure corresponds to an entity in the Internet. A name given to a node in the structure is called a domain label. All nodes are given labels with one exception: the root node, as shown at the top of Figure 1, which has no label. The domain name of an entity (node) is a sequence of node labels starting from itself up to the root, and where labels are separated by periods. As to the length, a domain label should not exceed 63 octets and an entire domain name should not be longer than 255 octets.

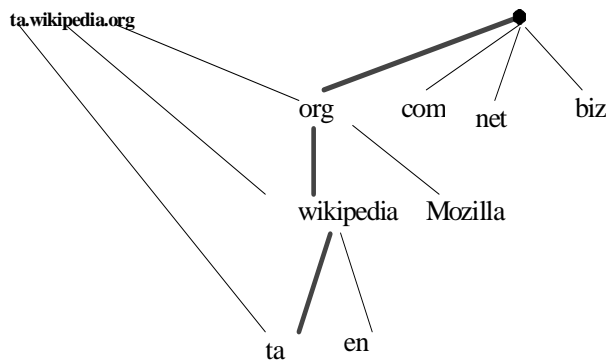


Figure 1 – The structure of domain names

There are several solutions available from Client Side and Server Side to address multilingual issues:

Client-side versus server-side solutions

As regards the question of *where* non-ASCII codes should be recognized, approach to the solution of this problem is typically based on one of the following methods:

a) Client-side solution

In a client-side solution, translation between the multilingual script and the ASCII-compatible representation is performed in the user applications (e.g. a web browser). The client application translates multilingual scripts into ASCII strings, which can then be processed in the current Internet: the domain names are subsequently processed as ASCII domain names throughout the Internet. This category includes the case of an application that consists of both client-side and server-side software. But for the sake of convenience, the term "client-side" is used in the interest of consistency with the ICANN survey report⁴.

Technically, a client-side solution is needed regardless of which approach is chosen. It is unlikely that an ASCII-only application will work immediately with multilingual domain names. Some form of upgrade will be necessary, either through provision of fonts, input methods or additional technical functionality to support internationalization.

b) Server-side solution

In a "server-side" solution, domain names are sent natively over the Internet by the client application in local encoding, such as UTF-8⁵, GB or BIG5⁶, or Unicode. Applications and services communicate with each other using non-ASCII domain names all the way along the path between them (sometimes referred to as "on the wire"). Note that the original implementations of IDN were actually proxy server solutions that intercepted local encoding from client applications and converted the encoding into an ASCII-compatible encoding so that the DNS server remained unaltered.

IETF (Internet Engineering Task Force) is inclined towards client Side solution for following reasons:

- **Stability**
 - DNS is a huge distributed database
 - DNS is working on a delicate balance

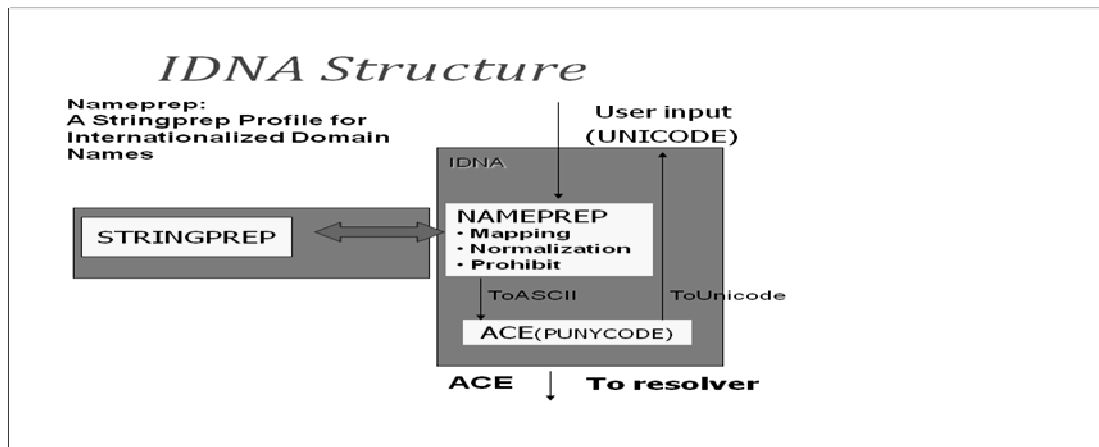
⁴ See <http://www.icann.org/committees/idn/final-report-28aug01.htm>.

⁵ See <http://www.ietf.org/rfc/rfc2279.txt>.

⁶ GB and BIG5 are coding schemes for Chinese characters.

- Substantial change of DNS is dangerous
- **Deployment speed**
Changing all the servers takes long time
- **Consistency**
Partial deployment of server-side solution may lead to separation of the Internet

Internationalizing Domain Names in Applications (IDNA) is a mechanism defined in 2003 for handling internationalized domain names containing non-ASCII characters. Now, let us see how multilingual characters are converted into ASCII Strings. This is accomplished by two algorithms **toASCII** or **toUnicode**. For more details about these algorithms please refer to **RFC3490**.



Below is the process flow to achieve Non ASCII to ASCII:

- If given a domain label containing at least one non-ASCII character, ToASCII will apply the Nameprep algorithm, which converts the label to lowercase and performs other normalization, and will then translate the result to ASCII using Punycode. Punycode is a simple and efficient transfer encoding syntax designed for use with Internationalized Domain Names in Applications (IDNA). It uniquely and reversibly transforms a Unicode string into an ASCII string before prepending the four-character string "xn—". This four-character string is called the ASCII Compatible Encoding (ACE) prefix, and is used to distinguish Punycode encoded labels from ordinary ASCII labels.
- The function ToUnicode reverses the action of ToASCII, stripping off the ACE prefix and applying the Punycode decode algorithm.

6. Influence of Tamil in IDN

In the year 2005 CE, Indic IDN being considered was Unicode 3.0. But now, Unicode 5.2 is the current version, and the following Tamil letters need implementation in IDN.

(a) **Tamil OM sign (U+0BD0)**

If Devanagari OM sign (U+0950) is allowed for IDN, Tamil OM sign (U+0BD0) is needed in IDN also.

For blocking any confusability problems, the letter ழ should be blocked after Tamil OM sign (U+0BD0). Note that when Om is written as a sequence of two letters, i.e., the letter ழ will always be present. So, the blockage of ழ following U+0BD0 will distinguish between OM sign and the word, Om as a sequence of 2 letters.

- This is similar to situation in Devanagari script. And if Devanagari OM sign is allowed, Tamil OM sign should be allowed in IDN also. While Grantha loan conjunct, Shrii (Section (a)) will be allowed, Tamil OM sign is important

in the native religions of India. Also, note a graphic variant of Tamil OM sign contains Vel "spear" also. It is very popular form among Tamils not just in India but also in Malaysia, Singapore and Sri Lanka. If you add that glyph (with vel) as a requirement for Tamil OM sign for IDN, it will be easy to distinguish visually even in a small-screen PDA.

7. Summary

To make Tamil domain names and Language fully usable on the Internet, we need to take following steps:

- a) Standardization of technology;
- b) Policy and coordination of registration and management rules;
- c) Deployment of applications and name servers.

As mentioned here, the base technical standards for IDN have been approved by IETF. However, as all languages of the world have yet to be considered, the specifications of the standard may need to evolve further. In addition, as the DNS itself is evolving, longer-term solutions such as server-based solutions or additional software layers may emerge (e.g. keywords) and prove to offer better solutions.

The deployment of applications and name servers must rely on the dynamics of the business sector. In order to achieve satisfactory usage, it is important to promote deployment of both servers and applications. It is vital that application development be catalyzed and widely promoted. As one practical example, the Japanese Domain Names Association (JDNA), which was established in July 2001, has Japan-based members such as application vendors, network service providers and domain name registries. Likewise, we need to have Tamil-based members to provide more inputs on the usage of complex terms and terminologies on the internet world.

To summarize, there is substantial market and user demand for multilingual domain names. To satisfy this demand, the whole environment needs to be developed to take into account technology standardization, policy and administrative arrangements, as well as new applications. The future of multilingual Internet names is imminent. We should not underestimate the significance of this activity, as it is part of a far nobler goal: the ongoing internationalization of the Internet.

8. References:

- a) www.w3.org
- b) www.ietf.org
- c) <http://www.unicode.org/L2/L2005/05185-kssa-zwj.pdf>
- d) Wikipedia (IDN)

வலைமச் சொற்கள்

இலக்குவனார் திருவள்ளுவன் thiru2050@gmail.com

கணிப்பொறி தொடர்பாகவும் பிற அறிவியல் தொடர்பாகவும் மிகுதியான கலைச் சொற்கள் தமிழில் உருவாக்கப்பட வேண்டி உள்ளன. அறிவியல் கலைச்சொற்கள் அவ்வப்பொழுது சொல்லாக்க ஆர்வலர்களால் வெளிப்படுத்தப் படுகின்றன. மேலும் பல்லாயிரக் கணக்கான கலைச்சொற்கள் தேவை. இதுவரை அறிமுகப் படுத்தப்பட்ட கலைச்சொற்களைக் கட்டுரையாளர்களும் நூலாசிரியர்களும் பயன்படுத்தினால்தான், இவற்றால் பயன் விளையும். இல்லையேல் விழலுக்கு இறைத்த நீர்தான். இனியேனும் படைப்பாளர்கள் தமிழ்க்கலைச் சொற்களையே பயன்படுத்தும் வேண்டுகோளுடன் கட்டுரையைத் தொடருகிறேன். கணிணிச் சொற்களில் வலைப்பணி (Network) சார்ந்த கலைச் சொற்களைக் காணலாம்.

கலைச்சொல்லாக்க நெறிமுறைகள் குறித்துக், 'கணிணியியலில் தேர்பெயர்ப்புச் சொற்களும் ஒலிபெயர்ப்புச் சொற்களும்' என்னும் (செருமனியில் நடைபெற்ற இணையத் தமிழ் மாநாட்டுக்) கட்டுரையிலும் <<http://thirupadaippugal.blogspot.in/2010/06/1.html>> 'கணிணியியலில் தமிழ்ப் பயன்பாடு' என்னும் (கோவையில் நடை பெற்ற இணையத் தமிழ் மாநாட்டுக்) கட்டுரையிலும் குறித்துள்ளேன். எனவே சுருக்கமாகச் சிலவற்றை மட்டும் குறிப்பிட விழுகிறேன்.

கலைச்சொற்கள் என்ற பெயரில் அயற் சொற்களையும் அயல் எழுத்துகளையும் புகுத்துவது அடியோடு நிறுத்தப்பட வேண்டும். முன்னரே நடைமுறையில் உள்ள கலைச்சொற்களை அறிந்து, தக்கனவாய் இருப்பின், அவற்றையே பயன்படுத்த வேண்டும். ஒரே சொல்லிற்கு ஒவ்வொருவரும் புதுப்புதுச் சொற்களை அறிய முற்படுவது காலக்கேடும் பயன்கேடும் ஆகும். அதே நேரம் பயன்பாட்டில் உள்ள சொற்கள் தவறாக அல்லது வேறு பயன்பாட்டில் இருப்பின், வேறு சொற்களை அறிய முற்படுவதே சரியானதாகும். கலைச்சொற்களைத் தொடர்களாக அமைக்கக் கூடாது; இவை சுருக்கமாகவும், செறிவாகவும், தெளிவான பொருள் உடையனவாகவும், பொருள் மயக்கம் இல்லாதனவாகவும் இருத்தலே சிறப்பு. சுருக்கக் குறியீடுகள் அல்லது தலைப்பெழுத்துச் சொற்கள் தமிழிலேயே இருக்க வேண்டும். கலைச்சொற்களைப் பாடநூல்களிலும் பயன்படுத்த வேண்டும். இதன் அடிப்படையில் நாம் வலைமச் சொற்கள் குறித்துக் காண்போம்.

வலைமச் சொற்கள் எனப் பார்க்கும் பொழுது கணிணி தொடர்பான பிற பிரிவுகளில் பலவும் தொடர்புடையனவாக உள்ளமையைக் காணலாம். எனவே ஆயிரக்கணக்கில் உள்ள கலைச்சொற்கள் முழுமையையும் காணக் கால வரையறை இடம் தராது. எனவே, கால வரையறைக்கு உட்பட்ட எண்ணிக்கையில் வலைமக் கலைச்சொற்களைக் காணலாம்.

அ) வலைமம் (Network)

வலைப்பணி என்பதற்கு மாற்றாகச் சிலர் பிணையம் எனக் கையாளுகின்றனர். பிணையம் என்பது ஒருவரை விடுவிப்பதற்கு அளிக்கப்படும் பிணை அல்லது காப்பிணைக் (bail or security) குறிப்பதாக அமையும். பிணைக்கப்படுவது என்னும் பொருளில் கையாள எண்ணினாலும் இத்துறையில் அச்சொல் பொருந்தாது. குறுக்கு நெடுக்குக்கட்டமைப்பு, கணிணிக்கட்டமைப்பு என்பன விளக்கமாக அமைந்தாலும் கலைச்சொல்லாக அமையவில்லை. நிகரம் என்பது, நிகரத் தொகை என்பது போன்ற இடங்களில் சரியாக இருக்கும். மருத்துவத்தில் பத்தியம் என்னும் பொருளிலும் இச்சொல் கையாளப்படுகிறது. சிலந்தி வலை, மீன் வலை முதலான இடங்களில் உரியவாறு கையாளப்படுகிறது. இங்கே நாம் வேறு வலைகளில் இருந்து வேறுபடுத்துவதற்காக வலைப்பணியைச் சுருக்கமாக வலைமம் எனலாம்.

வலைமம் - Network

வலைமச் சொற்களில் சில வருமாறு :-

தே.வா.வா.ப. வலைமம் நாசா என்பது National Aeronautics and Space Administration என்பதன் ஆங்கிலத்தலைப்பெழுத்துச் சொல்லாகும் .தமிழில் தேசிய வானூர்தி -வான்வெளி பணியாட்சி எனலாம் .சுருக்கமாகத் தே.வா.வா.ப .எனலாம். ஆதலின் தே.வா.வா.ப. வலைமம்.	Nasa communications network (nascom)
முனைப்பு மின் வலைமம்	Active electric network
அகப்பரப்பு வலைமம் /அ.ப.வ	Local Area Network / LAN
அகல்பரப்பு வலைமம்/ அக.ப.வ	Wide Area Network / WAN
அகலக்கற்றை வலைமம்	Broad band network
அழை நிலை வலைமம்	Dial-up networking

இணக்க வலைமம்	Compromise network
இணை நோக்கு வலைமம்	Peer- to- peer network
இணைக் குறுக்கு வலைமம் (H-வடிவ வலைமம்)	H network
இணைப்பிடை வலைமம்	Inter connected network
இணைப்பிலி வலைமம் பணி இ.இ.வ.ப.	Connection less network service/ CLNS
இணைப்புசார் வலைமம் பணி / இ.சா.வ.ப.	Connection oriented network service/ CONS
இணைவு நீக்கு வலைமம்	Decoupling network
இருபுற வலைமம்	Bilateral network
இருமுனைய வலைமம்	Two terminal network
இருமை வலைமம்	Dual network
ஈடுசெய் வலைமம்	Compensating network
உயர்மட்ட வலைமம்	High-level network
ஒத்தியக்க வலைமம்	Synchronous network
ஒப்புமை வலைமம்	Analog network
ஒருங்கு பணிகளின் எண்ம வலைமம்	Integrated services digital network (ISDN)
ஒருங்கு வலைமம்	Integral network
ஒருமுக வலைமம்	Linear network
கண்ணி வலைமம்	Mesh network
கணினி வலைமம்	Computer networking
கல்வியிய வலைமம்	Academic network
கலப்பின வலைமம்	Hybrid network
கானகச் சூழலமைப்பு வலைமம்-கா.சூ.வ	Forest ecosystem network-FEN
கிளைமை வலைமம்	Star network
சமனி வலைமம்	Bridge network
சமனுறு வலைமம்	Balanced network
செயற்கை வலைமம்	Artificial network
ஞாலப் பரப்பு வலைமம்/ ஞாபவ	Global area network / GAN
ட வடிவ வலைமம்	L network
தகவல் வலைமம்	Information network
தகவல்தொடர்பு வலைமம்	Communication network
தன்னியக்க எண்ம வலைமம்	Automatic digital network
தனிப்படுத்து வலைமம்	Isolation network
தனிப்பரப்பு வலைமம்/த.ப.வ.	Personal area network / PAN
திரள் மாறிலி வலைமம்	Lumped-constant network
திருத்து வலைமம்	Corrective network
நடுவண் வலைம அமைவம்	Centralized network configuration
நான்முனை மின் வலைமம்	Four terminal network
நுண்ணலை வலைமம்	Microwave network
பகிர்வு வலைமம்	Distributed network
பங்கிடு வலைமம்	Dividing network
படிநிலை வலைமம்	Hierarchical network
படிமுறை வலைமம்	Ladder network
பல் திவலை வலைமம்	Multi drop network
பல்துறை வலைமம் பகுப்பாய்வி	Multiport network analyzer
பல்பெருக்க அணுக்க வலைமம்	Multiple access network
பல்முறைமை வலைமம்	Multisystem network
பல்வகை வலைமம்	Heterogeneous network
பல்வரிசை வலைமம்	Multitier network
பிரிப்பு வலைமம்	Crossover network
பின்னல் வலைமம்	Lattice network
பின்னாட்ட வலைமம்	Feed back network
புறவெளி வலைமம்/ புவெவ	Deep-space network (DSN)
பெருநகர்ப் பரப்பு வலைமம் / பெ.ப.வ.	Metro area network / MAN
பெரும வலைமம்	Maximum network
பொருத்து வலைமம்	Fixer network
மறுதலை வலைமம்	Inverse network
முக்கோண வலைமம்	Delta network
வலிசுன்று வலைமம்	Attenuation network
கிளைப்பி வலைமம்	Derivative network
வலு நீக்கு வலைமம்	Deemphasis network
வலைம ஒருங்கமைப்பு	Network organization
வலைம வரிவரை	Network diagram
வலைம அட்டை	Network card
வலைம இயக்க முறைமை	Network operating system

வலைம இயக்கக் கட்டுப்பாட்டு மையம் –வ.இ.க.மை.	Network Operations Control Center (NOCC)
வலைம நிரல்படம்	Network chart
வலைம மாறிவி	Network constant
வலைம வரிப்படம்	Network diagram
வலைமக் கட்டுமானம்	Network architecture
வலைமச் சுற்றம்	Network neighborhood
வலைமப் பகுப்பாய்வு	Network analysis
வலைமப் பணியாட்சியர்	Network administrator
வளாகப் பரப்பு வலைமம்/ வ.ப.வ.	Campus area network / CAN
விளம்பர வலைமம்	Advertising Network

ஆ) பொதியம் (Packet)

பாக்கெட்டு (packet) என்பது தமிழில் இடத்திற்கேற்றவாறு, பொட்டலம், பொதி, பொட்டணம், சிப்பம், சிறுபொதியம் எனப் பலவாறாகக் குறிக்கப்படுகின்றது. இங்கே தரவுகளைப் பொதிந்து வைப்பதைக் குறிப்பதால் பொதியம் எனலாம்.

பொதியம் - Packet

பொதிய இழப்பும் மீளனுப்புகையும் - Packet Loss And Retransmission

மீ விரைவு புவிஇணைப்பு பொதிய அணுக்கம்/ மீ.வி.பு.பொ.அ. -

High-Speed Downlink Packet Access / HSDPA

இ) ஆவளி - Sequence

array, order, queue, row, tier, ordinal, sequence என எல்லாமே வரிசை என்றே குறிக்கப்படுகின்றன. எனவே, வேறு தனிச்சொல்லைக் குறிப்பதே பொருத்தமாக இருக்கும். தீபங்கள் வரிசையாக அடுக்கப்படுவதைத் தீபாவளி என்றும் பட்டம் பெற்ற குருமாரின் தலைமுறை வரிசையைப் பட்டாவளி என்றும் வரிசை என்பதைக் குறிக்கும் ஆவளி என்னும் சொல்லை இணைத்துக் குறிக்கின்றனர். எனவே, இந்த இடத்தில் வரிசை என்பதற்கு மாற்றாக ஆவளி எனக் குறிக்கலாம்.

பொதிய ஆவளி எண் - Packet Sequence Number

ஈ) ஒற்றி (sniffing)

முகர்வு, மோப்பம் பிடித்தல் என sniffing என்பதைக் குறிப்பிடுகின்றர். இவ்வாறு நேர் பொருளில் குறிப்பதைவிடக் கணிணிக்குப் பொருத்தமான சொல்லைக் குறிப்பதே சிறப்பாகும். அனைத்து விவரங்களையும் ஒற்றி எடுத்தல் என்னும் பொருளில் ஒற்றி எனலாம். மோப்பம் பிடிப்பதும் தடத்தை ஒற்றி அறியத்தானே!

ஒற்றி- Sniffing

பொதிய ஒற்றி - Packet Sniffing

உ(ஒம்பி - host

ஃகோசுட் கம்பியூட்டர் /host computer என்பதற்கு விருந்துக் கணினி/ஏற்புக் கணினி எனத் தவறாகக் குறிக்கின்றனர் . தருகணினியை ஏற்புக் கணினி என்பது தவறல்லவா? பிற அறிவியல் துறைகளில் ஃகோசுட்/host என்பதற்கு ஒம்புயிரி, ஒம்பி, விருந்து ஒம்புநர், புரவன், ஒம்புநன், தருநர் என்கின்றனர் .ஒம்பி என்பதையே பொதுச் சொல்லாக எடுத்துக் கொண்டு இத்துறையிலும் கையாளலாம் .எனவே, ஒம்பிக்கணினி (host computer) எனலாம்.

ஒம்பிக்கணினி - host computer

ஊ(கோரி (client)

கிளையன்ட்டு client என்பதற்கு வாடிக்கையாளர், கட்சிக்காரர் என்றெல்லாம் சொல்லப்படுகின்றன .இவற்றுக்கு முறையே customer, partymen ஆகிய சொற்கள் உள்ளன .பொதுவாக ஒரு பணியை முடித்துத் தருமாறு கோருபவரே கிளையன்ட் எனப்படுகிறார் .எனவே, கோருநர் என்பதே சரியாக இருக்கும் .அஃறிணையில் கோரி எனல் வேண்டும் .

கோரி/கோருநர்- client

வலைமக் கோரி -network client

எ(விளம்பி (server)

அதுபோல் server என்பதற்கு உயர்திணையில் விளம்புநர் என்றும் அஃறிணையில் விளம்பி என்றும் சொல்வது பொருத்தமாக அமையும் .நான், பல ஆண்டுகளுக்கு முன்பு கணினி மொழியை விளம்பி எனச் சொல்லாம் என எழுதியிருந்தேன் .மனிதர்களால் மொழியப்படுவதுதான் மொழி எனவே, கணினிக்கு இது பொருந்தாது எனக்

கருதியிருந்தேன் .ஆனால், கணினி மொழி என்பது பரவலாகப் பயன்பாட்டிற்கு வந்து விட்டது .ஆதலின் விளம்பி என்ற சொல் செர்வர் என்பதற்கு மட்டும் பொருந்தும் எனக் கருதலாம்.

விளம்பி/ விளம்புநர் - server

கோரி - விளம்பி வலைமம் - client - server network

ஏ.) நிரலி (protocol)

சீர் மரபு, செம்மை நடப்பு வழக்கு, செம்மை நடப்பொழுங்கு, செய்மை நடப்பொழுங்கு, நெறிமுறை, மரபு பேணுகை, மரபு முறை, மரபுச்சீர் முறைமை என வெவ்வேறு வகையாக இப்பொழுது குறித்து வருகின்றனர்.

மின் குழுவும் ஒன்றில் பேரா. செல்வகுமார்,

“நெறிமுறை எதிர் விதிமுறை/ எதிர் வரைமுறை - protocol, எது சரி? இரண்டுமே சரியாக இருக்கும். தொடர்பாடல் துறையில், கணினித் துறையில் இரண்டிலுமே இரு சொற்களும் பொருந்தும். இல்லை சுருக்கமாக நெறி என்றே சொல்லலாம். நெறி என்றால் ஒரு குறிப்பிட்ட திசைமுறைப்படி செல்லுதல்/நடத்தல். ‘புரோட்டோக்கால்’ என்பது பணி அல்லது தொடர் பாடலுக்கான நெறிமுறை. விதிமுறை என்பதும் அதுவே. நெறி என்ன என்றால் புரோட்டோக்கால் என்ன. வகுநெறி, தொடர்நெறி, தொடக்குநெறி என்று தேவைக்கும் இடத்துக்கும் ஏற்றாற்போல செய்து கொள்ளலாம்” எனத் தெரிவித்து உள்ளார்.

career, norm, moral, principle, என்பனவற்றிற்கெல்லாம் நெறி என்றே குறிக்கின்றனர். எனவே, நெறி என்பதும் நெறிமுறை என்பதும் புரோட்டோக்கால் என்பதற்குப் பயன்படுத்தும் பொழுது குழப்பம் வரும். ஆதலின் தக்க கலைச்சொல்லையே பயன்படுத்த வேண்டும்.

நிரல் என்பதற்கு வரிசை, முறைப்படி ஒழுங்குபடுத்தல், வரிசைப்படுத்தல் எனப் பொருள்கள் உள்ளன.

பொதுநோக்கான வேந்தன் வரிசையா நோக்கின்

அதுநோக்கி வாழ்வார் பலர் (குறள் எண்: 528) என்கிறார் தெய்வப்பலவர் திருவள்ளுவர். பாராட்டிற்கும் மதிப்பிற்கும் உரியவர்களைப் பொதுவாக நோக்காமல் தகுமுறைக்கேற்ப நோக்க வேண்டும் என்பதையே அவர் குறிப்பிடுகிறார். அதற்கு அவர் குறிப்பிடும் சொல்லே வரிசை என்பதாகும். புரோட்டோக்கால்/protocol என்பதும் வரிசை முறைப்படி அமைவதையே குறிக்கிறது. இரண்டு அல்லது இரண்டிற்கு மேற்பட்ட தகவல் தொடர்பிகளிடையே பரிமாறப்படுவது உரிய வரிசையமைப்பிற்கேற்ப நிகழ்வதைக் குறிக்கும் நெறிப்பாட்டை, வரிசைநெறி என்னும் பொருளிலான நிரல் நெறி எனலாம். நிரல் உடைய நெறி என்பதால் சுருக்கமாக நிரலி எனலாம்.

நிரலி - Protocol

வலைம நிரலி	Network protocol
ஓர்மி* ஒத்திசை நிரலி	Bit synchronous protocol
தகவல்தொடர்பு நிரலி	Communications protocol
இணைப்புசார் நிரலி	Connection oriented protocol
கோப்பு மாற்று நிரலி	File transfer protocol
மீஉரை மாற்ற நிரலி	Hyper text transfer protocol (http)
இணைய நிரலி	Internet protocol
அடுக்கு நிரலி உத்தி	Layered-protocols technique
பல்தொகுப்பு முனைக்கு முனை நிரலி	Multilink point-to-point protocol
வலைம நிரலி	Network protocol
வலைமப் போக்குவரத்து நிரலி	Network transport protocol
முனை-முனை நிரலி	Point-to-point protocol
நிரலி உடன்படிக்கைகள்	Protocol agreements
நிரலித் தொடர்	Protocol sentence
நிரலி அடுக்கம்	Protocol stack
நிரலித் தொகுமம்	Protocol suite
நிரலி நிலைக் கடிகை	Protocol-level timer
பயனர் தரவுவரி நிரலி	User data gram protocol

(*binary digit என்பதன் சுருக்கம்தான் bit என்னும்பொழுது, மூலச்சொற்களை மறந்துவிட்டுச் சுருக்கச் சொல் அடிப்படையில் துண்டு என மொழி பெயர்த்தால் தவறான பொருள்தானே கிடைக்கும். சிலர் பிட்-சிறு துண்டு, பைட்-பெருந்துண்டு; மேலும் முறையே சிறுதி (சிறு துண்டு என்பதன் சுருக்கமாம்), பெருதி; துண்டம், எண் துண்டம்; விள்ளல், எட்டு விள்ளல்/அட்ட விள்ளல்; துணுக்கு, எட்டியல் என்றெல்லாம் குறிக்கின்றனர். அனைத்திற்கும் அடிப்படைக் காரணம் 'பிட்' என்பதை ஒற்றைச் சொல்லாகக் கருதி மொழி பெயர்ப்பதே ஆகும். எனினும் 1, 0 என்னும் இரண்டு எண்களை மட்டும் அடிப்படையாகக் கொண்டு இயங்குவதைக் குறிப்பதால் இருமை எண்கள், ஈரிலக்க எண்கள், இருநிலை எண்கள், இரட்டை எண்கள், ஈரடிமானம் என்று வெவ்வேறு வகையாக இச் சொல் விளக்கப்படுகிறது. எனவே சுருக்கமாக இருமம் என்று குறிப்பிடலாம். எனினும் binary என்பதன் பொருள்தான் இவ்விளக்கங்கள். 'பிட்' ஓர் அலகு என்ற அளவில்தான் பயன்படுகிறது. இவ்வலகில் 0, 1 ஆகிய ஈரெண்கள் இருந்தாலும் பயன்பாட்டில் ஏதேனும் ஒன்றைத்தான் குறிக்கும். எனவே, அளவீடாகிய அலகைக் குறிக்கையில் 'இருமம்' என்பது தவறாகிறது. எனவே, ஓர் அலகு என்னும் நடைமுறைக் கிணங்க, 'ஓர்மம்' > 'ஓர்மி' என்று சொல்லலாம். அப்படியானால் பொருள் மாறுபடாதா என்றால் சொல் இடத்திற்கேற்ற பொருளைத்தானே பெற வேண்டும். எடுத்துக்காட்டாக 'back' என்றால் 'பின்' எனப் பொருள். backside - பின்பக்கம்; ஆனால், back file - முன் கோப்பு என இடத்திற்கேற்ற பொருளைப் பெறுகிறது. அதுபோல்தான் 'பைனரி' என்னும் பொழுது இருமத்தைக் குறிக்கும் சொல் பைனரி யூனிட் என்னும் பொழுது ஓர்மி ஆகிறது. இல்லாவிடில் 'பிட்' என்பதை இருமம் என்றால் அதன் எண்மடங்காகிய 'பைட்' என்பது எண்ணிரண்டு பதினாறு என்றாகும். எனவே, தவறான பொருள் உண்டாகும். மாறாக, எட்டு 'ஓர்மி' (பிட்) சேர்ந்த 'பைட்' என்பதை 'எண்மம்' > 'எண்மி' என்று சொல்லலாம். Bit – ஓர்மி; Byte – எண்மி எனச் சுருக்கமாகவே குறிப்பிடலாம். இவற்றின் அடிப்படையில் பிற அளவைகளையும் காணலாம். (கணினியியலில் நேர்பெயர்ப்புச் சொற்களும் ஒலிபெயர்ப்புச் சொற்களும் : இலக்குவனார்திருவள்ளுவன்<<http://thin-podaiappugal.blogspot.in/2010/06/2.html>>)

ஐ.) இலி/இன்மை (null)

நல் (null) என்றால் வெறுமை அல்லது ஏதுமற்ற என்று பொருள். நல் என்னும் தமிழ்ச்சொல்லிற்கு நல்ல என்று பொருள். ஆனால், ஆங்கிலச் சொல்லையே பயன்படுத்தினால் நல்ல என்னும் பொருள்தாப் பொருள் அல்லவா தோன்றும். நல் என்று தமிழ் வரிவடிவிலேயே பயன்படுத்திவிட்டு நல் என்றால் நல்ல அல்ல என்று சொல்வதால் பயனில்லை. அந்த இடத்தில் வேண்டுமென்றால் அதைப் புரிந்து கொள்ளலாம். ஆனால் கலைச் சொற்களைப் பயன்படுத்தும் இடங்களில் படிப்போர் தவறாகவே எண்ணுவர். எனவே, நல்மோடம் எனத் தமிழில் பெயர்த்து எழுதுவதைவிடத் தமிழ்க்கலைச்சொல்லையே பயன்படுத்த வேண்டும்.

மோடம் (Modem - Modulation/Demodulation) என்பது தரவை மாற்றித் தரும் கருவி; அஃதாவது, கணினியில் இருந்து தொலைபேசிக் கம்பிகள் வழியாக அனுப்புவதற்காக அல்லது தொலைபேசிக் கம்பியில் இருந்து கணினிக்கு வந்துசேருவதற்காகத் தரவினை மாற்றித் தரும் கருவி; 'தருவி' யின்றி இரு கணினிகளை வடம் (கேபிள்) வழி இணைப்பதும் உண்டு. இவ்வடமே தருவியிலி (null modem) எனப்படுகிறது.

இலி- null ;

தருவியிலி - null modem

தருவிச்சொற்கள் சில :

மாறு நிலை தருவி (சமச்சீரற்ற என்பதைவிட மாறுபட்ட வேகங்களை உடையதால் மாறுநிலை என்பது சரியாக அமையும்.)	Asymmetrical modem
தொலைவு வரம்பு தருவி	Limited-distance modem
நேரிணை தருவி	Direct connect modem
ஒருங்கிணை தருவி	Integral modem
தொலைபேசித் தருவி	Telephone modem
குரல் தகைவு தருவி	Voice-capable modem
குரல் உணர் தருவி	Voice recognition modem
அகநிலை தருவி	Internal modem
அழை நிலை தருவி	Dial-up modem
வட நிலை தருவி	Cable modem
தொடர்முறை தருவி	Analog modem
ஒலிக்குறிப்பு தருவி	Acoustic modem

ஒ.) விசைப்பி - switch

ஆளி, இணைப்பி, இணைப்புமாற்றி, சுற்றிணை, சொடுக்கி, தாவி, திறப்பான், திறப்பி, தூண்டுவிசை, தொடர் மாற்றி, தொடு சாவி, நிலைமாற்றம், நிலைமாற்றி, பாதைமாற்றி, மடை, மின் பொருத்தி, மின்விசை மாற்றுக்குமிழ், விசை இயக்கி, விசை,

என சுவிட்ச்/switch என்பதற்கு ஒவ்வொரு துறையில் ஒவ்வொரு வகையாகக் கையாள்கின்றனர். பயன்முறை அடிப்படையில் பெரும்பாலும் சரியாக இருந்தாலும் பொருள் மயக்கம் தரும் சொற்களைப் பயன்படுத்தக்கூடாது. ஒவ்வொரு சொல்லும் சரியான பொருளை வெளிப்படுத்தவில்லை என ஒரு சாரார் கருதுவதாலேயே புதுப்புதுச் சொற்கள் பயன்பாட்டிற்கு அறிமுகப்படுத்துகின்றன.

விசை என்னும் சொல், கீ/Key என்னும் பொருளில் பரவலாகக் கையாளப்படுவதால், இதற்குப் பொருந்தாது. எனினும் விசை அடிப்படையில், விசைப்பி எனலாம். விசைப்பு என்பதை switching எனலாம்.

விசைப்பி- switch

தன்னியக்க வலைம விசைப்பி- Automatic network switching

ஒ.) சீர்மம் (ideal)

இலக்கியலான, இலட்சியம், உயர் குறிக்கோள், உயர்வான, எடுத்துக்காட்டாக விளங்குவது, கற்பனைச் சார்பான, குறிக்கோள் நிலை, குறிக்கோள் வடிவான, கனனவியலான, குறிக்கோள், குறைவிலா நிறைசெப்பம், சால்பு சீர்மை, செயல்துறை சாராத, நிலையான, கருத்தியலான, புனைவியலான, பொருத்தமான, முழுநிறைநலம், முன்மாதிரி,

என ஐடியல்/ ideal என்பதற்கு வெவ்வேறு சொற்கள் கையாளப்படுகின்றன.

சீர்மை அடிப்படையில் சீர்மம் எனக் குறித்து சீர்ம வலைமம் எனலாம்.

சீர்மம் –ideal

சீர்ம வலைமம் – ideal network

இவ்வாறு வலைமச் சொற்களைத் தமிழில் எண்ணித்தமிழிலேயே வெளிப்படுத்தலாம். முதலில் குறிப்பிட்டவாறு காலச்சுருக்கத்தால் சில மட்டுமே பார்த்துள்ளோம். எனினும் அனைத்துக் கலைச் சொற்களையும் தமிழில் பயன்படுத்தலாம். ஆனால், நடைமுறையில் கணிணிநூல்களிலும் கட்டுரைகளிலும் தமிழ்வரிவடிவில் ஆங்கிலமே பரவலாகப்பயன்படுத்தப்படுகின்றது.

தமிழ்க்கலைச் சொற்களை உருவாக்குவதிலும் பயன்படுத்துவதிலும் ஆர்வம் மிக்கவர்களே, அவற்றைப் பயன்படுத்தாமை அல்லது புதிய கலைச்சொற்களைத் தேடுவதில் நாட்டம் காட்டாமை போன்றவற்றால் இவ்வாறு, ஆங்கிலக் கலைச் சொற்களைத் தமிழ் எழுத்து வடிவில் பயன்படுத்தும் அவலத்தை அரங்கேற்றுகின்றனர். சான்றாகச் சொல்ல வேண்டுமெனில் வலைமம் குறித்து 960 பக்கத்தில் ‘நெட்வொர்க் தொழில்நுட்பம்’ என்னும் நூலை எழுதி உள்ள நூலாசிரியர், பலரின் தமிழார்வத்திற்கு உந்துதலாக உள்ளவர்.

ஆனால், இந்நூலைத் தமிழ் நூல் என்று கூறுவதற்கு இயலவில்லை. அருவினை ஆற்றும் ஆர்வத்தில் பெருமுயற்சி மேற்கொண்டு நூலை உருவாக்கி உள்ளார் ஆசிரியர். ஆனால், அருவினை அருந்தமிழ்க்கொலை புரிய அல்லவா துணை நின்றுள்ளது. “மார்க்னிங் எய்ட் ஓ கிளாக் பிரேக் ஃபாசுட் முடித்துவிட்டு நைன் எம் வண்டியில் ஆஃபீசுக்குப் புறப்பட்டு டென் எம் போய்ச் சேர்ந்தேன்” என்று அன்றாடம் மொழிக் கொலை புரிவதற்கும் முழுக்க முழுக்க ஆங்கிலக் கலைச் சொற்களையே பயன்படுத்தி நூலை எழுதுவதற்கும் வேறுபாடு இல்லை. இவ்வாறு எழுதுவது மொழிக் கொலை மட்டுமல்ல அறிவுக் கொலையுமாகும். ஆங்கிலத்தில் கணிணியியல் படித்தவர்கள் தமிழறிந்தவர்களாக இருப்பின் அவர்களுக்கு இந்நூல் எளிமையாக அமையலாம். ஆனால் முதலிலேயே தமிழில் படிப்பவர்களுக்கு ஆங்கிலக் கலைச்சொற்களை முழுக்க முழுக்கப் பயன்படுத்தி எழுதி உள்ளமையால், இந்நூல் குழப்பத்தைத்தான் ஏற்படுத்தும். உளவியல் அடிப்படையில் இது சரியானதல்ல. சொல்லை அறிமுகப்படுத்தும் பொழுதே தமிழில் அறிமுகப்படுத்தினால் எளிதில் புரியும். ஆதலின் தமிழ்க்கலைச் சொற்களைப் பயன்படுத்தி மறுபதிப்பு வெளியிடுவது ஆசிரியருக்கும் பெருமை சேர்க்கும்; படிப்போர்க்கும் பயன்பெருக்கும்; தமிழுக்கும் வளமை நல்கும்.

இவ்வாறே அனைத்துப் படைப்பாளர்களும் நல்ல தமிழிலேயே எழுத வேண்டும். தெரியாச் சொல்லிற்கான உரிய சொல்லை அறிவதில் சோம்பல், முயற்சியின்மை, நாட்டமின்மை போன்றவற்றால் தமிழுக்குக் கேடு செய்வதைப் படைப்பாளிகள் உணர வேண்டும். இத்துறையாயினும் வேறு எத்துறையாயினும், பாடநூல்களும் தமிழ்க்கலைச் சொற்களைப் பயன்படுத்தியே உருவாக்கப்பட வேண்டும்.

அறிவியல், தமிழ்வழி இருப்பின் அறிவியலும் வளர்ந்தோங்கும்! தமிழும் சிறந்தோங்கும்!

பார்வையில் பட்டவை :

1. Computer Dictionaryகணிப்பொறி அகராதி (English - Tamil) - இராம்குமார் ; திருநெல்வேலி தெ.சை.சி.நா.கழகம்
2. மருத்துவ அறிவியல் தொழில்நுட்ப கலைச்சொல் களஞ்சிய அகராதி - மணவை முசுதபா ; மணவை வெளியீடு
- 3.நெட்வொர்க் தொழில்நுட்பம் - மு.சிவலிங்கம் ;பாரதி பகத் பதிப்பகம்
4. சென்னைப் பல்கலைக்கழகத்தின் தமிழ்ப்பேரகராதி
5. தமிழ்இணையக்கல்விக்கழகத்தின் கலைச்சொற்கள் தொகுப்பு
6. ஐரோப்பிய அகராதி (<http://eudict.com/>)
7. தமிழ் கியூபு அகராதி (<http://tamilcube.com/>)

திறவூற்றுத் தமிழ் மென்பொருள்கள்:

மதிப்பீடும் எதிர்காலத் தேவையும்

முனைவர் இரா. பன்னிருகைவடிவேலன், முதுநிலைத் தொகுப்பாளர்,

தமிழ்ப் பேரகராதித் திருத்தப் பணித்திட்டம், சென்னைப் பல்கலைக்கழகம், சென்னை 5, Email: pann1973@gmail.com

திறவூற்று மென்பொருள் - வரையறை

இணையத்திலிருந்து இலவசமாகப் பதிவிறக்கம் செய்து கொள்வதற்கு உரியதும் மூலக் குறியீடுகளை ஒளிவு மறைவின்றித் தருவதுமான மென்பொருள் திறவூற்று மென்பொருள் ஆகும்.

திறவூற்று மென்பொருளுக்கு முதலில் 1998இல் ரேமண்ட் (Raymond) ஓர் அமைப்பை உருவாக்கினார். இதன் பிறகு பல நிறுவனங்கள் தமது மூலக்குறியீடுகளின் சில பகுதிகளை வெளியிடத் தொடங்கின. திறவூற்று மென்பொருள் வளர்ச்சியில் தொழில் நுட்பம் இருந்தும் வளர்ச்சி குறைவாகவே உள்ளது. திறவூற்று மென்பொருள் இயங்கு தளங்களாக ஜி என் யூ, (GNU Operating System) லினக்ஸ், உபுண்டு (Ubuntu Operating System), ஓபன் ஆபிஸ் (www.openoffice.org) போன்றவை உள்ளன.

திறவூற்று மென்பொருள்களாக மொழில்லா ஃபையர் பாக்ஸ் (Mozilla Firefox), கூகிள் குரோம் (Google Chrome), அண்ட்ராய்டு (Android) மற்றும் அப்பாச்சி (Apache) ஓபன் ஆபிஸ் சூட் (OpenOffice Suite) போன்றவை உள்ளன.

பயனர் அனைவருக்கும் ஏற்றதாகத் திறவூற்று மென்பொருள் அமைக்கப்பெறவேண்டும். திறவூற்று மென்பொருளைப் பயன்படுத்துவோர் அனைவருக்கும் பயன்படும் வகையில் மூலக் குறியீடு (Source Code) அமைந்திருக்க வேண்டும். இம்மென்பொருள் பயனர் அனைவரும் ஓர் இயக்கமாக அமைவர். இவ்வியக்கம் என்பது பரந்த தொலைநோக்கினை உடையதாகும்.

இணையத்திலிருந்து, திறவூற்று மென்பொருளை இலவசமாகப் பதிவிறக்கம் செய்து பயன்படுத்த முடிவதால் இம்மென்பொருளிற்குத் தொழில்நுட்ப ஆதரவு கிடைக்கிறது. திறவூற்று மென்பொருள் கற்றல் கற்பித்தலில் தெளிவான புரிதலுக்கு உறுதுணையாக அமைகிறது.

திறவூற்றுத் தமிழ் மென்பொருள் இயக்கத்தின் நோக்கம்

- கணினித் தமிழ் மொழியை மனித வள மேம்பாட்டிற்கு ஏற்பவும் பன்னாட்டுச் சமூக அமைப்பிற்கு ஏற்பவும் வளர்ச்சி அடையச் செய்தல்.
- பன்னாட்டு மொழிச் சூழலுக்கு ஏற்பக் கணினித் தமிழ் மொழியை வளர்த்தெடுத்தல்.
- கணினித் தமிழ் வளர்ச்சிக் கொள்கையை, உலகக் கணினி மொழி வளர்ச்சிக்கு இணையாக உருவாக்குதல்.
- கணினித் தமிழ் வளர்ச்சி வாரியம் அமைத்திட அரசை அறிவுறுத்தல்.

திறவூற்று மென்பொருள் உருவாக்கத்தில் உள்ள சிக்கல்கள்

- பயனாளர்-பயன்பாட்டுக்கு ஏற்றபடி மென்பொருள் உருவாக்கம்
- மொழிபெயர்த்தல்
- மென்பொருள் தேவைக்குத் தகுந்தபடி வடிவமைத்தல்

திறவூற்று மென்பொருளின் மொழி சார் சூழலில் (Language localization) மொழிபெயர்ப்புக் கருவி, பல மொழிகளின் அகரவரிசை, பல்வேறு விசைப்பலகைகள், நாள், தேதி, நேரம், எண் மற்றும் நாணயமதிப்பு இவை அனைத்தும் உள்ளீடு செய்யப்பட்டிருக்கும்.

மொழி மற்றும் மொழியியல் சார்ந்த சிக்கல்கள்

திறவூற்று மென்பொருள்களில் மொழிசார்கருவிகள் மிகவும் இன்றியமையாதவை. அவையாவன

- சொல்திருத்தி (Spell Checker)
- இலக்கணப் பிழை திருத்தி (Grammar Checker)
- சந்திப் பிழை திருத்தி (Junction Checker)
- பனுவல் பகுப்பாய்வுக் கருவி (Text Analyzer)
- வாக்கியப் பகுப்பாய்வுக் கருவி (Syntax Analyzer)
- பனுவல் உரை மாற்றி (Text to Speech Converter)
- உரை பனுவல் மாற்றி (Speech to Text Converter)

மேலே குறிப்பிட்ட இம் மொழி சார் கருவிகள் செயல்படுவதற்கு அடிப்படையாகத் தரவு மொழியியல் (Linguistic Courts) அமைகிறது. இது முக்கியமாகப் பயனர் இடைமுகம், ஆவண மொழிபெயர்ப்பு ஆகியவற்றை உள்ளடக்கியது.

- பண்பாடு சார்ந்த சிக்கல்கள்
- தொழில் நுட்பச் சிக்கல்கள்
- மென்பொருளின் தன்மை மற்றும் நோக்கம்
- சந்தை மற்றும் நுகர்வோர் அளவு
- சந்தை ஏற்பு
- மென்பொருளின் வாழ்நாள் மற்றும் மேம்பாட்டுக் காலம்
- போட்டியாளர்
- மாநில மற்றும் தேசியச் சட்டம், பகிர்வு, விற்பனை உரிமை

திறவூற்றுத் தமிழ் மென்பொருள்களின் முதன்மைக் கூறுகள்

- தரப்படுத்துதல்
- எழுத்துத் தொகுப்புகள் மற்றும் குறியீட்டு முறை
- ஒரு கீற்று (Byte) மற்றும் ஒன்றுக்கு மேற்பட்ட கீற்றுகள் (Byte) குறியீடுகளுக்கான இட ஒதுக்கீட்டு முறை
- பல்வேறு தரப்பட்ட குறியாக்க அமைப்புகள்
- குறியாக்கங்கள் மற்றும் உள்ளூராக்கல்
- எழுத்துருக்கள் மற்றும் வெளியீட்டு முறைகள்
- எழுத்துருக்களும் (Characters) கீற்றுகளும் (Bytes)

திறவூற்று மென்பொருள்களின் சாப்புகள்

உலகப்பொதுமை உடையது (Universality), மூலக்குறியீட்டின் உதவியால் மென்பொருளை மேலும் செம்மைப்படுத்த, மாற்ற, விரிவுபடுத்த, புதியன படைக்க உதவுவது, வணிக நோக்கமற்றது, பயன் பாட்டுக்கு ஊக்கமளிப்பது, மென்பொருளுலகில் காணலாகும் ஆதிக்கச் சக்திகளை ஒடுக்குவது, சுதந்திரம், நெகிழ்வுத் தன்மை, வெளிப்படைத்தன்மை கொண்டது கணினியிலிருந்து முக்கிய தகவல்களைக் களவாட முடியாதபடி அமைவது; நம்பகமானது, பாதுகாப்புத்தன்மை மிக்கது உள்ளிட்ட சாப்புகள் பல கொண்டவை திறவூற்று மென் பொருள்கள் ஆகும்.

கணினிப் பயன்பாட்டில் பன்னாட்டு நிறுவனங்களின் ஆதிக்கத்தைக் கட்டுப்படுத்த வேண்டும் என விரும்பும் தன்னார்வத் தொண்டர்கள் 'திறவூற்று' (open source) என்னும் இலட்சியத்தின் அடிப்படையில் பல மென்பொருள்களை உருவாக்கி இலவசமாகத் தருகிறார்கள் (www.opensource.org). இவை தமிழிலும் கிடைக்கின்றன. இத்தகு மென்பொருள்கள் வாயிலாகத் தன்னார்வத் தொண்டர்களும் கணினித் தமிழ் பரவ பெரும் உற்சாகத்தோடும் உறுதிப்பாட்டோடும் செயல்பட்டு வருகின்றனர்.

“இ- கலப்பை2.0 வின் வளர்ச்சியாக இ- கலப்பை 3.0 உருவாக்கப்பட்டுள்ளது. இது திறவூற்று மென்பொருளாக (Open Source) படைக்கப் பட்டுள்ளது. இதில் தமிழ் 99, ஒலிமுறை, தட்டச்சு, பாமினி, இன்ஸ்கிரிப்ட் ஆகிய விசைப்பலகைகளைப் பயன்படுத்தமுடியும்”.

லினக்ஸ் மென்பொருளின் மூலக் குறியீடும் வெளிப்படையாக்கப்படுவதால், எவரும் தங்கள் தேவைக்கேற்ப மாற்றியமைத்துப் பயன்படுத்த முடியும். சீன மொழியில் லினக்ஸைக் கொண்டு வரும் முயற்சி தீவிரமாக நடந்து வருகிறது. தமிழிலும் Mandrake என்ற பெயரில் லினக்ஸ் வரப்போகிறது. லினக்ஸ் மற்றும் விண்டோஸ் பயன்பாடு இணைந்த விண்டோஸ் ஏற்கெனவே வந்துவிட்டது. “விண்டோஸுக்கு மாற்றாக ஒரு தன்னார்வப் படைப்பை உண்டாக்கும் முயற்சியில் சீனா, ஜப்பான், தென்கொரியா ஆகிய மூன்று நாடுகளும் ஏற்கெனவே கைகோத்துவிட்டன. உலக நாடுகளின் அரசுகள் இப்போது மைக்ரோசாஃப்ட் நிறுவனத்தின் தயாரிப்புக்களையே பெரிதும் சார்ந்திருக்கின்றன. அடுத்த ஐந்தாண்டுகளில் இந்த இடத்தைத் தன்னார்வப் படைப்புக்கள் பிடித்துவிடும் என்று உறுதியாகச் சொல்கிறார் கொலம்பியா பல்கலைக்கழகப் பேராசிரியரும் தன்னார்வப் படைப்பியக்கத்தின் பொதுச் செயலருமான எபென் மாக்லென்”. (<http://www.columbia.edu/kernit/ftpscripts.html>)

உலகளவிலான கணினிச் சந்தை சில பன்னாட்டு நிறுவனங்களின் பிடியில் சிக்காமல் தடுப்பதுடன், நவீன தொழில்நுட்பத்தின் பலன்களைக் குறைந்த விலையில் எல்லோருக்கும் பரவலாக்கும் தன்னார்வப் படைப்புக்கள் உலகம் முழுவதும் பெரும் வரவேற்பைப் பெற்று வருகின்றன. இந்தியக் கணினிக் கட்டமைப்பை லினக்ஸ் சார்ந்து மாற்றுவதற்கான பணியை விரைவாகச் செய்ய வேண்டும் என்று புதிய பாரதத்துக்கான கனவு ததும்பிய அழைப்பை முன்னைக் குடியரசுத் தலைவர் கலாம் விடுத்திருக்கிறார். திறவூற்றுமென்பொருள் (Open Source Software) - தொழில்

நுட்பப் படைப்பாளிகளுக்கு ஒரு சமூக முகம் தேவை என்பதைப் பறைசாற்றுகிறது. இவ்வகையில் ஒருங்குறி (Unicode) நிறைய நற்பலன்களைக் கொண்டிருக்கிறது.

இது மரபு குலையாமல், உள்ளது உள்ளபடி தமிழைப் பாதுகாக்கும் அதிநவீன ஏற்பாடு. இணையத்தில் தமிழின் விரைவான முன்னேற்றத்துக்கு இம்முயற்சி வரப்பிரசாதமாக அமைந்திருக்கும். தன்னார்வப்படைப்புக்கள், பல்வேறு நாடுகளைச் சார்ந்த கணினி வல்லுநர்கள் தாங்களாகவே முன்வந்து ஒருவருக்கொருவர் மின்னஞ்சலில் கலந்துரையாடிப் பொதுப் பயன்பாட்டுக்காகத் தயாரித்தளிக்கும் மென்பொருள்களே திறவூற்று மென்பொருள்கள் (Open Source Software) எனப்படுகின்றன.

உலகம் முழுவதும் மைக்ரோசாஃப்ட் நிறுவனத்தின் விண்டோஸ் இயக்கத்தளமே பெருமளவில் பயன்படுத்தப்பட்டு வருகிறது. அதைத் தொடர்ந்து யூனிக்ஸ் இயக்கத்தளம் புழக்கத்தில் உள்ளது. யூனிக்ஸ் தயாரிப்பு நிறுவனத்திலிருந்து வெளியேறிய கணினி வல்லுநர்கள் லினக்ஸ் (Linux) என்ற புதிய இயக்கத்தளத்தைத் தயாரித்துப் பொதுப் பயன்பாட்டுக்காகப் புழக்கத்தில் விட்டார்கள். குறுகிய காலத்தில் மைக்ரோசாஃப்ட் மற்றும் பிற முதல் வரிசைப் பன்னாட்டு நிறுவனங்களை அசர வைக்கும் அளவுக்கு லினக்ஸின் பயன்பாடு மிக வேகமாகப் பரவியது. குறிப்பிட்ட இணையத் தளத்திலிருந்து யார் வேண்டுமானாலும் எடுத்துப் பயன்படுத்தக்கூடிய மென்பொருள் லினக்ஸ். விண்டோஸ் மற்றும் பிற நிறுவனத் தயாரிப்புக்கள் போலப் பெரும் விலை கொடுக்க வேண்டியதில்லை.

முக்கிய முடிவுகள் (Key Findings)

நிறுவனங்களில் பரந்துபட்ட நிலையில் உள்ள தரவகங்கள், அலைபேசி பயன்பாட்டில் திறவூற்று மென்பொருளைப் பயன்படுத்தல் என்பது குறைவாகவே உள்ளது. திறவூற்று மென்பொருளின் வளர்ச்சியை மேம்பாடு அடையச் சில கூறுகள் உள்ளன. அவை, அடுத்த ஐந்து ஆண்டுகளில், சற்றேறக்குறைய 50% வரை திறவூற்று மென்பொருள்கள் கையகப்படுத்திவிடும். 2015ஆம் ஆண்டிற்குள் இரண்டு வகையான நிலைகளைத் திறவூற்று மென்பொருள்கள் அடையும். முறைசார தொழில் நுட்பப் பிரிவு, உதரணமாக சுகாதாரம், பாதுகாப்பு, வாகனம் மற்றும் அரசாங்கம் போன்ற துறைகளில் திறவூற்று மென்பொருள்களின் வளர்ச்சி இருக்கும். மற்றும் 2012 முதல் தரவு மேலாண்மை (Data Management) பிரிவிலும் திறவூற்று மென்பொருளின் ஆதிக்கம் நிலவி வருகிறது.

திறவூற்று மென்பொருள்களைப் பயன்படுத்த முக்கிய காரணிகளாக உள்ளவை, இத்திட்டத்தின், பயன்பாட்டுத் தன்மைகளில் எளிமையும், காலத்திற்கு ஏற்ப மென்பொருளின் புதுப்பித்தல் (Update) மற்றும் சேவைகளுக்காகவும் (Service), திறவூற்று மென்பொருள்கள் வளர்ச்சி அடைகிறது. திறந்த மூல மென்பொருள் (திறந்த மூலக் குறியீடு கொண்ட மென்பொருள்) கடந்த சில ஆண்டுகளில் மிகவும் பிரபலமாக உள்ளன. திறவூற்று மென்பொருள் வர்த்தக நிறுவனங்களுக்கு எதிரான நிலையில் இருப்பதாகக் கருதிய நிலை மாறி, திறவூற்று மென்பொருள்கள் விளம்பரங்களிலும், போட்டியாளர்கள் மற்றும் வணிகப் பகுதியாகச் சிறந்து விளங்குகிறது.

இந்த ஆர்வத்தில் கூட திறவூற்று மென்பொருள்கள் உலகில் பரவி வருகிறது. இது ஏகபோக உருவாக்கத்தை எதிர்க்கும். அது, மென்பொருள் தயாரிப்பாளர்கள் சுதந்திரம் வகைப்படுத்தப்படும். திறந்த மூல மென்பொருள் இப்போது தான் அதிகரித்து, வர்த்தக நோக்கங்களுக்காகப் பயன்படுத்தப்படும், மற்றும் அதை ஒரு 'தடையற்ற பொறுப்புக் கொள்கையாக' (Free of Charge Principle) வகைப்படுத்தப்படும் என்பதால், திறந்த மூல மென்பொருள் தேசிய மற்றும் சர்வதேச அளவில், அரசியலாக்கப்பட்டுப் பெரிய தாக்கத்தை ஏற்படுத்துகிறது.

1. <http://www.projectmadurai.org/>
2. <http://www.tamilheritage.org>
3. <http://infitt.org/>
4. <http://repository.nie.edu.sg/jspui/simple-earch?query=thinnappan&submit=Go>
5. <http://www.openoffice.org/>
6. <http://www.linux.org/>
7. <http://www.ubuntu.com/>

தமிழ் இணையத்தில் புதிய அலை

முனைவர் அண்ணாகண்ணன்
(annakannan@gmail.com | 08553004475)

தமிழ் இணைய வெளியின் பரப்பும் அடர்த்தியும் தொடர்ந்து அதிகரித்து வருகின்றன; ஆழம் கேள்விக் குறியே ஆயினும், இதன் முக்கியத்துவம் பெருகி வருகின்றது.

இணைய இதழ்களின் பரிணாம வளர்ச்சி

இணையப் பக்கங்கள், இணைய இதழ்களாக வளர்ந்தன; இணைய இதழ்கள், ஒற்றைத் தன்மையிலிருந்து, பன்முகத்தன்மை கொண்டவையாக மாறின. அத்துடன் நில்லாமல், வலைப்பதிவுகளும் நுண் வலைப்பதிவுகளும் பிறந்தன. இந்த நுண் வலைப்பதிவுகள் பின்னர், சமூக ஊடகங்கள் எனப் பெயர் பெற்றன. இவை, தனி மனித இதழ்கள் எனக் கருதத் தக்கவை. எழுத்து மட்டுமே பெரும்பான்மையாக இருந்த இணையத்தில், நிழற்படங்களும் ஒலி - ஒளித் தரவுகளும் பெருகத் தொடங்கி, பல்லாடக உள்ளடக்கங்கள் வளர்ந்து வருகின்றன. இந்தப் பின்னணியில் இணைய இதழ்களை விரிவான கண்ணோட்டத்துடன் வரையறுத்தல் வேண்டும்.

எவர் ஒருவர் உள்ளடக்கத்தை உருவாக்குகின்றாரோ, அவர் உள்ளடக்க உற்பத்தியாளர் எனக் கருதத் தக்கவர். இவற்றைப் பெறுவோர், பயனர் ஆவர். இந்த உள்ளடக்கம் வெளியாகும் இடங்கள், இணையத் தளங்களாகவோ, இணைய இதழ்களாகவோ, வலைப்பதிவுகளாகவோ, சமூக ஊடகங்களாகவோ அமைந்திருக்கலாம். பெயரளவுக்கு இவை வெவ்வேறாக இருப்பினும் இவை ஒன்றுக்கொன்று தொடர்புடையவை. ஒன்றின் தொடர்ச்சியாக, இன்னொன்று தோன்றியவை. எனவே இவை அனைத்தையும் உள்ளடக்க வெளி எனக் குறிப்பிடுவது பொருந்தும்.

தமிழ் இணையத்தில் புதிய அலை

இந்தியாவில் உள்ள இணையப் பயனர்களின் எண்ணிக்கை, 15 கோடியாக உள்ளது. இது, 2013 இறுதிக்குள் 18 கோடியாக வளரக் கூடும் (2013 India Internet outlook - <http://techcircle.vccircle.com/2013/02/01/2013-india-internet-outlook/>). மேலும், ஃபேஸ்புக் தளத்தின் இந்தியப் பயனர்களின் எண்ணிக்கை, 7.8 கோடியாக உயர்ந்துள்ளது (Facebook India Active Users Up 50% To 78M In Q1-2013 - <http://www.medianama.com/2013/05/223-facebook-india-78-million-users/>). இந்தியப் பயனர்களின் எண்ணிக்கையில் சரி பாதிக்கும் மேல், ஃபேஸ்புக் தளத்தில் உறுப்பினர்களாக உள்ளனர். இது, சமூக வலைத்தளங்களின் மீது மக்களுக்கு உள்ள ஈர்ப்பினைத் தெள்ளத் தெளிவாகக் காட்டுகிறது. இதில் தமிழர்களின் பங்களிப்பும் குறிப்பிடத்தக்க அளவு உள்ளது.

இணைய வெளியில் பல்வேறு தேசிய, அரசியல், சமூக, தொழில், கல்விப் பின்புலங்களைக் கொண்ட படைப்பாளர்களும் பயனர்களும் தொடர்ந்து தோன்றிய வண்ணம் உள்ளனர். இதன் நல்விளைவாக, தமிழுக்குப் புதிய புதிய ஆக்கங்கள் அதிகம் கிடைத்து வருகின்றன. இது வரை வெளிப்படாத புதிய பார்வைகளும் அனுபவங்களும் நுணுக்கமான கருத்துகளும் பதிவாகி வருகின்றன. அன்றாட உலக நடப்புகளும் செய்திகளும் அவற்றின் மீதான விமர்சனங்களும் உடனுக்குடன் பதிவாகின்றன.

தமிழின் முக்கியமான எழுத்தாளர்கள், கலைஞர்கள், அரசியல் புள்ளிகள் எனப் பலரும் இணையப் பரப்பிற்கு வந்துள்ளதோடு, தங்களுக்கெனத் தனித் தனி வலைப்பதிவுகளையும் இணையத்தளங்களையும் உருவாக்கி வருகின்றனர். இணையப் பரிச்சயம் இல்லாத மூத்த எழுத்தாளர்கள் பலருக்கும் அவர்கள் சார்பில் யாரேனும் அவர்களின் ஆக்கங்களை இணையத்தில் வெளியிட்டு வருகிறார்கள். இது, இணையத்தில் தங்கள் இருப்பை உறுதி செய்ய வேண்டிய கட்டாயம், தமிழ்ப் படைப்பாளர்களுக்கு ஏற்பட்டிருப்பதை எடுத்துக் காட்டுகிறது.

பிரபலமானவர்கள் மட்டுமின்றி, இதுவரை வெளியுலகிற்கு வராத பெண்களும் மாணவர்களும் எளிய பணிகளில் இருப்போரும் இணைய வெளிக்கு வந்துள்ளனர். இந்த வகையில் தமிழ் இணையம், இது வரை காணாத புதிய பயனர்களையும் புதிய துறைகளையும் காணத் தொடங்கியுள்ளது. இது, பல்வகையான பரிமாணங்களுடன் நாளும் விரிந்து வருகிறது.

இலவசம் என்னும் கவர்ச்சி

தமிழ் இணையப் பயனர்களிடையே இவ்வளவு ஆர்வமும் துடிப்பும் எழுந்துள்ளதற்கு முக்கிய காரணங்களுள் ஒன்று, இணையத் தளங்களில் இலவசமாகக் கணக்குத் தொடங்கி, எவ்வளவு வேண்டுமானாலும் பயன்படுத்தலாம் என்ற வாய்ப்பு.

வலைப்பதிவுகள் தோன்றிய காலத்தில் பயனர்கள் பலரும் அதில் தங்களுக்கெனத் தனி வலைப்பதிவுகளை உருவாக்கினர். ஒருவரே பல வலைப்பதிவுகளை உருவாக்கினார். அந்த அலை, இப்பொழுது அப்படியே சமூக ஊடகங்களில் எழுந்துள்ளது. இணையத்திற்குப் புதிதாக அறிமுகமாகும் ஒருவர், மின்னஞ்சல் கணக்குத் தொடங்குவதோடு, ஃபேஸ்புக் போன்ற சமூக வலைத்தளங்களிலும் கணக்குத் தொடங்குகின்றனர். செல்பேசிகளில் படங்கள் எடுத்து, அவற்றை ஃபிளிக்ர் போன்ற புகைப்படத் தளங்களில் பகிர்ந்து வருகின்றனர். விழியங்கள் எடுப்போர், அவற்றை யூடியூப் போன்ற விழியத் தளங்களில் வெளியிடுகின்றனர். இவ்வாறாக, எழுத்து, புகைப்படம், விழியம் ஆகிய கூறுகளில் பல்லாடக உள்ளடக்கம், இணையத்தில் பெரும் பாய்ச்சலைக் கண்டு வருகின்றது.

புகழ் வேட்கை

தமிழ் இணையப் பயனர்களைத் தொடர்ந்து உந்திச் செலுத்தும் முக்கிய விசைகளுள் ஒன்று, , அவர்களின் புகழ் வேட்கை. இணையத்தில் கிடைப்பது 15 நிமிடப் புகழ் என முன்பு எழுத்தாளர் சுஜாதா குறிப்பிட்டார். ("இப்போது புதிதாக பிலாக்ஸ் என்று வந்திருப்பது ஒரு விதத்தில் சின்ன வயசில் நாங்கள் எல்லோரும் நடத்திய கையெழுத்துப் பத்திரிகையின் மறுவடிவம்தான். 'இதோ பார் என் கவிதை' 'இதோ பார் என் கருத்து' 'இதோ பார் உலகுக்கு என் உபதேசம்' என நானும் இருக்கிறேன் நண்டுவளையில் என்று. ஒவ்வொரு தனிமனிதனும் தன்னைப் பிரகடனப்படுத்த கைகளைக் குவித்து ஏதோ ஒரு திசையில் குரல் கொடுத்துவிட்டு, யாராவது பதில் தருகிறார்களா என்று அவரவர் பதினைந்து நிமிஷப் புகழுக்குக் காத்திருப்பதுதான் இது! பிரபஞ்சத்தில் வேற்று கிரகங்களில் உயிரினங்கள் இருக்கிறதா என்பதை இப்படித்தான் தேடுகிறார்கள்" - சுஜாதா - http://agaramuthala.blogspot.in/2004/08/blog-post_24.html) அது, இப்போது, ஓரிரு நிமிடங்கள் தாக்குப் பிடிப்பதே அரிதாக இருக்கிறது. ஏனெனில் நொடிக்கு நொடி, புதிய புதிய உள்ளடக்கங்கள் பெருகி வரும்போது, ஒவ்வொன்றையும் ஆழ்ந்து படிப்பதற்குப் பயனர்களுக்கு நேரமும் பொறுமையும் இருப்பதில்லை. எனவே, அந்தந்தப் படைப்பாளர்களின் புகழ் ஈட்டும் அவகாசம் குறைந்துவிடுகிறது.

ஆயினும், தான் கவனிக்கப்பட வேண்டும் என்ற எண்ணம் அவர்களிடையே வலுப்பெறுகிறது. இதுவே, அவர்கள் இணையத்தில் மிக அதிகமான நேரத்தைச் செலவிடத் தூண்டுகிறது. நட்பையும் உறவையும் புதிய செய்திகளையும் வாய்ப்புகளையும் தேடி வருகின்ற பயனர்கள், இயல்பாகவே, புகழை விரும்புகின்றனர். தாங்கள் கவனிக்கப்பட வேண்டும் என்பதில் இவர்களுக்கு உள்ள ஆர்வம், நாளுக்கு நாள் உச்சமடைந்து, ஒரு கட்டத்தில் இணைய வெளிக்கு இவர்களை அடிமையாக்கிவிடுகிறது.

பயனராகத் தன் பயணத்தைத் தொடங்கும் ஒவ்வொருவரும் தாங்களே படைப்பாளராகப் பரிணமிக்கின்றனர். இதன் மறுபக்கமாக, உள்ளடக்கம் தீர்ந்துவிடுகிற நேரத்திலும் எதையாவது வெளியிட வேண்டிய உந்துதல் இவர்களுக்குப் பிறக்கிறது. இது, ஆழமற்ற உள்ளடக்கத்துக்கு இட்டுச் செல்வதோடு, மொக்கைகளும் மரண மொக்கைகளும் பின்னுட்டக் கும்மிகளுமாக வளர்ச்சி அடைகிறது.

இணையத்திலிருந்து அச்சுக்கு

அச்சில் வெளியானவற்றை இணையத்தில் ஏற்றுவது, ஒரு புறம் நடந்துகொண்டிருக்கிறது. அதே நேரம், இணையத்தில் வெளியானவற்றிலிருந்து தேர்ந்தெடுத்து, அச்சில் வெளியிடுவதும் பெருகி வருகிறது. ஆனந்த விகடன், குமுதம், குங்குமம், பாக்யா உள்ளிட்ட பல்வேறு அச்சிதழ்களில் இந்தப் போக்கினைக் காணலாம். அத்துடன் இணையத்தில் எழுதத் தொடங்கிய பலரும், அச்சிதழ்களிலும் எழுதி வருகின்றனர். மேலும் டிவிட்டர், பேஸ்புக் உள்ளிட்ட சமூக ஊடகங்களிலும் இணைய இதழ்களிலும் பலர் வெளியிடும் கருத்துகள், பரபரப்பை ஏற்படுத்துவதோடு, அச்சிதழ்களில் அவை செய்திகளாக இடம் பிடிக்கின்றன. இவை, இணையத்தளங்கள், சமுதாயத்தின் மீது செலுத்தும் தாக்கத்திற்கு ஒரு சான்றாகும்.

குறையும் நம்பகத்தன்மை

இணையப் பரப்பில் வெளியாகும் ஆக்கங்கள், பெரும்பாலும் தாங்கள் முன்வைக்கும் செய்திக்கும் தகவலுக்கும் உரிய சான்றுகளை முன்வைப்பதில்லை. அந்தச் செய்தியை வழங்கிய செய்தியாளர் யார், முகாமை நிறுவனம் எது, தகவலைத் தெரிவித்தவர் யார்? போன்ற அடிப்படைத் தகவல்கள் கிடைப்பதில்லை. எனவே, யூகத்தின் பேரிலும் கற்பனையாகவும் செவிவழி அறிந்தவையாகவும் செய்திகள் வெளியாகும் சாத்தியங்கள் அதிகரித்துள்ளன. நடுநிலைத் தன்மையுடன் செய்திகளை அணுகாமல், ஒரு செய்தியில் தங்களுக்குத் தேவையான பகுதியை மட்டும் வெளியிடும் வாய்ப்பும் உள்ளது. இதனால் வதந்திகள் பரவிடவும் சமூக இயல்பு நிலை கெடுவதற்கும் அதிக வாய்ப்புகள் உள்ளன.

அவதூறுகள் - தாக்குதல்கள்

இணைய வெளியில் அவதூறுகளுக்கும் தாக்குதல்களுக்கும் அதிகப் பார்வையாளர்கள் கிடைக்கின்றனர். எனவே, பிரபலம் அடைய வேண்டும் என விரும்புவோரும் அடைந்த பிரபலத்தைத் தக்கவைக்க விரும்புவோரும் இவற்றில் அதிக முக்கியத்துவம் அளித்து வருகின்றனர். இது, பல்வேறு குழுக்களை உருவாக்கி வருவதோடு, நேரத்தைப் பெருமளவு வீணடிப்பதாகவும் மன உளைச்சலைப் பெருக்குவதாகவும் அமைகின்றது.

கருத்துச் சுதந்திரத்துக்குப் பாதிப்பு

இணையப் பரப்பில் வெளியாகும் பல்வகை ஆக்கங்களைப் பலரும் மவுனமாகத் துய்ப்பதோடு நின்றுவிடுவர். துய்த்த அனைவரும் பின்னூட்டம் இடுவதில்லை. ஆயினும் பயனருள் சிலர், சற்று சுதந்திர உணர்வுடன் தங்கள் கருத்துகளை வெளிப்படுத்தி வருகின்றனர். சர்ச்சைக்கு உரிய ஆக்கங்களில் கார சாரமான பின்னூட்டங்கள் வரத் தொடங்கி, பின்னர் அது, பரவத் தொடங்குகிறது. இதன் நீட்சியாக, தரக் குறைவான, பொருத்தமில்லாத பின்னூட்டங்கள் மட்டுமின்றி, எரித (ஸ்பாம்) மறுமொழிகளும் மிகுதியாக வருகின்றன. இதனால், இணைய வெளியில் பின்னூட்டங்கள் பெரும்பாலும் மட்டுறுத்தப் பெறுகின்றன. மேலும் பல இணையத்தளங்களில் மறுமொழி இடும் வாய்ப்பு, முடக்கப்பெற்றுள்ளது. இத்தகைய நடவடிக்கைகள், இணைய இதழ்களில் கருத்துப் பரிமாற்றத்தை மட்டுப்படுத்துகின்றன. மேலும் வெளிப்படாததன்மை குறைந்து, வாசகர்களின் சுதந்திரம் குறைந்துவிடுகிறது. இணைய இதழ்கள், இருவழிப் பாதையாக இல்லாமல், ஒருவழிப் பாதையாகக் குறுகி வருகின்றன.

பொதுவான இணைய இதழ்களிலும் பிறரின் வலைப்பதிவுகளிலும் பக்கங்களிலும் கருத்துகளைப் பதிவதற்குக் கட்டுப்பாடுகள் உள்ளன. எனவே, ஒருவர் தன் தளத்தில், வலைப்பதிவில், பக்கத்தில் தாம் விரும்புவதைப் பதிய முனைகின்றார். ஆனால், இங்கும் எதை வேண்டுமானாலும் எழுதலாம், பகிரலாம் என்ற நிலை இல்லை. பொய்யாகவும் அவதூறாகவும் அடுத்தவர் மனம் புண்படும் வகையிலும் பதிவாகின்ற உள்ளடக்கங்களை வெளியிடுவோர், வழக்குகளைச் சந்திக்க வேண்டிய நிலை உள்ளது. இதற்கென இந்தியாவின் தகவல் தொழில்நுட்பச் சட்டத்தில் 66A என்ற திருத்தம் மேற்கொள்ளப்பெற்றது. இந்தச் சட்டத்தை நடைமுறைப்படுத்துவதில் பல்வேறு சர்ச்சைகள் எழுந்தவண்ணம் உள்ளன.

அறிவுச் சொத்துகளுக்கு ஆபத்து

இணைய இதழ்களில் செய்திகள் தொடங்கி, பல்வகை படைப்பிலக்கியங்களும் புகைப்படங்களும் பல்லாடக உள்ளடக்கங்களும் இடம்பெற்று வருகின்றன. இவற்றைத் துய்க்கும் பயனர்கள், தங்கள் தளங்களிலும் சமூக ஊடகங்களிலும் அவற்றைப் பகிர்ந்து வருகின்றனர். ஆனால், மூலத் தரவுகளை எங்கிருந்து பெற்றோம் என்ற அடிப்படைத் தகவலைக் கூட அறிவிக்காமல், அவரவர் விருப்பத்துக்கு வெளியிடுகின்றனர். சிலர், அவற்றைத் தங்கள் பெயரிலேயே வெளியிடுவதும் நிகழ்கிறது. இதனால், இணையப் படைப்பாளிகளின் அறிவுச் சொத்துகளுக்கு ஆபத்து ஏற்பட்டுள்ளது.

போலி அடையாளங்கள்

இணையத்தில் பொய்யான பெயருடன், படத்துடன், போலி அடையாளத்துடன் வலம் வருவோர், அதிகரித்து வருகின்றனர். பெண்களின் பெயர்களில் / படங்களில் கணக்குத் தொடங்குவோர் பலர். திரை நட்சத்திரங்களின் பெயர்களில் உள்ள கணக்குகள் பலவும் அவர்களுடையவை அல்ல. இது தொடர்பாகப் பலரும் காவல் துறையில் புகார் அளித்துள்ளனர். பண மோசடி உள்பட பல்வேறு குற்றங்களும் இணையத்தில் நடைபெற்று வருகின்றன. எனவே, பயனர்கள், இத்தகைய விவகாரங்களில் மிகுந்த எச்சரிக்கையுடன் செயல்பட வேண்டும். தங்கள் தனிப்பட்ட விவரங்களை அளிப்பதற்கு முன், அவர்கள் நம்பகமானவர்களா என உறுதி செய்துகொள்ள வேண்டும்.

பயனர்களின் கவனத்திற்கு

ஆண்டுதோறும் மாதம்தோறும் நாள்தோறும் பொழுதுதோறும் இணையப் பயனர்கள் அதிகரித்து வருகின்றனர். முழு மக்கள் தொகையும் இணையத்தில் தங்களை இணைத்துக்கொள்ளும் காலம், வெகு தூரத்தில் இல்லை. இலவச மடிக்கணினிகளும் இணையக் கட்டணங்கள் குறைப்பும் இணைய வேகம் அதிகரித்து வருவதும் சமூகத்தில் பெரும் மாற்றங்களை விரைவில் நிகழ்த்த உள்ளன.

இந்தச் சூழ்நிலையில், இணையப் பயனர்களின் உரிமைகளையும் கடமைகளையும் பொறுப்புணர்வையும் எச்சரிக்கை உணர்வையும் அவர்களுக்கு உணர்த்த வேண்டிய தேவை எழுந்துள்ளது. செய்யத் தக்கவை, செய்யத் தகாதவை ஆகியவற்றில் அவர்களுக்குத் தெளிவு தேவை. தங்கள் செயல்பாடுகள் ஏற்படுத்தக் கூடிய விளைவுகளை இணையப் பயனர்கள் அவசியம் உணர வேண்டும். இணையத்தில் தங்கள் நேரத்தை வீணடித்துவிடாமல், அதே

இணையத்தைக் கொண்டு, தங்கள் நேரத்தைச் செம்மையாகப் பயன்படுத்துவது எப்படி என்பதை அவர்கள் கற்க வேண்டும்.

இதகைய பண்புகள் கொண்ட தமிழர்கள், தமிழ் இணையத்திற்கு அளப்பரிய பங்காற்றிட முடியும். அதன் மூலம், தமிழ் மொழியும் தமிழ்ச் சமூகமும் புதிய உயரங்களை எட்ட முடியும். உலகம் முழுவதும் பெரும் மறுமலர்ச்சியையும் உயிரோட்டத்தையும் தோற்றுவிக்க முடியும். இவை அனைத்தும் அலை அலையாய்ப் புறப்படும் புதிய பயனர்களின் கைகளிலேயே உள்ளது.

மின்தமிழ் வரலாற்றில் தமிழ் மரபு அறக்கட்டளையின் சுவடுகள்

நா.கண்ணன், சுபாஷினி ட்ரம்மெல்

தமிழ் மரபு அறக்கட்டளை (www.tamilheritage.org)

முன்னுரை:

கணினி பாவிப்பு பரவலாகும் போது ஆக அது குறித்த கனவுகளும் பரவலாகி வருவது காணக்கிடைக்கிறது. உதாரணமாக, ஒளி நகல் எனும் தொழில்நுட்பம் ஒளி இலக்க நகல் என்று உருமாறி காகிதம் என்பதைத் தவிர்த்து ஒரு பௌதீகப் பொருளின் நகலை இலக்க வடிவில் நேரடியாக கணியில் சேமித்து வைத்து, வேண்டிய போது வேண்டிய போது மூலத்தைப் பயன்படுத்துவது போல் பயன்படுத்தலாம் என்ற நிலை வரும்போது தனி மனிதரிலிருந்து பல்கலைக்கழகங்கள் வரை தம்மிடமுள்ள மரபுச் செல்வங்களை இலக்கவடிவில் சேமித்து வைக்கும் முயற்சியில் ஈடுபடலாயினர். சில வருடங்களுக்கு முன் சிந்தனையாக இருந்த ஒன்று இன்று திட்டங்களாக மாறி பல்வேறு வடிவமெடுத்திருப்பது கண்கூடு. உதாரணங்கள், மதுரை மின்பதிப்புத்திட்டம் (<http://www.projectmadurai.oeg>), இணையக் கல்விக்கழகம் (<http://www.tamilvu.org>), இந்திய இலக்க நூலகம் (<http://www.dli.gov.in/>), தமிழ் மரபு அறக்கட்டளை (<http://www.tamilheritage.org>), போன்றவை.

இதில் முதற் சொன்ன மூன்று திட்டங்களும் தமிழ் எழுத்தை (e-text) முதன்மையாக மின்னாக்கம் செய்வதில் கருத்தாயுள்ளன. ஆயின் தமிழ் மரபு அறக்கட்டளை மரபு சார்ந்த ஆவணங்கள் எவ்வடிவில் இருந்தாலும் அவைகளை மின்னாக்கம் செய்ய வேண்டுமென்று எண்ணுகிறது. 'புத்தக-லட்சம்' (<http://www.archive.org/texts/collection.php?collection=millionbooks>) எனும் திட்டம் செயல் வடிவு கொண்டவுடன் அதன் நீட்சியாக இந்திய இலக்க நூலகம் உருவாகிறது, சுவடிகள் சேமிப்பு உள்ள காஞ்சி மடம், அது நடத்தும் சாஸ்திராக் கல்லூரி மற்றும் சில தொழில்நுட்பக் கல்லூரிகள் இத்திட்டத்தில் பங்கேற்கின்றன. தனியார் முயற்சியாக தமிழம்.வணி (www.thamizham.net), சென்னைவலை.வணி (<http://www.chennainetwork.com>), நூலகம் (<http://www.noolaham.org/>) போன்றவை உருவாகியுள்ளன. இது தவிர வணிக இதழ் பதிப்பகத்தார் தங்கள் இதழ்களை மின்பதிப்பாக்கத் தொடங்கியுள்ளனர். இது தவிர குறைந்த அளவில் தனியார் வலைப்பக்கங்களிலும், வலைப்பதிவுகளிலும், தமிழ் மின்னாடற்குழு பிரத்தியேக பக்கங்களிலும் தமிழ் இலக்கிய மின்பதிவுகள் காணக்கிடைக்கின்றன.

சமகாலச் சூழலும். பின்னடைவுகளும்:

இவையெல்லாம் வரவேற்கப்படவேண்டிய முயற்சிகள். ஆயினும் இம்முயற்சிகள் பற்றிய சேதிகள் இன்னும் பரவலாக தமிழ்மக்களை அடையவில்லை. இத்திட்டங்களில் தமிழர்களின் பங்கேற்பு என்பது மிகக்குறைவாகவே உள்ளது. இன்னும் கணினி பாவிப்பு இந்தியத்தமிழர்களிடம் ஆழமாய் பரவவில்லை. அதன் அளப்பரிய பயன் குறித்த சிந்தனைகளும் பரவவில்லை. இதனால் பாரிய அளவில், இத்திட்டங்கள் எதிர்பார்க்கும் அளவில் பொதுமக்களின் பங்கேற்பு என்பது இல்லை. ஆகையால், பாதுகாக்கப்பட்டிருக்க வேண்டிய புத்தகங்கள் பரணில் வீணாக பூச்சிக்கு இரையாகிக் கொண்டிருக்கின்றன. அமெரிக்க இந்தியரான முனைவர் நா.கணேசன் சமீபத்தில் ஒரு வேண்டுகோள் விடுத்தார், “தமிழில் 6 லட்சம் விருத்தம் அழியும் நிலையில் உள்ளது, படிப்பாரும் சுவைப்பாரும் இன்றி. 3-4 ஆயிரம் பிரபந்தங்கள், 700-1000 தலபுராணம் - கருத்துக்கள் பழைமை, பௌராணிகம் என்றாலும் எத்தனை அறியாச் சொல்லாட்சி, பொருளாட்சி இருக்கும். அகராதிகளில் ஏற்ற வேண்டியவை எத்தனை? எண்ணிப்பாருங்கள். 1835-1940 வரை அச்சான புத்தகங்கள் அப்பளம் போல் நொறுங்கியும், போற்றாமையாலும் அழிபடுகின்றன. கணினி ஒரு வரப்பிரசாதம். வருடி அலகிட்டு, தட்டிடித்தும், சுவைத்து உரையாடவும், பேணவும் புதுப்பொருள், விஞ்ஞானம் போன்றவற்றைப் பெருக்கவும் உதவுகிறது அன்றோ?” என்று. நிலமை இவ்வாறு இருப்பினும் புத்தகங்கள் உள்ளவர்கள் கூட துணிந்து தங்கள் புத்தகங்களை இலாக்கப்பதிவாக்குவதில்லை. உதாரணமாக புத்தக-இலட்சம் திட்டம் எதிர்பார்த்த அளவு தமிழ் புத்தகங்கள் இலக்கப்பதிவிற்கு வந்து குவியவில்லை. காரணங்கள்?

1) புத்தக-இலட்சம் திட்டம் பல்கலைக்கழகங்கள் வழியாக நடைபெறுகிறது. காலம் காலமாக பல்கலைக்கழகங்கள் பொது நீரோட்டத்தில் கலந்து கொள்ளாமல் ஒரு உயர் மட்டத்தில் தங்களை வைத்துக்கொண்டுள்ளன. இவர்களுக்கும் பொது மக்களுக்கும் உள்ள உறவு தாமரை இலைத்தண்ணீர் போன்றது. எனவே இவர்களை நம்பி பொது மக்கள் தங்கள் புத்தகங்களைத் தருவதில்லை. மேலும், பல்கலைக்கழகப் பேராசிரியர்களும் இதற்காக பிரயத்தனப்படுவதுமில்லை. உ.வே.சா போன்று களப்பணி செய்த பேராசிரியர் கூட்டம் வெகுவாகத்தமிழ் மண்ணில் குறைந்துவிட்டது. எல்லோரும் குளிர் பதனப்படுத்தப்பட்ட அறைகளிலிருந்து தமிழாராய்ச்சி செய்யவே ஆசைப்படுகின்றனர்.

2) அரசு நூலகங்களில் ஆயிரக்கணக்கான ஓலைச் சுவடிகளும், புத்தகங்களும் உள்ளன. ஆனால், அங்குள்ள நூலகக்காப்பாளர்கள் “பூதம் காத்த பொக்கிஷம்” போல் இவைகளை ‘இருந்தது, இருந்தபடி’ வைத்திருக்கவே ஆசைப்படுகின்றனர். இவையெல்லாம் இலக்கப்பதிவாகி பொது மக்கள் பார்வைக்கு வந்துவிட்டால் எங்கே தங்கள் வேலைக்கே உலை வந்துவிடுமோ என்ற பயமாகவும் இருக்கலாம். இக்கட்டுரை ஆசிரியர் வித்தியாசமான ஒரு உளவியலைக் கண்டு இ-சுவடி மடலாடற்குழுவில் எழுதினார் (<http://groups.yahoo.com/group/esuvadi/>). அதாவது, பண்டைய ஓலைச் சுவடிகளிலுள்ள இரசவாதச் சூத்திரங்களை வாசித்துவிட்டு, பித்தளையைத் தங்கமாக்கி பணக்காரரான வாசகர்கள் உண்டென நம்பி, இந்த நூலகக் காப்பாளர் பல அரிய சுவடிகளை யாருக்கும் வாசிக்கத்தருவதே இல்லை! அதாவது, ‘நாய் காத்த வைக்கோல்போர்’ போல் தனக்கும் உதவாமல். பிறர்க்கும் உதவாமல் இந்த சுவடிகளை இவர்கள் ‘பாதுகாத்து’ வருகிறார்கள்!

3) இவர்கள் பொதுவாகத்தரும் சாக்கு, இந்நூலகங்களே அவைகளைப் பதிப்பித்துவிடும் என்பது. அவர்கள் பதிப்பிக்கிறார்கள் என்பதில் சந்தேகமில்லை, ஆயின், அரசு ஊழியர்களுக்கேயுள்ள மெத்தனத்துடன் ஆண்டுகள் பல எடுத்துக்கொண்டு படியெடுத்து, படியெடுத்து (அதாவது ஒரு மேசையிலிருந்து ஒரு மேசைக்கு நகர சில மாதங்கள் என்ற கணக்கில்) இவர்கள் பதிப்பிப்பதற்குள், பதிப்பிக்கப்பட வேண்டிய நூல்கள் பரலோகம் போய்விடுகின்றன! அப்படிப் பதிப்பிக்கப்படும் நூல்களும் மிக மட்டமான தாளில், அசிரத்தையாக மை கசிய பதிப்பிக்கப்பட்டு அப்புகளை பதிப்பித்த சில ஆண்டுகளுக்குள் பாதுகாக்க வேண்டிய நிலைக்குத்தள்ளி விடுகிறார்கள்.

4) இவர்களிடமிருந்து ஒரு நூலை வாங்குவதற்குள் ஆய்வாளன் மனம் தளர்ந்துவிடுவதாக பேராசிரியை சே.பிரேமா அவர்கள் இக்கட்டுரை ஆசிரியருக்கு அளித்த பேட்டியில் கூறுகிறார் (இப்பேட்டி, தமிழ் மரபு அறக்கட்டளை வெளியிட்டுள்ள ‘தமிழகம் கண்ட மருத்துவமுறை’ எனும் திகிரி (CD)யில் உள்ளது - <http://bharani.dli.ernet.in/thf/SALES/cdprod.html>). அரசு அதிகாரிகள் காலம் காலமாக உயர் அதிகாரிகளுக்கு மட்டுமே செவி சாய்ப்பதாக ஒரு பழக்கத்தை (பயம் காரணமாக) ஏற்படுத்திக்கொண்டிருப்பதால் இலக்கப்பதிவுத் திட்டங்களுக்கு அரசின் அதிகாரிகளின் துணையுடனே செல்ல வேண்டியுள்ளது. இவர்கள், ஆய்வாளர், பேராசிரியர்கள் பேச்சைக் கேட்பதில்லை. ஆனால், ஆய்வாளர்களுக்கோ அரசியல் எட்டிக்காய் போல் கசப்பதால் பல நூல்கள் பாதுகாக்கப்படாமல் அழியும் நிலையிலுள்ளன. அரசியல்வாதிகளுக்கோ இம்மாதிரித்திட்டங்களினால் தமக்கு ஆதாயம் அதிகமில்லை என்பதால் கண்டு கொள்வதே இல்லை என்பதும் யதார்த்தம்! ஆனால், அரசு நூலகங்களில் புதிய புத்தகங்கள் வரும் போது பழைய புத்தகங்களை வைப்பதற்கு இடமில்லையென்று அவைகளை எரித்துவிடுவதாகப் பத்திரிக்கைச் சேதிகளைக் கண்ணுறும் போது கண்ணீர் சிந்தாமல் இருக்க முடிவதில்லை! இதே போல், அரசு சார்ந்த வானொலி நிலையங்களில் பழைய ஒலிநாடாக்கள் அழிக்கப்படுவதால் மிக அரிய அறிஞர் உரைகள் மீண்டும் கேட்கமுடியா வண்ணம் அழிந்துவிடுவதாக சிட்டி சிந்தரராஜன் அவர்கள் ‘கண்டெடுத்த கருவூலங்கள்’ எனும் புத்தகத்தில் குறிப்பிடுகின்றார்.

5) சரி, தனியார் நூலகங்களைப் பாதுகாக்கலாமென புறப்பட்டால் வேறு மாதிரியான பிரச்சனைகளை எதிர்கொள்ள வேண்டியுள்ளது. உதாரணமாக, தனியார் நூலகங்கள் தனிப்பட்ட உழைப்பில் கட்டமைக்கப்படுபவை. அவர்கள் ஆண்டாண்டு காலமாக உழைத்து சேகரித்ததை ஒரு எளிய தொழில் நுட்பத்திற்கு பரிசாக அளிக்க விரும்புவதில்லை. அவர்களுக்கு ஈட்டுத்தொகை வழங்கி இலக்கப்பதிவாக்கும் நிலையில் இப்போதுள்ள திட்டங்கள் நிதி வசதியுடன் இல்லை. உதாரணமாக, அடையாறிலுள்ள உ.வே.சா நூலகத்தை நவீனப்படுத்தி இலக்கப்பதிவாக 45 இலட்சம் ரூபாய் ஆகுமென்று நூலக பாதுகாவலர் இக்கட்டுரை ஆசிரியருக்குச் சொன்னார்கள். இத்தகைய காரணங்களால் தனியார் நூலக நூல்களும் பாதுகாப்பற்ற நிலையில் உள்ளன.

6) நூலகமென்று வைக்கும் அளவிற்கு இல்லையெனினும் பல நபர்கள் பல அரிய நூட்களை சேகரித்து வைத்துள்ளனர். அவைகளை நல்ல விலைக்கு விற்க வேண்டுமென்பதே இவர்கள் அவா. எனவே இத்தகைய நூற்களை இலக்கப்பதிவாக்க வேண்டுமென்றால் நமது திட்டத்தில் இதற்கென தனி நிதி ஒதுக்கீடு அவசியம். பாரதியாரின் எழுத்து பொதுவுடமையான பின்பும், அவரின் எழுத்தை சேகரித்து வைத்து கொஞ்சம், கொஞ்சமாகப் பதிப்பித்து வரும் ஒரு சேகரிப்பாளரை இக்கட்டுரை ஆசிரியர் அணுகிய போது அவர் போட்டோ எடுக்கக்கூட அனுமதிக்கவில்லை. காரணம், பாரதியின் எழுத்து பாரதிக்கு சோறு போட்டதோ என்னவோ! இவருக்குச் சோறு போடுகிறது. எனவே அவைகளை இலக்கப்பதிவாக்க இவர் விரும்பவில்லை. இம்மாதிரிப் போக்கினால் அடுக்கில் சேகரமான பழைய புத்தகங்கள் அழியும் நிலைக்கு தள்ளப்படுகின்றன.

7) இதே போன்றதுதான், வெளிநாட்டு நூலகங்களிலிருக்கும் நமது புத்தகங்களை இலக்கமயமாக்கும் முயற்சியும். உதாரணமாக, இலண்டனில் உள்ள பிரித்தானிய நூலகத்திலிருந்து ஒரு நூலை பதிவாக்க கீழ்க்காணும் செல்வினங்களுண்டு: அ) நேரடிக் கைப்பதிவு. அதாவது கணினியில் தட்டச்சுமுறையில் நூலகத்தில் அமர்ந்து பதிவது. நூல் இலவசம். ஆனால் பதிவு கூலிக்கு செலவு அதிகம். இதற்கென ஒருவரை வேலைக்கு வைக்கவோ அல்லது ஒரு ஆய்வாளருக்கு உதவித்தொகை வழங்கவோ நிதி வைத்துக்கொள்வது அவசியம். அப்படி

இல்லையெனில், ஆ) கருப்பு/வெள்ளை ஒளி நகலுக்கு ஒரு கட்டணம்; மூலத்தின் நிறத்துடன் ஒளி-இலக்கப்பதிவாக்க: குறைந்த தெளிவு, நடுத்தரத்தெளிவு, மிகத்துல்லிய பதிவென ஒவ்வொன்றிற்கும் ஒரு கட்டணமுண்டு. இதைக் கொடுப்பதற்கு நமது திட்டத்தில் பண வசதி இருக்க வேண்டும். ஐரோப்பா முழுவதுமுள்ள பல்வேறு நூலகங்களில் தமிழ்ச்சுவடிகளும், புத்தகங்களுமுள்ளன. அவையெல்லாம் முறையாகப்பதிவாக பல மில்லியன் டாலர் செலவாகும். அதற்காக மெகாத்திட்டங்களை நாம் இனிமேல்தான் உருவாக்க வேண்டும்!

8) தமிழகப் பல்கலைக் கழகங்கள் மத்திய, மாநில நிதி உதவியுடன் இத்தகைய ஆய்வுத்திட்டங்களை செயல்படுத்தலாம். சமீபத்தில் சென்னைப் பல்கலைக்கழகம் மத்திய அரசின் நிதியுதவியுடன் மருத்துவ நூல்களை இலக்கப்பதிவாக்கப் போவதாக ஒரு செய்தி வந்துள்ளது (<http://groups.yahoo.com/group/esuvadi/message/4497>). அதேபோல், திருப்பதிப் பல்கலைக்கழகமும் புத்தக-இலட்சம் திட்டத்துடன் இணைந்து சென்னையில் ஒரு நூல் சேகரிப்பு மையமொன்றை நிறுவியிருப்பதாக செய்தி வந்துள்ளது. (<http://groups.yahoo.com/group/esuvadi/message/4450>).

9) தமிழ் இணையக் கல்விக்கழகம் 'மதுரைத்திட்டத்தின்' நூல்களைப் பயன்படுத்திக்கொள்வதாக ஆரம்பித்து நூல்களில் அச்சுப்பிழைகள் உள்ளன என்று கூறி அனைத்து நூல்களையும் மீண்டும் அச்சேற்றும் வேலையில் ஈடுபட்டுள்ளது. இது எவ்வளவு கால விரயம் என்பது புரியும். அதற்குப்பதில் தேர்ந்த தமிழ் அறிஞர்களிடமோ, அச்சகங்களில் வேலை செய்த அநுபவம் உள்ளவர்களிடமோ கொடுத்து பிழை திருத்தி வெளியிட்டிருக்கலாம்.

10) பிழையற்ற பதிப்பு என்பது ஒரு பிரச்சனைதான். மதுரைத்திட்டம் போன்ற தொண்டுமியத்தை நம்பி இயங்கும் திட்டங்களில் 100% பிழையற்ற பதிப்புகளை எதிர்பார்க்க முடியாது. காரணம், தட்டச்சு செய்பவர் தமிழுக்கு சம்மந்தமில்லாத ஏதாவதொரு வேலையில், வெளிநாட்டில் இருப்பார். தமிழைப் பாதுகாக்க வேண்டுமென்ற ஒரே உந்துதல் காரணமாக தட்டச்சு செய்வார். அவருக்கு கிடைத்திருக்கும் புத்தகத்திலேயே பிழைகள் இருக்க வாய்ப்புள்ளது. பின் அவர் சேர்க்கும் பிழைகள் வேறு! இணைய உலகில் சுயநலமற்று தட்டச்சு செய்பவர் அரிது. அதைவிட அரிது அதைப்பிழை திருத்துவோர்!

11) இணையப் பரப்பில் தமிழறிந்த அறிஞர்கள் இல்லாமல் இல்லை. ஆனால், அவர்கள் எல்லோரும் தனித்தனி தீவாக செயல்படுகின்றனர். அவரவர் சேகரிப்பு அவரவர்க்கு என்று. அவர்கள் சேகரங்களை பொதுவான இடத்தில் வைக்கும் அளவிற்கு அவர்கள் 'கீழிறங்கி' வருவதில்லை அல்லது இருக்கும் திட்டங்களுடன் தங்களை இணைத்துக்கொள்வதில்லை. இது பொது இழப்பிற்கு வழியிடுகிறது! வீர வணக்கம் (ஹீரோ ஒர்டிப்) என்பதை ஒரு மரபாகக்கொண்ட தமிழர் வாழ்வியலில் பல திட்டங்கள் இம்மாதிரி வீரவணக்கத் திட்டங்களாக, தனித்தீவுகளாக அமைந்து வருவதையும் காண முடிகிறது. 'என் வழி தனி வழி' என்பது சினிமா வசனமாக இருந்தாலும் அது தமிழர் உளவியலை மிகச்சரியாக பிரதிபலிப்பதாக உள்ளது.

12) தமிழ் ஆர்வமுள்ள பிறதிறை சார்ந்தவர்கள், உதாரணமாக, திருக்குறளையே மீண்டும், மீண்டும் பதிப்பிக்கின்றனர். சென்ற தமிழிணைய மாநாட்டில் எழுத்தாளர் சுஜாதா அவர்கள் திருக்குறள் பட்டும் இணையத்தில் குறைந்தது 40 படிகளாவது இருக்குமென்றார். பிரபலமானவையே மீண்டும், மீண்டும் பிரபலமடைகின்றன அல்லது கண்ணில் படுகின்றன என்பது பொதுவிதியாயினும் சேமிப்பிற்கு ஏங்கி நிற்கும் பாதுகாப்பற்ற பல நூல்கள் இன்னும் பல்லாயிரமுண்டு.

தமிழகத்திலிருக்கும் பல்வேறு நூலகங்களிலுள்ள புத்தகங்கள் என்று, யாரால் மின்னாக்கம் பெறும்? அவைகள் தமிழ் மரபு அறக்கட்டளை போன்ற தன்னார்வச் சங்கங்களுக்கு மின்னாக்க இடம் தருமா? அனுமதி கிடைக்குமா? இது தவிர தமிழ் அறிஞர்களின் நூற்சேர்புகள் கொண்ட நூலகங்களை யார் மின்பதிப்பு செய்வது? எவ்வாறு? அவைகளை முறையாகப் பதிப்பிக்க இந்திய மத்திய அரசின் நிதி பெற்றுச் செயல்படும் பல்கலைக்கழகத்துறைகள் என்ன செய்யவுள்ளன? தமிழ் மீது ஆர்வமுள்ள பிற நாட்டு நூலகத்துறைகள் என்ன செய்யவுள்ளன (குறிப்பாக மலேசியா, சிங்கப்பூர்)? இத்தகைய திட்டங்களுக்கான ஒரு பொது நிதி உருவாக்கவியலுமா? திட்டமிட்ட வடிவம், இலக்கு, முடிவு தேதிகள், இணைப்பு என்று அங்கு பலரால் செயல்பட முடியுமா? அவை ஏன் என்று மனம் திறந்து பேச வேண்டிய ஒரு காலக்கட்டத்தில் இருக்கிறோம்.

இணையம் என்பது இணைக்கும் ஊடகம், தொடர்பு படுத்தும் ஊடகமென்றாலும் பலதிட்டங்களுக்குள் ஒரு ஆழமான தொடர்பு இல்லை. இதில் பல திட்டங்கள் தமிழ் இணைய மாநாடுகள் மூலம் பிரபலமடைந்திருப்பதால் இது குறித்த ஒரு 'சாலை வரைபடம்' இம்மாநாட்டில் விவாதிக்கப்பட வேண்டியது அவசியமென்று தோன்றுகிறது. எனவே திட்டமிட்ட இலக்கை இத்திட்டங்கள் எதுவும் குறித்த காலக்கட்டத்தில் அடைந்துள்ளதாகக் கூறமுடியாது. ஒவ்வொரு திட்டமும் ஆரம்பிக்கும் காலத்தில் பெருந்த வரவேற்பையும், எதிர்பார்ப்பையும் கொண்டிருந்தாலும் அவை செயல்படத்தொடங்கிய பின் ஒரு தோய்வு நிலையை அடைகின்றன. ஒருங்கிணைப்பு என்பது இம்மாதிரித் திட்டங்களின் அகவமைப்பில் ஒரு அங்கமா? என்பதும் கேட்க வேண்டிய கேள்வி.

இலக்குகள்:

இலக்க தொழில் நுட்பம் அறிந்த அனைவரும் ஒத்துக்கொள்வது, தமிழ்ப் பாரம்பரியப்பாதுகாப்பிற்கு இத்தொழில் நுட்பம் ஒரு வரப்பிரசாதமென்பது. எனவேதான் விவரமறிந்தவர்கள் தனியாகவோ, குழுவாகச் சேர்ந்தோ அல்லது நிறுவனம் மூலமாகவோ தமிழ்ப்பண்பாட்டு முதுசொமைப் பாதுகாக்க முனைந்துள்ளனர். இலக்கு தெளிவாக இருக்கிறது, ஆனால் நிறைய நடைமுறைச் சிக்கல்கள் உள்ளன. அவைகளைக் களையும் முறை பற்றி ஒரு மனம் திறந்த கலந்துரையாடல் அவசியம். இதற்கென ஒரு கூடலையோ, கருத்தரங்கையோ தமிழ்நாட்டிலோ, வெளிநாட்டிலோ பல்கலைக்கழக சார்புடனோ அல்லது உத்தமம் வழியாகவோ நடத்த வேண்டும்.

தமிழ் முதுசொம் மின்வெளியில் நிரந்தரப்பாதுகாப்பில் இருக்க வேண்டி கீழ்க்காணும் முன்வரைவுகளை முன்வைக்கலாம்:

1. முதலில் தோன்றிய தமிழ் இணையக் கல்விக்கலைக் கழகம், அதன் பின் தோன்றிய தமிழ் மரபு அறக்கட்டளை, புதிதாகத் தோன்றியுள்ள இந்திய இலக்க நூலகம் இவை நோக்கில் ஒன்று பட்டு இருப்பதால் ஒருங்கிணைந்து செயல்பட வேண்டும். ஒன்று மாநில அரசு, மற்றது மத்திய அரசு சார்ந்த நிறுவனம். த.ம.அ ஈட்டுமுனைப்பற்ற பொது நிறுவனம். இவை சேர்ந்து செயல்பட்டால் அது நாட்டிற்கே முன்னுதாரணமாகப்படும்.

2. தமிழ் நாட்டில் நூற்றுக்கும் மேலான கல்லூரிகள் உள்ளன. பொதுவாக எல்லாக்கல்லூரிகளிலும் கணினி வசதி வந்துவிட்டது. இக்கல்லூரிகளிலுள்ள தமிழ்த்துறையும், கணினித்துறையும் சேர்ந்து செயல்பட்டு தொண்டாழியத்தில், தமிழில் ஈடுபாடுள்ள மாணவர்களுக்கு முறையான இலக்கப்பதிவு செய்யும் முறைகளை கற்றுத்தர வேண்டும். இதற்கான பயிற்சியை மேலே சொன்ன மூன்று நிறுவனங்களும் அளிக்கலாம். இப்படிப் பதிப்பிக்கப்படும் நூல்களை பிழைத்திருத்தம் செய்ய கல்லூரித் தமிழ்ப் பேராசிரியர்கள் முன்வரவேண்டும். எல்லாம் ஆங்கிலப்படுத்தப்பட்டுவரும் தமிழகத்தில் 60 களில் காணப்பட்ட அபரிதமான தமிழ் உணர்ச்சியை- தமிழ் மாணவர்களுக்கு மீண்டும் ஊட்ட இம்மாதிரி முயற்சிகள் துணை போகும். இதில் கணினி வசதியுள்ள பள்ளிகளையும் சேர்த்துக் கொள்ளலாம். அமெரிக்காவில் நடந்த தமிழிணைய மாநாட்டில் 50, 000 பள்ளிகளுக்கு கணினி வசதி செய்யப்போவதாக ஒரு திட்டம் முன் வைக்கப்பட்டது. அம்மாதிரி கணினி வசதியைப் பள்ளிக்குத்தரும் போது ஒவ்வொரு பள்ளியும் குறைந்தது 100 புத்தகங்களையாவது இலக்கப்பதிவாக்கித்தர வேண்டும் என்ற அன்புக் கட்டளையை வைக்கலாம். இப்படிப் பதிவாகும் நூற்களை இணையத்தில் வைத்து பராமரிக்க மேற்சொன்ன மூன்று நிறுவனங்களும் முன்வரலாம்.

3. தனியார் வணிக நிறுவனங்கள் தமிழிணைய மாநாட்டில் தொடர்ந்து பங்கேற்றாலும் தங்களது வலைப்பக்கங்களில் இம்மாதிரியான சமூகப்பொறுப்புள்ள திட்டங்களுக்கு ஒரு இணைப்பு கூடத்தருவதில்லை. இணைப்பு என்பதுதான் இணையத்தின் அடிக்கோடு. ஆனால் 'சேதி' என்பதை விலைக்கு விற்று வரும் பழக்கம் இன்னும் மக்கள் மனதிலிருந்து அகலாத காரணத்தால் இணைப்புக் கொடுப்பது பண இழப்பு என்பது போன்றதோர் மனோநிலை இவ்வணிக நிறுவனங்களுக்கு இருப்பதாகப்படுகிறது. இது மாறி, இத்திட்டங்களை பிரபலப்படுத்துவதில் இணைய இதழ்கள் முன்னிற்க வேண்டும். வெறும் கிசு, கிசுப்பு, சுயபுராணம் பாடும் மடலாடற்குழுக்களுக்கு இணைப்புத்தர உடனே முன்வரும் இ-தழ்கள், சமூகப்பொறுப்புள்ள திட்டங்களைக் கண்டு கொள்ளாமல் விடுவது ஆரோக்கியமற்ற போக்கையே சுட்டுகிறது.

4. சுயநலமற்ற திட்டங்களுடன் செயல்படும் பொறுப்புள்ள கூட்டமொன்று உருவாக வேண்டும். எல்லோரும் பின் பெஞ்சில் அமர்ந்து கொண்டு என்ன நடக்கிறது? பார்க்கலாம் என்ற மனோநிலையிலேயே இருக்கிறார்கள். நான்கு பேர் சேர்ந்து ஒரு தேரை இழுப்பதற்கும், நாலாயிரம் பேர் சேர்ந்து இழுப்பதற்கும் பாரிய வித்தியாசமுண்டு. அதிகம் சாதிக்கவில்லை என்று வக்கணையாக முறையிடுவதைத் தவிர்த்து சாதனை என்பது கூட்டு முயற்சியால் மட்டுமே நிகழும் ஒன்று என்று எண்ண வேண்டும்.

5. தமிழ் மரபு அறக்கட்டளை போன்ற நிறுவனங்கள் பொது நிதியை நம்பி இருப்பவை. உலகமெங்கும் பரவியிருக்கும் தமிழ் கூறும் நல்லுக்கம் நிதியுதவி செய்ய வேண்டும். மலேசியா-சிங்கப்பூர் அரசுகள் முன்வந்து ஆதரவு தர வேண்டும். மேலும் இம்மாநாட்டில், இம்மாதிரித்திட்டங்களுக்கான ஒரு பொது நிதியை உருவாக்கும் முனைப்பு பற்றி சிந்திக்க வேண்டும். இருக்கின்ற நிதியைப் பகிர்ந்து கொள்ளும் வழியையும் கண்டறிய வேண்டும்.

6. வெளிநாட்டுப் பல்கலைக்கழகங்கள், நூலகங்கள், பிற மரபு சார் நிறுவனங்களுடன் சேர்ந்து செயல்படும் முன்வரைவு ஒன்றைப் பற்றியும் சிந்திக்கலாம்.

7. மிக முக்கிய வளச்சேர்ப்பு என்றால் அது மனித வளச்சேர்ப்பே! எனவே தொழில்நுட்பமறிந்த தமிழர்கள், பிறநாட்டினர் இந்நூலக வளர்ச்சியில் பங்கேற்க வேண்டும். உதாரணமாக சேமிப்புக்களை எப்படிச்

செந்தரமாக்குவது? அட்டவணைப்படுத்துவது? சேகரிப்பிலுள்ள விஷயங்களை எப்படி எளிதாகத்தேடி எடுப்பது போன்றவை நாம் கொள்ளும் சவால்கள். இதற்கான ஒரு மூளை வங்கி காலத்தின் தேவை.

8. முடிந்தால் கூகுள், மைக்ரோ சாஃப்ட் இவைகளுடன் நட்புறவை ஏற்படுத்தி நமது சேகரிப்பை உலகமயமாக்க முடியுமாவெனவும் சிந்திக்க வேண்டும்.

9. பெரும்பாலான தமிழர்களுக்கு இம்மாதிரித்திட்டங்களை உலகமயமாக்குவதில் ஒரு தயக்கம் உள்ளது. அது காரணமில்லாத தயக்கமில்லை. அது என்னவெனில், நம்மவர் பல்லாண்டு அனுபவத்தில் பெற்ற அறிவை பொதுவில் வைக்கும் போது மற்றவர் (பிற நாட்டவர்) அவ்வறிவு கொண்டு தொழில்முறையில் ஆதாயம் பெற்றுவிட முடியாதா? என்பதே! இது ஒரு வகையில் அறிவுக்காப்புரிமை பற்றியது. உண்மையில், நமது அறிவுச் செல்வங்களை இணையத்தில் வைப்பதின் மூலமே நமது இந்தியக் கட்டுபிடிப்புகளுக்கு காப்புரிமை வழங்கவியலும். உதாரணமாக, மஞ்சள் கிழங்கை பல்லாயிரமாண்டு நாம் பயன்படுத்தி வந்தாலும் அது பற்றிய நூற்குறிப்பை நாம் சுவடிகளிலிருந்து எடுத்து இணையத்தில் வெளியிடும் போது மஞ்சள் பயன்பாட்டின் இந்தியக்கண்டுபிடிப்பு காப்புரிமை பெறுகிறது. மேலும் சுவடிகள் வெளியிடப்பட்டவுடன் அதுகுறித்த ஒரு உலகம் சார்ந்த கருத்துப்பரிவர்த்தனை இடம் பெற வேண்டும். சுவடி வாசிக்கத் தெரிந்த தமிழர்கள் உடனே அச்சுவடிச் செய்திகளை நவீன தமிழில் மாற்றித் தரவேண்டும். அவை, உடனுக்குடன் ஆங்கில மற்றும் பிற மொழிகளில் மொழி பெயர்க்கப்பட வேட்டும் அப்போதுதான் நமது சிந்தனைகளுக்கு காப்புரிமை கிட்டும். இது குறித்த ஒரு கலந்துரையாடல் கூட 2004 தமிழிணைய மாநாட்டில் இடம் பெற வேண்டும்.

10. பல்கலைக்கழகங்கள் பங்கேற்பதுடன் தமிழ் மரபு அறக்கட்டளை மற்றும் பிற திட்டங்களின் சேகரிப்பை ஆக்கமுறையில் பயன்படுத்தவும் முன்வரவேண்டும். உதாரணமாக, சான்பிரான்சிஸ்கோவில் நடந்த மாநாட்டில் மாணவர்களுக்கான ஒரு தமிழறிவுப் போட்டியில் எமது சேகரிப்புகளை 'இணையவளம்' (internet resource) என்று குறிப்பிட்டு மாணவர்களை இவ்வளத்தைப் பயன்படுத்துமாறு செய்தனர். இதே முறையில் தமிழ்நாட்டுப் பல்கலைக்கழகங்கள், பிற மாநில, பிற நாட்டு தமிழ்த்துறைகள் இவ்வளத்தைப் பயன்படுத்திக்கொள்ள வேண்டும். இதை ஊக்குவிக்கும் வண்ணம் மலேசியப்பல்கலைக்கழகத் தமிழ்த்துறையுடன் எமது அறக்கட்டளை பேச்சு வார்த்தை நடத்தி வருகிறது. பல்கலைக்கழகங்கள், கல்லூரிகள் இவ்வளப்பயன்பாட்டில் ஊக்கமுடன் ஈடுபடவேண்டும்.

இதற்கான முன்வரைவாக சென்ற மாநாட்டுக்கட்டுரைகளில் கட்டுரை வாசிக்கப்பட்டுள்ளது. 1) சேகரிப்பிலுள்ள நூல்களை மாணவர்கள்/ஆசிரியர்களை வாசிக்க ஊக்குவித்தல், 2) அது குறித்த ஒரு அறிவு பூர்வமான கலந்துரையாடலை 'மின்தமிழ்' (<http://groups.google.com/group/minTamil>) போன்ற இணைய மன்றத்தில் வைப்பது, 3) தொடர்ந்து கூட்டமைப்பாகச் செயல்பட 'வலைப்பதிவு' தொழில் நுட்பத்தைப்பயன்படுத்துவது (http://our_legacy.pitas.com/)

தமிழ் மரபு அறக்கட்டளையின் தொடர் வளர்ச்சி:

தமிழ் மரபு அறக்கட்டளை 2001 ஆண்டு கோலாலம்பூரில் நடந்த தமிழிணைய மாநாட்டில் துவக்கப்பட்டது. அம்மாநாட்டில் முன்வைக்கப்பட்ட ஒரு கட்டுரையின் பின்விளைவாகவே இத்திட்டம் தோற்றுவிக்கப்பட்டது. அம்மாநாட்டில், அறக்கட்டளைக்கான விதைக்காசு தரப்பட்டு, அது தொடர்ந்து மலேசிய இந்தியக் காங்கிரஸால் பேணப்படும் என்று பத்திரிகைகளில் செய்தி வெளியாகியது. அது தந்த உத்வேகத்தில் தமிழ்மரபு அறக்கட்டளை இன்றுவரை ஊக்கமாக செயல்பட்டு வருகிறது. தனது தொழில்முறை வேலைகளுக்கிடையே இக்கட்டுரை ஆசிரியர்கள் இதற்காக நேரம் ஒதுக்கி பலமுறை இந்தியாவில் களப்பணி செய்துள்ளனர். இக்களப்பணியில் நூல் சேகரிப்பு என்பது முக்கிய நோக்கமாக இருந்தாலும், இத்திட்டத்தை பொது மக்களிடம் எடுத்துச் செல்ல தொலைக்காட்சிப் பேட்டிகள், வானொலிப் பேச்சுகள், இதழ்களில் கட்டுரைகள் என்று பல அளித்துள்ளனர். அது மட்டுமின்றி ஒவ்வொரு முறையும் தனியார் நூலகங்களுக்குச் சென்று அவர்களது ஆதரவை வேண்டியுள்ளனர். இதுவரை சந்திந்த தனியார் நூலகங்கள் எல்லாமே இலக்கப்பதிவிற்கு ஒப்புதல் அளித்துள்ளன. அதற்கான ஆள் பலம், நிதி பலம் இவற்றை வரும் ஆண்டுகளில் திரட்ட வேண்டும். தமிழ் மரபு அறக்கட்டளை தனது சேகரிப்புகளை கீழ்க்காணும் பகுதிகளில் மின்வடிவாக வைத்திருக்கிறது:

1) மின் நூல் சேகரிப்பு

அ) <http://www.tamilheritage.org/old/text/ebook/ebook.html>;

ஆ) <http://bharani.dli.ernet.in/thf/text/text.htm>

இ) http://www.tamilheritage.org/uk/bl_thf/bl_thf.html

2) இலத்திரன் சுவடிகள் (இ-சுவடி)

அ) <http://bharani.dli.ernet.in/thf/palmleaf/plindex.htm>

ஆ) http://www.tamilheritage.org/thfcms/index.php?option=com_content&view=article&id=336&Itemid=448

- 3) தமிழ் முதுசொம் களஞ்சியம்
<http://www.tamilheritage.org/old/monument/oodu/sangkam.html>
- 4) தமிழ் மரபு விக்கி
<http://www.heritagewiki.org/>
- 5) தமிழ் புகைப்படக் களஞ்சியம்
<http://image-thf.blogspot.com/>
- 5) தமிழ் ஓவியக்களஞ்சியம்
<http://www.tamilheritage.org/old/art/artist.html>
- 6) முதுசொம் இசையரங்கம்
அ) <http://bharani.dli.ernet.in/thf/audio/luvmusic.html>
ஆ) <http://voiceofthf.blogspot.com/>
- 7) முதுசொம் ஒளியரங்கம்
அ) <http://bharani.dli.ernet.in/thf/video/moindex.html>
ஆ) <http://video-thf.blogspot.com/>
- 8) முதுசொம் வளச்சாரளம்
<http://www.tamilheritage.org/thfcms/>

இலண்டன் வாழ் தமிழர்களின் ஆதரவுடன் இலண்டன் கிளையொன்று ஆரம்பிக்கப்பட்டு பிரித்தானிய நூலகத்தைச் சார்ந்த 'இந்திய அலுவலகம்' சேகரிப்புகளை மின்னாக்கம் செய்துள்ளது (http://www.tamilheritage.org/uk/bl_thf/bl_thf.html). ஆங்கிலக் குடியரசில் முதன்முறையாக ஒரு இனம் தனது பாரம்பரியச் சொத்துக்களை இலக்கவடிவில் சேகரிக்கும் முதல் முயற்சியாகும். இதற்கு மிகுந்த உற்சாகத்துடன் பிரித்தானிய நூலகம் ஆதரவளித்துள்ளது. இத்திட்டத்தின் வெற்றி ஆங்கிலக் குடியரசில் வாழும் பிற இன மக்களை ஊக்குவித்துள்ளது. தமிழ் மரபு அறக்கட்டளையின் தமிழ் மின்னாக்கத்திட்டம் முன் உதாரணமாகக் காட்டப்பட்டுள்ளது. தமிழ் இனம் இதில் முன் நிற்பது நமக்கு பெருமையளிக்கும் விஷயமாகும்.

இக்கட்டுரை ஆசிரியர்கள் பாரதக் குடியரசுத்தலைவர் டாக்டர் அப்துல் கலாம் அவர்களைச் சந்தித்ததாலும், 'புத்தக-இலட்சம்' இந்தியத்துறைத்தலைவர் பேராசிரியர் பாலகிருஷ்ணனை சந்தித்ததாலும் தமிழ் மரபு அறக்கட்டளைக்கு இந்திய இலக்க நூலகத்தில் நிரந்தர இடம் கிடைத்துள்ளது. இனிவரும் சேகரிப்புகள்

<http://bharani.dli.ernet.in/thf/index.html>

என்ற இணையத்தளத்தில் பொது மக்கள் நுகர்விற்கு வைக்கப்படும்.

தமிழ்மரபு அறக்கட்டளை பரிட்சார்த்தமாக மின்னலகில் (cyberspace) இதுவரை இயங்கி வந்தாலும் 2004ம் ஆண்டு கோடை முதல் இந்தியாவில் பதிவு பெற்ற சங்கமாக உருவாகியிருக்கிறது. இதன் முதல் செயற்குழுவில் உலகம் தழுவிய முறையில் பல தமிழ் அறிஞர்கள் இடம் பெற்றுள்ளனர்.

<http://www.tamilheritage.org/old/ec/ec.html>

மின்னிட நெருக்கடியால் மின்னாக்கம் பெற்ற பல நூற்கள் இன்னும் இணையத்தில் வலையேற்றம் பெறாமல் இருந்து வந்தன. இனிவரும் காலங்களில் அவையெல்லாம் இந்திய இலக்க நூலக வலைத்தளத்தில் வலையேற்றம் காணும்.

தமிழ் மரபு அறக்கட்டளை 12 ஆண்டு காலத்தில் மிகக்குறுகிய மனித வளத்துடன் மிகக்குறுகிய பொருள் வளத்துடன் இவைகளை சாதித்து இருப்பது குறிப்பிடத்தக்கது. தமிழ் உணர்வுடைய பலர் தொடர்ந்து ஆதரவு அளித்து வருகின்றனர். தமிழ் இணைய மாநாடு நமது சங்கமாக இருக்கிறது. இங்கு வருபவர் ஆதரவு தொடர்ந்து இருந்து வருகிறது.

இனி வரும் காலங்களில் நாம் எப்படி நமது இலக்கை நோக்கி நடை போடப் போகிறோம்? மக்களிடம் எதிர் கொள்ளும் பிரச்சனைகளை எப்படி சமாளிக்கப் போகிறோம், மனித வளத் தேக்கத்தை எப்படி எதிர் கொள்ளப் போகிறோம்?, தொழில் நுட்ப அறியாமை/போதாமை இவைகளை எப்படி சமாளிக்கப் போகிறோம்? தொடர்ந்து செயல்பட பொருள் வளத்தை எப்படிக் கூட்டப்போகிறோம்? இவையெல்லாம் நம் முன்னே நிற்கும் கேள்விகள். கேள்விகள் அசௌகரியமானவை என்றாலும், கேள்விகளின்றி முன்னேற்றமென்பது இல்லை. இப்போதுள்ள இணைய வளங்களை எப்படிப் பேணுவது, பகிர்வது. நாளை சவால்களுக்கு இப்போதே திட்டமிடுவது. நாளை ஒருங்கிணைப்பிற்கான சாலையை இன்றே போட ஆரம்பிப்பது என்பதே இக்கட்டுரை தமிழிணைய மாநாட்டில் விவாதிக்க உள்ள கேள்விகள்.

பார்வை மாற்றுத் திறனாளிகள் பயன்பாட்டில் தமிழ் கணினியும் செல்பேசியும்

சே. பாண்டியராஜ், ஆய்வியல் நிறைஞர் பட்ட மாணவன், தமிழ்த் துறை, பச்சையப்பன் கல்லூரி, சென்னை – 30.

நெறியாளர்: முனைவர். வா.மு.சே. முத்துராமலிங்க ஆண்டவர், தமிழ் இணைப் பேராசிரியர், தமிழ்த் துறை,
பச்சையப்பன் கல்லூரி, சென்னை – 30.

கணினி மயமாக்கப்பட்ட இன்றைய உலகில் தமிழ் பல்வேறு நிலைகளில் வளர்ச்சி அடைந்துள்ளது. அதிலும் தமிழ், கணினியிலும் செல்பேசியிலும் பெற்றுள்ள வளர்ச்சி மிகவும் குறிப்பிடத்தக்கதாகும். தமிழில் ஒருங்குறி முறை பயன்பாட்டிற்கு வந்தபிறகு, தமிழ் கணினி வளர்ச்சி பல்வேறுநிலைகளில் விரிவடைந்துள்ளது. ஒருங்குறி முறை வந்த பிறகு தான், பார்வை மாற்றுத் திறனாளிகள் கணினியையும் செல்பேசியையும் பயன்படுத்துவதற்கான திரை வாசிப்பு மென்பொருளில் (screen reading software) தமிழினை எளிய முறையில் கொண்டுவர முடிந்தது. பார்வை மாற்றுத் திறனாளிகள் கணினியையும் செல்பேசியையும் பயன்படுத்துவதற்கு உதவும் திரைவாசிப்பு மென்பொருள் குறித்தும், அதில் ஏற்பட வேண்டிய வளர்ச்சிக் குறித்தும் இவ்வாய்வு கட்டுரையில் காண்போம்.

பார்வை மாற்றுத் திறனாளிகள் கணினியை பயன்படுத்துவதற்கு உதவும் திரைவாசிப்பு மென்பொருளினை பல நிறுவனங்கள் உருவாக்கிவருகின்றன. உதாரணமாக, அமெரிக்காவில் உள்ள ஃபிரீடம் சைன்டிஃபிக் (Freedom Scientific) என்ற நிறுவனம் ஜாஸ் (JAWS) என்ற திரைவாசிப்பு மென்பொருளினை ஆங்கிலம் உள்ளிட்ட மேலை நாட்டு மொழிகளில் கணினியை பயன்படுத்த, 1991 முதல் வெளியிட்டு வருவதனை குறிப்பிடலாம்.

தமிழிலும் பல்வேறு நிறுவனங்கள் தமிழ் திரைவாசிப்பு மென்பொருளினை உருவாக்கி வருகின்றன. குறிப்பாக தற்பொழுது அனைவராலும் பயன்படுத்தப்பட்டு வரும் இஸ்பீக் (Espeak) என்ற மென்பொருளைக் குறிப்பிடலாம். இந்த மென்பொருளானது இங்கிலாந்து நாட்டைச் சார்ந்த திரு. ஜோனத்தான் டட்டிங்டன் (Jonathan Duddington) என்பவரால் 1995 முதல் வெளியிடப்பட்டு வருகிறது. இந்த மென்பொருளானது முற்றிலும் இலவசமாகும். இந்த மென்பொருளினை மற்ற திரைவாசிப்பு மென்பொருளில் இணைத்து மட்டுமே பயன்படுத்த முடியும். ஏனென்றால், இது ஒரு ஒலிப்பான் (Speech synthesizer) ஆகும். இஸ்பீக் மென்பொருளின் 1.30 வெளியீடு 31.12.2007 அன்று வெளியானது. இந்த வெளியீட்டில் தான் முதன் முதலில் தமிழ் மொழி சேர்க்கப்பட்டது. அதனைத் தொடர்ந்து பல வெளியீடுகளில் தமிழ் பல்வேறு வளர்ச்சியுடன் வெளியானது. பார்வை மாற்றுத் திறனாளிகள் தமிழ் மொழியில் கணினியையும், செல்பேசியையும் பயன்படுத்த உறுதுணையாக இருப்பது இந்த மென்பொருளே ஆகும்.

இலவச மென்பொருளான இஸ்பீக் ஒலிப்பானை பயன்படுத்தி, என்விடிஏ (NVDA) என்ற இலவச மென்பொருள் 2006 ஆம் ஆண்டு முதல் திரு. மைக் கரன் (Mr. Mike curren), திரு. ஜேம்ஸ் டே (Mr. James teh) என்பவர்களால் உருவாக்கப்பட்டு, கணினியில் பல்வேறு பயன்பாட்டுக்கு சிறந்த மென்பொருளாக தொடர்ந்து வெளியாகி வருகிறது. இந்த மென்பொருளை உருவாக்கி வெளியிட்டுவரும் இருவருமே பார்வை மாற்றுத் திறனாளிகள் என்பது குறிப்பிடத்தக்கதாகும். மேற்கூறப்பட்ட இஸ்பீக் (Espeak) மற்றும் என்விடிஏ (NVDA) மென்பொருட்களில் தமிழினை மேம்படுத்த உறுதுணையாக இருந்துவருபவர் திரு. தினகர் அவர்கள் ஆவார். இவர் தெற்கு இரயில்வேவின் சென்னை அலுவலகத்தில், துணை முதன்மை திட்ட அலுவலராக உள்ளார். இவரும் ஒரு பார்வை மாற்றுத் திறனாளி என்பது குறிப்பிடத்தக்கதாகும். மேலும் இவர், மைக்ரோசாஃப்ட் இயங்குதளத்தில் கொடுக்கப்பட்டுள்ள தமிழ் ஒருங்குறி விசைப்பலகை முறையை அடிப்படையாக வைத்து வள்ளுவன் பார்வை என்ற விசைப்பலகை முறையை உருவாக்கியுள்ளார். இவருடைய இந்த விசைப்பலகை முறையை அடிப்படையாக கொண்டு, எழில்த்தமிழ் என்ற விசைப்பலகையை முனைவர். சரவணன் என்பவர் உருவாக்கியுள்ளார். இவரும் ஒரு பார்வை மாற்றுத் திறனாளி ஆவார்.

மேலே குறிப்பிடப்பட்டுள்ள மென்பொருட்களைப் பயன்படுத்தி பார்வை மாற்றுத் திறனாளிகள் தமிழ்க் கணினியில் பெற்றுள்ள வளர்ச்சி மிகவும் குறிப்பிடத்தக்கதாகும். பொதுவாக கணினியில் பார்வையுள்ளவர்கள் மேற்கொள்ளும் அடிப்படைப் பயன்பாடுகள் அனைத்தையும் பார்வை மாற்றுத் திறனாளிகளும் திரைவாசிப்பு மென்பொருட்கள் உதவியுடன் எளிதாக கையாளுகின்றனர். “தேமதுரத் தமிழோசை உலகமெல்லாம் பரவும் வகை செய்தல் வேண்டும்”, என்ற, மகாகவி பாரதியாரின் கனவு, இணையத் தமிழால் இன்று நிறைவேறியுள்ளது. இந்த இணையத் தமிழில் பார்வையற்றோரும் தங்களது பங்களிப்பை செவ்வனே மேற்கொண்டுள்ளனர். தமிழ் இலக்கியங்களையும், தங்களது அனுபவங்களையும், கருத்துக்களையும் உலகத்தமிழரோடு பகிர்ந்துகொள்ளும் வகையில், தமிழ் மின்மடல் குழுமங்களையும், வலைதளங்களையும், வலைப்பூக்களையும் நடத்தி வருவது குறிப்பிடத்தக்கதாகும்.

பார்வை மாற்றுத் திறனாளிகள் தமிழ் கணினியில் மேலும் பயன்பெற வேண்டுமானால், உருவாக்கப்படும் தமிழ் மென்பொருட்கள் அனைத்தும் திரைவாசிப்பு மென்பொருள் உதவியுடன் பயன்படுத்தும் வகையிலும், முழுக்க

முழுக்க விசைப்பலகை மூலமாகவே பயன்படுத்தும் வகையிலும் உருவாக்கப்பட வேண்டும். தமிழ் OCR, சொற்பிழைத் திருத்தி போன்ற மென்பொருட்களை உதாரணமாக குறிப்பிடலாம்.

பார்வை மாற்றுத் திறனாளிகள் செல்பேசியை தமிழில் இயக்க, மேலே குறிப்பிடப்பட்டுள்ள இஸ்பீக் (Espeak) என்ற ஒலிப்பாணை செல்பேசியில் நிறுவி, டாக்பேக் (talkback) அல்லது ஸ்பீல் (spiel) என்ற மென்பொருளுடன் இணைத்து பயன்படுத்துகின்றனர். இந்த மென்பொருட்கள் தொடுதிரை வசதியுடன்கூடிய செல்பேசிகளையும் எளிய முறையில் பயன்படுத்த உதவுகின்றன. பார்வையுள்ளவர்கள் தொடுதிரை செல்பேசிகளில் தாங்கள் விரும்பிய பட்டியலுக்குள் செல்ல, திரையில் தோன்றும் குறிகளின் மீது விரலினை வைத்தால் போதும், உடனே பட்டியல் தெரிவாகி காட்சிக்கு வந்துவிடும். பார்வை மாற்றுத் திறனாளிகள் தொடுதிரை செல்பேசியை பயன்படுத்த வேண்டுமானால், முதலில் திரையில் இருப்பதை அறிந்த பின்னரே தெரிவு செய்ய இயலும். இதற்கு மேலே குறிப்பிடப்பட்டுள்ள மென்பொருட்கள் பயன்படுத்தப்படுகின்றன.

இம்மென்பொருட்கள் நிறுவப்பட்ட தொடுதிரை செல்பேசியின் திரையில் விரலினை வைத்தவுடன், விரல் எவ்விடத்தில் வைக்கப்பட்டதோ அவ்விடத்தில் உள்ளதை சொல்லும். இதன் மூலம் திரையில் என்ன உள்ளது என்பதை அறிந்துகொள்ளலாம். உடனே மீண்டும் அதே இடத்தில் தொடும்போது, அதற்குரிய பட்டியல் தெரிவாகிவிடும். உதாரணமாக, தொடர்பு பட்டியலினை தெரிவு செய்ய வேண்டுமானால், திரையின் மீது விரலினை வைத்து, வசதிக்கேற்ப நகர்த்தி வர வேண்டும். அவ்வாறு நகர்த்தும் போது, விரலில் படும் குறிகளை சொல்லிக்கொண்டே வரும். தொடர்பு பட்டியல் என்று சொன்னவுடன், மீண்டும் ஒரு முறை தொட்டால் தொடர்பு பட்டியல் தோன்றும். இவ்வாறு தொடுதிரை செல்பேசியினை பார்வை மாற்றுத் திறனாளிகள் பயன்படுத்திவருகின்றனர்.

இதன் மூலம் பார்வை மாற்றுத் திறனாளிகள் தமிழிலேயே செல்பேசியில் தகவல்களைத் தட்டச்சு செய்கின்றனர், குறுஞ்செய்தி அனுப்புகின்றனர், இணையத்தில் தகவல்களை பெறுகின்றனர், பகிர்ந்துக்கொள்கின்றனர், மின் அஞ்சல் அனுப்புகின்றனர், தங்களுக்கு வேண்டிய தகவல்களையும் பாடப் புத்தகங்களையும் வேர்டு கோப்புக் களாகவும் (word files) ஒலி கோப்புக்களாகவும் (audio files) பதிவு செய்து பயன்படுத்துகின்றனர். இந்த அளவிற்கு பார்வை மாற்றுத் திறனாளிகளின் விழியாகவே செயல்படும் செல்பேசியினை மேம்படுத்துவது அவசியமாகும்.

பார்வை மாற்றுத் திறனாளி ஒருவர் ஒரு தமிழ் புத்தகத்தை படிக்க வேண்டுமானால், அந்த புத்தகம் பிரெய்லிலோ (braille) ஒலிக்கோப்பாகவோ இருக்க வேண்டும், அல்லது பார்வையுள்ளவர்கள் படித்துக்காட்ட வேண்டும். இது எல்லா நேரங்களிலும் சாத்தியம் இல்லை. இதற்கு பதிலாக செல்பேசியையே புத்தகங்களை படித்துக்காட்ட செய்யலாம். ஒருவர் தான் படிக்க வேண்டிய ஒரு தாளின் மேலே செல்பேசியை குறிப்பிட்ட இடைவெளியில் வைத்து புகைப்படம் எடுத்துக்கொள்ள வேண்டும். பின்னர் செல்பேசியில் நிறுவப்பட்டுள்ள OCR மென்பொருள் உதவியுடன் அதனை படிக்கலாம். இவ்வகையில் தற்பொழுது nokia போன்ற செல்பேசிகளில் ஆங்கில மொழிக்காக KNFB reader என்ற மென்பொருள் பயன்படுத்தப்படுகிறது. மேலும், தற்பொழுது அதிக அளவில் பயன்படுத்தப்பட்டு வரும் Android போன்ற சாதனங்களில் மேம்படுத்தப்பட்ட பல OCR மென்பொருட்கள் ஆங்கிலத்தில் பயன்படுத்தப்படுவது குறிப்பிடத்தக்கதாகும். இப்படிப்பட்ட மென்பொருட்கள் தமிழில் மேம்படுத்தப்பட வேண்டும்.

பார்வை மாற்றுத் திறனாளி ஒருவர் ஓர் இடத்தில் இருந்து மற்றொரு இடத்திற்கு செல்ல வேண்டுமானால் பல்வேறு இடையூறுகளை சந்திக்க நேரிடுகிறது. இந்த இடையூறுகளை GPRS வசதியுடன் கூடிய செல்பேசிகள் மூலம் நீக்கலாம். பார்வை மாற்றுத் திறனாளி ஒருவர் தான் செல்ல நினைக்கும் இடத்தை செல்பேசியில் குறிப்பிட்டால் அது GPRS உதவியுடன் அவர் செல்ல வேண்டிய இடத்திற்கு வழிகாட்ட வேண்டும். அவர் நடந்து செல்லும்போது வழியில் ஏதாவது தடைகள் இருந்தால் அறிவிக்க வேண்டும். இந்த முறை தற்பொழுது மேலை நாடுகளில் வாகன போக்குவரத்தில் பயன்படுத்தப்பட்டு வருவது குறிப்பிடத்தக்கதாகும்.

மேலே குறிப்பிடப்பட்டுள்ள தொழில்நுட்ப வசதிகளை பார்வை மாற்றுத் திறனாளிகள் பெற வேண்டுமானால், தமிழ் மென்பொருட்களிலும் திரைவாசிப்பு மென்பொருட்களிலும் தொழில்நுட்ப வளர்ச்சிக்கேற்ப மாற்றங்கள் அவசியமாகும். இந்த மாற்றங்களை அவ்வப்போது ஏற்படுத்தினால் கணினியும் செல்பேசியும் பார்வை மாற்றுத் திறனாளிகளின் விழிகளாகவே செயல்படும் என்பது உறுதி.

மேற்கோள்

- <http://espeak.sourceforge.net/>
- <http://en.wikipedia.org/wiki/ESpeak>
- www.nvda-project.org
- www.freedomscientific.com
- www.knfbreader.com
- http://eyesfree.googlecode.com/svn/trunk/documentation/android_access/services.html

‘Raspberry PI’ Future Outlook for Tamil

Saravanan Mariappan

Nexus NGN Sdn. Bhd., Malaysia. Email: nexusngn@gmail.com

Abstract. With the fast developing IT World and broadband penetration expanding rapidly from urban to rural, drastic economical, financial and educational improvement within the community were observed in developed nation. Hence many government institution have dedicated arm focusing mainly on the development of broadband facilities in rural area. With the made available broadband facilities, further challengers lays ahead on how to disseminate information and knowledge. Basic computing tools at low cost can be made available to the rural communities, with the introduction of mini computer “Raspberry PI” information and opportunity to be online can be made quick and cost effective. This paper addresses how Raspberry PI can contribute to the development of Tamil Community and Tamil Language. The mini computer can be used to develop teaching tools such as Mini USB Projector and simple Android Mobiles with internet access. It can also be used to be integrated with server as thin clients bringing the cost to implement ICT labs at remote rural sites cost effective. This paper will also discuss the current influence of the mini computer “Raspberry PI” in the Tamil community and how Tamil Language and the community can develop in the future.

Keywords : Raspberry PI, Thin Client, ICT Lab, Open Source

1.0 INTRODUCTION

Raspberry PI is a credit-card-sized single-board computer developed in United Kingdom by Raspberry PI foundation with the intention of promoting the teaching of basic computer. Since the development and touch of Raspberry Pi mini computer in 29th February 2012 the sales of Raspberry PI reached beyond 1million units. The used of this mini computer have gone into various industries, replacing expansive programmable controls. Creating new era of programmable device development. Some of the industries that uses Raspberry PI are:

- Nuclear industries
- Real-time monitoring devices with web interfaces
- Cloud storage and management
- Industries depend on Thin Client
- Home automation and monitoring
- Laptops
- File sharing portal
- Attach to camera and take picture on command.
- Digital signature
- Personal computers and many more

What interest us for the development of Tamil is the usage of Raspberry PI for teaching and the usage of broadband. For this purpose we can use Raspberry PI as personal computer and Thin client.

2.0 RASPBERRY PI REQUIREMENT FOR TEACHING

The basic requirement for affordable teaching devices are that the device should be able to operate using open source operating system and able to work on LAN.

Some of the criteria of open source operating system and software are:

- ⤴ Licenses are not restricted to any party
- ⤴ Licenses are allowed to be modified
- ⤴ Modifications are allowed by adding on “patch files”
- ⤴ No discrimination against person or groups
- ⤴ Redistribution of program can be done without the requirement for additional licenses
- ⤴ License must not be specific to a product
- ⤴ License must not restrict other software- it's not required for other programs to be open source.

The type of open source programs that Raspberry PI should able to execute are:

1. Open office



5. Video conversion



Media Corder

6. Video player

2. Web browsing



Miro



Miro

3. Instant messaging



Pidgin

7. Graphics editing



GIMP



Paint.net



Inkscape

4. Video playback



VLC

8. Email



Mozilla Thunderbird

9. Sound recording



Audacity

3.0 RASPBERRY PI AS THIN CLIENT

An ideal thin client is the computer hardware without hard drive and well equipped with dedicated mini motherboard which depends on the central server for processing data and connects to the server through network to run application, access files, print, and perform all services available in on ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.



Plate 1 : Raspberry PI as Thin Client.

A cost-effective similar solution can be provided by using Raspberry PI as thin client. It can be supported by open source edubuntu server platform or LTSP server and run applications on central server.

The Raspberry PI that operates as thin clients are called Berry Terminals. The Berry Terminals linux base, designed to turn Raspberry PI mini computer into a low-cost thin clients. Figure 1 shows the working principal of Raspberry PI as thin client.

3.0 RASPBERRY PI AS THIN CLIENT

An ideal thin client is the computer hardware without hard drive and well equipped with dedicated mini motherboard which depends on the central server for processing data and connects to the server through network to run application, access files, print, and perform all services available in an ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.



Plate 2: Centralized Raid Server

4.0 ADVANTAGES OF OPEN SOURCE COMPUTING SYSTEM

The thin client system in school computer labs are capable of providing affordable server-based open source computing solution. The main advantages of using this system in schools are to increase reliability and consistency of technology over time. Through a password-accessible accounts, students, teachers, and administrators can store and can access saved documents and personal settings. Since all files and programs are stored centrally, users can access their work from any computers within the network.

Eg. when a teacher or student “logs in”, the server provides them with their “desktop configuration”. Users can even access their “desktop” from home or other remote location.

The other advantages of open source computing with thin client system are as below:

- ▲ less administration – central management of users, patches, software, data and backups.
- ▲ Higher security – elimination of viruses, Trojans or other vulnerabilities on the user desktops.
- ▲ Hardware independence – support of virtually all client devices and computer hardware and very low system

requirement.

- ⤴ Easy access – teacher and students can access their documents and applications from any PCs in the local area network.
- ⤴ Reduction in TCO – Total Cost of Ownership reduction by up to 50%.

Benefits for school in using open source computing thin client system:

- ⤴ Lowers cost of technology over time
- ⤴ Secure data and equipment
- ⤴ Less downtime and greater efficiency
- ⤴ Reduces administrator staffing costs
- ⤴ Lessen the risk of data theft
- ⤴ Disaster recovery: Data is more secure
- ⤴ Reduced time for technical support
- ⤴ Lower power consumption: save electricity
- ⤴ Zero licensing management
- ⤴ Minimum maintenance
- ⤴ Highly trained individuals are not required

At present we have successfully done about 70 ICT labs in Tamil schools in Malaysia using thin client and server base computing.

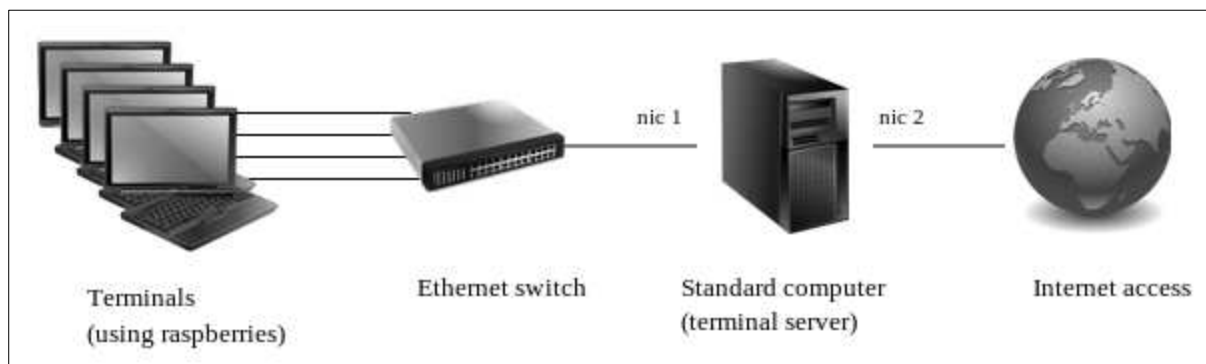


Figure 1.0 shows the working principal of Raspberry PI as thin client.

The labs at present uses refurbish computers of Pentium 3 and 4 or above. These systems are venerable to breakdown without any signs of defects due to aged motherboard, cooling fan, power supply, and other moving parts. Hence with the introduction of Raspberry PI the longevity and warranty period of the thin client can be extended beyond two to three years as Raspberry PI do not have any moving parts.

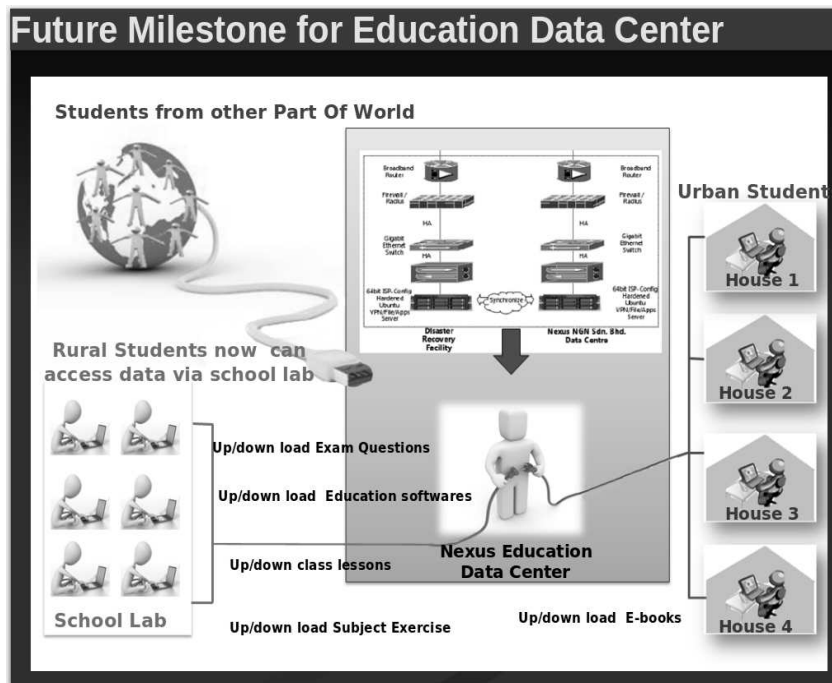
5.0 CONCLUSION

The open source thin client computing environment currently implemented in Malaysian Tamil schools have been well received with 70 over schools been teaching ICT and Tamil since 2009. With the use of Raspberry PI as thin client, the performance of the lab can be enhanced further. The cost to implement the lab can also be reduced substantially. The server based open source technology making it easier for students and teachers to organize and access education materials and multimedia content. At the same time the thin client network enables the schools to use this technology more effectively to monitor students work. With open source applications and thin client technology it was possible to decrease the cost of installation and the cost of maintaining the computer lab. With servers installed at each and every schools, the setting up of centralized school management system is now possible. With centralized school management system in place, teaching and training materials can be deployed to every schools at realtime. This will enhance greater sharing of knowledge between schools, universities, individuals and even between countries, pushing the envelop to further height for the growth of Tamil language and Tamil community.

Plate 3 : Tamil school students using open source Thin Client in an ICT lab.



Figure 2 : Shows the future of Thin Client and centralized management allowing students and schools to have full access to centralized server from anywhere at anytime.



முகநூலில் (facebook) தமிழர்கள் பயன்படுத்தும் மொழிவகைகள் Language Variations in Facebook Used by Tamils

மலர்விழி சின்னையா (kovaimalar@yahoo.com)

பரமசிவம் முத்துசாமி (param@fbmk.upm.edu.my)

கிருஷ்ணன் இராமசாமி (krishnan@um.edu.my)

புத்ரா பல்கலைக்கழகம் (Universiti Putra Malaysia)

ஆய்வுச் சாரம்

காலத்திற்குக் காலம் மொழி நடையும் மொழி ஊடகங்களும் மாறிவருகின்றன. அம்மாற்றத்திற்கேற்ப காலத்திற்கேற்ற ஊடகங்களில் பயன்படுத்தப்படாத மொழிகள் வழக்கிறந்து போவதற்கான வாய்ப்புகள் அதிகம். அதனால்தான் தமிழ்மொழிப் பயன்பாடு நீடித்து நிலைக்க வேண்டும் எனக் கருதிய கணிஞர்களும் தொழில்நுட்ப வல்லுநர்களும் இன்றைய மொழிக்கருவியாகக் கருதப்படும் கணினியிலும் இணையத்திலும் தமிழ்மொழியை ஏற்றியுள்ளனர். இருப்பினும் மொழியைப் பயன்படுத்துவோரே ஒரு மொழி நிலைப்பதற்கும் அழிவதற்கும் வழிகோலுகின்றனர். அதனடிப்படையில் இக்கட்டுரையானது முகநூல் (facebook) எனும் வலைத்தளத்தில் தமிழர்கள் பயன்படுத்தும் மொழி வகைகளை ஆராயும். தொடர்ந்து, தமிழறிந்த (எழுத-படிக்க-உரையாட) தமிழர்கள் எந்த மொழிவகையினை (language variation) அதிகமாகப் பயன்படுத்துகின்றனர் என்பதையும் தமிழ் எழுத்துருக்களைக் குறைந்த அளவிலேயே பயன்படுத்துவதற்கான காரணங்களையும் இக்கட்டுரை விவரிக்கும்.

முன்னுரை

ஒரு மொழி முழுமையாக வெளிப்படுவதற்கு உதவும் கருவிகள் பற்பலவாகும். தமிழ்மொழி, தன் நெடிய வரலாற்றில் பல்வேறு ஊடகங்களின் வழி வளர்ந்து வந்திருப்பதை அறியவியல்கிறது. ஓலைச்சுவடி, காகிதம், தொலைபேசி, வானொலி, தொலைக்காட்சி என்னும் கருவிகளின் வரிசையில் இன்று இணையம், கைபேசி ஆகியவற்றின் மூலமாகவும் தமிழ் தொடர்ந்து வெளிப்பட்டு வருகின்றது. வெவ்வேறு கருவிகளின் வழி வளருவதால் தமிழ்மொழியின் நடை, வடிவம் ஆகியவற்றிலும் தொடர்ந்து மாற்றங்கள் நிகழ்ந்து வருகின்றன.

ஓலைச்சுவடிகளில் எழுதிய போது, அதன் அளவு தன்மை போன்றவற்றிற்கேற்ப புள்ளிகள் இடாமல், சுருக்கமாக, செய்யுள் வடிவிலான மொழிநடை பயன்படுத்தப்பட்டது. காகிதங்களும் அச்ச இயந்திரங்களும் புழக்கத்திற்கு வந்ததும், உரைநடை வடிவில் விரிவாகவும் விளக்கமாகவும் எழுதும் வழக்கம் பரவலாகப் பின்பற்றப்பட்டது. இன்றும் அஞ்சலட்டைகள், தந்தி போன்றவற்றில் எழுதும்பொழுது, அவற்றின் அளவும் தேவையும் கருதிச் சுருக்கமாகவும் செறிவாகவும் எழுதுகிறோம். ஆனால், கடிதம், உரை, கட்டுரை போன்றனவற்றில் தெளிவாகவும் விரிவாகவும் எழுத முனைகிறோம். எனவே, ஒரு மொழியின் அமைப்பையும் நடையையும் நிர்ணயிப்பதில் அம்மொழியைப் பயன்படுத்துகின்ற ஊடகம் முக்கியக் காரணியாகின்றது எனலாம். அவ்வகையில் இக்கட்டுரை சமூக வலைத்தளங்களின் வருகையால் தமிழ்மொழிப் பயன்பாட்டில் ஏற்பட்டுள்ள மாற்றங்களை ஆராய்கிறது.

ஆய்வின் பிற்புலம்

இன்றைய அறிவியல் துறைகளின் வளர்ச்சி முழுக்க முழுக்க மேனாட்டார்களால் உருவாக்கப்பட்டு வளர்க்கப்பட்டது என்றும் அவற்றிற்கான ஆங்கிலக் கலைச்சொற்களே ‘உலகப் பொதுமொழி’ போன்று உலகத்தாரால் பயன்படுத்தப்படுகின்றன என்றும் மணவை முஸ்தபா (2010) கருத்துரைத்துள்ளார். அவரது கருத்துக்குக் கணினியும் இணையமும் விதிவிலக்கல்ல. இருப்பினும், தொழில்நுட்பம் அறிந்த கணிஞர்கள், மொழி ஆர்வலர்கள் ஆகியோரின் முயற்சியால், உலகப் பொதுமொழி போல் திகழும் ஆங்கிலம் தவிர, பிற மொழிகளிலும் கணினியை இயக்கவும் இணையத்தளங்களில் உலாவவும் முடிகிறது. தமிழ் ஆர்வலர்களும் தமிழ்க்கணிஞர்களும் செய்த சேவையின் பயனாகவே கணினியிலும் இணையத்திலும் தமிழைப் (தமிழாகவே) பயன்படுத்துவதற்கான வாய்ப்பும் வசதியும் ஏற்பட்டுள்ளன. மேலும், இன்று இணையத்தளங்களில் தமிழ்மொழி சார்ந்த வலைத்தளங்களும் அகப்பக்கங்களும் மலிந்து கிடப்பதற்கும் இவர்கள் ஆற்றிய அரும்பணியே காரணம் என்பதையும் மறுக்கவியலாது.

தமிழில் எழுத-படிக்க-பேசத் தெரிந்த தமிழர்களைத் தவிர, தமிழில் எழுத-படிக்கத் தெரியாத ஆனால், தமிழில் பேசி உரையாடத் தெரிந்த தமிழர்களும் உளர். மேலும் தமிழில் எழுத-படிக்க-பேசத் தெரியாத ஆனால், தமிழில் பேசுவதைப் புரிந்துகொள்ளும் ஆற்றல் மட்டுமே கொண்ட தமிழர் சிலரும் உளர். இப்படி மூன்று நிலையிலான தமிழர்களும் சமூகவலைத்தளங்களில் தங்களின் கருத்துகளைப் பகிர்ந்துகொள்கின்றனர். எனவே, தமிழிலேயே தட்டச்சு செய்தல், ரோமன் எழுத்தில் தமிழைத் தட்டச்சு செய்தல், ரோமன் எழுத்தில் தட்டச்சுச் செய்யப்பட்ட தமிழ்மொழி கருத்துகளுக்கு ஆங்கிலத்தில் பதிலளித்தல் எனச் சமூகவலைத்தளங்களில் இடப்பட்டுள்ள கருத்துகளை

வகைப்படுத்தலாம். இவற்றைத் தவிர பன்மொழிச்சூழலில் வாழும் தமிழர்களின் அன்றாட வழக்கில் கலப்பு மொழி என்பது தவிர்க்க முடியாத ஒன்றாகிவிட்டமையால் முகநூல் உரையாடலிலும் கலப்பு மொழியினைக் காண முடிகின்றது. இப்படியாக இக்கட்டுரை முகநூலில் தமிழர்கள் பயன்படுத்தும் மொழி வகைகளை (language variations) ஆராய்கிறது.

மலேசியர்களின் விருப்பத்திற்குரிய சமூக வலைத்தளமான முகநூலில், தமிழ் படிக்கத் தெரிந்தவர்கள் தமிழிலேயே கருத்திடுகின்றனர் என்பதை ஆய்வது இக்கட்டுரையின் பொது நோக்கமாக இருப்பினும் அவர்கள் பயன்படுத்தும் மொழிவகைகளை ஆய்வதே இக்கட்டுரையின் முதன்மை நோக்கமாகும். அதோடு தமிழ் எழுத்துருக்களைக் குறைந்த அளவிலேயே பயன்படுத்துவதற்கான காரணங்களையும் இக்கட்டுரை விவரிக்கும்.

ஆய்வுச் சிக்கல்

இளஞ்செழியன் (2011), கட்டற்ற கலைக்களஞ்சியமான ‘விக்கிபீடியாவில்’ (www.ta.wikipedia.org) 19, 638 தமிழ்க் கட்டுரைகள் இடம்பெற்றுள்ளன. அதுவே, ஆங்கிலமொழி கட்டுரைகளின் எண்ணிக்கையோ 30.7 இலட்சத்திற்கும் மேல் உள்ளன என்ற சான்றை முன் வைக்கிறார். அவரே, ஆங்கிலம் பேசுவோரின் எண்ணிக்கை தமிழ் பேசுவோரைவிட ஏறத்தாழ 15 மடங்கு அதிகமாக இருந்தாலும், இணையத்தில் இவ்வேறுபாடு 100 மடங்குக்கும் மேலாக உயர்கின்றது என்கிறார். இவ்வளவு பெரிய இடைவெளிக்குக் காரணம் என்ன? என்ற கேள்வியே இக்கட்டுரை உருவாவதற்குக் காரணமாக விளங்குகிறது. தமிழ் பேசுபவர்கள் அனைவரும் இணையத்தில் தமிழைப் பயன்படுத்துவதில்லையா? இவர்கள் பயன்படுத்தும் மொழிவகைகள் யாவை? என்பனவற்றை அறியும் ஆவல் தலைதாக்கியது. எனவே, இணையத்தில் குறிப்பாக அதிகமானோர் ஈடுபடுகின்ற சமூகவலைத்தளங்களில் தமிழர்கள் பயன்படுத்தும் மொழிவகைகளை ஆராய்ந்தால் மேற்சொன்ன வினாக்களுக்கு விடைகாண முடியும் என்ற எதிர்பார்ப்பில் இந்த ஆய்வு மேற்கொள்ளப்பட்டது.

ஆய்வு முறை

தமிழ் படிக்கத் தெரிந்த தமிழர்கள், முகநூலில் பயன்படுத்தும் மொழிவகைகளைக் கண்டறிவதே இக்கட்டுரையின் நோக்கம் என்பதால் தமிழ் மொழியில் எழுதப்பட்ட பனுவல்கள் (text) தொடர்பான கருத்துகள் சேகரிக்கப்பட்டன. முகநூலில் இடம்பெற்ற தமிழ்ப்பனுவல்கள் தொடர்பான கருத்துகளை ஆய்ந்து, ஆறு பனுவல்கள் மட்டும் ஒரு குறிப்பிட்ட முறையின்றி (random sampling) மாதிரிகளாகத் தேர்ந்தெடுக்கப்பட்டன. அவை முறையே ‘ஓம்தமிழ்.காம்’-இன் “தமிழர் பரம்பரை” எனும் பட்டியல், “வணக்கத்தின் வலிமை” எனும் உரைவீச்சு, பாரதியாரின் “தேடிச்சோறு நிதம் தின்று” எனும் பாடல் பகுதி, இறையருளை வேண்டுகின்ற பிரார்த்தனை வரிகள், மலேசியப் படைப்பாளர் தயாஜியின் “தேவதையானவன்” எனும் உரைவீச்சு, தமிழாசிரியர் சுப. நற்குணனின் ஆசிரியர் பணியைப் பற்றிய கூற்று ஆகியனவாகும்.

அடையாளங்கண்ட பனுவல்களில் கருத்திட்டோரின் பெயர்கள் நீக்கப்பட்டுத் தகவலாளர்களாக எண்ணிடப்பட்டன (ப1த1, ப1த2...). கருத்திட்ட நாள், நேரம் ஆகியன ஆய்வுக்குத் தேவையற்றவை என்பதால் அவையும் நீக்கப்பட்டு, கருத்துகள் மட்டுமே ஆய்வுக்கு எடுத்துக்கொள்ளப்பட்டன.

அட்டவணை 1: ஆய்வின் தகவலாளர்கள்

பனுவல்	தகவலாளர்கள் (sample)	கருத்துகள்
ப1	9	11
ப2	12	14
ப3	13	13
ப4	16	32
ப5	18	34
ப6	33	34
மொத்தம்	100	138

அட்டவணை 1 தேர்ந்தெடுக்கப்பட்ட பனுவல்களில் ஈடுபட்டோரின் எண்ணிக்கையையும் அவர்கள் பதிவேற்றிய கருத்துகளின் எண்ணிக்கையையும் காட்டுகின்றது. தேர்ந்தெடுக்கப்பட்ட ஆறு பனுவல்களையொட்டி 100 பேர் கருத்திட்டுருந்தனர். சிலர் ஒரு முறைக்கு மேல் கருத்திட்டிருந்தமையால் மொத்தம் 138 கருத்துகள் இடம்பெற்றிருந்தன. இக்கட்டுரையின் மற்றொரு நோக்கமான, தமிழைத் தமிழிலேயே தட்டச்சு செய்யாததற்கான

காரணத்தை அறிய, தகவலாளர்களிடம் ஒரு கருத்துக் கணிப்பு மேற்கொள்ளப்பட்டது. இக்கருத்துக் கணிப்பில் 100 தகவலாளர்களில் 68 பேர் மட்டுமே கருத்திட்டிருந்தனர்.

சமூகவலைத்தளங்களில் தமிழர்கள் பயன்படுத்தும் மொழி வகைகள்

ஆய்வுக்கு உட்படுத்தப்பட்ட 100 பேரும் தமிழ்ப் படிக்கத் தெரிந்தவர்கள் என்றாலும் அவர்கள் அனைவரும் ஒரே மொழிவகையைப் (variation) பயன்படுத்தவில்லை. தமிழைத் தமிழ் எழுத்துருக்களிலும் ரோமன் எழுத்துருக்களிலும் தட்டச்சு செய்துள்ளனர். மேலும், பேச்சுமொழி எழுத்துமொழி ஆகிய இரு வழக்குகளிலும் தமிழைப் பயன்படுத்தியிருக்கின்றனர். தமிழ்மொழியைத் தவிர ஆங்கிலம், மலாய் எனப் பிற மொழிகளின் பயன்பாடும் இவர்களது கருத்துகளில் அடையாளங்காணப்பட்டன. பிறமொழிகளைப் பயன்படுத்தும் பொழுது அவற்றில் சொற்சுருக்கம், குறியீட்டுமொழி முதலியனவும் இடம்பெற்றிருந்தன.

தமிழ்மொழியில் இடப்பட்ட கருத்துகள்

தேர்ந்தெடுக்கப்பட்ட பனுவல்கள் அனைத்தும் தமிழ்மொழியிலேயே அமைந்திருந்த போதிலும், அடையாளங்காணப் பட்ட 138 கருத்துகளில் 78 கருத்துகள் மட்டுமே தமிழ் மொழியில் இடப்பட்டிருந்தன. தமிழில் இடப்பட்டிருந்த 78 கருத்துகளில் இருவேறு எழுத்துருக்கள் பயன்படுத்தப்பட்டிருந்தன.

காட்டாக: ப6த20: நன்றி நண்பரே. தங்களுக்கும் எங்களின் இனிய ஆசிரியர் தின வாழ்த்துகள்
ப5த15: *nandri nanbare...*

ஆசிரியர் பணியைப் பற்றிய ஆறாவது பனுவலில் 20ஆம் தகவலாளர் “நன்றி நண்பரே” என்று கருத்திட்டுள்ளார். அதே கூற்றை, “தேவதையானவன்” எனும் ஐந்தாவது பனுவலின் 15ஆம் தகவலாளர் ‘*nandri nanbare*’ என்று கருத்திட்டிருக்கிறார். இவ்வாறு தமிழ் எழுத்துருக்களாலும் ரோமன் எழுத்துருக்களாலும் தமிழில் இடப்பட்ட கருத்துகளின் எண்ணிக்கை பனுவல் வாரியாக அட்டவணை 2இல் காட்டப்பட்டுள்ளது.

அட்டவணை 2: தமிழில் இடப்பட்ட கருத்துகள்			
பனுவல்	கருத்துகளின் எண்ணிக்கை		
	தமிழ் எழுத்துரு	ரோமன் எழுத்துரு	மொத்தம்
ப1	5	2	7
ப2	-	6	6
ப3	5	4	9
ப4	1	10	11
ப5	13	10	23
ப6	18	4	22
மொத்தம்	42	36	78

இவ்வட்டவணையிலிருந்து ஆய்வுக்கு எடுத்துக்கொள்ளப்பட்ட 138 கருத்துகளில் 57% கருத்துகளே தமிழில் இடப்பட்டுள்ளன என்பது தெளிவாகின்றது. அவற்றிலும் தமிழ் எழுத்துருக்களைக் கொண்டு தமிழிலேயே இடப்பட்ட கருத்துகள் 42 மட்டுமே. இது மொத்த கருத்துகளோடு ஒப்பிடுகையில் 30% மட்டுமே. 26% அதாவது 36 கருத்துகள் ரோமன் எழுத்துருக்களால் இடப்பட்டிருந்தன. எஞ்சிய 60 கருத்துகளில் 41 கருத்துகள் அதாவது 30% கருத்துகள் ஆங்கிலமொழியிலும், 19 (14%) கருத்துகள் கலப்பு மொழியிலும் இடப்பட்டிருந்தன.

இருநிலைவழக்கில் (diglossia) இடப்பட்ட கருத்துகள்

பேச்சுவழக்கு, எழுத்து வழக்கு எனும் இருநிலைவழக்கு (diglossia) கொண்டது தமிழ்மொழி. முகநூலில் இடப்பட்ட கருத்துகளிலும் இந்த இருநிலைவழக்கையும் காண முடிந்தது.

காட்டாக: ப1மா1: என்னோட பூட்டன் வரை நினைவு இருக்கு. பூட்டி பேரு என்னனு தெரியலை.
ப6மா27: ஆசிரியர் தின நல்வாழ்த்துக்கள் ஐயா...
இப்படிக்கு உங்கள் முன்னாள் மாணவன்.

இவ்விருநிலை வழக்குகளில் ரோமன் எழுத்துருக்களைக் கொண்டு எழுதப்பட்ட தமிழ்மொழி கருத்துகள் பெரும்பாலும் பேச்சு வழக்கிலேயே இடப்பட்டுள்ளன. தமிழ் எழுத்துருக்களைக் கொண்ட கருத்துகளில் இருவழக்கும் கலந்த நிலையே அதிகமாக உள்ளது.

ஆங்கில மொழியில் இடப்பட்ட கருத்துகள்

முகநூலில் இடம்பெற்ற தமிழ்ப் பனுவல்களை வாசித்துக் கருத்திடுகின்றனர் என்றால் ஆய்வுக்கு உட்படுத்தப்பட்ட தகவலாளர்கள் அனைவரும் தமிழ் அறிந்தவர்கள் என்பது புலனாகிறது. இருப்பினும், இவர்கள் தங்கள் கருத்துப் பரிமாற்றத்தில் ஆங்கிலத்தை அதிகமாகவே பயன்படுத்தியுள்ளனர் (30%).

காட்டாக: ப1த5: tamil is grate (great)

“தமிழர் பரம்பரை” எனும் முதல் பனுவலின் 5ஆம் தகவலாளர் “tamil is grate” என்று கருத்திட்டுள்ளார். தமிழ் சிறந்த மொழி என்பதை ஆங்கிலத்தில் கருத்திட்டுள்ளார். அதுவும் “great” என்பதை “grate” என்று எழுதியுள்ளார். இது போல் உச்சரிப்புக்கேற்பவும் சுருக்கமாகவும் எழுத வேண்டும் என்பதற்காக ஆங்கிலத்தை வேறுவிதமாக எழுத்துக் கூட்டியிருப்பதையும் பிறரது இடுகைகளிலும் காண முடிகிறது.

காட்டாக: ப2த4 : exsillant (excellent) ப4த5 : tanx (thanks)
ப5த10: tats (that's) ப6த10: hpy (happy)

கலப்பு மொழியில் இடப்பட்ட கருத்துகள்

தமிழ்மொழியில் இடப்பட்ட கருத்துகள், ஆங்கிலமொழியில் இடப்பட்ட கருத்துகள் என்பனவற்றோடு தமிழ், ஆங்கிலம், மலாய் ஆகிய மொழிகள் கலந்து இடப்பட்ட கருத்துகளும் அடையாளங்காணப்பட்டன. அவற்றில் ரோமன் எழுத்துருவிலான தமிழ்மொழியோடு ஆங்கில மொழிச் சொற்களையும் கலந்து தட்டச்சு செய்திருந்தனர்.

காட்டாக: ப3மா4: Sema Lines
ப4மா6: thanks da nanba

பாரதியார் பாடலைக் கொண்ட மூன்றாவது பனுவலின் 4ஆம் தகவலாளர் “செம்மை” என்பதை ரோமன் எழுத்துருவில் “Sema” என்று குறிப்பிட்டுள்ளார். இத்தகைய மொழிக்கலப்பில் அமைந்த 17 கருத்துகள் அடையாளங்காணப்பட்டன. அதே வேளையில், தமிழ் எழுத்துருவோடு ஆங்கில மொழி கலந்து இடப்பட்ட கருத்துகள் இரண்டு மட்டுமே அடையாளங்காணப்பட்டன.

காட்டாக: ப5த3: thx தோழியே... ப5த18: ஜந்துகளுக்குத் தெரியுமா
நான் பேனாக்காரன் மட்டுமல்ல நட்பின் ஆழம்!!!
கொஞ்சம் இதுக்கள் தாயைக்கூட
நான் கோவக்காரனும் கூட... சங்கடப்படுத்தும்
கேவலங்கள்!!!!... TRUE

“தேவதையானவள்” எனும் ஐந்தாவது தலைப்பில் 3ஆம் தகவலாளரும் 18ஆம் தகவலாளரும் இப்படித் தமிழும் ஆங்கிலமும் கலந்து கருத்திட்டுள்ளனர்.

அதே பனுவலின் 10ஆம் தகவலாளர் ஆங்கிலமும் மலாய்மொழியும் கலந்து கருத்திட்டுள்ளார்.

காட்டாக: ப5த10: tats rite.....i sokong u..... fren

இப்படி ஆங்கிலமும் மலாய்மொழியும் கலந்த கருத்து இஃது ஒன்று மட்டுமே. மொத்தத்தில் கலப்புமொழியில் 19 (14%) கருத்துகள் இடப்பட்டிருந்தன.

குறியீடு (emitecon) இடப்பட்ட கருத்துகள்

கைப்பேசிகளில் அனுப்பப்படும் குறுந்தகவல்களில் பயன்படுத்தப்படும் குறியீடுகள் முகநூல் உரையாடல்களிலும் இடம்பெறக் காணலாம்.

காட்டாக: ப2த9 : Nandri jin.: ♥ :)
ப5த15: ippadi oru tholan enaku illaye endru varunthugiren...: (

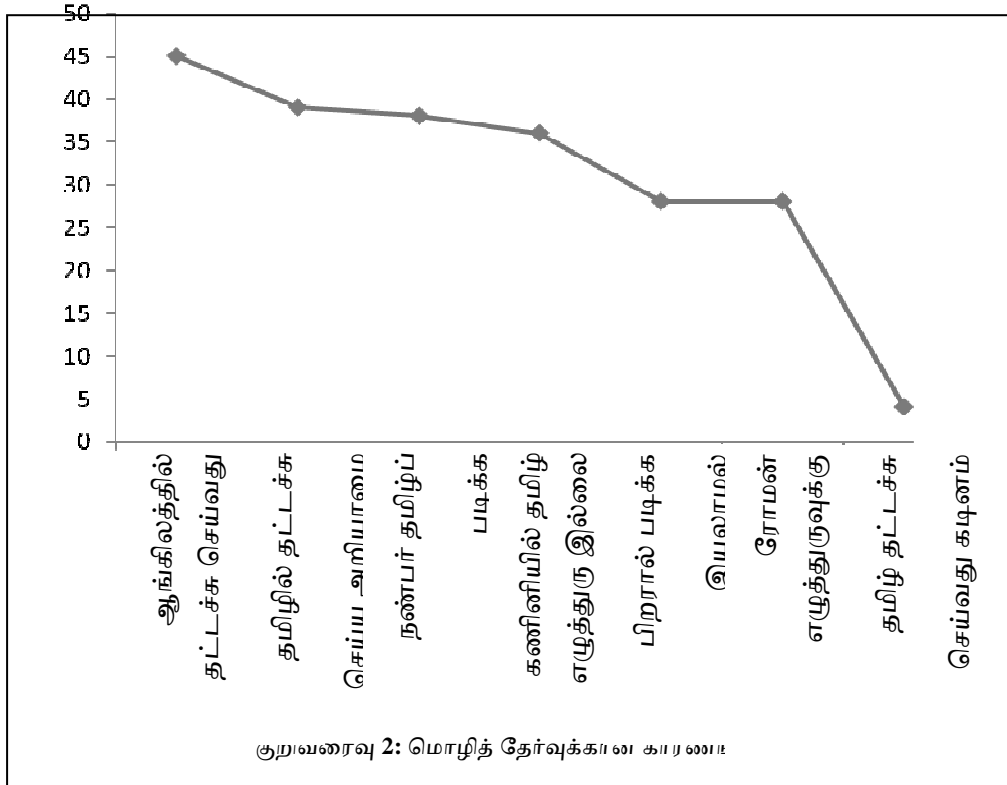
காட்டுகளில் இடம்பெற்றுள்ள : ♥, :), : (எனும் குறியீடுகள் ஒவ்வொன்றும் ஓர் உணர்ச்சியை வெளிப்படுத்துவதற்காகப் பயன்படுத்தப்பட்டுள்ளன.

♥ - விருப்பம்,
:) - மகிழ்ச்சி,
: (- வருத்தம்

இத்தகைய குறியீடுகள் 10 கருத்துகளில் இடம்பெற்றிருந்தன. அவற்றுள் 9 கருத்துகள், ரோமன் எழுத்துருவில் அமைந்த தமிழ் கருத்துகளாகவும் ஆங்கிலத்தில் இடப்பட்ட கருத்துகளாகவும் அமைந்திருந்தன. தமிழ் எழுத்துருவில் அமைந்த ஒரே ஒரு கருத்தில் மட்டுமே இத்தகைய குறியீடு பயன்படுத்தப்பட்டிருந்தது.

தமிழ் எழுத்துருக்களைக் குறைந்த அளவில் பயன்படுத்துவதற்கான காரணங்கள்

இருவகை எழுத்துருக்களைக் கொண்ட தமிழ், இருநிலைவழக்குத் தமிழ், ஆங்கில மொழி, கலப்பு மொழி, குறியீட்டு மொழி என முகநூலில் தமிழர்கள் பயன்படுத்தும் ஐந்து மொழிவகைகள் கண்டறியப்பட்டன. ஆய்வுக்குட்படுத்தப்பட்ட மொத்த தகவலாளர்கள் 100 பேரில் 56 விழுக்காட்டினர் ரோமன் எழுத்துருவில் அமைந்த தமிழையும் ஆங்கில மொழியையும் பயன்படுத்துகிறார்கள் என்பதும் அறியப்பட்டது. எனவே, தமிழறிந்த தமிழர்கள் முகநூலில் தமிழ் எழுத்துருவைப் பயன்படுத்தாமைக்கான காரணங்களைத் தொடர்ந்து காண்போம்.



கணிஞர்கள் சிலரோடும் தமிழ் ஆசிரியர்கள் சிலரோடும் கலந்துரையாடியதில் கண்டறியப்பட்ட ஆறு காரணங்களைப் பதிவேற்றி கருத்துக் கணிப்பு மேற்கொள்ளப்பட்டது. “தமிழில் தட்டச்சு செய்தல் கடினம்” எனும் ஏழாவது கருத்தைத் தகவலாளர்களே பதிவேற்றியிருந்தனர். 100 தகவலாளர்களுள் 68 பேர் மட்டுமே இக்கருத்துக்கணிப்பில் பங்கு கொண்டனர். அவர்களுள் சிலர் ஒன்றுக்கும் மேற்பட்ட காரணங்களுக்கு அடையாளமிட்டிருந்தனர்.

தமிழைக் காட்டிலும் ஆங்கிலத்தில் தட்டச்சு செய்வது எளிது என்ற காரணத்திற்கு 45 பேர் அடையாளமிட்டிருந்தனர். தமிழில் தட்டச்சு செய்ய அறியாதவர்களின் எண்ணிக்கை 39 ஆகும். 38 பேர் தங்கள் நண்பர்கள் தமிழ் படிக்க அறியாதவர்கள் என்று அடையாளமிட்டனர். தங்கள் கணினியில் தமிழ் எழுத்துருக்கள் இல்லாமல் இருப்பவர்கள் 36 பேர். தலா 28 பேர், பிறரால் தங்கள் கருத்துகளைப் படிக்க இயலாமல் போகலாம், தமிழில் தட்டச்சு செய்து

கொண்டிருக்கும் பொழுது அதுவாகவே, இடையிடையே ரோமன் எழுத்துருவுக்கு மாறிவிடுகிறது எனும் இரண்டு காரணங்களைத் தெரிவு செய்துள்ளனர். தமிழில் தட்டச்சு செய்வது கடினம் எனும் காரணத்துக்கு 4 பேர் மட்டுமே அடையாளமிட்டிருந்தனர்.

பரிந்துரைகள்

ஆங்கில மொழியில் தட்டச்சு செய்வது எளிது என்பதையே, ஆய்வுக்குட்படுத்தப்பட்ட தகவலாளர்கள் தங்களின் மொழிவகைத் தேர்வுக்கு முதற்காரணமாகச் சுட்டியுள்ளனர். இணையப் பயணர்கள் பெரும்பாலோர் இத்தகைய எண்ணமுடையவர்களாகவே இருக்கின்றனர். இத்தகையோர் தங்களது எண்ணத்தை மாற்றிக்கொண்டாலொழிய இணையத்தளங்களில் தமிழ்மொழிப் பயன்பாடு எதிர்பார்க்கின்ற வளர்ச்சியை அடையாது. சித்திரமும் கைப்பழக்கம் என்பது போல, பழகப் பழகத்தான் எதுவும் எளிமையாகும் என்பதனை ஏற்று, இவர்கள் தமிழில் தட்டச்சு செய்ய முனைய வேண்டும்.

தமிழில் அமைந்த பணுவல்கள் தொடர்பான கருத்துப் பரிமாற்றங்கள் மட்டுமே இந்த ஆய்வுக்குட்படுத்தப்பட்டன. எனவே, ஆய்வுக்குட்படுத்தப்பட்ட அனைவரும் தமிழ் படிக்க அறிந்தவர்கள் என்று உறுதியாகக் கூறலாம். இருப்பினும், தங்கள் நண்பர்கள் தமிழ் படிக்க அறியாதவர்கள் எனும் காரணத்தினாலேயே தாங்கள் ரோமன் எழுத்துருவில் தமிழைப் பயன்படுத்துவதாகவும் ஆங்கிலத்தில் தட்டச்சு செய்வதாகவும் 39 தகவலாளர்கள் அடையாளமிட்டுள்ளனர். இவர்களது கருத்து, ஆய்வுக்குட்படுத்தப்பட்ட தகவலாளர்களைக் குறிக்காது என்பது கவனிக்கத் தக்கது. பல நாடுகளில் பன்மொழிச் சூழலில் வாழுகின்ற தமிழ் பேசு அறிந்த தமிழர்கள் பலர், தமிழில் படிக்க அறியாமல் இருக்கின்றனர் என்பது வருத்தத்திற்குரிய உண்மை. தமிழை இரண்டாம் மொழியாகக் கற்பதற்கென அமைக்கப்பட்டுள்ள வலைப்பக்கங்களை அவர்களுக்கு அறிமுகப்படுத்தி, அவர்களைத் தமிழ் படிக்க ஊக்கப்படுத்த வேண்டும்.

தமிழில் தட்டச்சு செய்ய அறியாமை எனும் கருத்துக்கும் தங்கள் கணினியில் தமிழ் எழுத்துரு இல்லை எனும் கருத்துக்கும் 50 விழுக்காட்டுக்கும் மேற்பட்டோர் அடையாளமிட்டுள்ளனர். 4 தகவலாளர்கள் தமிழில் தட்டச்சு செய்வது கடினம் எனும் காரணத்தையும் குறிப்பிட்டுள்ளனர். பெரும்பாலான இணையப் பயணர்கள் தங்களாகவே தட்டுத்தடுமாறிதான் (trial and error) தமிழில் தட்டச்சு செய்யக் கற்றுக் கொள்கின்றனர். எனவே, தமிழ் எழுத்துருவைப் பெறுவதற்கான வழிகளையும் தமிழில் தட்டச்சு செய்யும் முறையினையும் பள்ளிகளிலேயே கற்பிக்க வேண்டும். தமிழாசிரியர்கள் இதில் முக்கியப் பங்காற்ற வேண்டும். தங்களிடம் தமிழ் கற்கின்ற மாணவர்களுக்குக் கற்றல் கற்பித்தல் துணைப்பொருளாக இணையத்தளத் தகவல்களைப் பயன்படுத்தி இணையத்தில் தமிழைப் பயன்படுத்தும் முறையையும் கற்பிக்க வேண்டும்.

தம் நண்பர்களின் கணினியில் தமிழ் எழுத்துரு இல்லாமையால் அவர்களால் படிக்க இயலாமல் போகலாம், தமிழில் தட்டச்சு செய்து கொண்டிருக்கும்பொழுது, தாமாகவே ஆங்கில எழுத்துருவாக அல்லது புரியாத குறியீடாக மாறிவிடுகின்றது எனும் இரண்டு காரணங்களையும் ஒரே அளவிலான தகவலாளர்கள் (28 பேர்) தெரிவு செய்துள்ளனர். இச்சிக்கலைக் களைவதில் தமிழ்க் கணிஞர்கள்தாம் உதவ வேண்டும். தமிழ் எழுத்துருக்களைப் பணம் கொடுத்துத்தான் பெற வேண்டும் என்ற நிலை இணையத்தில் தமிழ்மொழியைப் பயன்படுத்துவோரின் எண்ணிக்கையை அதிகரிக்கத் துணை செய்யாது. எனவே, தமிழ் எழுத்துருக்களை விற்பனை செய்பவர்கள் சற்றுப் பொது நோக்கோடு சிந்தித்து, தமிழறிந்த அனைவரும் தமிழ் எழுத்துருவைப் பயன்படுத்த வகை செய்ய வேண்டும். எந்தக் கணினியை வாங்கினாலும் எந்த மென்பொருளைப் பயன்படுத்தினாலும் ஆங்கிலம் போல் தமிழையும் இலகுவாகப் பயன்படுத்தலாம் எனும் நிலையைக் கணிஞர்கள் உருவாக்கித் தர வேண்டும்.

நிறைவு

இணையத்தில் தமிழ்மொழியைப் பயன்படுத்துவோரின் எண்ணிக்கை அதிகரிக்கவும், தமிழ் இணையப் பயன்பாடு நீடித்து நிலைக்கவும் தமிழாசிரியர்களும், தமிழ்க் கணிஞர்களும் ஒன்றிணைந்து செயல்பட வேண்டும். தமிழறிந்த இணையப் பயணர்கள் அனைவரும் தமிழரோடு உரையாடும் பொழுது தமிழில் தட்டச்சு செய்வதைக் கட்டாயமாக்கிக் கொள்ள வேண்டும். அனைவரும் ஒன்றிணைந்து செயல்பட்டால், இனிய தமிழ் இணையத்தில் இலகுவத் தமிழாக உலா வரும்.

மேற்கோள் நூல்கள்

1. இளஞ்செழியன், வே. (2011). மலேசியப் பன்மொழிச் சூழலில் தமிழ்; கணினி தமிழ் பேசுமா? கோலாலும்பூர்: மலாயாப் பல்கலைக்கழகம்.
2. திருவாசகம், க. (2010). உலகத் தமிழ் செம்மொழி மாநாட்டுச் சிறப்பு மலர்; கணினித் தமிழ் வளர்ச்சி. தமிழ்நாடு.

3. மணவை முஸ்தாபா. (2010). உலகத் தமிழ் செம்மொழி மாநாட்டுச் சிறப்பு மலர்; கணினிக் கலைச்சொல்லாக்கம். தமிழ்நாடு.
4. Courtenay Honeycutt. (2011). *Examining The Diffusion of CMC Technologies in Minority Languages: A Case Study of Welsh Language Blogging and Twittering*. Bloomington: Indiana University.
5. Maslina Binti Osman (2008). *Kajian Penggunaan Bahasa Melayu Dalam Semang Siber*. Kuala Lumpur: Universiti Malaya.
6. Supyan Hussin. (2007). *Korpus Bahasa SMS: Antara Kreativiti dan Ancaman Kepada Bahasa* dlm. Linguistik Teori dan Aplikasi. Bangi: Penerbit Universiti Kebangsaan Malaysia.

கலாசாரப் பெயர்ப்பு

Ilangkumaran Sivanadhan

Sultan Idris Education University, Malaysia

மொழிபெயர்ப்பு என்பது ஒரு மொழியில் இருக்கும் தரவுகளை இன்னொரு மொழிக்கு மாற்றம் செய்வது என்பதோடு நின்றுவிடுவது அல்ல. மாறாக, ஒரு கலாசார தரவுகளை இன்னொரு கலாசாரத்துக்குக் கொண்டு செல்வதிலும் மொழிபெயர்ப்பாளர்களின் பணி அடங்கியிருக்கின்றது. பெருவாரியான மொழிபெயர்ப்பாளர்கள் உட்பட பெரும்பாலானோர் இதனை கருதுவதில்லை என்பதால் இதற்கு முக்கியத்துவம் அளிக்கப்படவில்லை என்றே கூற வேண்டும். ஆகையால் 'கலாசார பெயர்ப்பு' என்பது பல காலங்களாக மொழிபெயர்ப்புப் பணியில் ஈடுபட்டிருப்போருக்கும் கூட கேட்பதற்குப் புதுமையான ஒன்றாக இருக்கும். கேட்பதற்குப் புதுமையானதாக இருக்குமே தவிர, அவர்கள் இதனைக் கண்டிப்பாக கடந்து வராமல் இருந்திருக்க முடியாது. ஏனெனில் எப்போது அவர்கள் ஒரு மொழியில் இருக்கும் இலக்கியங்களை இன்னொரு மொழிக்கு மொழிபெயர்க்க எத்தணித்தனரோ அப்போதே இதனை அவர்கள் செய்தே இருக்க வேண்டும். ஆக, அத்தகையோர் அறியாமல் செய்து; செய்துவரும் கலாசார பெயர்ப்பை அவர்களுக்கும் இனி வரும் மொழிபெயர்ப்பாளர்களுக்கும் அறிமுகப்படுத்துவதே இவ்வாய்வுக் கட்டுரையின் நோக்கமாகும்.

மக்களில் பெரும்பாலானோர் தங்கள் தாய்மொழியிலேயே சிந்திக்கின்றனர் என ஆய்வுகள் கூறுகின்றன. அவர்களுக்குப் பல மொழிகள் தெரிந்திருப்பினும், பல்வேறு மொழிகளைப் பேசும் மக்கள் வாழும் சூழலில் வாழ்ந்து, அவர்களோடு பிற மொழிகளிலும் சரளமாக உரையாடும் திறம் கொண்டவர்களும் தத்தமது தாய்மொழியில் சிந்தித்தே சம்பாஷனையில் ஈடுபடுகின்றனர். சாதாரணமாக பார்க்கும் வேளையில் இதனை ஒரு பொருட்டாக நாம் கருதுவதில்லை என்றபோதும், கூர்ந்து நோக்குகையில், ஒரு மொழியில் சிந்தித்து அதனை சிற்சில விநாடிகளிலேயே பிற மொழிகளில் மாற்றி அந்தந்த மொழி பேசும் மக்களுக்குப் பதிலளிப்பது என்பது நமது மூளையின் மகத்துவத்தைக் காடுவதாக இருக்கின்றது. இதிலிருந்து நமக்கு அறியவரும் இன்னொரு தகவல் என்னவெனில், ஒவ்வொரு மனிதனும் மொழிபெயர்ப்பாளன்தான் என்பதே ஆகும்.

ஒவ்வொரு மொழிக்கும் அதன் கலாசார பின்னணி கண்டிப்பாக இருக்கும். அவை சில வேளைகளில் ஒரு சில மொழிகளோடு ஒத்துப்போகலாம். இவ்வாறு ஒத்துபோகும் மொழிகளுக்கிடையிலான மொழிபெயர்ப்புகளைச் செய்யும்போது, கலாசார பெயர்ப்பில் சிக்கல்கள் மிகக் குறைவாகவே இருக்கும். உதாரணத்திற்கு, ஆங்கில மொழியும் பிரஞ்சு மொழியும் வெவ்வேறு மொழிக்குடும்பத்தைச் சார்ந்தவை. ஆனாலும், இவ்விரு மொழிகளின் (இந்த மொழி பேசுவோரது கலாசாரம்) கலாசாரங்களும் ஏறக்குறைய ஒன்றாகவே இருக்கும். ஆதலால், இவ்விரு மொழிகளுக்குமிடையே மொழிபெயர்ப்பு செய்யும் வேளையில் கலாசார பெயர்ப்பு குறித்து அதிகம் கவலைப்படுவதற்கில்லை. அவை இயல்பாகவே வந்துவிடுகின்றன.

ஆனால், பெரும்பாலான மொழிகளின் கலாசாரப் பின்னணிகள் ஒன்றுக்கொன்று ஒத்துப்போகாமல் தான் இருக்கும். இப்படிப்பட்ட மொழிகளுக்கிடையில் செய்யப்படும் மொழிபெயர்ப்பில் மிகுந்த அக்கறை காட்ட வேண்டியது அவசியமாகும். வெவ்வேறான மொழிக்குடும்பங்களில் இருக்கும் மொழிகளில் இத்தகைய விஷயங்கள் ஏற்படுவது சகஜமான ஒன்று. ஆனால், ஒரே மொழிக்குடும்பத்தில் இருக்கும் மொழிகளுக்கிடையேயும் இத்தகைய சூழல் இருப்பது கவனிக்க வேண்டிய ஒன்று. இத்தகு வேளையில்தான் மொழிபெயர்ப்பாளர்கள் விழிப்புடன் செயல்பட வேண்டும். இதனை விளக்க உதாரணம் வேண்டுமெனில், இராமாயணத்தை மொழிபெயர்க்கும் வேளையில் கம்பர் இந்தக் கலாசார பெயர்ப்புக்கு முக்கியத்துவம் காட்டியிருப்பதைச் சொல்லலாம். வால்மீகியின் இராமாயணத்தை அப்படியே வெறும் மொழி பெயர்ப்பை மட்டுமே செய்யாமல் கலாசார பெயர்ப்பு செய்த திரத்தைப் பாராட்ட வார்த்தைகளே இல்லை. ஆக, மொழிபெயர்ப்புக்கும், கலாசார பெயர்ப்புக்கும் வித்திட்டவர் கம்பரே என்று கூறினால் அது மிகையில்லை.

மலேசியச் சூழலைப் பொருத்தவரையில், மலாய் மொழியின் மூலம் என வருணிக்கப்படும் அரபு மொழி உண்மையில் ஆய்ந்து நோக்குகையில் வெவ்வேறு மொழிக்குடும்பத்தைச் சார்ந்தன எனும் தகவல் சுவாரசியமானதாக இருக்கும். இவ்விரு மொழிகளுக்கிடையிலான இலக்கணங்களும் பெரிதும் வேறுபட்டவை. இருப்பினும், இவ்விரு மொழிகளுக்கிடையிலான மொழிபெயர்ப்புகளும் மற்றும் அது குறித்த ஆய்வுகளும் மிக அதிகமாக எழுந்துள்ளன. இதற்குக் காரணம் அவை இரண்டிற்கும் உள்ள கலாசார ஒற்றுமையே ஆகும். இவ்விரு கலாசாரங்களும் ஒருமித்து இருக்குங்காலை, இவ்விரு மொழிகளுக்கிடையிலான மொழிபெயர்ப்பு இலகுவாகவும் இயல்பாகவும் அமைந்து விடுகின்றன. ஆதலால் மொழிபெயர்ப்புகளின் எண்ணிக்கையும் அதிகமாக இருக்கின்றன.

ஒரு மொழியின் குணாதிசயம் அம்மொழி பேசுபவர்களின் குணாதிசயங்களை ஒத்திருக்கும் என்பதனை ஆய்வுகள் காட்டுகின்றன. (Booth, A.D 1958 : 78). இரு மொழிகளைச் சரளமாகப் பேசும் ஒருவரது நடவடிக்கைகள் அம்மொழி பேசப்படும்போது மாற்றம் காண்பது அறியப்பட்டுள்ளது. அமெரிக்க நாட்டைச் சேர்ந்த சீனப் பெண்மணி ஒருவர், ஆங்கில மொழி பேசும்போது மேலாதிக்கப் போக்கையும் சுயநலப் போக்கையும் காட்டும் வகையில் பேசுகின்றார். அதே பெண்மணி சீன மொழியைப் பேசும்போது. பணிவாகவும் விட்டுக்கொடுக்கும் விதமாகவும் பேசுவதும் அறியப்பட்டிருக்கின்றது. இதே போன்று அமெரிக்காவைச் சேர்ந்த மெக்ஸிகோ வழிதோன்றலான ஆடவர் ஒருவர், ஸ்பேனிஷ் மொழி பேசும்போது தன் நண்பர்களுடன் மிகுந்த நட்புறவோடு பேசுவதாகவும், அவரே ஆங்கிலம் பேசும்போது தன்னைப் பற்றி மட்டுமே சிந்தித்துப் பேசுவதாகவும் ஆய்வுகள் காட்டுகின்றன. (Booth, A.D 1958 : 79).

மொழி பெயர்ப்பு

சிறந்த மொழிபெயர்ப்பாகக் கருதப்படுவது எதுவெனில், மூல மொழியிலிருந்து இலக்கு மொழிக்கு மொழி பெயர்க்கும்போது, இலக்கு மொழியைப் பேசுபவர்கள் எவ்வாறான மொழிப் பயன்பாட்டைக் கொண்டுள்ளனரோ அவ்வாறே மொழிபெயர்க்கப்படுவதே சிறப்பு. மொழிபெயர்க்கப்பட்ட அப்பகுதி, மொழிபெயர்க்கப்பட்ட ஒன்று என்ற எண்ணம் ஏற்படாமல் அம்மொழியிலேயே இயற்றப்பட்ட படைப்பு என்ற எண்ணத்தைக் கொடுப்பதாக அமைதல் சிறப்பு. உதாரணத்துக்கு, மலாய் மொழியில்,

Siapakah nama awak?

என்று கேட்கும் பாணியும் உண்டு. உண்மையில் கேட்க வருவது,

Apakah nama awak?

இந்தச் சூழலில் நேரடியாக மொழிபெயர்ப்பு செய்தால், உங்கள் பெயர் யார்? என கேட்க வேண்டி இருக்கும். தமிழ் மொழியைப் பொருத்தவரை உங்கள் பெயர் என்ன? என்று கேட்பதே மரபு. இவ்விடத்தில், தமிழ் மொழி இலக்கு மொழியாக இருக்கும்பட்சத்தில் தமிழின் இலக்கணங்களை நன்கு அறிந்திருந்தால் மட்டுமே சிறப்பான மொழிபெயர்ப்பைக் கொடுக்க இயலும்.

கலாசார பெயர்ப்பு

கலாசார பெயர்ப்பு குறித்து இவ்வளவு நேரம் ஆய்ந்து வந்தாலும், எப்போதுமே கலாசார பெயர்ப்பு செய்தல் சரியான கேள்வி எழும்பட்சத்தில் இல்லை என்றே பதிலுரைக்க வேண்டியிருக்கிறது.

பொதுவாக, சிறுகதை, நாவல், கவிதைகள், கட்டுரைகள் போன்ற இலக்கியப் படைப்புகளை மொழிபெயர்க்கும்போது கலாசார பெயர்ப்பு மிகவும் வரவேற்கப்படுகின்றது. இவ்வாறு கலாசார பெயர்ப்பைச் செய்யும்போது, ஒரு கட்டுரையின் முன்னுரையிலோ அல்லது புத்தகத்தின் முகப்பிலோ இன்ன மொழியிலிருந்து தழுவப்பட்டது என்ற வாசகத்தைப் பதிப்பது இலக்கிய ஒழுக்கமாகக் கருதப்படுகின்றது. ஏனெனில், அவற்றை வாசிக்கும் வாசகர்களுக்குச் சம்பந்தப்பட்ட கலாசாரத்தைப் பற்றி குழப்பம் ஏற்படாமல் இருக்க இது மிக அவசியம்.

மறுபுறத்தில், ஒரு கலாசாரத்தை விளக்கும் விளக்கக் கட்டுரைகள், புராதன சிறப்புமிக்க இலக்கிய வகைகள், சில நாடுகள் குறித்த கையேடுகள் ஆகியவற்றை மொழிபெயர்க்க வேண்டிய சூழலில் கண்டிப்பாக கலாசார பெயர்ப்பு செய்யப்படக்கூடாது என்பது மொழிபெயர்ப்பின் விதியாகும். இதனை மீறுதல் ஏற்றுக்கொள்ள முடியாததாகும். ஏனெனில், இத்தகைய படைப்புகள் அவர்களது (குறிப்பிட்ட நாடு அல்லது மக்கள்) தனிப்பட்ட பெருமையையும் சிறப்புகளையும் உணர்த்தும் வகையில் வரையப்பட்டவையாகும். உதாரணத்திற்கு:

மூல மொழி	: He only had plain bread and black coffee for breakfast
கலாசார பெயர்ப்பு	: அவன் பசியாறுவதற்குத் தோசை மற்றும் தேநீரை உட்கொண்டான்.
மொழிபெயர்ப்பு	: அவன் பசியாறுவதற்கு வெறும் ரொட்டி மற்றும் கருங்காப்பியை உட்கொண்டான்.
மூல மொழி	: He had no time to put on his <i>pyjama</i> before going to bed.
கலாசார பெயர்ப்பு	: படுக்கைக்குச் செல்லும் முன் அவனுக்குக் கைலியை மாற்றக்கூட நேரமில்லை
மொழிபெயர்ப்பு	: படுக்கைக்குச் செல்லும் முன் அவனுக்கு பஜாமாளை மாற்றக்கூட நேரமில்லை.

இவ்விரு உதாரணங்களிலும் கலாசார பெயர்ப்பு மற்றும் மொழிபெயர்ப்பு ஆகியவற்றுக்கிடையிலான வேற்றுமைகளை நாம் காண முடிகின்றது.

இவ்வாறு பல சிக்கல்கள் நிறைந்த இந்த கலாசார பெயர்ப்பைச் செய்வதற்கு மொழி அறிஞர்கள் சில அனுகுமுறைகளை விளக்கியிருக்கின்றனர். அவற்றைக் கடைபிடிப்பதன் வாயிலாக இப்பிரச்சனைகளுக்கு ஓரளவு விடிவு கிடைக்கின்றது என்றாலும், எந்த நேரத்தில் எவ்வாறான முடிவெடுப்பது என்ற பொறுப்பு மொழிபெயர்ப்பாளர்களையே சார்ந்தது. உதாரணத்தின் வழியே இந்த அனுகுமுறைகளை விளக்குதல் எளிமையாக இருக்கும், எடுத்துக்காட்டாக, ஆங்கிலத்தில் பயன்படுத்தப்படும் *snow white* என்ற வார்த்தையை

எடுத்துக்கொண்டோமானால், ஆங்கிலேயர்களின் வாழ்க்கைச் சூழலில் பனி இருப்பதனால் அவர்கள் இத்தகைய உவமையைப் பயன்படுத்துகின்றனர். அந்த வார்த்தையை மொழிபெயர்க்கும் சமயத்தில் கீழ்வரும் வகையாகப் பிரயோகிக்கலாம்:

கலாசார குறிப்புகளை அப்படியே பயன்படுத்துதல்

இந்த அணுகுமுறையின் வாயிலாக *snow white* என்ற வார்த்தையை அப்படியே பயன்படுத்தலாம். நமது வாழ்க்கைச் சூழலில் பனி கிடையாது என்றாலும், பனி எனப்படுவது எப்படி இருக்கும் என நம்மில் மிகப் பலருக்குத் தெரியும் மற்றும் அவற்றை இணையம், தொலைகாட்சி, பத்திரிகைகள் ஆகியவற்றின் வாயிலாகப் பார்த்து அறிந்திருக்கிறோம் என்பதனால் அதனைப் பிரயோகித்தால் வாசகர்களுக்கு குழப்பங்கள் ஏதும் வாராது என்ற கருத்தின் அடிப்படையில் இவ்வணுகுமுறை பின்பற்றப்படுகின்றது.

கலாசார குறிப்புகளை மறைத்து விடுதல்

அதாவது இவ்வணுகுமுறையின் வாயிலாக *snow white* என்ற வார்த்தைப் பிரயோகத்தை அறவே தவிர்த்துவிட்டு, அச்சொற்றொடருக்கு ஏற்ற விளக்கத்தை மொழிபெயர்ப்பில் தருவது என முன்மொழியப்பட்டுள்ளது. தூய்மையான வெள்ளை, பரிசுத்தமான வெள்ளை போன்ற வேறு ஏதும் சொற்களைக் கொண்டு இதன் பொருளை ஈடுகட்டச் செய்வதாக இவ்வணுகுமுறை குறிப்பிடுகின்றது.

மூல மொழி கலாசார குறிப்புகளுக்கு ஏற்ற இலக்கு மொழி கலாசார குறிப்புகளைத் தருதல்

இவ்வணுகுமுறையின் வாயிலாக, மூல மொழியில் இருக்கும் கலாசார குறிப்புகளுக்குப் பதிலாக இலக்கு மொழியில் உள்ள கலாசார குறிப்புகளைத் தருவது என பொருள்படுகின்றது. பெரும்பாலும் இந்த அணுகுமுறை பழமொழிகளை மொழிபெயர்க்க உறுதுணையாக இருக்கின்றது. உதாரணத்திற்கு ஆங்கில பழமொழிகளுக்கு அதே பொருளுடைய தமிழ் பழமொழிகளைக் கொடுக்க இயலும்,

Known is a drop. Unknown is an ocean
கற்றது கைமண் அளவு, கல்லாதது உலகளவு
Like Father like son
தாயைப் போல பிள்ளை, நூலப் போல சேலை
A quarter for the berry, three quarters to deliver it
சுண்டைக்காய் கால் பணம், சுமைக்கூலி முக்கால் பணம்

ஆக, மொழிபெயர்ப்பாளர்கள் வெறும் இரு மொழிகளைத் தெரிந்து வைத்துக்கொண்டு வெறுமனே மொழிபெயர்த்து விடுவது கிடையாது. மாறாக, மூல மொழி மற்றும் இலக்கு மொழி ஆகிய இரு மொழிகளும், அதன் இலக்கணங்களும், கூடவே அதன் கலாசாரங்களும் தெரிந்திருந்தால்தான் சிறப்பான மொழிபெயர்ப்பைக் கொடுக்க முடியும். இதுவரை அப்படி செயல்படாத மொழிபெயர்ப்பாளர்களும் தங்களின் பொறுப்புணர்ந்து இனி செயல்பட வேண்டும். இந்தக் கலாசார பெயர்ப்பில் என்னதான் சில அணுகுமுறைகள் வழங்கப்பட்டிருந்தாலும் அவற்றைப் பொருத்தமான இடத்திலும் சூழலிலும் பயன்படுத்துவது மொழிபெயர்ப்பாளர்களின் கையில்தான் உள்ளது; அவற்றை அறிந்து புரிந்து செயல்படுவதின் வாயிலாகத்தான் அவர்கள் மிகச் சிறந்த மொழிபெயர்ப்பாளர்கள் என்ற சிறப்பும் புகழும் அடைய இயலும்.

மேற்கோள் நூல்கள்:

1. Booth, A.D.et. al., 1958 Aspects of Translation. Secker and Warburg
2. Nida, E.A. 1964. Toward A Science of Translation. Leiden: E. J. Brill

E-Governance Initiatives in Tamil Nadu

E.Iniya Nehru

National Informatics Centre, Chennai nehru@nic.in

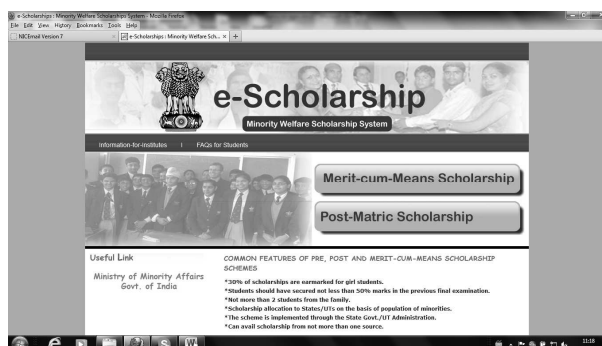
Introduction

NIC, Tamilnadu State Centre has implemented a large number of e-governance projects across many departments of the Tamilnadu State Government. Some of the successful projects rolled out in the state are e-District Scholarship system, e-District Social welfare, e-District Revenue, Online VAT registration for Commercial Taxes Department, ePension for Directorate of Treasuries and Accounts, Dr. Muthulaksmi Reddy Maternity Benefit Scheme for Directorate of Public Health, e-Services for Transport, Scheme monitoring System for Rural Development Department, CCTNS project for the Tamilnadu Police, Online Allotment for PDS, Pre-Examination Monitoring System for Anna University etc.

National Knowledge Network (NKN) connectivity is established at more than 1000 locations across the country covering all IITs, NITs and major Research and Educational Institutes. The NKN will enable scientists, researchers and students from different backgrounds and diverse geographies to work closely for advancing human development in critical and emerging areas.

All the Districts in Tamil Nadu have minimum of 34 Mbps connectivity. NIC Data Centre, Chennai has 55 servers with a total capacity of 21 TB SAN space with an online tape library. Video Conferencing facility has been provided to the State headquarters and to all the Districts and is being extensively used by the State & Central Government Departments in the State.

Minority Welfare Scholarship System (<http://www.momascholarship.gov.in/>)



The Objective of the Merit cum Means Scholarship Scheme of Ministry of Minority Affairs is to provide financial assistance to the poor and meritorious students belonging to the Minority communities for studying Professional courses. A web based system has been designed and developed as part of the eScholarship project. The system handles the complete processing of applications till disbursement of scholarship.

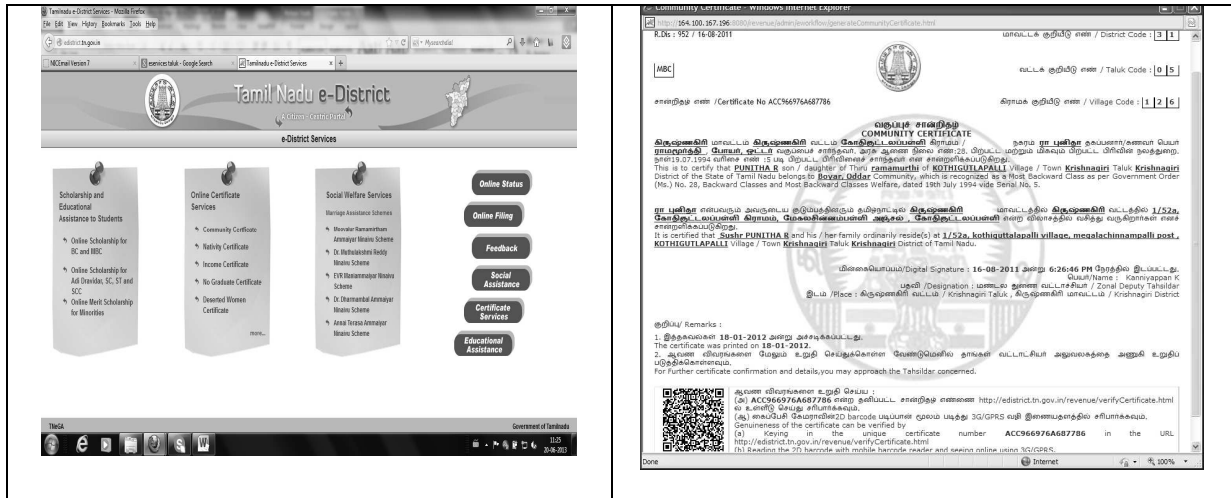


e-Services for Commercial Taxes (<http://www.tnvat.gov.in/>)

The basic objective is to facilitate the Dealers to file online VAT monthly Return along with facility for e-Payment and also track the status of applications submitted online so as to avoid visit to Commercial Tax Offices. The following services are also available to the Dealers:

- Online filing of e-Request for saleable forms
- Online submission of Form-W refund Claims
- Online submission of Goods movement through Checkpost for quick clearance at Checkpost (Fast Track Clearance System)
- Online submission of New registration application
- Online issue of C and F forms
- e-Payment facility with five nationalized banks
- Filing online and issue of Transit pass

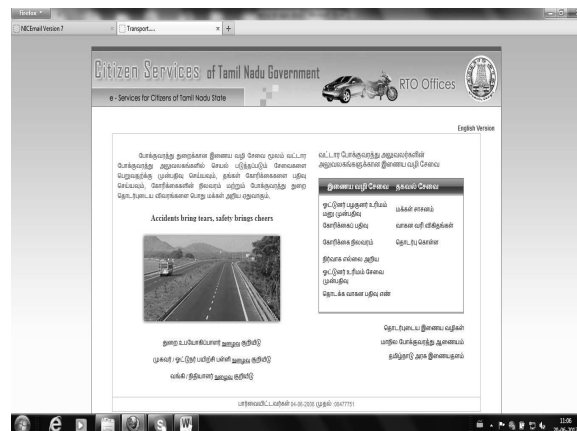
eDistrict – Revenue (<http://edistrict.tn.gov.in/>)



The objective of the eDistrict project is to computerize the workflow system and internal process of the administration of the Districts with the help of Information and Communication Technologies (ICT), thereby enhancing the delivery of the following services to the citizens:

- To bring in more transparency and accountability in the system of processing.
- To facilitate citizens to apply for various certificates through CSCs
- Built in electronic workflow to the system for processing the application
- To create an efficient delivery mechanism of services from the District administration to the doorstep of citizens

The services that are covered are No Graduate Certificate, Deserted Women Certificate, Income Certificate, Community Certificate, Nativity Certificate etc.



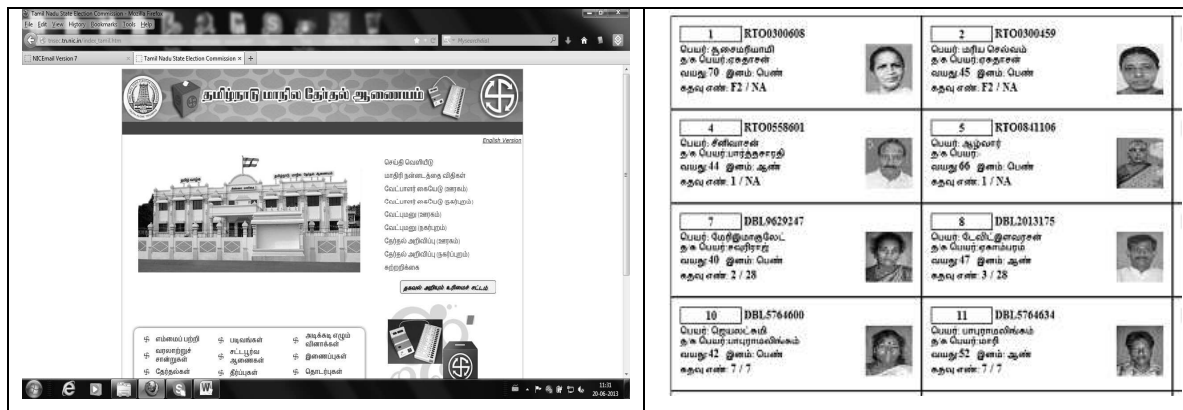
e-Services for Transport (transport.tn.nic.in/transport/)

The following services are enabled through e-Services for Transport

- Online filing of New Vehicle Registration Applications

- **e-Payment facility enabled for dealer to pay fees and taxes related to New Vehicle applications with the banks.**
- **Online filing of Learner's License**
- An average of 1.57 Lakh New Vehicle Registration applications is being filed online every month.. More than 80, 000 learner's license applications are filed online every month.

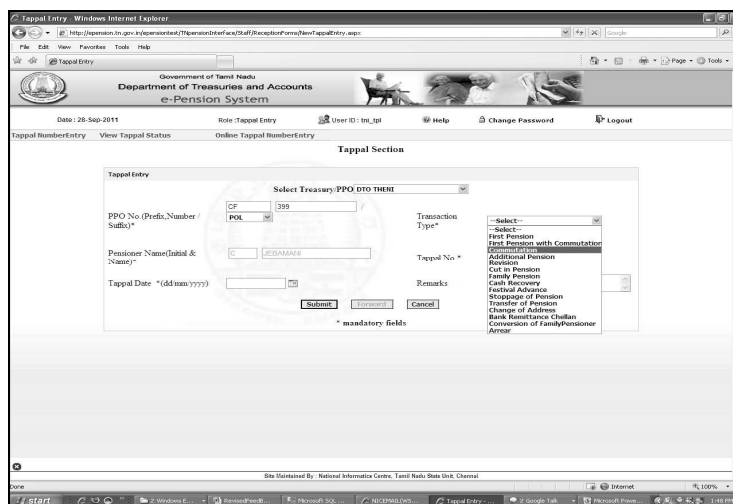
Tamil Nadu Local Body Elections (<http://tnsec.tn.nic.in/>)



1	2
1 RTO0300608 பெயர்: துரைமணி தக.பெயர்: துரைமணி வயது: 70 இனம்: பெண் தகவல் எண்: F2 / NA	2 RTO0300459 பெயர்: மீனாட்சி செல்வம் தக.பெயர்: துரைமணி வயது: 45 இனம்: பெண் தகவல் எண்: F2 / NA
4 RTO0558601 பெயர்: சீனிவாசன் தக.பெயர்: பாத்திரசாமி வயது: 44 இனம்: ஆண் தகவல் எண்: 1 / NA	5 RTO0841106 பெயர்: அழகன் தக.பெயர்: அழகன் வயது: 66 இனம்: பெண் தகவல் எண்: 1 / NA
7 DBL0629247 பெயர்: சீனிவாசன் தக.பெயர்: சீனிவாசன் வயது: 49 இனம்: பெண் தகவல் எண்: 2 / 28	8 DBL2013175 பெயர்: சீனிவாசன் தக.பெயர்: சீனிவாசன் வயது: 47 இனம்: ஆண் தகவல் எண்: 3 / 28
10 DBL5764600 பெயர்: செல்வம் தக.பெயர்: செல்வம் வயது: 42 இனம்: பெண் தகவல் எண்: 7 / 7	11 DBL5764634 பெயர்: பாத்திரசாமி தக.பெயர்: பாத்திரசாமி வயது: 52 இனம்: ஆண் தகவல் எண்: 7 / 7

Voter's database is created from Local Body Election from Electoral database and the Electoral Roll printing was done. Result Dissemination is also done through web.

- **Printing Draft Electoral Rolls**
- **Printing Photo Electoral Rolls**
- **Printing of Booth Slips**
- **Voters Reservation entry for various Local Body Posts**
- **Nomination entry of Candidates**
- **Dissemination of results**



ePension System (epension.tn.gov.in/)

The system handles all pension related activities of the Department of Pensions by providing better quality service to TN Govt. Pensioners and terminal benefits to pensioners.

Pregnancy and Infant Cohort Monitoring Evaluation System

The screenshot shows a web browser window with the URL <http://intra.tn.nic.in/dph/jcm/ndPage.asp>. The page title is "PREGNANCY COHORT MONITORING SYSTEM AN Mother General Information". The form is for the month and year "November-2009". It contains several sections: "PHC" (PHC1, HSC, PHC2), "Name of VEH" (malathi, VEH Phone No. 9445137651), "Modifying AN Mother" (ANC ID No. 54, Para 1), "Name of AN mother" (Ratha, LMP 12/04/2008), "Name of Husband" (Kumaran, EDD 17/01/2009), "Phone" (9941141438, Date of AN Registration 10/10/2008), "Address" (Lakshimpuram, Height (in cm) 156), "Age of mother" (21, Blood Group A+), "Community" (SC, VDRL Non-reactive), "Education Status of AN Mother" (dropdown), "HIV Status of AN Mother" (Negative), "Education Status of Husband" (dropdown), "HIV Status of Husband" (dropdown), "Gravida" (1), and "Remarks" (empty). There are "Save" and "Back" buttons at the bottom right.

The DPH&PM annually registers about 12.46 lakh Ante Natal Cases and renders various services to pregnant mothers. An online reporting system has been developed to monitor the health status of Pregnant Women registered with any PHC and Infants in the rural areas of Tamil Nadu. The system tracks the health of the women during Ante-Natal, Delivery and Post Natal care. Infants are linked to the Mother by Registration ID.

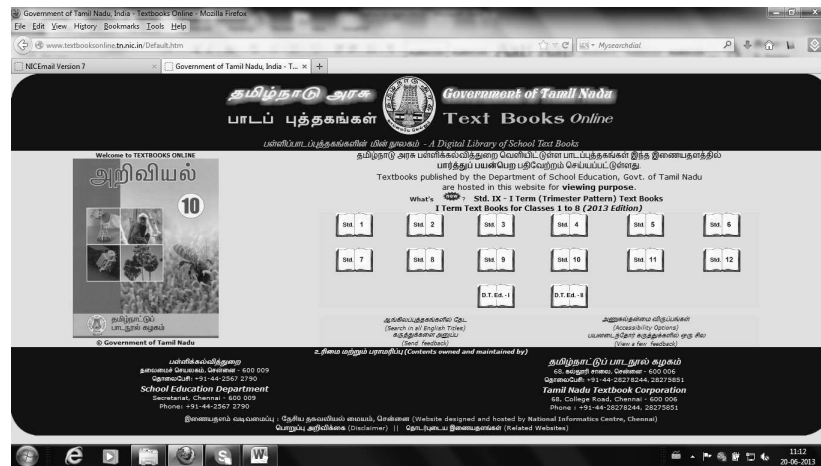
Government e-Procurement System (<http://tntenders.gov.in>)

The screenshot shows the "Government eTenders - Mozilla Firefox" web page. The URL is <https://tntenders.gov.in/nicges/app>. The page title is "Tenders TamilNadu The TamilNadu Government Tenders Information System". The date is "19-Jun-2013". The page has a navigation bar with links: "Search", "Latest Active Tenders", "Tenders by Closing Date", "Corrigendum", "Results of Tenders", "Home", "Contact Us", "Shop". The main content area is divided into two columns. The left column has a sidebar with links: "Reports", "Tenders by Location", "Tenders by Organisation", "Tenders by Classification", "Tenders in Archive", "Tender Notices/Displays", "Tender Related Links", "Tenders Status", "Cancelled Tenders", "Downloads", "Announcements". The right column has a "Welcome to eTendering System" message, a "Latest Tenders" table, a "Latest Corrigendums" table, a "User Login" section, and a "Tender Search" section. The "Latest Tenders" table has columns: "Tender Title", "Reference No", "Closing Date", "Bid Opening Date". The "Latest Corrigendums" table has columns: "Corrigendum Title", "Reference No", "Closing Date", "Bid Opening Date". The "User Login" section has fields for "Login ID" and "Password", and a "Login" button. The "Tender Search" section has a "Search with Tender Reference no/tender id/Title/Description" field and a "Tender Search" button. The page footer shows the date "21:15 19-06-2013".

The eSubmission Tender System of Tamil Nadu enables the Tenderers to download the Tender Schedule free of cost and then submit the bids online through this portal.

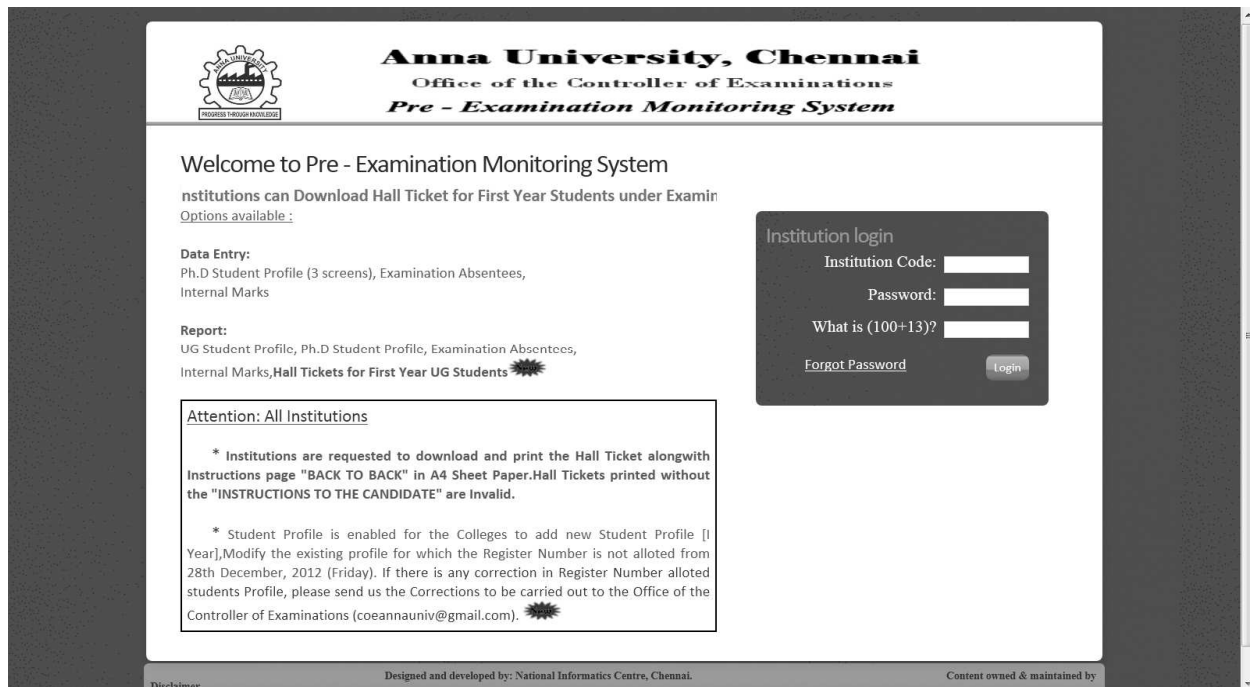
Tamil Nadu Textbooks Online (www.textbooksonline.tn.nic.in)

The entire textbooks for the school students of class 1 to Class 12 published by the School Education Department, Government of Tamilnadu are available in pdf format for the General public.



EXAMINATION MONITORING SYSTEM FOR ANNA UNIVERSITY, CHENNAI

The primary objective of this web application is to enable Anna University Examination Section to monitor all the 'Pre Semester Examination Processes' effectively.

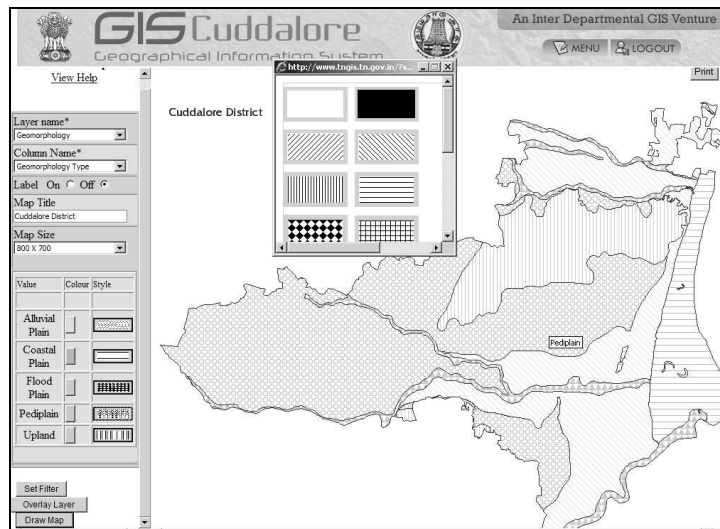


Web Enabled District GIS for Cuddalore & Nagapattinam Districts (<http://tngis.tn.gov.in/>)

District GIS has been initiated for two pilot districts viz. Cuddalore and Nagapattinam under the guidance of Tamil Nadu State Planning Commission. The Spatial and Non-spatial data were collated from varied Government departments. More than 80 layers for Cuddalore district and 60 layers for Nagapattinam district have been prepared.

Salient features of the web based framework designed and developed include the following:

- Any layer may be dynamically linked to any kind of attribute dataset for querying and thematic mapping;
- Unique Symbol mapping using Non-numeric based data, such as, Soil, Geomorphology, etc.
- Spatial Search on Administrative Units
- Dynamic spatial queries including Buffer generation, distance based queries, villages without schools / health facilities, etc.



Dynamic Generation of Thematic / Unique Symbol map

Chennai Utility Mapping Project

Union Planning Commission initiated “Computer Aided Digital Mapping Project” in Chennai. Digital base map of 192.5 sq. km. area covering more than 80 layers, using aerial photography under stereoscopic vision has been created with the same origin and accuracy covering Building structures, Financial Service, Railways, Water Bodies, Recreational, Religious, Transportation, etc. The following departments have been provided with sufficient infrastructure such as hardware, software, LAN with training to utilize the base layers generated under this project:

The following organizations are stakeholders in the project.

1. Corporation of Chennai
2. Tamil Nadu Electricity Board (TNEB)
3. Chennai Metropolitan Development Authority (CMDA)
4. Chennai Metro Water Supply and Sewerage Board (CMWSSB)
5. Chennai City Police
6. Bharat Sanchar Nigam Limited (Chennai Telephones)



Distances from Base Station Centre to Base Tower Stations

Location Based Services using SOA (GPS based)

An online mobile based system has been implemented for Rural development and Panchayat Raj Department, Government of Tamil Nadu, India. This system is used by the department for monitoring the execution of various road projects executed in rural area of Tamil Nadu.

Viewing the Photos of a Road Work on the Google Map

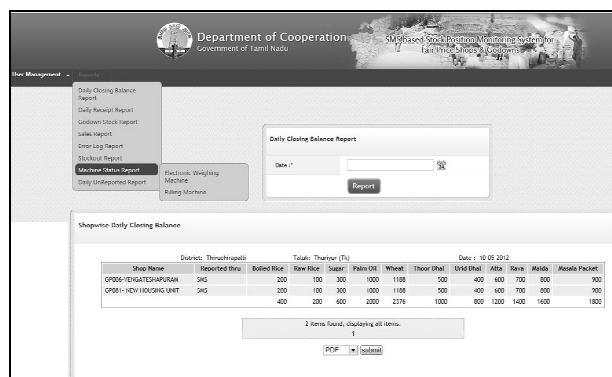
View in the Mobile Device



eGovernance Portal for DoTE (<http://www.tndte.com/>)



The Citizen gets the basic details of all the **450+ Polytechnic Colleges**. The Diploma Students studying in these Polytechnic Colleges can also get to know the Attendance details for the current Academic Year, SMS based monitoring system for Fair Price Shops of the Cooperative Department (<http://www.tncsc.tn.gov.in/html/pds.htm>)



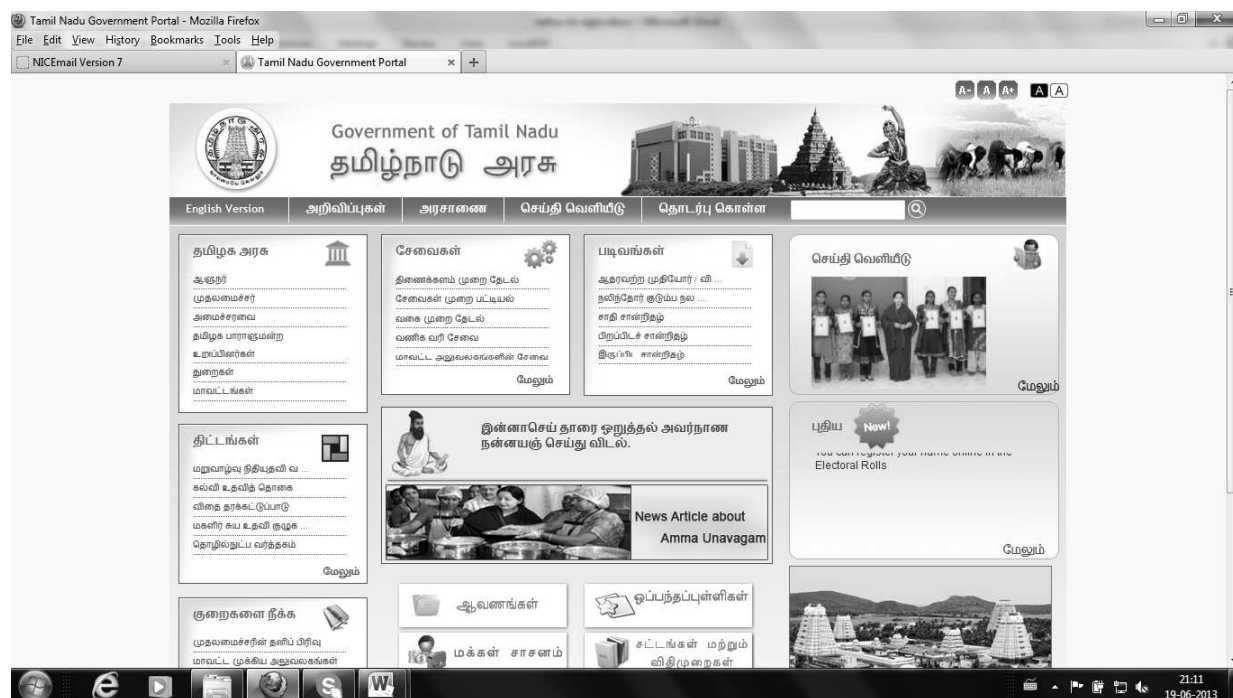
An SMS based system has been developed for collecting stock details from FPSs & Godowns on a daily basis, for planning timely movement of stocks of PDS commodities. The online system will make the process of data collection, compilation & collation more efficient.

Intranet Based Monitoring & Management System for Noon Meal Scheme

[illegible]

An Intranet based Monitoring and Management system is developed to assist in effective implementation, tracking and planning of activities related to the Noon Meal Programme.

TN Government State Portal & CMS (www.tn.gov.in)



‘Any Time, Any Where’ access to the information and services of the State Government at all levels from State Government to District Administration or Panchayat level for the Citizens and Business. **Consolidation of all the information** presently available at existing departmental websites & portals are done. Development of State Portal is based on **State Portal Framework** and National Portal content structure.

National Knowledge Network

- **NKN provides a common platform to the scientists, researchers, doctors, scholars and students to work together for advancing human development.**
- **NKN shall be a knowledge base for revolutionizing the education system of country**
- **The architecture of NKN has been designed for reliability, availability & scalability**
- **The network consists of an ultra-high speed core, starting with multiple 2.5/10 G and progressively moving towards 40/100 Gigabits per Second (Gbps)**
- **The core is complimented with a distribution layer covering all districts at appropriate speeds**
- **The participating institutions at the edge would seamlessly connect to NKN at Gigabit per second speed.**

Conclusion

During the last three decades, NIC has implemented many "network centric" application software for Programme implementation in various ministries and departments, using state-of-the-technology software tools. During 1980s and early part of 1990s, the policy thrust was on creating "Management Information System (MIS)" and "Decision Support System (DSS)" for development, planning and responsive administration in governments which led to the genesis of present day "e-Governance" / "e-Government". "Bridging the Digital Divide", "Social and Financial Inclusion through ICT" and "Reaching-the-Unreached" concepts were tried and made operational in the late nineties.

NIC has vast expertise and experience in the design, development and operationalisation of various e-Government projects in the areas of Public Administration and Governance like Agriculture & Food, Animal Husbandry, Fisheries, Forestry & Environment, Industry, Health, Education, Budget and Treasury, Fiscal Resources, Transport, Water Resources, Court Management, Rural Development, Land Records and Property registration, Culture & Tourism, Import & Exports facilitation, Social Welfare Services, Micro-level Planning, etc. With increasing awareness leading to demand and availability of ICT infrastructure with better capacities and programme framework, the governance space in the country witnessed a new round of projects and products, covering the entire spectrum of e-Governance including G2C, G2B, G2G, with emphasis on service delivery.

மலேசியத் தமிழ் இளைஞர்களிடையே மொழித் தாவலும் புதிய தொழில்நுட்பங்களின்வழி மின்னணுக் கலாசாரமும்.

செல்வஜோதி இராமலிங்கம் -மலேசியா (selvajothi29@gmail.com)

முன்னுரை

தமிழ்மொழியின் சிறப்பு அதன் தொன்மையிலா? தொடர்ச்சியிலா? என்ற வினா தமிழ்க்கூறு நல்லுலகிற்கு ஒரு சவாலாக இருந்து வருகிறது. தமிழ்மொழியில் பல்வேறு சிறப்புகள் புதைந்து கிடப்பினும் இன்றைய காலகட்டத்தின் தொழில்நுட்ப வளர்ச்சியுடன் இணைந்து செயல்படாவிடில் அதன் தொடர்ச்சிக்குப் பெரும் பாதகமாக அமைந்துவிடும். தற்போதைய காலகட்டத்தில் மொழி ஆர்வலர்களும் மொழியியல் ஆய்வாளர்களும் பரவலாகப் புதிய தொழில்நுட்பங்களைப் பயன்படுத்திவருகின்றனர். அதாவது புதிய மொழியியல் வடிவங்களும் சமூக செயல்பாடுகளும் ஒருங்கிணைப்பதன் வாயிலாக சிறுபான்மை மக்களைச் சார்ந்த, மொழியின் பயன்பாட்டையும் பயனர்களையும் அதிகரிக்க வழிக்கோலும் (King, 2001).

இந்த அடிப்படையில் மலேசியத் தமிழ் இளைஞர்களிடையே மின்னணு சார்ந்த நடைமுறைகளில் அதாவது இணைய பயன்பாட்டில் மொழித் தாவலும் மொழி நிலைநிறுத்தலும் எவ்வாறு நடைபெறுகிறது என்பதை ஆராய்வதே இவ்வாய்வின் நோக்கமாகும். ஒரு மொழியைத் தொடர்ந்து பயன்படுத்துவதையும் அதனைக் கைவிடுவதையும் இந்தப் புதிய தொழில்நுட்பங்களின் பயன்பாடு பெரும் தாக்கத்தை ஏற்படுத்துகின்றது. புதிய தொழில்நுட்பங்கள் ஒரு தொடர்பு சாதனமாக இருந்து சமூகப் பண்பாட்டுக் கூறுகளை உரைக்கோவையோடு இணைக்கிறது. மொழித் தாவலும் மொழி நிலைநிறுத்தலும் என்ற துறை, இனமொழியியல் (ethnolinguistics) துறையோடு தொடர்புடையது. ஆனால் உலகமே இன்று கணினிமயமாகிக் கொண்டிருக்கும் சூழலில் இலக்கியல் (டிஜிட்டல்) தொழில்நுட்பத்தில் ஒரு மொழியின் பயன்பாடு குறைந்தால் மொழித் தோய்விற்கு இது வித்திடும் என்ற செய்தி இங்குப் பதிவு செய்யப்பட வேண்டிய ஒன்று.

மொழித் தாவலும் மொழி நிலைநிறுத்தலும்

மொழித் தாவலும் மொழி நிலைநிறுத்தலும் என்ற இந்த இரண்டு நிலைப்பாடுகளும் ஒரு நாணயத்தின் இரண்டு பக்கங்களாக சித்தரிக்கப்படுகிறது (Fasold, 1984). ஒரு மொழி எவ்வாறு அழிவை எதிர்நோக்குகிறது; பல சவால்களுக்கிடையே தொடர்ந்து பயன்பாட்டில் உள்ளது என்பது பற்றிய ஆய்வுதான் மொழித் தாவலும் மொழி நிலைநிறுத்தலும் என்ற துறையானது. குறிப்பாகச் சிறுபான்மை மக்கள் வாழும் நாடுகளில் சில மொழிகள் தொடர்ந்து நிலைநிறுத்தப்பட்டு வருகிறது மாறாக சில நாடுகளில் அழிவை எதிர்நோக்கி சென்று கொண்டிருக்கிறது. ஒரு நாட்டில் தங்கள் தாய்மொழியை நிலைநிறுத்தும் சூழலில் அதே பின்னணியைக் கொண்ட இன்னொரு நாட்டில் தாய்மொழி புறக்கணிக்கப்பட்டு வேறு மொழிக்குத் தாவல் ஏற்படுவது அந்த மொழியைப் பயன்படுத்தும் பயனர்களின் முனைப்பும் தங்கள் தாய்மொழியின் மீது அவர்களுக்கு இருக்கும் பற்றும் மிக அடிப்படையாகக் கருதப்படுகிறது (Baker, 1992). அதாவது எந்தவொரு வாய்ப்பும் வசதியும் இல்லாத சூழலிலும் சில நாடுகளில் தங்களின் பண்பாட்டைக் கடைபிடிக்கவும்; தங்களின் அடையாளத்தை நிலைநாட்டவும் தங்கள் தாய்மொழியைத் தொடர்ந்து பல தலைமுறைகளாகக் கட்டி காத்து வருகின்றனர்.

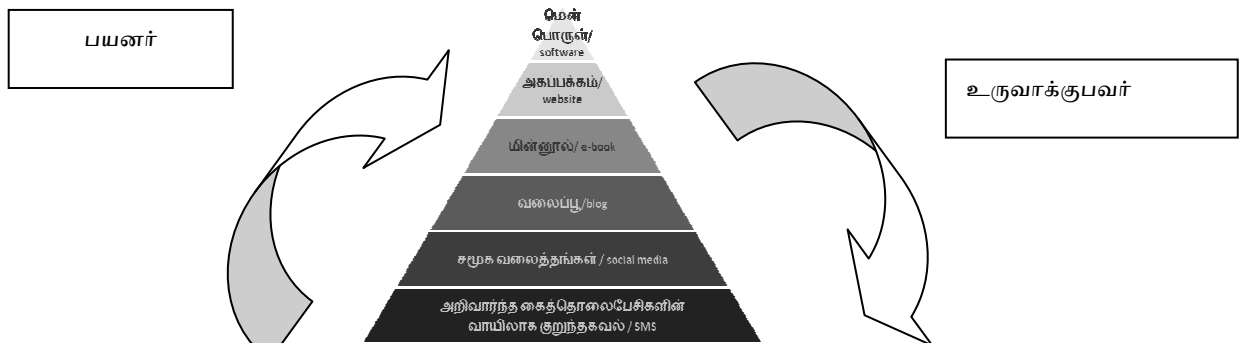
ஒரு மொழியைப் பேசும் சிறுபான்மை மக்கள் அவர்கள் நாட்டில் பயன்படுத்தப்படும் மேலாதிக்க மொழியோடு (dominant language) தொடர்பில் இருக்கும் போது ஒரு பகுதியோ முழுமையாகவோ அந்த மேலாதிக்க மொழியைப் பயன்படுத்தும் சூழலில் மொழித்தாவல் ஏற்படுகின்றது. குறிப்பாக கணினியுலகைப் பொருத்த மட்டில் மேலாதிக்க மொழியாக விளங்குவது ஆங்கிலமே. நாளடைவில் மொழி பேசுபவர்களிடையே மொழித் திறன் மறைந்து, அழிவை நோக்கிச் செல்லும் வாய்ப்பு உள்ளது (Fishman, 1985). இந்த மொழித்தாவல் என்பது ஒரே இரவில் ஏற்படக்கூடிய மாற்றம் அல்ல. பொதுவாக மூன்று தலைமுறைகளில் இந்த மாற்றம் ஏற்படக் கூடிய சாத்தியங்கள் இருப்பதாக ஆய்வாளர்கள் குறிப்பிடுகின்றனர். ஆனால் கணினியில் அல்லது இணையத்தில் ஏற்படுகின்ற மொழித்தாவல் ஒரே தலைமுறையில் ஏற்படுவது அபாயகரமானதாக சித்தரிக்கப்படுகிறது. இது ஒரு குறிப்பிட்டத் தனிமனிதரைச் சார்ந்ததல்ல. மாறாக ஒரு ஒட்டு மொத்த சமுதாயமே தங்களது தாய்மொழியைப் புறக்கணித்து வேறொரு மொழியைத் தங்களது குடும்ப மொழியாகவோ பிற பயன்பாட்டுச் சூழல்களுக்குப் பயன்படுத்தும் மொழியாகவோ ஏற்றுக்கொள்ளும் சூழலில் இது அமைகிறது. பொதுவாகவே இது பல்லின மக்கள் வாழும் நாடுகளில் இவ்வாறான சூழலைக் காணலாம் (Fishman, 1991). ஒரு மொழியைச் சார்ந்த சமுதாயம் பிறரிடையே தொடர்பு கொள்வதற்கு மேலாதிக்க மொழியைப் பயன்படுத்தும் சூழலில் இந்த மொழித்தாவல் ஏற்படுகின்றது. இது

அந்த நாட்டின் தேசிய மொழியாகவோ அல்லது பன்னாட்டுச் சுழலுக்கேற்ற உலக மொழியான ஆங்கிலமாகவோ இருக்கலாம்.

மொழி நிலைநிறுத்தல் என்று பார்க்கும் போது ஒரு மொழியைப் பேசும் சமூகம், எந்தவொருச் சூழலிலும் எத்தனை சவால்களுக்கிடையிலும் தங்களது தாய்மொழியைத் தக்கவைத்துக் கொள்வது ஆகும். மேலாதிக்க மொழிகளின் தாக்கம் அதிகமாக இருப்பினும் தங்களது மொழியைத் தொடர்ந்து நிலைநாட்டுவர். தங்களின் பண்பாட்டை வெளிப்படுத்துவதற்கும் தங்களின் அடையாளத்தைத் தக்கவைத்துக் கொள்வதற்கும் தொடர்ந்து தங்கள் தாய்மொழியைக் குடும்ப மொழியாகப் பயன்படுத்தி வருவர்; அடுத்த தலைமுறைக்குக் கொண்டுச் செல்வதில் அல்லது பரிமாற்றம் செய்வதில் வெற்றியும் அடைவர். இந்தச் சங்கிலித் தொடரில் ஏதேனும் ஒரு பகுதியில் அல்லது ஒரு தலைமுறையில் தடை ஏற்பட்டால் கூட தங்கள் குடும்பங்களில் தாய்மொழி பயன்பாட்டின் தொடர்ச்சிக்கு ஒரு முட்டுக்கட்டையாக இது அமைந்திடும்.

மொழித் தாவலும் நிலைநிறுத்தலும் தொடர்பான ஆய்வுகளில் மின்னியல் ஊடகங்கள் மிக முக்கியமான பங்கினை வகித்து வருகிறது. குறிப்பாக சிறுபாண்மை மக்கள் வாழ்கின்ற நாடுகளில் பொதுவாகவே அவர்களின் தாய்மொழி பயன்பாடு ஒரு குறிப்பிட்ட எல்லையோடு நின்றுவிடும். இன்னும் சொல்லப்போனால் அவர்களின் தாய்மொழியை நிலைநிறுத்துவதற்கு இது மிகப்பெரிய அச்சுறுத்தலாகவே அமைந்து வருகின்றது. தொடர்ந்து சிறுபாண்மை மக்களாக ஒரு நாட்டில் வாழ்கின்ற சூழலில் அவர்களின் தொடர்புநிலை சில குறிப்பிட்ட சூழமைவுகளிலேயே மையமாகக் கொண்டிருக்கின்றது. அதாவது குடும்பம், நண்பர்கள், ஆலயம்/சமயம், வணிகப் பரிவர்த்தன, பொழுதுபோக்கு, உறவினர், அலுவலகம் (Amy Sobrielo, 1985) போன்றவற்றில் எத்தனை சூழமைவுகளில் தாய்மொழி இடம்பெறுகின்றது என்பது ஒரு கேள்விக்குறியாகவே அமைகிறது. காரணம் மலேசியா போன்ற பல்லின மக்கள் வாழும் நாடுகளில் சிறுபாண்மை மக்கள் தங்கள் தாய்மொழியை ஒரு சில சூழமைவுகளில் மட்டுமே பயன்படுத்துவதற்கு வாய்ப்புகள் உள்ளன. மேலும் மக்கள் வீடமைப்புச் சூழலும் மொழியை நிலைநிறுத்துவதற்கு ஒரு சவாலாகவே உள்ளது. ஒரு குறிப்பிட்ட வட்டாரத்தில் ஒரே மொழியைப் பேசுகின்றவர்கள் வாழும் சூழல் ஏற்பட்டால் அங்கு மொழியைத் தற்காக்கும் சூழல் சாதகமாக அமைகிறது (Giles, 1977). இல்லையேல் அது பாதகமாக அமைந்துவிடும். குறிப்பாக நகர்ப்புறங்களில் வீடமைப்புப் பகுதிகளில் பல்லின மக்களுக்கு மத்தியில் வீடுகள் அமைந்திருக்கும் சூழலில் ஒருவர் ஒருவரோடு தொடர்பு கொண்டு உரையாடும் சூழல் இன்னும் அரிதாகவே அமைந்துள்ளது (Roksana Bibi, 2002). இவ்வாறான கட்டுப்பாடுகளையெல்லாம் உடைத்தெரிந்து எந்த நேரத்திலும் எந்தச் சூழல்களிலும் ஒருவர் இன்னொருவரோடு உரையாட உறவாட உற்றதொரு கருவியாகத் திகழ்வதுதான் தகவல் தொழில்நுட்ப மின்னணு கருவிகள். மொழியியலின் புதிய வடிவமாக, சமூக செயல்பாடுகளாகத் தகவல் தொழில்நுட்ப மின்னணு கருவிகள் ஊடகமாகச் செயல்படுவது ஒரு முக்கியச் சவாலாக ஏற்றுக்கொள்ள வேண்டும். இதன் முக்கிய நோக்கம் சிறுபாண்மை மக்கள் பேசுகின்ற மொழியின் பயன் அல்லது பயனர்கள் அதிகரிக்க வேண்டும்.

அழிவை நோக்கி சென்று கொண்டிருக்கும் மொழிகளுக்குப் புத்துயிர் கொடுக்கும் ஒரு காரணியாக தகவல் தொழில்நுட்ப மின்னணுக் கருவிகள் அமைகிறது என்றால் அது மிகையாகாது (Woolard & Schieffelin, 1994). தொடக்க காலத்தில் குறிப்பிட்ட ஒரு மொழியில் புலமை இல்லாததால் தகவல் தொழில்நுட்ப மின்னணுக் கருவிகளைக் கையாளுவதில் ஏற்படுகின்ற சிக்கல் ஒரு குறைபாடாகவே இருந்து வந்தது. இன்று நம் தாய்மொழியிலேயே கணினி அறிவைப் பெற்று அதை இயக்கும் அளவுக்கு முன்னேறியுள்ளோம். மற்றொரு சூழல், எந்தவொரு அடிப்படை கணினியறிவு இன்றியும் கணினியை ஒரு சில காரணங்களுக்குப் பயன்படுத்தக்கூடிய நிலையையும் நாம் காண முடிகின்றது. தமிழ்க்கணினி வளர்ச்சி என்பது இரு நிலைகளில் பார்க்கலாம். ஒன்று படைத்தல் (உருவாக்குதல்) மற்றொன்று பயன்படுத்துதல்.



வரைபடம் 1

பொதுவாகவே படைப்பாளிகளைக் காட்டிலும் அதிகமானோர் மேற்காணும் பிரிவுகளில் பயனர்களாகவே இருக்கின்றனர். எத்துணைச் சிறந்த; பிரமாண்டமான மென்பொருள்களை உருவாக்கித் தந்திருந்தாலும் பயனர்கள் அவற்றைப் பயன்படுத்தாவிடில் அத்துணைச் சிரமப்பட்டு படைத்த கருவிகளினால் பயன் ஏதும் இல்லை. இவ்வாறான சூழலை மலேசிய நாட்டில் இடைநிலைப்பள்ளிகளில் பயிலும் தமிழ்ப் படித்த மாணவர்களிடையே பரவலாகக் காணமுடிகின்றது. இது ஒரு அபாயகரமான சூழலாக உருவெடுத்துள்ளது. இதனைக் கருத்தில் கொண்டு இந்த மாணவர்களிடையே இணையத்தின் வாயிலாக தாய்மொழிப் பயன்பாட்டை எவ்வாறு நடைமுறைப்படுத்துகின்றனர் என்பதை அடிப்படையாகக் கொண்டு இவ்வாய்வு நடத்தப்பட்டது.

கணினி அல்லது தகவல் தொழில்நுட்ப பயன்பாட்டின் மிகவும் அடித்தள நிலைகள் இரண்டை மட்டும் எடுத்துக் கொண்டால் ஒன்று அறிவார்ந்த கைத்தொலைப்பேசிகளின் வாயிலாக குறுந்தகவல் அனுப்புவது மற்றொன்று சமூக வலைத்தளங்களில் கருத்துப் பரிமாற்றம் செய்வது. இவற்றில் இவ்வாய்வுச் சமூக வலைத்தளங்களில் ஒன்றான முகநூலை (facebook) மையமாகக் கொண்டுள்ளது.

பொதுவாக ஒருவரின் அடிப்படை மொழித்திறன் கேட்டல்-பேச்சு, எழுத்து, வாசிப்பு ஆகிய கூறுகளை அடிப்படையாகக் கொண்டுள்ளது. சமூக வலைத்தளங்களான முகநூல். டுவிட்டர் (Twitter) ஆகியவற்றின் பயனர்கள் தங்களோடு தொடர்பில் உள்ளவர்களிடம் தொடர்பு கொள்வதற்குப் பயன்படுத்துகின்ற மொழி; பிறரிடேத் தொடர்பு கொள்வதற்கும் பயன்படுத்துகின்ற மொழி, தங்களின் தாய்மொழியாக இல்லையெனில் மொழித்தாவல் ஏற்படுவதற்கான அறிகுறியாக இது அமைகிறது. மொழித்தாவல் ஏற்படுகின்ற சூழலில் மிக விரைவாக; மிகப் பரவலாக பாதிப்பை ஏற்படுத்துகின்ற சூழமைவாக இது அமைகிறது.

இந்த ஆய்வு, ஆரம்பக் கல்வி தமிழ்ப்பள்ளியில் பயின்று பின் இடைநிலைப்பள்ளிகளில் பயிலும் இந்திய மாணவர்கள்; அதாவது 13 வயதிலிருந்து 18 வயது வரை உள்ள மாணவர்களின் முகநூல்களில் அவர்கள் பயன்படுத்திய; பயன்படுத்தி வரும் மொழி எதுவாக இருக்கிறது என்பதை ஆராய்வதாகும். அதாவது ஆறு ஆண்டுகள் தமிழ்ப்பள்ளியில் படித்து அதில் பெரும்பாலோர் இடைநிலைப்பள்ளியிலும் தமிழ்மொழியை ஒரு பாடமாக எடுத்துப் படிக்கும் மாணவர்கள் மத்தியில் தங்கள் முகநூல்களில் எந்த மொழியைத் தொடர்பு மொழியாக பயன்படுத்துகின்றனர் என்பதை ஆராயும் முயற்சியாக இக்கட்டுரை அமைந்துள்ளது. இந்த மாணவர்களின் குடும்பங்களில் தமிழ்மொழியே தொடர்புமொழியாக இருக்கிறது; நண்பர்களோடு உரையாடும்போதும் பெரும்பாலோர் தமிழ்மொழியிலேயே உரையாடுகின்றனர். மேலும் இவர்கள் தமிழ்மொழியில் சரளமாகப் பேசும்; எழுதும் ஆற்றலும் கொண்டவர்களாக இருக்கின்றனர். இருப்பினும் முகநூல்களில் தமிழ்மொழியின் பயன்பாடு வருத்தம் அளிக்கும் வகையிலேயே அமைந்துள்ளது. இதனைக் கருத்தில் கொண்டு முந்நூற்று இருபது மாணவர்களின் முகநூல்களில் அவர்கள் பயன்படுத்திய மொழியை அடிப்படையாகக் கொண்டு இவ்வாய்வு நடத்தப்பட்டது. இந்த ஆய்விலிருந்து வெளிக்கொணரப்பட்ட சில தகவல்களைக் கீழ்க்காணும் அட்டவணை வாயிலாகக் காணமுடியும்.

அட்டவணை 1: முகநூலில் மொழி பயன்பாடு

	தமிழ்		ஆங்கிலம்	மலாய்	கலவை		மொத்தம்
	ஆ	த			ஆ	ம	
பயனர்	115	5	284	6	7	78	
விழுக்காடு	36%	2%	89%	2%	2%	24%	320

அட்டவணை:1

குறிப்பு:

ஆ - ஆங்கிலம்

த - தமிழ்மொழி

ம - மலாய் மொழி

மேற்காணும் ஆய்வில் மாணவர்கள் தங்களின் முகநூலில் அவர்கள் அனுப்பிய கருத்துகள்; தகவல்கள்; வாழ்த்துதல்கள்; சங்கிலித் தகவல்கள் ஆகியவற்றில் முதல் பத்தை மட்டும் எடுத்துக் கொண்டு ஆய்வு நடத்தப்பட்டது. அவை 2013-ஆம் ஆண்டு ஜூலை மாதம் 5- ஆம் திகதி முதல் 8-ஆம் திகதி வரை நடத்தப்பட்டது. முந்நூற்று இருபது மாணவர்களில் நூற்று இருபது மாணவர்கள் மட்டுமே தமிழ்மொழியைப் பயன்படுத்தி உள்ளனர். அவர்களில் நூற்று பதினைந்து மாணவர்கள் ஆங்கில எழுத்துருக்களைக் கொண்டு தமிழ்ப்படுத்தி எழுதியுள்ளனர். ஐவர் மட்டுமே தமிழ் எழுத்துகளைப் பயன்படுத்தி தமிழிலேயே எழுதியுள்ளனர். இங்கு இரண்டு காரணிகளை முன்வைக்கலாம். ஒன்று மாணவர்களின் தமிழில் தட்டச்சு செய்யும் திறனும் ஒரு சவாலாகவே அமைந்துள்ளது அல்லது தமிழ் எழுத்துருக்களைத் தங்கள் கணினிகளில் பதிவிடுக்கம் செய்து வைத்திருப்பதில்லை. அதாவது ஆறு ஆண்டுகள் தமிழ்ப்பள்ளிகளில் தமிழ்மொழியில் உரையாடிய; உறவாடிய மாணவர்கள் இடைநிலைப்பள்ளிகளைச்

சென்றடைந்த வேளையில் தமிழ்மொழியைத் தங்களது கணினியில் ஏற்றினாலே தகவல் தொழில்நுட்ப தொடர்பு கருவிகளில் மொழித் தாவல் ஏற்படுகின்ற சூழல்களைத் தவிர்க்க இயலும். ஆனால் பெரும்பாலானோர் ஆங்கில மொழியையே பயன்படுத்துகின்ற சூழலை மாணவர்களின் உரைக்கோவையில் காண முடிகின்றது. அதாவது மொத்தம் 284 பேர் (89%) ஆங்கில மொழியைப் பயன்படுத்துகின்றனர். இது மிகவும் வருத்தம் அளிக்கும் வகையில் அமைந்துள்ளது.

மேலும் பல்லின மக்கள் வாழும் சூழலில் மாணவர்கள் தங்களின் நண்பர்கள் வேறு இனத்தைச் சேர்ந்தவர்களாக இருப்பின் ஆங்கிலமோ மலாய் மொழியோ பயன்படுத்தியிருந்தால் ஏற்றுக்கொள்ளலாம். ஆனால் ஆங்கில மொழியைப் பயன்படுத்தும் அந்த 284 பேரில் 85 (30%) பேரின் முகநூல் பக்கத்தைச் சென்று பார்வையிட்டதில் 68 பேரின் முகநூல் நண்பர்களில் 90 விழுக்காட்டினர் இந்தியர்களாகவே இருக்கின்ற நிலையைப் பார்க்க முடிகின்றது. ஆங்காங்கே ஒரு சில வேறின நண்பர்கள் மட்டுமே இருக்கின்றனர். ஆகையினால் தமிழ்மொழியைப் பயன்படுத்த தடை ஏற்படுவதற்கு இது ஒரு காரணமாக அமையாது.

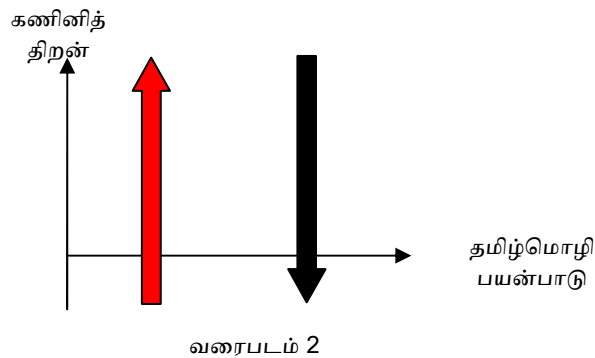
தொடர்ந்து அந்த முந்நூற்று இருபது மாணவர்களில் ஐம்பது மாணவர்களைத் தேர்ந்தெடுத்து அவர்கள் மைக்ரோசாப்ட் (Microsoft word), எக்ஸ் எல் (excel), பவர் பாயிண்டு (power point) மற்றும் முகநூல் ஆகியவற்றைப் பயன்படுத்துவதில் அவர்களிடம் காணப்படும் திறன் எவ்வாறு அமைந்துள்ளது என்பதை கீழ்க்காணும் அட்டவணை மூலம் அறிய முயற்சித்தபோது சில தகவல்கள் வெளிக்கொணரப்பட்டது. அவற்றைக் கீழ்க்காணும் அட்டவணையின்வழி காண்போம்.

	1	2	3	4	5	பதில் இல்லை	மொத்தம் பேர்	
மைக்ரோ சாப்ட்		5	14	26	5	-	50	90%
எக்ஸெல்	5	10	16	9	3	7	50	70%
பவர் பாயிண்டு	3	10	19	18	-	-	50	74%
முகநூல்	-	-	8	13	29	-	50	100%

1:நிச்சயமாக இல்லை
2:இல்லை
3:நடுநிலை
4:ஆம்

மாணவர்களின் கணினித் திறன்

மேற்காணும் அட்டவணை மைக்ரோசாப்ட், எக்ஸெல், பவர் பாயிண்டு மற்றும் முகநூல் ஆகியவற்றில் மாணவர்கள் எந்தளவிற்குத் திறன் பெற்றவர்களாக இருக்கின்றனர் என்பதைப் புலப்படுத்துகின்றது. இங்குக் கூர்ந்து கவனிக்கப்பட வேண்டிய ஒரு செய்தி என்னவெனில் அடிப்படை கணினித் திறன் உடைய மாணவர்கள் தங்கள் முகநூல்களில் தமிழ்மொழியைப் பயன்படுத்துவது மிகவும் குறைவாகவே உள்ளது.



மாணவர்கள்- கூட்டு நோக்காணல்

மாணவர்கள் தங்களது முகநூல்களில் தமிழ்மொழியைப் பயன்படுத்த இயலாத நிலைக்குப் பல காரணங்களை முன்வைத்தனர். இந்த ஆய்வின் நோக்கத்திற்கேற்ப சில காரணங்களை மட்டும் காண்போம். அவையாவன;

1. தமிழ் எழுத்துருக்கள் தங்களிடம் இல்லை.
2. தமிழில் தட்டச்சுச் செய்வதற்குக் கடினமாக இருக்கிறது.

3. தமிழில் தட்டச்சு செய்யப் பயிற்சி இல்லை.

அதாவது தமிழ் மொழியில் சரளமாகப் பேசும், கருத்துகளை முன்வைக்கும் ஆற்றல் இருக்கும் மாணவர்கள்கூட தமிழில் தட்டச்சு செய்ய சிரமப்பட்டு எளிதில் வேறு மொழிக்குத் தாவி விடுகின்றனர். மாணவர்கள் தங்களின் சிந்தனையில் உதித்த கருத்துகளை எழுத்து வடிவில் கொண்டுவர அதாவது தமிழிலேயே தட்டச்சுச் செய்து வெளியிட தட்டச்சுச் செய்வது ஒரு தடையாகவே இருப்பது இங்கு உணர முடிகின்றது. ஆகையால்தான் தங்களுக்கு எளிதாக இருக்கும் வண்ணம் ஆங்கில எழுத்துருக்களைப் பயன்படுத்தி அல்லது ஆங்கில மொழியோ; மலாய் மொழியோ பயன்படுத்தி தாங்கள் கூறவரும் செய்தியைப் பதிவு செய்கின்றனர்.

பரிந்துரை

ஆறாண்டுகள் தமிழ்ப்பள்ளியில் ஆரம்பக் கல்வியைப் பயின்று இடைநிலைப்பள்ளியிலும் தமிழ்மொழியை ஒரு பாடமாக எடுத்துப் பயின்று, பள்ளி வாழ்க்கையை முடித்து வெளியேறும் ஒரு மாணவனுக்குத் தமிழ்மொழியில் தட்டச்சு செய்யும் சிக்கலை எதிர்நோக்குகிறான் என்றால் வருத்தமளிக்கும் வகையில் இருக்கிறது. இச்சிக்கலைக் களைய உலகளாவிய நிலையில் தமிழ்த் தட்டச்சு உள்ளீட்டு முறையை ஒருமுகப்படுத்த வேண்டும். அனவரும் ஏற்றுக்கொள்ளும் பொதுவான ஒரு உள்ளீட்டு முறை இருந்ததென்றால் பயனர்களுக்கு அது மிகப்பெரிய பயனாக அமையும். மேலும் இதனைக் கருத்தில் கொண்டு பள்ளியில் ஆசிரியர்களும் மாணவர்கள் சுயமாக தட்டச்சு செய்யும் வகையில் தமிழில் தட்டச்சு செய்ய ஊக்குவிக்கவும் முறையான அடிப்படை பயிற்சிகளும் வழங்க வேண்டும்.

முடிவு

தமிழ்மொழி பல்வேறு பரிணாமங்களைக் கடந்து வந்து இன்று ஆறாம் திணையான இணையத்தை கொண்டாடி வருகிறோம். உலக வளர்ச்சிக்கு ஏற்ப தகவல் தொழில்நுட்பம்வழி தமிழ்மொழியை அடுத்தக் கட்டத்திற்குக் கொண்டுச் செல்ல வேண்டிய சூழலில் நாம் இருக்கிறோம். புதிய மொழியியல் வடிவங்களில் ஒன்றான தகவல் தொழில்நுட்ப மின்னணுக் கருவிகளில் தமிழ்மொழி இடம்பெறத் தவரினால் சிறுபான்மை மக்களிடையே மொழி பயன்பாடும் பயனர்களின் எண்ணிக்கையிலும் மிகப்பெரிய தோய்வு ஏற்படும். இதனைக் களைய தக்க நேரத்தில் தக்க நடவடிக்கை எடுத்து அன்னை தமிழ்வழி கணினியுலகை வெல்வோம்.

மேற்கோள் நூல்கள்

Fasold, R. (1987). *The sociolinguistics of society*. Oxford, UK: Blackwell.

Fishman, J.A. (1985). *Language maintenance and ethnicity*, M.H. Gertner

Giles (ed.) (1977), *Language, ethnicity and intergroup relations*. London: Academic Press

King K.A. (2001). *Language Revitalization proses and Prospects: Quichuain the Ecuodarian Andes*. Clevedon, UK

தகவல் தொழில்நுட்பத்தில் தமிழ் - ஒரு பார்வை

தங்கராஜா தவருபன்

தலைமை நிறைவேற்று அலுவலர், இணையத்தள/மென்பொருள் வடிவமைப்பாளர்

Speed IT net. யாழ்ப்பாணம் இலங்கை, Moodle Administrator, யாழ்ப்பாணம் பல்கலைக்கழகம், இலங்கை மின்னஞ்சல்
thava@speedit.lk

முன்னுரை

தகவல் தொழில்நுட்பம் இன்று நமது வாழ்வியக்கத்தின் சகல நிலைகளிலும் ஆதிக்கம் செலுத்திக்கொண்டிருக்கும் நிலையில் அதன் நுகர்வு மற்றும் பிரயோகத்தில் மொழி மிகப்பெரிய செல்வாக்கு செலுத்தி வருவது யாவரும் அறிந்தது. தொழில் நுட்பம் என்றாலே அது ஆங்கிலமொழி மூலமானது அதன் மூலமே அதன் உண்மையான வடிவத்தினை காணமுடியும் என்றும் உலகமயமாதலில் ஆங்கிலம் இன்றி எதுவும் இல்லை என்றும் கூறி பலர் தொழில்நுட்பத்தில் தாய்மொழி பிரயோகம் குறித்து எள்ளி நகையாடுவதை இந்நாட்களில் அவதானிக்கலாம்.

சீனர்கள் யப்பானியர்கள் போல தமிழர்களாலும் தாய்மொழியில் தொழில்நுட்பத்தினை நுகரவோ பிரயோகிக்கவோ முடியும் என்பதை கடந்தகாலங்களில் நிரூபித்து வருகின்றோம். இருந்தபோதிலும் நம்மில் பெரும்பாலானோர்களிடம் உள்ள ஆங்கிலம் மீதான கவர்ச்சி, தாய் மொழி மீதான தாழ்வு மனப்பான்மை, தாய் மொழியான தமிழினை முன்னிலைப்படுத்தும் போது எம்மை ஆங்கிலப்பலமையற்றவர்களாக காட்டிவிடுமோ என்கின்ற அச்சம், ஆகியன தமிழ் மொழியினை தகவல் தொழில்நுட்பத்தில் தெரிவு செய்ய நமக்கு தடை போட்டு விடுகிறது. தற்காலத்தில் தமிழ்மொழிக்கல்வி தமிழ்மொழிமூலக்கல்வியின் தரக்குறைபாடுகள் அல்லது அவற்றின் அருகி வரும் நிலமையும் இதற்கு காரணம் எனலாம்.

எது எப்படியோ மொழியானது தொழில்நுட்பத்துக்கு ஒரு தடையாக இருக்கக்கூடாது என்பதற்காகக் குறிப்பிடத்தக்க அளவிலான மொழியியலாளர்களும் ஆராய்ச்சியாளர்களும் தொழில்நுட்பவியலாளர்களும் உழைத்துக் கொண்டிருப்பதையும் குறிப்பிடத்தக்களவு வெற்றிகளை அவர்கள் அடைந்து இருப்பதையும் இங்கு குறிப்பிட்டுத்தான் ஆகவேண்டும்.

இந்நிலையில் தற்போது தமிழ் தகவல் தொழில்நுட்பமானது பலஇடங்களில் தனது செல்வாக்கினை காட்டியிருக்கிறது. இருந்த பொழுதிலும் மேலே குறிப்பிட்ட காரணங்களின்றி வேறு பல காரணங்கள் கணினியிலோ, நடமாடும் கையடக்கசாதனங்களிலோ இயந்திர பொறிகளிலோ இணையத்திலோ தமிழ் ஒரு தெரிவாக இருப்பதற்கு சவாலாக இருக்கின்றதை பலரும் ஏற்றுக்கொண்டிருக்கின்றார்கள். அவை தர்க்கரீதியிலும் தொழில்நுட்பரீதியிலுமான காரணிகளாக காணப்படுகின்றன

இவற்றினை பட்டியலிட்டு அவற்றினை விளக்குவதும் அவற்றினை எவ்வாறு எதிர்கொள்ள முடிகிறது என்றோ முடியும் என்றோ விபரித்து விளக்குவதாக இந்த கட்டுரை அமைகிறது.

உண்மையில் தொழில்நுட்பம் ஆனது எந்தவொரு இனத்திற்கோ சமூக மட்டத்தினருக்கோ வயது பிரிவினருக்கோ வரையறுக்கப்பட்டதல்ல. அது சகலரையும் ஒரு கூரையில் ஒன்றிணைக்கும் அதேவேளை சகலரையும் ஒரே கண்ணோட்டத்தில் நோக்குகிறது. சகலருக்கும் இடையில் உள்ள இடைவெளிகளை இல்லாமல் செய்கின்றது. இதற்கு மொழி ஒன்றே முக்கிய தடை. அதனை வெற்றிகொண்டுவிடுவோமாகில் நிச்சயம் எம்முடைய சமூகம் நாடு இனம் அனைத்தும் முன்னேற்றப்பாதையில் சென்றுவிடும்.

தமிழ் தகவல் தொழில்நுட்பம்

இன்று உலகில் ஆங்கிலத்தில் தான் அநேக இணைய வழிச்செயற்பாடுகள் இருக்கின்றன. தொழில்நுட்பம் கணினிப் பயிற்சி பெற்றவர்களுக்கும் உயர், மத்தியதர வர்க்க மக்களுக்கும் மட்டுமல்ல என்பது தெளிபடுத்தப்பட்டு நிரூபணமாகிவிட்ட சூழ்நிலையிலும், இன்னும் தகவல்தொழில்நுட்பப்பாவனை குறித்த தரப்பினருக்கு மட்டும் தான் என்றும், நமக்குச் சரிப்பட்டு வராது என்றும் எண்ணிக்கொண்டிருப்பவர்களும், கணினிகளையும் இதர இலத்திரனியல் சாதனங்களின் பயன்பாட்டையும் கண்டு வியந்து பெருமூச்சு விட்டுக்கொண்டிருப்பவர்களும் நிறையவே இருக்கின்றார்கள். இருந்த போதிலும் தற்போது மக்களைத் தகவல் தொழில் நுட்பம் தேடிச் சென்று கொண்டிருக்கின்றது என்பதை மறுக்கமுடியாது.

அந்த வகையில்தான் இந்தத் தமிழ்த் தகவல் தொழில்நுட்பம் என்ற விடயம் பற்றியும் இன்று நம்மிடையே பேசப்படுகின்றது. இதன்போது தான் தமிழில் கணினி, தமிழ் இணையம் என்றெல்லாம் பேசுகின்றோம். இந்த விடயம் எம்மில் பலராலும் விவாதிக்கப்படுகின்றதும் பயன்பாட்டிற்குட்படுத்தப்படுவதும் இன்று சாதாரண

விடயமாகிவிட்டது.தகவல் தொழில்நுட்பத்துக்குள் தமிழ் புகுந்துகொண்டு அலங்கரிக்கும்போத தமிழ் தகவல்தொழில்நுட்பம் உயிர்பெறுகின்றது.

அடிப்படை காரணிகள்

தமிழில் தகவல் தொழில்நுட்பம், அதாவது நமது தாய் மொழியில் தகவல் தொழில்நுட்பம் என்னும்போது கணினி உள்ளிட்ட அனைத்து தகவல் தொழில்நுட்ப சாதனங்களிலும் அவற்றின் பயன்பாடு தொடர்பில் பின்வரும் காரணிகள் செல்வாக்கு செலுத்துகின்றன.

1. எழுத்துரு (Font)
2. உள்ளீடும் விசைப்பலகையும் (Input and Keyboard)
3. வெளியீடும் இடைமுகமும் (Output and GUI/Interface)
4. கலைச்சொற்கள் (Glossary)
5. தமிழ் மென்பொருட்கள் (Tamil Software)

இவற்றினைப்பற்றிய விளக்கங்கள் பல்வேறு கட்டுரையாளர்களால் காலத்துக்குக்காலம் விளக்கப்பட்டுவிட்டதனால் நேரடியாக சாவல்கள்பற்றி ஆயராயப்படுகிறது.

சவால்கள் - எழுத்துருக்கள்

தற்போது பரிந்துரை செய்யப்பட்டுள்ள *செந்தரங்கள் (Standard) மற்றும் கலைச்சொல்லாக்கங்கள் முழுமையாக பின்பற்றப்படாதிருக்கின்றன அல்லது அறியப்படாமல் இருக்கின்றது. யுனிக்கோட், எழுத்துருப்பாவனை ஏற்றுக் கொள்ளப்பட்டிருக்கின்ற போதிலும் பலர் பழைய சாதாரண எழுத்துருக்களை பயன்படுத்துகின்றனர்.

பலருக்கு “யுனிக்கோட்” எனப்படுகின்ற ஒருங்குறி குறியாக்கம் பற்றிய தெளிவான புரிதல் ஏற்பட்டிருக்கவில்லை. தொழில்நுட்பம் என்று வந்துவிட்டால் அதில் அடிப்படை அறிவின்றி பயன்படுத்த முடியாது என்பதை ஏற்றுக்கொண்டுதான் ஆகவேண்டும்.அதற்காக குறியாக்கம் (Encoding)குறித்த விளக்கம் தெரியாமல் யுனிக்கோட் எழுத்துருவை பயன்படுத்த முடியாது என்பதில்லை.இருப்பினும் சாதாரண எழுத்துருவுக்கும் யுனிக்கோட் எழுத்துருவுக்கும் வித்தியாசம் தெரியாமல் பயன்படுத்தும்போது பல நடைமுறைச்சிக்கல்களை எதிர்கொள்ளவேண்டிவரும்.

உதாரணமாக விரிவுரையாளர் ஒருவரால் ஏற்கனவே ASCII குறியாக்கத்தின் கீழமைந்த பாமினி எழுத்துருவில் தமிழில் தட்டச்சு செய்யப்பட்ட ஒரு கோப்பில் உடனடியாக ஒரு மாற்றம் செய்ய வேண்டியிருக்கின்றது.அதனை ஆய்வு செய்யும் பிறிதொரு விரிவுரையாளர் அதில் தனக்கு தெரிந்த புதிய யுனிக்கோட் குறியாக்கத்தின் கீழமைந்த எழுத்துருவில் மாற்றம் செய்கின்றார் இப்போது தான் சிக்கல் ஏற்படுகின்றது. பார்வைக்கு சகலதும் தமிழ் தான் ஆனால் குறியாக்கம் கலந்துவிட்டது.

இக்கோப்பினை பின்னாளில் கையாளும் இன்னும் ஒரு உதவி விரிவுரையாளர் மாற்றங்கள் செய்வதற்கும் வேறு ஒரு வடிவத்துக்கு மாற்ற விரும்பும்போதும் தீர்க்கமுடியாத பிரச்சனைக்கு உள்ளாவார்.

இங்குதான் குறியாக்கங்களிற்கிடையிலான அல்லது எழுத்துருக்களுக்கு இடையிலான வித்தியாசம் பற்றிய புரிதல் செல்வாக்கு செலுத்துகின்றது. இதற்கிடையில் TACE என்ற மற்றொரு குறியாக்கம் வேறு முளைத்துவிட்டது. தகவல் தொழில்நுட்பவியலாளர்களுக்கு இவை பெரிய பிரச்சனையே அல்ல. ஆனால் பயன்படுத்தும் சாதாரண பயனாளர்களுக்குத்தான் பிரச்சனை. தமிழ் மொழியில் தொழில்நுட்பத்தினை கையாளவேண்டும் என்று கொண்டுவரப்பட்ட சில மாற்றங்கள் பலவேளைகளில் பயனாளர்களை சலிப்புறச்செய்து உலகப்பொது மொழி ஆங்கிலத்துக்கு அழைத்துச்செல்கின்றது என்றால் அது மிகையில்லை.

சவால்கள் - விசைப்பலகை

விசைப்பலகையினை எடுத்துநோக்கினால் பலர் பழைய தட்டச்சுமுறையிலமைந்த பாமினி எழுத்துருக்கட்டமைப்பு விசைப்பலகையினை உபயோகப்படுத்துகின்றனர்.இது ஈழத்தமிழர் மற்றும் புலம்பெயர் தமிழர்களின் புழக்கத்தில் உள்ளது. வேறு சிலர் தமிழ் 99 விசைப்பலகையினையும் பயன்படுத்துகின்றனர். இன்னும் சிலர் திருத்தப்பட்ட சிறீலங்கா விசைப்பலகையினை உபயோகிக்கின்றனர். அதைவிடவும் சிலர் வேறு விசைப்பலகைமுறைகளையும் பயன்படுத்துகின்றனர். இதனால் விசைப்பலகை உற்பத்தியாளர்களும் விசைப்பலகை மென்பொருள் தயாரிப்பாளர்களும் தொடர்ந்தும் குழப்பத்தில் உள்ளனர்.

உதாரணமாக விசைப்பலகை மற்றும் குறியாக்கம் பற்றிய தெளிவில்லாத ஒரு வாடிக்கையாளர் விற்பனை நிலையம் ஒன்றிற்கு செல்கின்றார் எனக்கு தமிழில் தட்டச்சு செய்ய ஒரு விசைப்பலகை வேண்டும் என்று கேட்கின்றார்.அதற்கு கடைக்காரர் என்ன வகையான விசைப்பலகை வேண்டும் என்று திருப்பிக்கேட்டால், எதுவாயினும் பரவாயில்லை எனக்கு தமிழ் வேண்டும் என்கிறார்.இங்கு தான் பிரச்சனை ஆரம்பிக்கின்றது. இத்தருணத்தில் கடைக்காரர்

தன்னிடமுள்ள இருப்புக்கும் இலாபத்திற்கும் ஏற்ற வகையில் பயனாளரின் குறியாக்க முறைமை மற்றும் எழுத்துரு முறைமை ஆகியனவற்றினை தீர்மானிக்கப்போகின்றார்.

இன்றுவரை சிறந்த கையேடுகளுடன் சரியான முறையில் விளக்கம் தரப்பட்ட விசைப்பலகையினையோ விசைப்பலகை மென்பொருள்களையோ விற்பனை நிலையங்களில் காண முடியவில்லை. ஆனால் இலவசமான இணையத்தில் நிறைய விசைப்பலகை மென்பொருள்கள் கிடைக்கின்றன. அவற்றில் குறிப்பிடத்தக்கதாக NHM Writer மென்பொருள் விளங்குகின்றது.

கையடக்க சாதனங்களிலும் பிரச்சனை இருக்கிறது. தொழில்நுட்பம் நாளுக்குநாள் வளர்ந்து கொண்டிருக்க நமது தாய்மொழிப்பாவனைக்கான மென்பொருள்கள் விசைப்பலகைகள் போன்றவற்றிலும் சில மாற்றங்களை செய்யவேண்டியிருப்பதை நினைவில் கொள்ளவேண்டும்.

கையடக்க சாதனங்களில் ஆரம்பத்தில் "அன்றேராயித்" இயங்குதளத்திற்காக உருவாக்கப்பட்ட தமிழ் விசை போன்ற விசைப்பலகை மென்பொருள் இன்றுவரை பயன்பாட்டில் உள்ளது. அதைவிட சில மென்பொருள்கள் வந்தபோதிலும் அதை மிஞ்ச முடியவில்லை. இருப்பினும் அதன் பயன்பாடு அதனை நிறுவுதல் தொடர்பில் பலருக்கு தெரியாத நிலையும், இவை சாதனங்கள் மற்றும் இயங்குதளப்பதிப்புக்களுடன் ஒத்துபோகும்போது எழுகின்ற சிறுபிரச்சனைகளுக்கு தீர்வுகாண்பதில் உள்ள சிக்கல்களைத்தவிர்ப்பதற்காகவே, வேண்டாம் இந்த தமிழ் வழமைபோல தமிழ்க்கிளத்தில் ஆங்கில எழுத்துகளில் தட்டச்சு செய்துவிடுவோம் என்று தப்ப பலர் முயற்சிக்கின்றனர் . கட்டற்ற மென்பொருள் இலவசம் இந்த விடயங்களினால் நல்ல விடயங்கள் அரங்கேறியிருந்தாலும் சில வேளைகளில் வாடிக்கையாளரின் விருப்புக்குரிய சேவை மறுப்புக்குள்ளாவதையும் தாமதப்படுத்தப்படுவதையும் தவிர்க்கமுடியாமல் இருக்கிறது

iOS இயங்குதளத்தில் இயங்கும் சாதனங்களில் தமிழ் தட்டச்சுசெய்ய இதுவரை சிறந்த விசைப்பலகை மென்பொருள் உருவாக்கப்படவில்லை. முரசு செல்லினம் மட்டுமே முன்னுக்கு நிற்கின்றது.ஆயினும் அதில் பிறிதொரு பகுதியில் தட்டச்சுசெய்து பிரதி செய்து பயன்படுத்தவேண்டியிருப்பது எவருக்கும் இலகுவானதாக இருக்கவில்லை. iOS 7 பதிப்பில் அப்பிள் நிறுவனம் தமிழ் விசைப்பலகையினை இணைத்திருக்கிறது.எனினும் அதனை உய்யோகப்படுத்தும் சந்தர்ப்பம் உலகத்தமிழர்கள் அனைவருக்கும் இன்னும் வழங்கப்படவில்லை. இலகு முறைகள் தான் இனிவரும் தலைமுறையின் இதயத்தில் இடம்பிடிக்கும் என்பதை உற்பத்தியாளர்கள் கவனத்தில் கொள்ள வேண்டும்.இலவசங்களில் அதனை எதிர்பார்ப்பதும் கடினமே!

சவால்கள் - கலைச்சொல்லாக்கம்

கலைச்சொல்லாக்கம் கலைச்சொல்லகராதிப் பாவனைகளில் இன்னும் முரண்பாடுகள் இருந்துகொண்டே இருக்கின்றன.கலைச்சொற்கள் தான் உண்மையில் தமிழ் தகவல் தொழில்நுட்பத்தின் வளர்ச்சியை தீர்மானிக்கும் சக்தி என்றால் அது மிகையாகாது.கலைச்சொல்லாக்கத்தில் இன்னும் புரிந்துணர்வின்மை. கௌரவம், சமூகப்பொறுப்பு, மொழிப்பாதுகாப்பு, பழைமைவாதம், நேரடிமொழிபெயர்ப்பு போன்ற காரணிகள் செல்வாக்கு செலுத்திய வண்ணம் இருக்கின்றன.இன்றைய எவராலும் ஒரு பொதுக்கலைச்சொல்லகராதியினை உருவாக்க முடியவில்லை.

நாளுக்கு நாள் புதுச்சொற்கள் புதுத்தளங்கள் இதற்காக முளைக்கின்றன. பல்தேசியக்கம்பனிகளின் தளங்களை எதுவித தூரநோக்குமின்றி மொழிபெயர்க்க முற்படும் செயற்பாடுகளால் எமது வளம் வீணடிக்கப்படுகிறது. அதனை ஒரு பொது கலைச்சொல்லகராதியினை உருவாக்குவதற்கு பயன்படுத்தமுடியும். உள்ளூர் மயமாக்க தவறுகள் மீதான வெறுப்பு தமிழ் தகவல் தொழில்நுட்பத்தின் மூலமான வெறுப்புக்கு வித்திடுகின்றமை இங்கு நினைவில் கொள்ளவேண்டியிருக்கின்றது.

கலைச்சொல்லாக்கம் தான் மென்பொருட்கள் மற்றும் இணையத்தளங்களின் உள்ளூர்மயமாக்கலில் மிகப்பெரிய தாக்கத்தினை செலுத்துகின்றது . ஆங்கிலச்சொல்லை தமிழில் உச்சரித்து எழுதுவதைக்கூட கலைச்சொல் என்று கருதுகின்றமையும் ஏற்க முடியாதது. கலைச்சொல் ஒவ்வொன்றும் ஒரு குழந்தையின் பிரசவம் என்பதை நினைவில் கொள்ளாது நாமும் புதுச்சொல் கண்டுபிடிப்போம் என்ற வகையில் பலரால் கலைச்சொற்கள் உருவாக்கப்பட்டுவருகின்றன.

இருப்பினும் குறிப்பிடத்தக்களவு கலைச்சொற்கள் தகவல்தொழில்நுட்பத்திற்காக உருவாக்கப்பட்டு புழக்கத்தில் உள்ளமையினை மறுக்கமுடியாது. இலங்கை இந்திய அரசாங்கத்தின் கூட்டுமுயற்சியினால் உருவான தமிழ் தகவல் தொழில்நுட்ப கலைச்சொல்லகராதியை மீளவும் புதுப்பிக்கவேண்டிய அவசியம் தற்போது இருக்கின்றது.

சவால்கள் –மென்பொருள்கள்

எல்லாம் இருந்தாலும் கடைசியில் தகவல்தொழில்நுட்ப சாதனத்தினைப்பயன்படுத்தி தொழிற்பாடு ஒன்றை நிறைவேற்றவேண்டுமாயின் தமிழ் மென்பொருள்கள் அவசியம். அவற்றின் தரம் பயன்பாடு போன்றவற்றினைத்தான் மேற்குறித்த எழுத்துரு உள்ளீடு விசைப்பலகை கலைச்சொல்லகராதி ஆகிய காரணிகள் வழிநடாத்துகின்றன.

உதாரணமாக Hybrid கார் ஒன்றின் சாரதிக்கு முன்னால் உள்ள வழிகாட்டும் அறிவுறுத்தல் சாதனத்தின் மென்பொருளில் சகல அறிவுறுத்தல்களும் யப்பானிய மொழியில் மொழிபெயர்க்கப்பட்டு நிறுவப்பட்டிருப்பதை அவதானிக்கும்போது யப்பானியர்கள் எந்தளவுக்கு தமது மொழி மீது அக்கறை கொண்டிருக்கிறார்கள் என்பதும் அவர்களது தொழில்நுட்ப வளர்ச்சிக்கு எந்தளவு மொழி காரணமாக இருந்திருக்கிறது என்பதும் புலனாகின்றது. அதேபோல நாம் உபயோகிக்கும் சகல இலத்திரனியல் சாதனங்கள் இயந்திரங்களிலும் நமது தமிழ்மொழியினை உள்வாங்கினால் அதன் பயன்பாடுகள் அதிகரிக்கும் அதேவேளை சகல மட்டத்திற்கும் தகவல்தொழில்நுட்பத்தின் பயன்பாட்டினை தாவச்செய்யலாம்.

ஏற்கனவே குறிப்பிடத்தக்களவு மென்பொருள்கள் உருவாக்கப்பட்டிருக்கின்ற போதிலும்

- மொழிபெயர்ப்பு முறைகள்
- கலைச்சொல்லாக்க தவறுகள்
- இலகு தன்மை இன்மை
- அவற்றின் தரங்களில் குறைபாடு
- ஆங்கிலத்தின் மீதான இயல்பான மோகம்
- ஆங்கில மொழிப்பாவனை தமது சமூகத்தரத்தினை மற்றவர்களிடத்தில் உயர்த்திக்காட்டும் என்ற எண்ணம்
- தமிழ் மென்பொருட்களின் தொடர் பதிப்புக்கள் வெளியாவதில் உள்ள தாமதம்

போன்ற காரணிகள் தமிழ் மென்பொருட்களின் பாவனையினை கேள்விக்குட்படுத்துகின்றன. இருந்தும் திறவூற்று மென்பொருட்களாகவும் வர்த்தக மென்பொருட்களாகவும் நிறைய மென்பொருட்கள் உருவாக்கப்பட்டிருக்கின்றன. உருவாக்கப்படுகின்றன. இவை

- சரியான முறையில் சந்தைப்படுத்தப்படாமை
- ஒரு இடத்தில் மையப்படுத்தி வளங்களையும் ஒருமுகப்படுத்தப்படாமை
- அவை பற்றிய தகவல்கள் போதியளவு மக்களிடத்தில் கொண்டுசெல்லப்படாமை

பலவேளைகளில் வெற்றி பெற தவறிவிடுகின்றன. இருப்பினும் உத்தமம்போன்ற அமைப்புக்களின் செயற்பாடுகளினால் தமிழ் மென்பொருட்கள் முக்கியத்துவம் பெறும் என்பதில் ஐயமில்லை.

பல்தேசிய நிறுவனங்களினால் வெளியிடப்படும் மென்பொருள்களில் தமிழ் பயன்படுத்தப்படுவதற்கான ஏற்பாடுகள் அண்மைக்காலங்களில் மிகப்பெரிய வெற்றியளித்திருக்கின்றன என்பதையும் ஏற்றுக்கொள்ளவேண்டியிருக்கின்றது. அண்மையில் பதிப்புத்துறை மென்பொருளான Corel Draw மென்பொருளில் கூட தமிழ் யுனிக் கோட் பயன்படுத்தக்கூடிய நிலை ஏற்பட்டிருக்கின்றமை இங்கு குறிப்பிடத்தக்கது.

அதேவேளை எதிர்கால தொழில்நுட்பங்களைநோக்கிய தமிழ் மென்பொருட்களிலும் அதன் அபிவிருத்தியிலும் கூட நல்ல முன்னேற்றங்களை காணக்கூடியதாக உள்ளது. தமிழ் எழுத்துணரி மென்பொருட்கள் மற்றும் கூகிள் நிறுவனத்தின் தமிழ் எழுத்துக்களை ஒலிபெயர்க்கும் செயற்பாடுகளில் நல்ல முன்னேற்றங்களை காணக்கூடியதாக உள்ளது. எது எப்படியிருப்பினும் எமது மனங்களில் தான் தமிழ் மொழி மூல மென்பொருட்களின் பாவனை மற்றும் அதற்கான கேள்வி பெரிதும் தங்கியுள்ளது. இலவசங்கள் மீதான எதிர்பார்ப்பு தரமான தமிழ் மென்பொருட்களுக்கான சந்தையினை ஏற்படுத்தாது என்பது எனது வாதம்.

முடிவு

தகவல்தொழில்நுட்பத்தில் தமிழின் முக்கியத்துவத்தினை ஏற்படுத்துவதற்காக மக்கள் மனங்களில் மாற்றங்களை ஏற்படுத்தவும் ஒன்றுபட்ட செந்தரத்தினை பின்பற்றச் செய்வதற்கும் இதுவரை எடுக்கப்பட்ட முயற்சிகள் திருப்திகரமான இருக்கவில்லை. இது விடயத்தில் ஒன்றுபட்டு செயற்பட வேண்டியிருக்கின்றது. அரசு செயற்றிட்டங்களை மற்றும் மென்பொருள்களை செய்யும் தனியார் நிறுவனங்கள் மொழியில் தேர்ச்சி பெற்றவர்களை கொண்டிருக்க தவறியுள்ளமை தமிழுக்கு பெரும் பாதிப்பை ஏற்படுத்தி விடுகின்றது. கூட்டுப்பொறுப்புடன் தொழிற்படுவதற்கு நிறுவனங்கள் சங்கங்கள் தமிழ் அறிஞர்கள் தகவல்தொழில்நுட்பவியலாளர்கள் இன்னும் நிறைய அர்ப்பணிக்கவேண்டியிருக்கிறது. இதனை நமது வீட்டில் இருந்து தொடங்குதல் வேண்டும்.

தமிழ் மென்பொருட்களை வடிவமைப்பவர்கள் இது தொடர்பில் ஏற்கனவே உள்ள மென்பொருட்கள் பற்றி சிறந்த மீளாய்வு ஒன்றினை செய்திருக்கவேண்டியது அவசியமாகின்றது. விசைப்பலகை, ஒலிபெயர்ப்பி, எழுத்துணரி, பிழைதிருத்தி, கலைச்சொல்லகராதி போன்ற விடயங்களில் ஆராய்ச்சிகள் திரும்ப திரும்ப ஆரம்பத்தில் இருந்து ஆரம்பிக்கப்படுவது நிறுத்தப்பட்டு விட்ட இடத்தில் இருந்த தொடங்குவது ஆரோக்கியமானதாக இருக்கும். மாற்றங்களை செய்துகொண்டிருப்பதை மட்டும் செயற்பாடாக கருதாமல் மாற்றத்தினை மக்களிடம் கொண்டுசெல்லவேண்டிய பொறுப்பும் மாற்றங்களை ஏற்படுத்துவோருக்கு உள்ளது.

பல்கலைக்கழகங்களில் இறுதியாண்டு செயற்திட்டங்களில் தமிழ் மொழிமூல மென்பொருள்களுக்கான ஆர்வத்தினை ஏற்படுத்துவதன் மூலமும். வித்தகச்செருக்கினை ஒதுக்கிவைத்து தரமான கலைச்சொற்களை வழங்குவதன் மூலமும் பள்ளிக்கல்வியிலும் சரி பல்கலைக்கல்வியிலும் சரி தகவல்தொழில்நுட்பத்தினை கற்பிக்கும்போதும் தமிழின் பாவனை பற்றியும் ஒரு பாடம் அங்கே எடுத்து அடிப்படை தமிழ் தகவல்தொழில்நுட்ப அறிவை வழங்குவதன் மூலமும் மேற்குறித்த சவால்களை வெற்றி கொள்ளமுடியும்.மொழி தொழில்நுட்பத்திற்கு ஒரு தடையில்லை என்பதை அதன் மூலம் உணர்த்தலாம்.

திறந்தமூல தமிழ் மென்பொருட்கள் மற்றும் கணினிக்கு தமிழ் மொழியறிவு ஊட்டல் பற்றிய ஆய்வுகள்

ஜெயராஜன் யோகராஜ், இலங்கை

1. கணனி செயல்பாடுகளை உள்ளூர் மொழிகளில் மேற்கொள்வதற்கு பல நடவடிக்கைகள் எடுக்கப்பட்டுள்ளன. அந்த வகையில், திறந்த மூலங்களான உபுண்டு (Ubuntu OS), பையர் பொக்ஸ் (Fire Fox), தண்டர் பேர்ட் (Thunderbird), ஓபன் ஒப்பிஸ் (Open Office) உட்பட பல முக்கிய மென்பொருட்களை தமிழில் பயன்படுத்த நடவடிக்கைகள் மேற்கொள்ளப்பட்டுள்ளன.

2. இந்த மென்பொருட்கள் தமிழில் மொழிபெயர்க்கப்பட்டது மட்டுமல்ல, அவற்றை மொழிபெயர்க்கும் செயல்பாடுகளில் கணனி துறை நிபுணர்கள் மற்றும் தமிழ் அறிஞர்களும் உள்வாங்கப்பட்டுள்ளனர். அந்த வகையில், இந்த மொழிபெயர்ப்பாளர் சங்கம் ஒரு சமூகமாக விரிவடைவதற்கு ஆவண செய்யப்பட்டுள்ளது. இங்கு திறந்த மூல தயாரிப்புக்களுக்கான மொழிபெயர்ப்பு மிகவும் அவதானமாக மேற்கொள்ளப்பட்டது. அதாவது மொழிபெயர்ப்பு சொற்களுக்கான மீளாய்வு திட்டமொன்றும் அமுல்படுத்தப்பட்டுள்ளது. அதன் மூலம் குறிப்பிட்ட ஆங்கில சொல்லுக்கான திருத்தமான தமிழ் வடிவம் பெறப்பட்டது.

3. மேலும் இந்த மென்பொருட்களை பயன்படுத்துவதற்கு தேவையான அனைத்து ஊக்குவிப்பு நடவடிக்கைகளும் மேற்கொள்ளப்பட்டுள்ளன. இலங்கையில் தேசிய இலத்திரனியல் ஊடகங்கள், அச்சு ஊடகங்கள், ஒலி-ஒளி ஊடகங்கள் உட்பட தனியார் ஊடகங்களும் இந்த பயன்பாட்டை ஊக்குவிக்க ஆக்கபூர்வமான நடவடிக்கைகளை மேற்கொண்டுள்ளது.

4. இதன் ஊடாக சிறு பிள்ளை பராயம் தொடக்கம் பெரியோர் வரை அனைத்துத் தரப்பினரும் திறந்த மூலங்களுக்கான தமிழ் வடிவத்தை நன்கு புரிந்துகொள்ள வாய்ப்பு ஏற்பட்டுள்ளது. இதன் மூலம் திறந்த மூல பயன்பாடும் கணிசமான அளவு அதிகரித்துள்ளது. திறந்த மூல தயாரிப்புக்கள் மூலம் கணனி பாவனையாளர்களுக்கு பல நன்மைகள் கிட்டுகின்றன. இந்த மென்பொருட்களை பயன்படுத்துவதற்கான அனுமதி பத்திரங்கள் இலவசம். மேலும் மைக்ரோசொப்ட் தயாரிப்புக்களுக்கு நிகரான அனைத்து தயாரிப்புக்களையும் திறந்த மூலங்கள் கொண்டிருக்கின்றன.

Microsoft Products

Internet Explorer

Outlook Express

MS Office

Open Source Products (equivalent)

Fire Fox

Thunderbird

Open Office

மேற்குறிப்பிட்ட சில தயாரிப்புக்கள் இவற்றிற்கு சிறந்த உதாரணங்களாகும். எனவே, திறந்த மூல தயாரிப்புக்களில் உள்ள மற்றுமொரு நன்மை என்னவெனில், இவை வரைல் தாக்கங்களுக்கு உட்படாதவையாகும். இவ்வாறு திறந்த மூல தயாரிப்பு பாவனையாளர்களுக்கு பல நன்மைகள் உண்டு. இந்த நன்மைகளை அறிவுறுத்தும் வகையிலும், பல செயலமர்வுகள் மேற்கொள்ளப்பட்டுள்ளன. ஊடகங்களும் இவற்றிற்கு கணிசமான பங்களிப்பு செய்துள்ளன.

5. இதன் அடிப்படையில், திறந்த மூல தயாரிப்புக்களுக்கான மொழிபெயர்ப்பாளர்கள் மற்றும் பாவனையாளர் எண்ணிக்கையில், அண்மைக்காலத்தில் அதிகரிப்பு காணப்பட்டுள்ளது. இந்த செயல்பாடு ஒரு வரையறுக்கப்பட்ட செயல்பாடு அல்ல. மாறாக நீண்டகால செயல் திட்டமாக தொடர்ச்சியாக மேற்கொள்ளப்படுகிறது. எனவே, நீண்டகால நோக்கம் கருதி, மேற்கொள்ளப்பட்ட இந்த செயல்திட்டத்தின் மூலம், எதிர்காலத்தில் இலங்கை வாழ் தமிழ் பேசும் மக்கள் திறந்த மூல பயன்பாட்டை மேலும் விருத்தி செய்துகொள்வது என எதிர்பார்க்கப்படுகிறது. இது தேசிய ரீதியில் இலங்கைக்கு பல்வேறு அனுகூலங்களை பெற்றுத்தரும். வருடாந்தம் வெளிநாடுகளில் இருந்து மென்பொருள் அனுமதிப் பத்திரங்களுக்காக கோடிக்கணக்கான வெளிநாட்டு செலாவணி பயன்படுத்தப்படுகிறது. திறந்த மூல பயன்பாட்டின் மூலம் இந்த வெளிநாட்டு நாணய செலவினம் கணிசமான அளவு குறைவடையும் என எதிர்பார்க்கப்படுகிறது.

இதுவரை காலமும் கணினி துறைசார் மென்பொருட்களை தமிழில் மொழிப் பெயர்க்கும் பொழுது அதற்கான கூடுதல் செலவாக்கையும் ஆதிக்கத்தையும் இந்திய தமிழ் மொழிபெயர்ப்புச் சங்கங்களே கொண்டிருந்தன. அந்த வகையில் Ubuntu இயங்குதளத்திற்கான தமிழ் மொழிப் பெயர்ப்புச் செயற்பாடுகளையும் இந்திய கணினி தமிழ் குழுவொன்று மேற்கொண்டது. ஆனால், அந்த மொழிப் பெயர்ப்புக்கள் முழுமையாக பூர்த்தியடையவில்லை. மேலும் அவை இந்திய தமிழ் மொழி வழக்கில் காணப்பட்டன. இந்திய தமிழ் மொழி நடைக்கும் இலங்கை தமிழ் மொழி நடைக்கும் இடையில் சொற்கள், வாக்கிய அமைப்புக்கள், ஒத்த கருத்துச் சொற்கள் உட்பட பல விடயங்களில்

வேறுபாடுகள் காணப்படுகின்றன. எனவே, இவற்றை கருத்திற் கொண்டு இலங்கை தமிழ் எனும் இலங்கைக்கே உரித்தான தமிழ் மொழி வழக்கில் மொழிப் பெயர்ப்பதற்கு முயற்சிகள் மேற்கொள்ளப்பட்டன. இதற்கு ஏற்கனவே இந்தியாவில் உள்ள சர்வதேச அங்கீகாரம் பெற்ற Ubuntu மொழிப் பெயர்ப்பாளர்கள் எதிர்ப்பு தெரிவித்ததால் இதனை தீர்த்துக் கொள்ள Ubuntu இயங்குதளத்தின் மொழிப் பெயர்ப்பு தொடர்பான சர்வதேச அங்கீகாரம் மற்றும் அதிகாரங்களைப் பெற்ற திரு.டேவிட் பெனலாவை (David Panela) தொடர்பு கொள்ளவேண்டிய நிலை ஏற்பட்டது. தமிழ் மொழிப் பெயர்ப்புக்கு TAMIL (IN) எனும் இந்தியக் குழு இருப்பதனால் அந்தக் குழுவோடு இணைந்து செயல்படும்படி இலங்கைக் குழுவுக்கு வேண்டுகோள் விடுக்கப்பட்டது. ஆங்கிலம் ஒரு மொழியாக இருந்தாலும் English (UK), English (US) என்று இரு வேறு நாடுகளுக்கு ஆங்கில மொழிகள் இருக்கும் பொழுது, தமிழுக்கு ஏன் இருக்கக்கூடாது என்ற கேள்வி இலங்கைக் குழுவிடமிருந்து எழுப்பப்பட்டது. மேலும் சில சொற்கள் மற்றும் வசனங்களும் இங்கே ஆதாரமாக காட்டப்பட்டன. இதன் அடிப்படையில் Ubuntu, firefox, Thunderbird, Open Office உட்பட எதிர்காலத்தில் வெளிவரும் அனைத்து ஆங்கில மொழியில் அமைந்த மென்பொருட்களுக்கான தமிழ் மொழிப் பெயர்ப்புக்கள் கணினி இயந்திரங்களில் TAMIL (IN), TAMIL (SL) என்ற பிரிவுகளில் வெளியாகும். இதில் TAMIL (SL) என்பது இலங்கைக்கே உரித்தான தமிழ் மொழிப் பெயர்ப்பாகும்.

தேசிய மொழிகளை கணினியில் பயன்படுத்துவது குறித்து இலங்கையில் கடந்த ஏழு வருடங்களாக பல ஆக்கப்பூர்வ ஆய்வுகள் மேற்கொள்ளப்பட்டன. தேசிய மொழிகள் அல்லது உள்ளூர் மொழிகளாக சிங்களமும், தமிழும் காணப்படுகின்றன. இந்த மொழிகளை கணினி விசைப் பலகையின் ஊடாக பயன்படுத்தல், அதற்கான விசைப் பலகை வரைபட உருவாக்கம், இலங்கை கட்டளைகள் நிலையத்தின் விசைப் பலகை சட்ட விதிமுறைகள், இலங்கைக்கே உரித்தான நியம எழுத்துருக்கள், பிரபல மென்பொருட்கள் மற்றும் இயங்குதளங்களுக்கான சிங்கள மற்றும் தமிழ் மொழிப் பெயர்ப்புக்கள் உட்பட பல்வேறு கணினித்துறை சார்ந்த விடயங்கள் ஆராயப்பட்டன. இந்த விடயங்களை ஆராய்ந்து அவற்றுக்கு உரிய தீர்வுகளை வழங்கவென, இலங்கை தகவல் தொடர்பாடல் முகவராண்மை நிலையம் உள்ளூர் மொழி செயல்பாட்டுக் குழுவொன்றை ஏழு வருடங்களுக்கு முன்னர் உருவாக்கியது. அந்தக் குழு கடந்த ஏழு வருடங்களில் பல்வேறு ஆய்வுகளை மேற்கொண்டு தீர்வுகளையும் முன்வைத்துள்ளது. இதன் அடிப்படையில் உள்ளூர் மொழிகளை கணினியில் சிறந்த முறையில் பயன்படுத்துவதற்கென சில தயாரிப்புக்களும் உருவாக்கப்பட்டுள்ளன. இந்த முயற்சிகளுக்கு சில அரச தரப்பு அமைப்புக்களும் உதவி செய்தன.

தகவல் தொழிநுட்ப முகவராண்மை நிலைய தமிழ் நியம எழுத்துரு - ஸ்ரீ தமிழ் என்ற பெயரில் இந்த குழுவும் வெளியிடப்பட்டது.

Ubuntu தமிழாக்கத்தினால் இலங்கை வாழ் தமிழ் சமூகம் பெற்றுக்கொள்ளும் நன்மைகள்:

1. தமிழில் கணினி செயல்பாடுகளை சிறந்த முறையில் கற்றுக்கொள்வதற்கான வாய்ப்பு மேலும் அதிகரித்துள்ளது.
2. அதுவும் பிரபல திறந்த மூல தயாரிப்புக்களான Ubuntu OS, Firefox, Thunderbird, Open Office போன்ற மென்பொருட்களை இலகு தமிழில் புரிந்துக்கொள்ளக் கூடிய வாய்ப்பு ஏற்பட்டுள்ளது.
3. Microsoft தயாரிப்புக்களுக்கு இணையான பல தயாரிப்புக்கள் திறந்த மூலங்களில் வெளிவந்துள்ளன.செலவு குறைந்த அல்லது இலவச தயாரிப்புக்களாகவே திறந்த மூலங்கள் காணப்படுகின்றன.இவற்றை இலங்கைக்கே உரித்தான தமிழில் கற்றுக்கொள்ள கூடிய வாய்ப்பு இலங்கையர்களுக்கு கிடைத்துள்ளது.
4. இந்த தமிழ் மொழியாக்கம் அல்லது மொழிபெயர்ப்பை மேற்கொள்ள 'இலங்கைக்கே உரித்தான சர்வதேச அங்கீகாரம் பெற்ற மொழிபெயர்ப்பாளர் சங்கம்' உருவாக்கப்பட்டுள்ளது.
“Ubuntu Srilanka Tamil Translators” team <https://launchpad.net/~ubuntu-l10n-ta-lk>
5. இந்த சங்கத்தின் மூலம் இந்த மொழிபெயர்ப்பு செயல்பாடுகள் தடையின்றி இடம்பெற்று வருகின்றன.தேவைப்படும்பட்சத்தில் புதிய சொற்கள் உரிய ஆதாரங்களுடன் உருவாக்கப்படுகின்றன.
6. துறைசார் விடயங்களுடன் தொடர்புடைய பலர் இந்த சங்கத்தில் இணைந்துகொள்ள சந்தர்ப்பம் உண்டு.
7. இதனுடாக சமூகத்தில் இந்த செயல்பாட்டிற்கான பங்களிப்புகள் பலர் உருவாகி வருகின்றனர்.இதனால் சமூக விழிப்புணர்வு ஏற்படுகிறது.
8. இந்த மொழிபெயர்ப்பு செயல்பாடு முடிவில்லா செயல்பாடாக அமையவுள்ளது.எதிர்காலத்தில் இது போன்ற மேலும் பல திறந்த மூலங்களை தமிழாக்கம் செய்யக்கூடிய வாய்ப்பு ஏற்படும் என்று நாம் எதிர்பார்க்கிறோம்.
9. இலங்கை தமிழாக்க பாரம்பரியத்திற்கு அமைய, மொழிபெயர்ப்புக் கொள்கை விதிமுறை ஒன்றும் ஏற்கெனவே உருவாக்கப்பட்டுள்ளது.

https://docs.google.com/document/pub?id=16f-1Og_Y5luzFA0YC7W-BPx4E_wulTEJlmupDGGeHKHI

10. மொழிபெயர்ப்பு விடயத்தில் கூடுதல் கவனம் செலுத்தப்பட்டுள்ளது.ஏற்கெனவே உருவாக்கப்பட்ட கொள்கை விதிமுறை கோவைக்கிணங்க, மொழிபெயர்ப்பு மீளாய்வு திட்டமும் அமுல்படுத்தப்பட்டுள்ளது.
11. **குறிப்பு :** இங்கு பல இந்திய சொற்கள் உள்வாங்கப்பட்டாலும், அமுல்படுத்தல் வழிகாட்டல் ஆவணத்திற்கிணங்க இலங்கை தமிழ்சொற்களும் நடைமுறைகளும் தெளிவாக காட்டப்பட்டுள்ளன.
12. சில சொல் பரிமாற்றங்களையும் உதாரணங்களாக நான் இங்கே தந்துள்ளேன்.

ஆங்கில சொல்	இந்திய நடைமுறை	இலங்கை நடைமுறை (கொள்கை ஆவணத்திற்கிணங்க)
Yahoo!Mail	யாஹீ!மெயில்	Yahoo! மின்னஞ்சல்
Microsoft	மைக்ரோசொவ்ட்	Microsoft
Oracle	ஓராக்கிள்	Oracle
Yahoo	யாஹீ	Yahoo
Google	கூகிள்	Google
Gmail	ஜிமெயில்	Gmail
Firefox	பயர்பொக்ஸ்	Firefox
GIMP	கிம்ப்	GIMP
Debug		பிழை நீக்கு/வழு நீக்கு
Dialogs		உரையாடல்கள்/ உரையாடல் திரைகள்
Proxy	புரொக்ஸி	Proxy
Boot Option	துவக்க தெரிவு	இயக்க தெரிவு
Start	துவக்கு	ஆரம்பி
Main Server	மெயின் சர்வர்	முக்கிய சேவையகம்
Ubuntu	உபுண்டு	Ubuntu

இணையத்தள ஆதாரங்கள் (https://translations.launchpad.net/ubuntu/lucid/+lang/ta_LK)

1 — 75 of 1278 results

Template Name	Length	Status	Untranslated	Need review	Changed	Last Edited	By
bootloader	76		—	—	—	2011-08-30	Yogaraj
debian-installer	1614		179	—	—	2011-09-02	Jeyaranjan Yogaraj
debian-installer-help	177		163	—	—	2011-12-22	Jeyaranjan Yogaraj
ubiquity-debconf	217		128	—	—	2011-09-10	Janakan
ubiquity-desktop	3		—	—	—	2011-08-13	Jeyaranjan Yogaraj
ubiquity-slideshow-kubuntu	48		48	—	—	—	—
ubiquity-slideshow-ubuntu	49		45	—	—	2011-08-17	Jeyaranjan Yogaraj
ubiquity-slideshow-ubuntu-upgrade	3		—	—	—	2011-08-18	Jeyaranjan Yogaraj
ubiquity-slideshow-xubuntu	43		29	—	—	2011-08-27	Janakan
indicator-me	8		—	—	—	2011-08-16	Jeyaranjan Yogaraj
indicator-session	36		8	—	—	2011-09-01	Yogaraj
indicator-sound	5		—	—	—	2011-08-18	Jeyaranjan Yogaraj
rhythmbox-ubuntuone-music-store	11		—	—	—	2011-08-25	Janakan
indicator-messages	9		—	—	—	2011-08-25	Jeyaranjan Yogaraj
indicator-applet	16		—	—	—	2011-08-25	Janakan
language-selector	98		98	—	—	—	—

‘Raspberry PI’ Future Outlook for Tamil

Saravanan Mariappan

Nexus NGN Sdn. Bhd., Malaysia. Email: nexusngn@gmail.com

Abstract. With the fast developing IT World and broadband penetration expanding rapidly from urban to rural, drastic economical, financial and educational improvement within the community were observed in developed nation. Hence many government institution have dedicated arm focusing mainly on the development of broadband facilities in rural area. With the made available broadband facilities, further challengers lays ahead on how to disseminate information and knowledge. Basic computing tools at low cost can be made available to the rural communities, with the introduction of mini computer “Raspberry PI” information and opportunity to be online can be made quick and cost effective. This paper addresses how Raspberry PI can contribute to the development of Tamil Community and Tamil Language. The mini computer can be used to develop teaching tools such as Mini USB Projector and simple Android Mobiles with internet access. It can also be used to be integrated with server as thin clients bringing the cost to implement ICT labs at remote rural sites cost effective. This paper will also discuss the current influence of the mini computer “Raspberry PI” in the Tamil community and how Tamil Language and the community can develop in the future.

Keywords : Raspberry PI, Thin Client, ICT Lab, Open Source

1.0 INTRODUCTION

Raspberry PI is a credit-card-sized single-board computer developed in United Kingdom by Raspberry PI foundation with the intention of promoting the teaching of basic computer. Since the development and touch of Raspberry Pi mini computer in 29th February 2012 the sales of Raspberry PI reached beyond 1million units. The used of this mini computer have gone into various industries, replacing expansive programmable controls. Creating new era of programmable device development. Some of the industries that uses Raspberry PI are:

- Nuclear industries
- Real-time monitoring devices with web interfaces
- Cloud storage and management
- Industries depend on Thin Client
- Home automation and monitoring
- Laptops
- File sharing portal
- Attach to camera and take picture on command.
- Digital signature
- Personal computers and many more

What interest us for the development of Tamil is the usage of Raspberry PI for teaching and the usage of broadband. For this purpose we can use Raspberry PI as personal computer and Thin client.

2.0 RASPBERRY PI REQUIREMENT FOR TEACHING

The basic requirement for affordable teaching devices are that the device should be able to operate using open source operating system and able to work on LAN.

Some of the criteria of open source operating system and software are:

- ⤴ Licenses are not restricted to any party
- ⤴ Licenses are allowed to be modified
- ⤴ Modifications are allowed by adding on “patch files”
- ⤴ No discrimination against person or groups
- ⤴ Redistribution of program can be done without the requirement for additional licenses
- ⤴ License must not be specific to a product
- ⤴ License must not restrict other software- it's not required for other programs to be open source.

The type of open source programs that Raspberry PI should able to execute are:

1. Open office



5. Video conversion



Media Corder

6. Video player

2. Web browsing



Miro



Miro

3. Instant messaging



Pidgin

7. Graphics editing



GIMP



Paint.net



Inkscape

4. Video playback



VLC

8. Email



Mozilla Thunderbird

9. Sound recording



Audacity

3.0 RASPBERRY PI AS THIN CLIENT

An ideal thin client is the computer hardware without hard drive and well equipped with dedicated mini motherboard which depends on the central server for processing data and connects to the server through network to run application, access files, print, and perform all services available in on ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.



Plate 1 : Raspberry PI as Thin Client.

A cost-effective similar solution can be provided by using Raspberry PI as thin client. It can be supported by open source edubuntu server platform or LTSP server and run applications on central server.

The Raspberry PI that operates as thin clients are called Berry Terminals. The Berry Terminals linux base, designed to turn Raspberry PI mini computer into a low-cost thin clients. Figure 1 shows the working principal of Raspberry PI as thin client.

3.0 RASPBERRY PI AS THIN CLIENT

An ideal thin client is the computer hardware without hard drive and well equipped with dedicated mini motherboard which depends on the central server for processing data and connects to the server through network to run application, access files, print, and perform all services available in an ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.



Plate 2: Centralized Raid Server

4.0 ADVANTAGES OF OPEN SOURCE COMPUTING SYSTEM

The thin client system in school computer labs are capable of providing affordable server-based open source computing solution. The main advantages of using this system in schools are to increase reliability and consistency of technology over time. Through a password-accessible accounts, students, teachers, and administrators can store and can access saved documents and personal settings. Since all files and programs are stored centrally, users can access their work from any computers within the network.

Eg. when a teacher or student “logs in”, the server provides them with their “desktop configuration”. Users can even access their “desktop” from home or other remote location.

The other advantages of open source computing with thin client system are as below:

- ▲ less administration – central management of users, patches, software, data and backups.
- ▲ Higher security – elimination of viruses, Trojans or other vulnerabilities on the user desktops.
- ▲ Hardware independence – support of virtually all client devices and computer hardware and very low system

requirement.

- ⤴ Easy access – teacher and students can access their documents and applications from any PCs in the local area network.
- ⤴ Reduction in TCO – Total Cost of Ownership reduction by up to 50%.

Benefits for school in using open source computing thin client system:

- ⤴ Lowers cost of technology over time
- ⤴ Secure data and equipment
- ⤴ Less downtime and greater efficiency
- ⤴ Reduces administrator staffing costs
- ⤴ Lessen the risk of data theft
- ⤴ Disaster recovery: Data is more secure
- ⤴ Reduced time for technical support
- ⤴ Lower power consumption: save electricity
- ⤴ Zero licensing management
- ⤴ Minimum maintenance
- ⤴ Highly trained individuals are not required

At present we have successfully done about 70 ICT labs in Tamil schools in Malaysia using thin client and server base computing.

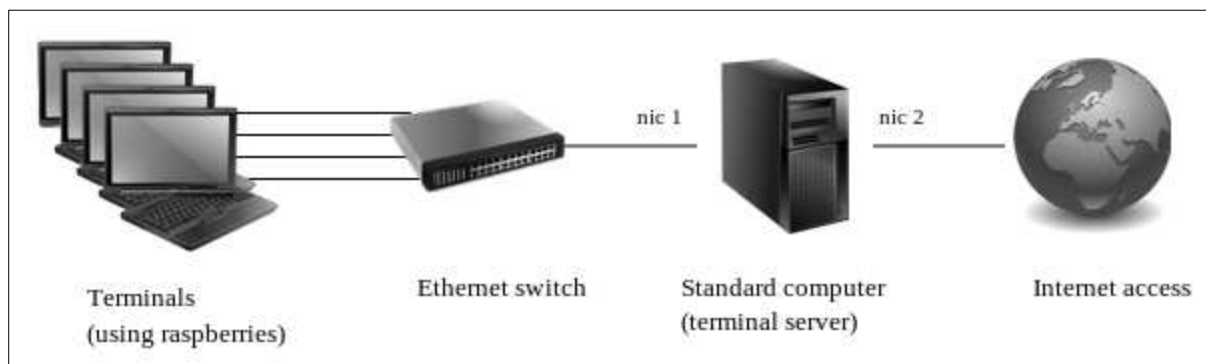


Figure 1.0 shows the working principal of Raspberry PI as thin client.

The labs at present uses refurbish computers of Pentium 3 and 4 or above. These systems are venerable to breakdown without any signs of defects due to aged motherboard, cooling fan, power supply, and other moving parts. Hence with the introduction of Raspberry PI the longevity and warranty period of the thin client can be extended beyond two to three years as Raspberry PI do not have any moving parts.

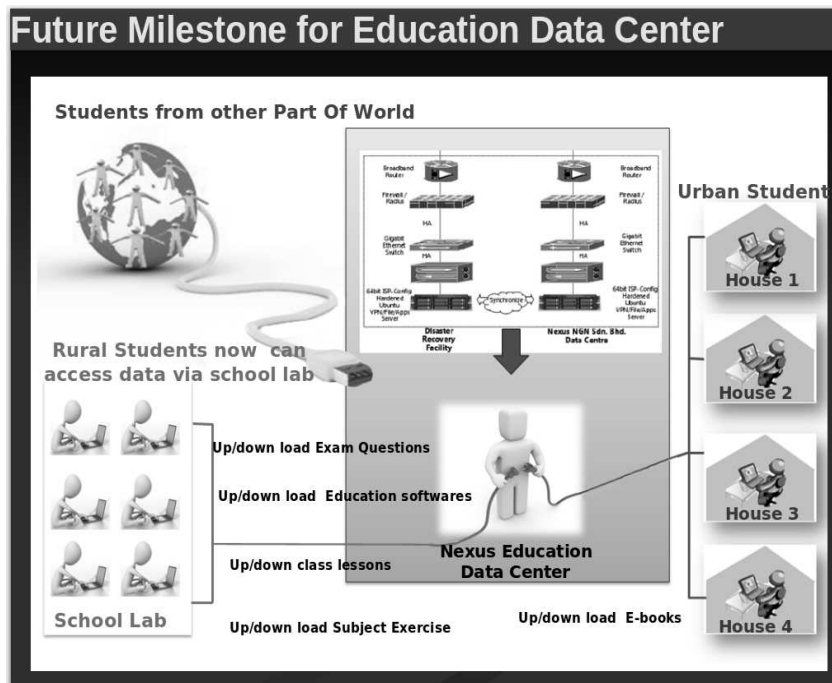
5.0 CONCLUSION

The open source thin client computing environment currently implemented in Malaysian Tamil schools have been well received with 70 over schools been teaching ICT and Tamil since 2009. With the use of Raspberry PI as thin client, the performance of the lab can be enhanced further. The cost to implement the lab can also be reduced substantially. The server based open source technology making it easier for students and teachers to organize and access education materials and multimedia content. At the same time the thin client network enables the schools to use this technology more effectively to monitor students work. With open source applications and thin client technology it was possible to decrease the cost of installation and the cost of maintaining the computer lab. With servers installed at each and every schools, the setting up of centralized school management system is now possible. With centralized school management system in place, teaching and training materials can be deployed to every schools at realtime. This will enhance greater sharing of knowledge between schools, universities, individuals and even between countries, pushing the envelop to further height for the growth of Tamil language and Tamil community.

Plate 3 : Tamil school students using open source Thin Client in an ICT lab.



Figure 2 : Shows the future of Thin Client and centralized management allowing students and schools to have full access to centralized server from anywhere at anytime.



TEDU-Tamil Educational Cloud for anytime anywhere Tamil Learning

G.Malini, Research scholar, malinigopalan@gmail.com

Dr.T.Mala, Assistant professor[S.G], malanehru@annauniv.edu

AnnaUniversity, Chennai, TamilNadu, India

Keywords: Cloud Computing- E learning services- Educational cloud.

Cloud computing is a computing model based on networks and on the Internet, which is highly scalable and creates virtualized resources that can be made available to users. Cloud also provides various services to users which can be accessed by them anytime and anywhere with high availability. With the advent of cloud based E-learning services an entire world of knowledge can now be made available to teachers and students that can be accessed anytime, anywhere, from any device. An educational cloud thereby helps an institution to make students learn their lessons and to access the services provided by the cloud. Therefore in this paper Tamil Education Cloud [TEDU-cloud] has been proposed and implemented for easy any time any where Tamil learning for students. TEDU-cloud brings learning and teaching Tamil together in single environment. It gives tools needed for the teacher to teach and also provide the students the resources for learning in a virtualized environment. It is the combination of education and cloud technology with ultimate twenty first century tools combined to support teaching and learning to the next level. The various services offered by TEDU-cloud are:

- EaaS (Education as a Service)-Using this service students can learn their lessons, submit assignments and write online tests in a shared environment.
- SaaS (Software as a Service) – Using this service students can access various software tools and compilers that are deployed in the cloud from anywhere and at anytime to run their applications.
- StaaS (Storage as a Service) – Students can use this service to personalize their contents and store them on the cloud.

The TEDU-cloud comprises of Tamil e-content materials such as online e-books, learning materials, assignments, lecture videos, online assignments and tests which are placed as e-learning resources and they can be accessed online anytime anywhere using any output device. Students are also allowed to submit their assignments and they can also store their e-materials on TEDU-cloud. Students can also use the tools that are available in TEDU-cloud to develop their own application and to personalize their materials. TEDU -cloud can be accessible even by the mobile devices, tablets, smart phones etc. Since cloud has the capability of high availability and fault tolerance with replicated content management TEDU-cloud uses the above capabilities to provide Tamil content immediately to the students with less delay and no loss. Scalability and elasticity features of TEDU-cloud helps in storing huge amount of Tamil E-content materials on the cloud and any number of students can simultaneously access TEDU-cloud. Thereby TEDU-cloud acts as anytime any where any device accessible Tamil e learning cloud for students and staff.

1. INTRODUCTION

The popularity of learning on the internet, the construction of perfect web-based learning environment has become one of the hot points on researching remote education. The massive proliferation of Internet broadband connectivity and rich education content[4] has created a global phenomenon in which information and communication technology is being used to transform education. Therefore, there is a need to redesign the educational system to meet the needs better.

It is envisioned that, in the near future, cloud computing [2]will have a significant impact on the educational and learning environment, enabling their own users (i.e., learners, instructors, and administrators) to perform their tasks effectively with less cost by utilizing the available cloud-based applications offered by the cloud based educational service providers.

The development of cloud based e-learning system allows the implementation of scalable, versatile, and customized e-Learning[6] systems, constructed on-demand. This allows more efficient use of computing resources, improving the revenue of the system and enhancing the Quality of Service (QoS) received by users while minimizing the power consumption of the machines .Some of the advantage of e-Learning on cloud are discussed below:

- Cost of Materials

Because the cloud allows all sorts of streamlining options, many educators and students could spend less on various materials[1]. A course instructor could upload all readings to the cloud, which would thus negate the need for an expensive textbook or for excessive printing costs. Students would only have to worry about the cost of a computer monitor, which they could use across all courses. The long term return on a computer is much better than on a set of textbooks or course packets that they'll only use once.

- Resource free

Devices with minimal hardware requirements (mobile phones, for example) could be successfully used as cloud clients. There is no need of heavy processing capabilities at the client side. Therefore the client is a thin client with minimal resources.

- Scalability and elasticity

Dynamic[5] "on-demand" provisioning of resources on a fine-grained, self-service basis in real-time, without users having to engineer for peak loads, thus paving way for availability.

- Device and location independence

Enable users to access systems using a web browser regardless of their location or what device they are using (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect from anywhere.

This paper implement an TEDU-cloud. TEDU-cloud brings learning and teaching together in single environment. It gives tools needed for the teacher to teach and also provides the students the resources for learning purposes in a virtualized environment. It is the combination of educators, cloud technology and the ultimate twenty first century tools combined to support teaching and learning to the next level.

The features of our TEDU-cloud comprise of online e-books, learning materials, assignments, lecture videos, online assignments and tests. The concept of cloud plays an important role here because of its significance in security and reliability. TEDU-cloud providing the following services.

EaaS (Education as a service)-Using this service students can learn their lessons, submit assignments and write online tests in a shared environment.

SaaS (Software as a Service) – Using this service students can access various software tools and compilers that are deployed in the cloud from anywhere and at anytime.

StaaS (Storage as a Service) – Students can use this cloud by personalizing their contents and storing them.

2. RELATED WORKS

Bo Dong et al describes E-Learning ecosystem[3] is facing challenges an optimizing resource allocations, dealing with dynamic concurrency demands, handling rapid storage growth requirements and cost controlling that due to huge growth of users, services education contents and resources[9]. The best solution for this is using new technologies such as integration between cloud computing and web 2.0.

Chun-C hia Wang et al describes a independent Learning Management System embedded in different E-Learning standards[7] to share the learning objects. Huge amount of learning objects can be accessed reliably, flexibly over the Internet by means of Cloud computing technology.

Luis M. Vaquero et al describes an Assessment Management System (AMS) which handles the student's testing process in order to evaluate their knowledge capabilities. It has Item Classification Module (ICM) which consists of the item classification algorithm and decision tree. This module accesses the item bank on the cloud database and then it generates the item-set pattern. The ICM module is also developed as a service thus it is called Question-as-a-Service (QaaS). In the next session detailed architecture of the TEdu cloud system has been described.

3. System Architecture

The architecture of the system is described in terms of its components and their functionalities. Figure 1 shows the overall system architecture of Educational cloud

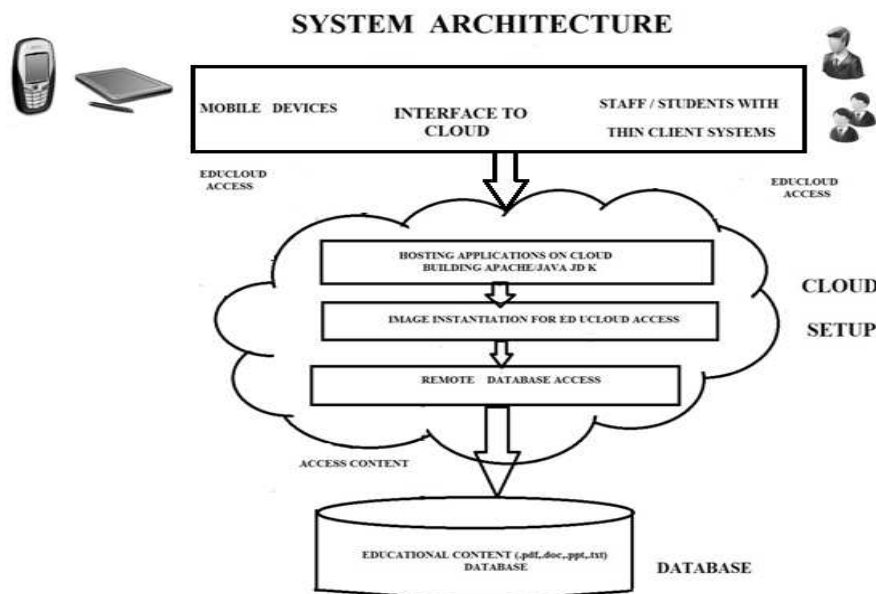


Figure1

The system consists of various modules namely setting of TEducloud, creating educational content and hosting & accessing applications on TEducloud, With various application functionalities. The various modules of the implemented system are explained in the following sessions.

3.1 Creating Tamil Educational content

The various educational contents needed for the system is collected. This ranges from txt, rtf, pdf, ppt, to html files. Various audio and video files are also included which make the learning experience effective and interactive. These are in various formats such as mp3, wav, mp4 and so on. This system is aimed at catering the needs of students for basic computer sciences courses. Hence the educational contents are also collected for the same. Educational contents are stored in database.

The content that have been collected is then stored inside the mysql database. This storage supports for large variety of files in the binary form. For easy retrieval purposes, the audio and video lectures are stored inside the cloud itself. Various operations such as insertion, deletion, updation of content can be performed on the database through SQL queries by starting mysql service.

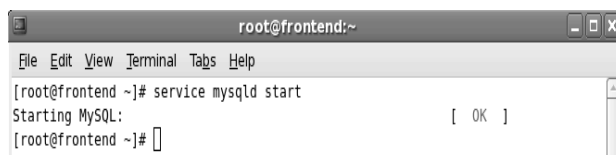


Figure2 Mysql database

Figure2 shows the educational contents are effectively stored mysql database. cloud setup discussed next session.

3.2 Tedu Cloud setup

Tedu setup consists of three setup :step1 Hosting \$Accessing application on eucalyptus cloud, STEP2 Instantiation of virtual machine of the machine of the image, STEP 3 Remote access to e-Learning content database.

3.2.1 Hosting & Accessing application on eucalyptus cloud

The eucalyptus private cloud is installed in the system, which consists of the cloud controller, cluster controller and the node controller. Xen is also installed in the system for creating the root filesystem and image of the guest OS.

A Eucalyptus image bundled with the necessary software such as apache software (PHP framework), JAVA JDK etc. is created. The PHP files and the video lectures are uploaded into the image. It is then uploaded and registered into eucalyptus cloud. Figure3 shows VM (Virtual Machine) for the image can be created using the euca-commands.

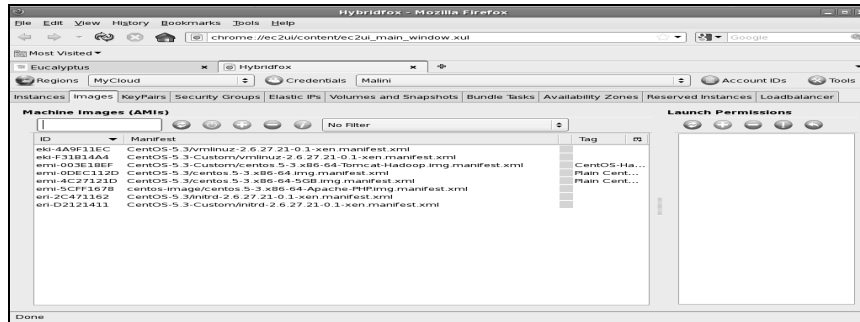


Figure3 Euca -commands

3.2.2 Instantiation of virtual machine of the machine of the image

Figure 4 shows the VM is instantiated as follows:

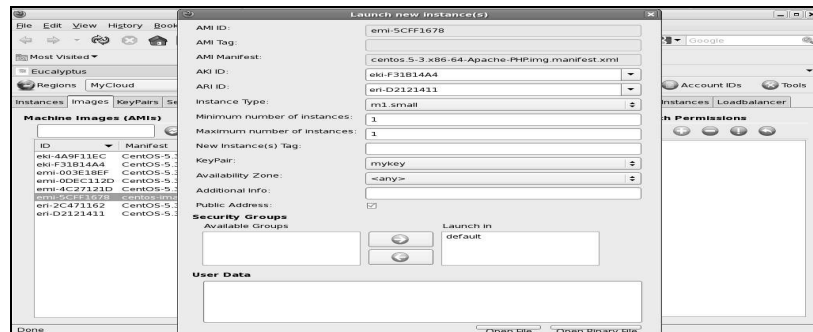


Figure 4 Instantiation of VM from image

The IP address of the Virtual Machine is obtained, using which we ssh into the image by providing the security key. Figure5 provide security key.



Figure 5 Login into the virtual machine

Figure 6 shows the Apache service is then started on the cloud as follows:

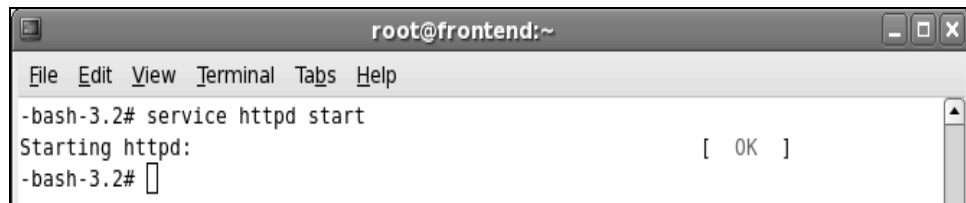


Figure 6 Apache Service start

3.4 Remote access to e-Learning content database

MySQL database which is running on the system should be accessible by the Virtual machine of the E-learning system image on the cloud. This requirement arises in order to access the educational content stored in the database.

By using the following command, the VM would be able to gain access over the tables of the database.

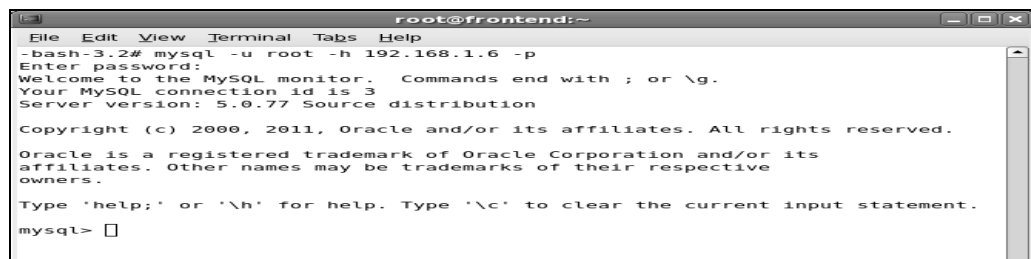


Figure 7 MySQL Remote Connection

Various application functionalities are explained the next session

3.5 Services provided on the cloud

The various services provided on the cloud as a result is as follows.

3.5.1 Education as a service (EaaS)

TEducloud provides education as a service to alternative classroom teaching and aids virtualized and distant learning.

3.5.1.1 Upload Assignment topics

The assignments that need to be completed for each course can be uploaded by the staff members, which can be accessed by the students. These are stored as files in the database.



Figure8 Upload Assignments

3.5.1.2 Review Students' Assignments

The assignments that are submitted online by the students can be reviewed by the instructors and awarded marks for the same. In this way the student-instructor interaction can be made more effective through this e-learning framework



Figure9 Review assignments

3.5.1.3 Upload and Manage Study Materials

The instructor can upload the study materials which can be in the form of text, rich text format, pdf, power point presentations files. Audio and video files can also be uploaded to make e-learning experience more friendly, interactive and personalized.

3.5.1.4 Online tests for students

The Instructors are given a provision to upload test questions of a given course and the topic. The students can take online tests of a particular topic of a course. The answers are then evaluated and they are awarded marks for the same.

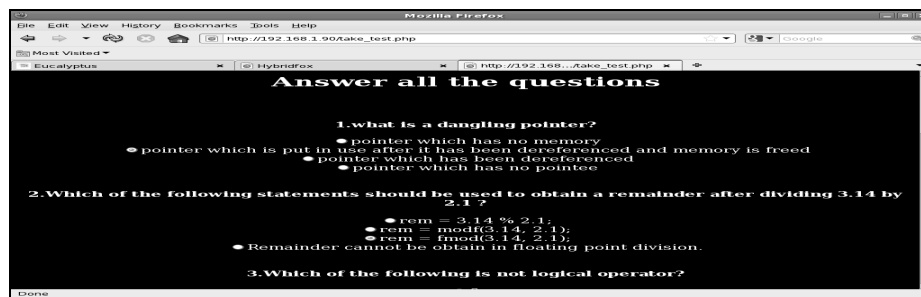


Figure 10 Online tests

Thus the above comprises of the functionalities of the instructors and the students of the e-learning environment which also helps in distance learning. It provides a complete package for effective e-learning.

3.5.2 Storage as a Service (StaaS)

The educational content and video lectures are stored in the image of the cloud, thus providing storage as a service on the cloud.

3.5.2.1 Student login

A level of personalization has been achieved by providing a separate login for each student. Accounts can be deleted and created as per the requirement the e-learning application is running successfully at the IP address of the Virtual Machine. Hence the instructors and students will be able to access the e-learning environment using the IP address of the VM. Figure 11&Figure12 provide separate student login.

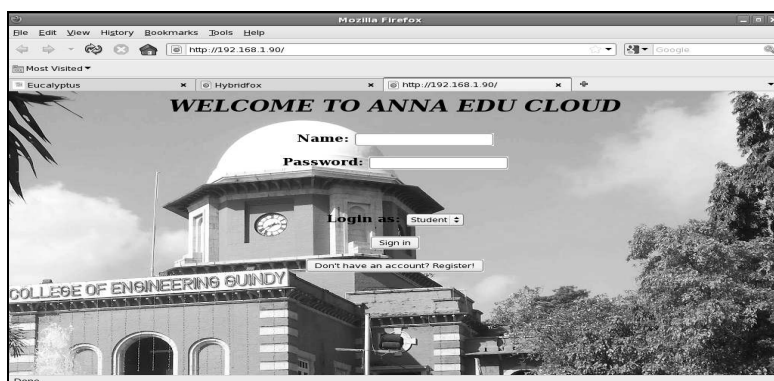


Figure11 Student login

Here, two case studies have been considered. One is the Micromax Funbook tablet and other is Samsung Galaxy Android phone. The mobile devices are made to be in the same network as that of the cloud machine such that each one is accessible by another. The e-learning framework is made available in the devices using the IP address of the VM of the image in the browsers.

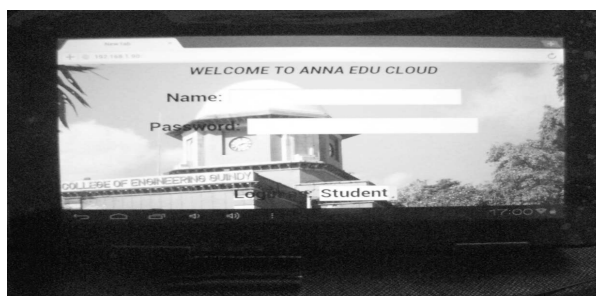


Figure 12 Student login

3.5.2.2 Access Study Materials

The students can view and download the study materials uploaded by the instructors. This makes learning experience more effective and interactive. The video lectures can be played in the browsers of the students in the e-learning environment through PHP framework.

3.5.3 Software as a Service (SaaS)

The JDK for JAVA installed in the image of the cloud helps a student to use a compiler to compile java programs though it may not be available in his/her own system, thus providing software as a service. The various conclusions and observations of this system and further areas of research and extensions are discussed in the next session.

4. Conclusion and future work

Thus a prototype of the educational cloud for the entire university has been successfully developed, thus enabling a virtualized e-learning environment which is an alternative cost-effective method to the conventional method. Various services such as Infrastructure, Education, Storage and Platform as Services have been made available with the help of the cloud. Support for mobile devices has also been established. The resources available in the cloud are thus sufficient to cater the educational needs of a thin client available. The system can further be enhanced by adding more features such as creation of learning objects for reusability. Walrus in Eucalyptus cloud can be used for storage of the learning objects.

Faster retrieval of educational content by efficient search of the database can be made possible. One could also make the system fully efficient by providing all possible services which can be availed through the technology of virtualization and cloud computing.

REFERENCES

- [1] Manop Phankokkruad, "Implementation of Cloud Computing for e-Learning System", Proceedings of International Conference on Computer & Information Science (ICCIS), 2012.
- [2] Guang Li, Guisheng Chen, "A Novel Enhanced Education Application Of Cloud Computing", Proceedings of IEEE CCIS2011.
- [3] Mohssen M. Alabbadi, "Cloud Computing for Education and Learning: Education and Learning as a Service (ELaaS)", 14th International Conference on Interactive Collaborative Learning (ICL2011).
- [4] Mohammed Al-Zoube, "E-Learning on the Cloud", International Arab Journal of e-Technology, Vol. 1, No. 2, June 2009.
- [5] Hui Ma, Zhongmei Zheng et al, "The Applied Research of Cloud Computing in the Construction of Collaborative Learning Platform under E-learning Environment", International Conference on System Science, Engineering Design and Manufacturing Informatization, 2010.
- [6] Luis M. Vaquero, "EduCloud: PaaS versus IaaS Cloud Usage for an Advanced Computer Science Course", IEEE Transactions On Education, Vol. 54, No. 4, November 2011.
- [7] Benjamin Hirsch, Jason W.P. Ng, "Education Beyond the Cloud: Anytime-anywhere learning in a smart campus environment", 6th International Conference on Internet Technology and Secured Transactions, December 2011.
- [8] Shinichiro Kibe, Minoru Uehara, Motori Yamagiwa, "Evaluation of Bottlenecks in an Educational Cloud Environment", Third International Conference on Intelligent Networking and Collaborative Systems, 2011.
- [9] Malini.G, Mala.T, "Semantic based for e-Learning content storage on cloud", Proceedings of IEEE International Conference on Computer Communication advanced network, 2013.

Challenges in Online learning towards better learning environment: A Case study on Bachelor of Business Management (BBM) Online Degree Programme at University of Jaffna, Sri Lanka

Rathirane, Y

Senior Lecturer, Department of Financial Management, University of Jaffna, Sri Lanka. rathi.yogen@yahoo.com

ABSTRACT

Sri Lanka is one of the countries to provide free education. The number of higher education opportunities is available to the students. However, a student still has to consider various factors and face many challenges, such as the lack of seats in national universities. The total number of Universities in Sri Lanka is fifteen (15) and nine (9) other Institutions that come under the University Grants Commission (UGC). Nearly 100, 000 students qualify each year to enter the Universities in Sri Lanka out of this around fifteen (15) percentage gain admission to follow the degree programmes. Only a small number among the left outs manage to continue studies at the tertiary level while majority of them are left with no future prospects. Even, the opportunity available to do the higher studies externally is negligible. The government has become increasingly willing to introduce online learning to deliver cost effective, easily accessible and current education to all ages and social backgrounds, on this regards of time, cost and geographical constraints in recent years. In the modern time of education, the rapid growth of Information and Communication Technology (ICT) leads to tremendously dynamic changes in a variety of tools and technologies applied to the learning process. Online learning has become an omnipresent concept in the time of ICT education and has also been identified as a major mode of learning where learners are able to select the occasion, location and the phase of learning. The University of Jaffna offering to contribute to the development of higher studies in its own way has initiated steps towards the provision of opportunities for higher studies at least to a small section, in the Tamil medium and in this regards the Bachelor of Business Management (BBM) Online Degree Programme has been introduced as an online mode delivery system. Now three batches are going on and around 300 students have been enrolled for this programme but statistics show that drop out rates from online learning courses are much higher than from traditional classroom based courses. In this paper it is argued that one reason for this is that the introduction of online learning and a more student-centred learning model involves a drastic shift for students who are brought up in very teacher-centred didactic educational cultures. This paper mainly focuses on challenges that were come across during the design, development and delivery stages of the BBM Online Degree programme at University of Jaffna, Sri Lanka. The Case study approach is used as the research methodology and semi structured interviews with students and subject matter experts and observations are used as the data collection techniques. Content analysis is used as the case analysis technique and results revealed the major challenges in online learning in perceptive Student support, Flexibility, Teaching and Learning Activities, Access, Academic confidence. In this paper these challenges will be discussed and solutions suggested. It is concluded that the solutions for these challenges lead to a better learning environment.

Key words: On line learning, Challenges, Student Support, Flexibility, Teaching and Learning Activities and better learning environment.

1 BACKGROUND OF THE STUDY:

Education is one of the most important factors for poverty alleviation and economic growth in developing countries. (UNDP, 2005, UNESCO, 2005,) Further it is believed that the education helps to empower poor people in developing countries. In the modern environment of education, the growth of Information and Communication Technology (ICT) leads to tremendously dynamic changes applied to the learning process. There are a variety of tools and technologies are used in the modern education. Online learning has become an omnipresent concept in the period of ICT education and has also been identified as a major mode of learning. In Sri Lanka only 15 % are admitted to the Universities under Ministry of Higher Education out of 100, 000 students who are eligible to enter the Universities with minimum requirements for entering universities due to lack of infrastructure facilities in the Universities (Central Bank, 2006). In recent years, the government has become increasingly keen to endeavour into the possibilities of online learning to deliver cost effective, easily accessible and ever-current education to all ages and social backgrounds, regardless of time, cost and geographical constraints. This is a

door to enrol the students who are unable to enter the universities due to lack of seats. Further, it is observed that many public as well as private sector entities engaged in higher education are moving towards online learning based educational systems in Sri Lanka. Within this context, the Ministry of Education with the help of the Asian Development Bank engaged the online courses to the students. In this regards University of Jaffna introduced a Bachelor of Business Management (BBM) Degree programme as an online education mode in Jaffna University in 2009. This is an external Tamil medium and highly demanded programme within the students. However, Online learning is facing a lot of obstacles and challenges in developing countries (Heeks, 2002, Dhanarajan, 2001, Rajesh, 2003) and drop out rates are much higher compare to traditional classroom-based teaching (Simpson, 2004, O'Connor et al., 2003, Eastmond, 2000). Now three batches are going on and 300 students enrolled for the BBM Online Degree Programme in University of Jaffna and now only 150 students are following the programme. Comparing with the traditional classroom courses these dropout rates are significant. This is a challenge facing by the BBM programme and when we analyse the reasons for these challenge we can observe the challenges and difficulties facing by the students and the Subject Matter Experts (SMEs). If the challenges to be eliminated it will be lead to better learning environment and may reduce the dropout rates in future. This study focuses on challenges facing by the students and the SMEs and how that can be led into better learning environment.

However, Sri Lanka is a developing country it gives free education to all people. The admission of the Universities under the University Grants Commission (UGC) is through the results of the GCE A/L Examination, which is a national examination conducted by Ministry of Education and a competitive examination because limited number of students are getting admissions to the universities due to lack of infrastructure facilities in the Sri Lankan Universities. These students follow courses full time on-campus (internal). In addition to these internal degrees they offered external degrees also. The UGC introduced the online programmes to manage the problems regarding limited number of students are admitted to the Universities. Sri Lanka is facing some challenges when it is to incorporate online learning in their education. Interactivity practices play a prominent role and to have computer literacy as precondition to implement online learning. There are difficulties when adapting online learning in to the learning community in Sri Lanka. To implement online practices in Sri Lanka, the target participants need to have knowledge in English Language, computer literacy, technology and design of the system should simple as possible (Gunawardana, 2005)

The BBM Online Degree Programme is conducted by the University of Jaffna (UOJ). It is a three (3) year external degree programme and the students have the option of continuing or leaving the programme on the completion of each year with certification. If the student achieves above a minimum criteria in their examination, who is awarded the Certificate at the end of the 1st year and the Diploma in Business Management at the end of the 2nd year and at the end of the final year, he/she will be awarded the degree of Bachelor of Business Management. Although, it faced difficulties it is delivering courses properly throughout the semester. However, in the past years since started the degree programme the dropout rates have been as high as 40% and compared with traditional mode system it is too high. The drop out rate has been identified as a major issue of the BBM online degree programme. The challenges facing by the students and staff have been identified as problem areas during the analysis of the BBM degree programme. In this regards the research problem of this study arises as the following research questions:

What do students in online degree programme perceive as the major challenges in learning activities?

What do academics (staff) in online degree programme perceive as the major challenges in teaching/delivering activities?

The prime objective of this study is to find out the challenges facing by the students and academics. In addition to the prime objective other objectives could be defined as follows:

- To investigate the problems and issues on student support, flexibility, teaching and learning activities, access and academic confidence.
- To find out the solutions to meet these challenges.

2 METHODS AND DISCUSSIONS:

This study based on case study on BBM online Degree Programme at University of Jaffna, Sri Lanka. The study mainly interpretative by being based on a qualitative analysis of interviews with respondents with semi structured questions, discussions and the observations. The study based on 90 students who are external students from online degree programme in the University of Jaffna from 2nd year and 3rd year onwards. Further the discussions had been made with the SMEs on this regards. In order to understand the major challenges of online learning, students were asked the questions on students

support, flexibility, teaching and learning activities, access and academic confidence etc., The all students have the traditional based experiences so that they were able to identify the differences between online and traditional system and they were asked to write the answers to these questions on blank papers. These responses are referred to in this study. Further, students were asked their opinions and ideas for the challenges which were summarized as solutions with the academics opinions.

3 DATA ANALYSIS:

Student face many challenges however only five (5) challenges which have been discussed in this study and they were directed to response only these. The responses have been collected and coded within five categories which are analyzed under the categories of students' support, Flexibility, Teaching and learning activities, access and academic self confidence. The selected responses have been given in the table format and students' responses have been given in the quotation marks and Subject Matter Experts (SMEs) who are the lecturers and tutor mentors' opinions also have been given in the columnar format. Further from the students and staff views the researcher's comments have been summarized as follows:

3.1 Students Support:

The students support is the important matter for the successfully completion of the programme. The student support services include tutoring mentoring, content development, technical support services, interaction with students, administrative activities, etc. From the researcher's past findings of BBM Online Degree Programme students have dissatisfaction on their support services. The following Table 1 illustrates the major challenges on students support identified by the researcher from the respondents' opinions.

Table 1: Challenges on Students support.

	Students' responses for the questions	Staff Opinions	Comments
Tutoring mentoring	<i>"We are unable to contact with the lecturers who are responsible for the respective courses. If the university arrange day classes frequently it will be useful to us"</i> (student, 2010)	The students' attendance is very low to the day classes only the students who come regularly gain the advantage others losing the opportunities and we are unable to satisfying all of them. (Lecturer 1)	Most of the students are working people and they have no time to attend to the day school and they will study at the examination time. They will have many doubts and they are unable to contact with the SMEs at that time.
Content development	<i>"If all course contents would be uploaded in advance it would be useful to study with planned learning activities because most of us are working people"</i> . (student, 2010) <i>"If there would be an announcement about the activities we can complete it within the deadline"</i> . (student, 2011)	All contents with activities are handed over to the tutor mentors and they will manage everything. (Lecturer 2) We have no specialized knowledge in all courses like Lecturers if we asked them on this regards they do not response immediately. (Tutor mentor 1)	SMEs are unable to change their traditional teaching system immediately to the modern online teaching system and they have no/little awareness on this regards.

Source : Survey collected data:

The student support activities are all those interactive processes that are intended to support and facilitate the learning process. The students are often confused as to what to learn and how to learn. They are unable to change their traditional learning system immediately as this system is a new one and they have been provided education on traditional mode so far. Further, University arranges the face to face session to the students but the attendance is very low as they have family and workplace commitments. In this situation all students won't come all sessions. The Learning Management System (LMS) Moodle is new one to the staff so that the workshops on LMS were held at University by the trained staff but the staff attendance was very low because of the local resource persons. If there would be outside resource person and nominated

participants attendance would be high. However, researcher's previous study also revealed that the students dissatisfied (50%) with the student support of BBM Programme.

3.2 Challenges on Flexibility:

The researches show that online students perform better if the course is flexible in the matter of timing and assignment choices, if they are allowed to work at their own pace and if they are provided with several modes of education delivery ie, CD, Web, print outs, SMS etc., (Patton, 2000, Delialioglu and Yildirim, 2007, Sankey, 2006). Students want flexible in all sphere. The following table 2 illustrates the challenges on flexibility of BBM Online degree.

Table 2: Challenges on Flexibility

	Students' responses for the questions	Staff Opinions	Comments
Anyone	"I was unable to enter the university due to lack of seats in the University because of Competitive G.C.E. (A/L) Examination, but I follow this programme while working". (student 2011)	No age limit to follow this course, but none can follow without minimum requirements". (Administrative Coordinator)	Students can benefited because when they are working or following other courses can follow this.
Anytime	"We have to wait for a long time (one year) for the repeat examination". (student 2010) If we are allowed to sit the examination one by one (subjects) it is easy to us to study while working". (2011)	The internal programme's rules and regulations have been followed to all. (Academic Coordinator)	The rules and regulations should be adopted in the University system, but examination results can be released as soon as possible.
Anywhere	"We can sit the examination only in Vavuniya Campus if the study centers will be established in various part of the country it will be easy to access. (student 2010)	It is difficult to manage study centers at various part of the country. When we conduct the examination at Vavuniya Campus we face several issues. (Administrative Coordinator)	It will be considered in near future.

Source : Survey collected data:

Flexibility is important because majority of the students are working or following other courses simultaneously following BBM Programme. But the poorer students are unable to follow or hardworking to finance their studies or studying several courses at the same time to get the maximum benefit in short time. It is difficult to them with inflexible deadline and a great help to them would be to allow repeat exams.

3.3 Challenges on Teaching and learning:

Distant learners feel very distant and their feelings of loneliness and the physical distance they experience. Much research shows that, students in distance mode misses social engagement and a feeling of being involved (Galusha, 1998, Bruckman, 2002, Schrum and Hong, 2002). Students being left to self-studies, feeling separated and isolated, is a commonly stated reason for not passing a course or dropping out. The table 3 illustrates the challenges on Teaching and learning of this programme.

Table 3: Challenges on Teaching and learning

	Students' responses for the questions	Staff Opinions	Comments
Interaction with students	<i>"If there are any questions, or doubts, that arise when studying online learning materials do not help to resolve the problems immediately as in a face to face lecture". (Student 2010)</i> "If we would like to contact with the	"They must attend to the face to face session and they can clear their problems at that time. We are busy at our internal programme and other research works. Tutor mentors can clarify these problems".	The students miss the interactivity like a classroom setting and, above all, the possibility of getting immediate feedback is necessary. Three Online tutors are unable to manage all 3years and they have to be

	<p>respective lecturers over the phone we are unable to reach them” . (Student 2010)</p> <p>We have no technological facilities (video calls/conference) to interact with teachers (Student 2010), (Student 2011)</p>	<p>(Lecturer 3)</p> <p>We don’t have specialized knowledge all over the subjects. We will clear some with the help of the SMEs but they do not help us immediately (Tutor mentor 2)</p> <p>We don’t have the video call facilities to interact with students. (Tutor mentor 3)</p>	<p>recruited more.</p> <p>The technological facilities are important for the interactivity. For the online learning setup it must be considered in the infrastructure facilities like video, audio etc.,</p>
Assessments	<p>“ Some time we are unable to upload the documents assignments with in the deadline due to connectivity problem and other issues on access” (Student 2011)</p> <p>Some time we will be out of the country for our workplace, in this situation we are unable to submit the assignments within the deadline. If we are informed the assignments dates before starting the semester it would be helpful to us to manage our activities properly. ” (Student 2010)</p>	<p>“Continuous assessment would greatly help achieve success at the examinations” But they do not submit all assignments. (Lecturer 4)</p> <p>“Everyone completes their assignments in the last minute so that they face problems” (Tutor mentor 1)</p>	<p>The continuous assessments will help to the students getting good results. Students should pay attention about this.</p> <p>The assignments marks must be displayed in each which helps them getting feedback about their progress.</p>

Source : Survey collected data:

Self studying students need to develop skills and get support in learning strategies and information literacy. Students should learn how to learn and how to manage time and how to complete their activities within the stipulated time.

3.4 Challenges on Access:

Infrastructure facilities and the connectivity are the important aspects for online learning. These are mostly challenges in the developing countries and until eliminate these challenges the online learning environment can’t be better improvement. The following table 4 illustrates the challenges on access.

Table 4: Challenges on Access

	Students’ responses for the questions	SMEs Opinions	Comments
Infrastructure	<p>“Some of the parts of our country still do not have electricity. In this situation how can we manage online learning system?” (student 2011)</p> <p>Our LMS system is interrupted frequently and last month also more than two weeks we were unable to access said reasoning that the NODES server is out’ (student 2010)</p>	<p>In our country as well, the rural areas don’t have electricity facilities and IT. How can we conduct online course? (Staff, External unit)</p> <p>This system depends on the Server from the National Online Distance Education Services (NODES), Ministry of Higher Education, Colombo. If there would be any issues our course interrupted by this. Last month also we face the difficulties and we were unable to serve the students more than 2 weeks and we have to receive so many telephone calls from the students and but the NODES did not give immediate response to this (Moodle Administrator).</p>	<p>Basically all students have access to a computer for their studies. The infrastructural problems concern those living in rural areas mainly refer to Internet not being available and bandwidth problems.</p> <p>Further, if we have own server it is easy to deliver courses to the students by the LMS without interruption.</p>

Connectivity	<p><i>“Bad thing is students in rural area cannot access internet because of the connection speed. And some times students will not able to follow the studies because of this.” (Student 2010)</i></p> <p><i>“But there are some things that make us fed up. Like sometimes we have to wait more than 45 minutes to download a study content” (Student 2010)</i></p>	<p>Connectivity is the big issue in the BBM online Degree course.</p> <p>We have to wait more than 30 minutes for downloading documents and the images can't say the time.</p> <p>If it will continuously be the drop outs will be high sometime online courses will be disappeared.</p>	<p>Some students access the Internet via an ADSL or broadband connection whereas others use dial-up (modem) connection. Still most students and staff say that it is the speed of the connection that is the major bottleneck. Connectivity also affects <i>where</i> students access the Internet. Even though most students have access to a computer at home, when accessing the Internet the picture is more varying. Here we get a rather equal distribution between the home, training institutes, Internet cafés, university, work places and offices.</p>
--------------	---	--	---

Source : Survey collected data:

Mostly in developing countries face these challenges and Sri Lanka has no exception. However most probably in Northern part of Sri Lanka faces not only these access problem but also face the electricity problem.

3.5 Academic confidence:

Self confidence is important thing to get better outcome in each and everything. In the BBM online degree most of the students have the better results in the G.C.E. (A/L) Examination. It does not matter in these challenges. The table 5 illustrates the challenges on academic confidence.

Table 5: Academic confidence

	Students' responses for the questions	SMEs Opinions	Comments
Quality of students	<p><i>In Sri Lanka G.C.E. A/L Examination is a competitive examination and however we get good results we are unable to enter the university. BBM is an opportunity to follow our higher degree in our mother tongue (Tamil) (student, 2011).</i></p> <p><i>Most of we are the Bank staff and our working time more than 8 hours and we have no time to attend the classes, in this situation we are able to study BBM after office hrs at home. If we get hardworking we can obtain good results (students 2011).</i></p>	<p>“We are getting the best A level students who are unable to enter the university due to lack of seats and some of them may earn overall GPA more than 3.7 and may obtain first class. (Academic Coordinator)</p>	<p>By opening up the BBM program for more students one will also receive students that attend the external mode because they did not qualify to the university. Whereas this is an admirable effort to disseminate knowledge to more than a selective few it will still be a challenge in future when we have the low infrastructure facilities.</p>
Subjects previous studied	<p><i>“In the business management programme we have to learn more accounting subjects. We have come from other disciplines (arts, science) it is difficult to follow the accounting subjects” (student, 2010).</i></p> <p><i>“It is an opportunity to study in Tamil Medium” (student 2011)</i></p>	<p>We have the management internal Degree programme in English medium. BBM is an opportunity for the students to study in Tamil medium. (Lecturer 5)</p> <p>All disciplines (Arts & Science students) can follow this degree programme (Administrative Coordinator).</p>	<p>This is an opportunity to the Tamil students to study their mother tongue in Sri Lanka. And any disciplines can be enrolled to this programme.</p>

Source : Survey collected data:

In this area the most challenge is the accounting subjects. If it is considered when making the revision of syllabus students can get more confidence with their learning capacity.

4. CONCLUSION

The problems with education for most developing countries are that admissions to universities are limited and that the people most in need of education are the ones most needed at home or at work. A solution to this problem would be to allow people to learn where they are. Providing better online learning environment is the way to enhance the online mode education as well as the education in the country. From the above analysis it is concluded that the feedback is very important for the better online learning environment. From the feedback it shows that learners prefer online learning mode rather than other modes of learning. Moreover, the results suggest that the tutor mentors/facilitators must acknowledge the learner views and the different behavior in order to provide better learning environment. It is not that far a time that our community will demand for more quality in learning processes that incorporates the online technologies that increase the level of interactivity. Academic institutions should be ready to meet this demand. Students complain about the poor organization and infrastructure facilities of the course which must be considered by the institutions favourably. Students being confused and rather lonely will need clear guidance on all practical issues in order to feel comfortable in distance mode. Students need a fixed schedule and exam dates and if changes are made they need to be informed with long notice. Moreover, the access and the electricity problem is the large barrier above at all which must be considered carefully in Sri Lanka. The implication of this study will help the researchers in future who are willing to do further study on this regards.

REFERENCES:

- Abeysekara, U. (2011), Critical Success factors of e-learning in Sri Lankan Universities, Colombo, MBA Thesis.
- Alias, N. A. & Rahman, N. S. N. A. (2005) The supportive distance learning environment: A study on the learning support needs of Malaysian online learners. *European Journal of Open and Distance Learning (EUODL)*, 2005.
- Andersson, A. (2007) Beyond Student And Technology: Seven Pieces to Complete The E-Learning Jigsaw Puzzle in Developing Countries. *30th Information Systems Research Seminar in Scandinavia (IRIS30)*. Tampere, Finland.
- Annual Report (2006), Central Bank.
- Beasley, N. & Smyth, K. (2004) Expected and Actual Student Use of an Online Learning Environment: A Critical Analysis. *Electronic Journal on e-Learning*, 2, 43-50.
- Bruckman, A. (2002) The future of e-learning communities. *Communications of the ACM*, 45, 60-63.
- Burn, J. & Thongprasert, N. (2005) A culture-based model for strategic implementation of virtual education delivery. *International Journal of Education and Development using Information and Communication Technology*, 1, 32-52.
- Delialioglu, O. & Yildirim, Z. (2007) Students' Perceptions on Effective Dimensions of Interactive Learning in a Blended Learning Environment. *Educational Technology & Society*, 10, 133-146.
- Dhanarajan, G. (2001) Distance Education: promise, performance and potential. *Open Learning*, 16, 61-68.
- Direct interviews and discussions with related staff and students and direct observations and experiences regarding the BBM Online Degree Programme.
- Eastmond, D. (2000) Realizing the promise of distance education in low technology countries. *Educational Technology Research and Development*, 48, 100-111.
- Galusha, J. M. (1998) Barriers to Learning in Distance Education. *Interpersonal Computing and Technology*. University of Southern Mississippi.
- Gunawardana, K. D. (2005), An empirical study of potential challenges and benefits of implementing E-learning in Sri Lanka. Proceedings of the 2nd international conference on e-Learning for knowledge based Society, Bangkok, Thailand. Retrieved from: <http://www.elearningap.comeLAP2005/Proceeding/PP33.pdf>
- Heeks, R. (2002) Information Systems and Developing Countries: Failure, Success, and Local Improvisations. *The Information Society*, 18, 101-112.

- O'Connor, C., Sceiford, E., Wang, G., Foucar-Szocki, D. & Griffin, O. E. (2003) Departure, Abandonment, and Dropout of E-learning: Dilemma and Solutions. IN CONSORTIUM, M. C. E.-L. (Ed.) *Human Resource Development*. James Madison University.
- Patton, M. A. (2000) The Importance of Being Flexible with Assignment Deadlines. *Higher Education in Europe*, 25, 417-423.
- Rajesh, M. (2003) A Study of the problems associated with ICT adaptability in Developing Countries in the context of Distance Education. *The Turkish Online Journal of Distance Education*, 4.
- Rathirane, Y. (2012), Learner Support Services and Online Distance Education: A Case Study on Bachelor of Business Management (Online) Degree Programme, University of Jaffna, Sri Lanka, Proceedings of the 11th international Tamil Internet Conference, Annamalai university, Sithamparam.
- Sankey, M. D. (2006) A neomillennial learning approach: Helping nontraditional learners studying at a distance. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, 2, 82-99.
- Schrum, L. & Hong, S. (2002) From the Field: Characteristics of Successful Tertiary Online Students and Strategies of Experienced Online Educators. *Education and Information Technologies*, 7, 5-16.
- Simpson, O. (2004) The impact on retention of interventions to support distance learning students. *Open Learning*, 19, 79-95.
- UNDP (2005) Information and Communication Technology (ICT) for Development.
- UNESCO (2005) United Nations Decade of Education for Sustainable Development.
- UNESCO (2006) Teachers and educational quality: monitoring global needs for 2015.

கற்றல் கற்பித்தலில் தகவல் தொழில் நுட்பத்தின் பயன்பாடு - UK

Siva Pillai

Principal Examiner – Cambridge yyyT Examination Tamil Language Cambridge-University. UK

Chief Examiner - London Edexcel - Tamil Language. UK

உடல் இல்லையெனில் உயிர் இல்லை. அது போல் தொடர்பாடல் இல்லாது இன்றைய கணினி உலகில் மக்கள் இல்லை. ஒவ்வொரு வயதினரும் ஒவ்வொரு தொடர்பாடலுடன் ஈடுபட்டுக் கொண்டிருக்கிறார்கள். நாளுக்கு நாள் தொழில் நுட்ப சாதனங்கள் வளர்ந்து கொண்டு போகின்றன. அன்று பெரிய குளிர் சாதன அறையில் கணினி யந்திரங்கள் இருந்தன இன்று அவற்றைவிட பன்மடங்கு ஆற்றல் கொண்ட கையடக்கக் கணினிகள் சந்தையில் உள்ளன.

ஆக்கத்தின் வழி கற்றல் கற்பித்தல்

1. ஆர்வத்தைத் தூண்டும் வகையில் கற்பித்தல்
2. செயற்பாடுகள் மூலம் கற்பித்தல்
3. தானே ஒலித்துக் காட்டும் கணினி மென் பொருள் மூலம் கற்பித்தல்
4. அவர்களது உள்ளார்ந்த ஆற்றலைக் கண்டறிந்து வாழ்த்துதல், வளர்த்தெடுத்தல், சுயமாக இயங்க அடித்தளம் அமைத்தல்

எந்த ஒரு குழந்தையும் தன்னுடைய தாய்மொழியை கற்கத்தொடங்கும் போது கேட்டல் திறனில் இருந்தே ஆரம்பிக்கின்றது. பிறந்ததில் இருந்து கேட்கத்தொடங்கும் குழந்தை தன்னுடைய தேவைகளுக்காக ஒரு வருட காலத்திற்குப் பின்னர் முதன் முதலில் சொற்களை பாவிக்கத்தொடங்குகின்றது.

தொடர்ந்து வரும் இரண்டு வருடங்களின் பின்னர் சிறிய வசனங்களைப்பேச ஆரம்பிக்கின்றது. நான்கு வருடங்களின் பின்னர் மகிழ்ச்சி ஊட்டக்கூடிய விதத்தில் கதை பாட்டுக்கள் என்பவற்றை மொழியூடாக அறிந்துகொள்கின்றது. இதேகாலப்பகுதியில் எழுத்துக்களை எழுதவும், வாசிக்கவும் கற்றுக்கொள்கின்றது. தொடர்ந்து வரும் காலங்களில் உலக அறிவை மேம்படுத்துவதற்காக மொழியைப் பயன்படுத்தி விடயங்களைக் கற்றுக்கொள்கின்றது.

அடுத்த அடுத்த வருடங்களில் மொழி மூலம் கேள்விகளுக்கு பதிலை அல்லது தன்னுடைய கருத்தை வெளிப்படுத்துகின்றது. மொழியில் தேர்ச்சி பெற்ற ஒரு பிள்ளை மொழியைப் பயன்படுத்தி கவிதை, சிறுகதை போன்ற ஆக்க இலக்கியச் செயற்பாடுகளில் ஈடுபடுகின்றது. இது எல்லாம் இயற்கையாக மொழிக்கல்விச் செயற்பாட்டில் ஒரு பிள்ளை செய்கின்ற செயல்கள். இங்கு பிள்ளையின் புறச்சூழல் என்பது பிள்ளை கற்கும் மொழி சார்ந்த சூழலாக இருக்கும் என்பது குறிப்பிடத்தக்கது. ஆனால் வெளிநாடுகளில் தமிழ்மொழி கற்க விளையும் குழந்தையின் புறச்சூழல் என்பது வேற்று மொழி உபயோகப்படுத்தும் நிலை காணப்படுவதால் பிள்ளைக்கு தமிழ்மொழியைப் கற்பிப்பதற்கு சூழ்நிலைக்கு ஏற்ப பல்வேறு உத்திகளை பின்பற்ற வேண்டிய நிலைக்கு ஆசிரியர்கள் தள்ளப்பட்டுள்ளனர்.

எந்த வயதினர் ஆயினும் தான் சார்ந்த சூழல் அல்லாத ஒரு மொழியைக் கற்கும் போது மொழியைக் கற்பிப்பவர் கற்பவரை பிறந்த குழந்தையாக எண்ணிச் செயற்பட வேண்டும். ஆகவே போதியளவு தமிழ்மொழியைக் கேட்பதற்கு ஏற்பாடுகள் செய்ய வேண்டும். இதற்காக குழந்தைகளின் மொழியில் அவர்கள் பயன்படுத்தும் சொற்களைக் கொண்டு உரையாடல்கள், மற்றும் கதை, பாடல்கள் போன்ற பொழுதுபோக்கு அம்சங்களை ஒலிப்பதிவு செய்து அவற்றின் அடிப்படையில் கேள்விகள், கேட்டு குழச்செயற்பாடுகள் மூலம் மொழியை அறிமுகம் செய்ய வேண்டும்.

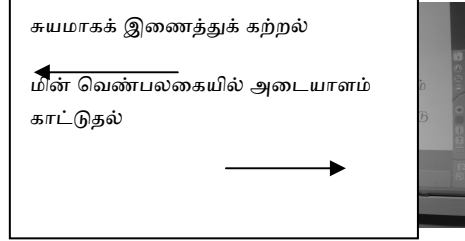
ஒலிப்பதிவுக்காக உரையாடல்களை தெரிவு செய்யும் போது ஒரு பிள்ளை இலங்கை, இந்தியா போன்ற முழுக்க முழுக்க தமிழ்மொழி பயன்படுத்தும் சூழ்நிலையில் தனக்குத் தேவையானவற்றை சுயமாகப் பூர்த்தி செய்வதற்கு ஏற்ற வகையில் இருக்குமாயின் கற்கும் விடயங்கள் அன்றாட வாழ்வில் பயன்படுத்தக்கூடியதாக இருக்கும்.

இரண்டாவது படிநிலையில் பிள்ளை எழுத்துக்களை அடையாளம் காணவும் உச்சரிப்பதற்கும் ஏற்ற வகையில் கற்பித்தல் இடம்பெற வேண்டும். ஒரு பிள்ளை ஒலிக்குரிய வடிவங்களை (எழுத்துக்களை) அடையாளம் காணவும்

உச்சரிக்கவும் முடியுமாயின் எதிர்காலத்தில் பிள்ளை இயல்பாக தனக்குத் தேவையான விடயங்களை தமிழ்மொழி மூலம் தேடி கற்றிந்து கொள்ள வசதியாக அமையும்.

அடிப்படை இலக்கணம் அறிமுகப்படுத்தப்பட்ட பின்னர் மொழியின் இலக்கணம் அறிந்து கொள்ள ஆசைப்படும் பிள்ளைகளுக்கு மட்டும் தமிழ் மொழி இலக்கணம் ஆழமாக கற்பிப்பது சிறந்தது. முதலில் இலக்கணம் பற்றிப் போதிப்பது மொழி கற்றலில் உள்ள ஆர்வத்தைக் குறைத்துவிடும்.

தற்போது வெளிநாடுகளில் உள்ள குழந்தைகள் தமிழ் மொழி கற்பதற்கு ஏற்ற பொருத்தமான ஆர்வமூட்டும் சிறுவர் கதைகள் தமிழில் இல்லை அல்லது குறைவு என்றே சொல்ல வேண்டும். இன்றைய காலத்திற்கு ஏற்ப சிறுவர்கதைகளை உருவாக்கி அவற்றை குறுந்தட்டு-சீடி-CD மூலம் அல்லது இணையத்தளங்கள் ஊடாக சிறுவர்களும் இணைந்து பங்குபற்றக்கூடிய வகையில் தயாரிக்க வேண்டும். தற்போது குறித்த தயாரிப்புக்களை தனியார்துறையினர் மேற்கொண்டு வருகின்றனர் இருந்த போதிலும் அவற்றுக்கான தரம் மற்றும் பொருத்தப்பாட்டை நிர்ணயம் செய்வதற்கு தமிழ் மொழி ஆற்றலும், ஆவர்வமும் உள்ளவர்கள் தனியார் துறையினருடன் இணைந்து செயற்பட வேண்டும். பொரும்பாலான மொழி கற்பிக்கும் இடங்களில் பிள்ளைகளின் சூழ்நிலையை புரிந்து கொள்ளாத நிலையே காணப்படுகின்றது என்பது வருத்தத்திற்குரியது.



கணினியை உபயோகித்து மாணவர்கள் இணைந்து செயற்படும் மென்பொருள்களைப்பயன்படுத்தி 'தமிழ் TamilBook 2000- கல்வி 2000' By Dr Periannan Kuppasamy PhD) போன்ற குறுந்தட்டு-சீடி-CD மூலம் மொழி கற்பித்தல் சாலச்சிறந்தது.

Effectiveness Of Computer Assisted Learning In Tamil Grammar

*Dr.G.Singaravelu, Associate Professor, UGC-Academic Staff College, & HODi/c Department of Education (SDE)
Bharathiar University, Coimbatore-641 046. Tamilnadu Email: singaravelu.bu@gmail.com*

ABSTRACT

The study illuminates the effectiveness of Computer Assisted Learning in Tamil Grammar for the learners at standard VIII. Computer Assisted Learning is particularly beneficial for any student learning a language. It promotes peer interaction, which helps the development of language and the learning of concepts with content. It is important to assign TLL (Tamil Language Learners) to different teams so that they can benefit from Tamil language role models. **Vasu Renganathan (2009) supported** enhancing the Process of Learning Tamil with Synchronised Media and he recommended the Computer Assisted Learning in enhancing Tamilgrammar learning . **Objectives of the study:** 1.To diagnose the problems of the learners in learning Tamil Grammar through conventional methods.2.To find out the significant difference in achievement mean score between the pre test of control group and post test of control group.3.To find out the significant difference in achievement mean score between the pre test of Experimental group and post test of Experimental group. 4.To find out the significant difference in achievement mean score between the post test of controlled group and the second post test of treatment given controlled group.5.To find out the impact of Computer Assisted Learning in learning Tamil Grammar. **Methodology:** Rotational group Experimental method was adopted in the study. **Participants:** Sixty students of studying in standard VIII from Maruthamalai Higher Secondary school, Vadavalli were selected as sample for the study. **Instrumentation:** Researcher's self-made achievement test was used as instrumentation for the study. **Findings:** Computer Assisted Learning is more effective than traditional methods in learning Tamil Grammar for the learners at standard VIII. **Educational implications:** It can be implemented in other levels and all other the languages learning.

Key words: Computer Assisted Learnin, Treatment given controlled group and

Experimental group. CALL-Computer Assisted Language Learning.

INTRODUCTION

Grammar is the basic of all languages and it may confirm that backbone of any language. Learning Tamil is difficult due to more letters in Tamil but English language has 26 letters. Remembering all the letters and learning grammar in Tamil is difficult to the learners of Tamil. Learning grammar is indispensable for learning Tamil language perfectly. Without understanding the grammar, the usage will not give clear meaning of sentences. Learning grammar from primary level is important for error free usage in Tamil. Even if Tamil is mother tongue of the many young learners, they commit mistakes in usage of spoken and written. Attractive methods of teaching grammar are the way to eliminate the problems of the learners in learning Tamil. Nowadays technology oriented learning attract the young learners. Maximum learners are engaging themselves by using Mobile, computer and internet for learning the language of Tamil. Learning by using computer bewitches the minds of learners as well as enthusiasm of the learners.

NEED AND SIGNIFICANCE OF THE STUDY

Parents encourage their children to study in medium of instruction in English. Teaching of Tamil in upper primary level was hurdle due to negligent of learning Tamil by the learners. Learners can talk Tamil less error free causes of unknown the grammatical rules. Another main reason of the less error free usage is conventional methods adopted in the classroom transaction. Hence the researcher prepared some activities through Computer Assisted Language Learning to learn Tamil Grammar.

OBJECTIVES

1. To diagnose the problems of the learners in learning Tamil Grammar through conventional methods.
2. To find out the significant difference in achievement mean score between the pre test of control group and the post test of control group.

3. To find out the significant difference in achievement mean score between the pre test of Experimental group and the post test of Experimental group.
4. To find out the significant difference in achievement mean score between the post test of controlled group and the second post test of treatment given controlled group.
5. To find out the impact of Computer Assisted Learning in learning Tamil Grammar

HYPOTHESES

1. Students of standard VIII have problems in learning Tamil Grammar through conventional methods.
2. There is no significant difference in achievement mean score between the pre test of control group and the post test of control group.
3. There is no significant difference in achievement mean score between the pre test of Experimental group and the post test of Experimental group.
4. There is no significant difference in achievement mean score between the post test of controlled group and the second post test of treatment given controlled group.
5. Computer Assisted Learning is more effective than conventional methods in learning Tamil Grammar

REVIEW OF LITERATURE

Lee, Cynthia (2013) investigated the effort to add to computer-Assisted language learning by extending a study on an essay critiquing system (ECS) feedback to secondary school language learners' writing. The study compared two groups of participants' performance, namely the treatment group which received both the system feedback and teacher feedback (i.e., blended learning mode), and the control group which received teacher feedback only. The study was conducted in a secondary school in Hong Kong in the form of an extra-curricular activity after school. Fifty-three students from the ages of 16 to 17 with different argumentative writing experience participated in a series of five writing workshops once a month voluntarily. The participants were assigned to a treatment and a control group based on the stratification approach. Their scores were compared and an interview with the treatment group was conducted. Analysing the total and content and organisation scores of the two groups, it was found that both groups demonstrated statistical significant gains. Nevertheless, the treatment group's gain appears to be more meaningful than the control group's in view of its group composition. The blended learning mode contributes to the encouraging results as revealed through the treatment group's scores, high rating and positive comments on both system and teacher feedback in the survey and interviews. The analysis also implies the need to improve the system feedback on paragraph coherence, workshop design and grading criteria.

VARIABLES

The independent variables namely Computer Assisted Learning and the dependent variable namely achievement test score were used in this study.

METHODOLOGY

Rotational Equivalent group Experimental method was adopted in the study. Pre test-treatment-Post test was used to perform the study. Computer Assisted Learning attracts the young students. Using computer in learning Grammar simplified the learning.

Sample

Sixty students of studying in standard VIII from Maruthamalai Higher Secondary school, Vadavalli were selected as sample for the study. Thirty students were considered as Controlled group and another thirty were considered as Experimental group.

Tool

Researcher's self-made achievement test was used as a tool for the study. Validity of the tool was established by the opinion of the juries. Reliability of the tool was established by the test-retest method.

Construction of tool:

The investigator's self made Achievement test was used for the pretests and post tests of both control groups and experimental groups. The same question was used for both pre and post tests to evaluate the effectiveness of Computer Assisted Learning in learning in Tamil through objective types of question which carried one mark for each question and contained 50 marks.

Reliability of the tool

Reliability had been computed using test-retest method and the calculated value was 0.89. The value was quite significant and implied that the tools adopted were reliable. Hence the reliability was established for the study.

Validity of the tool

Subject experts and experienced teachers were requested to analyse the tool. Their opinions indicated that the tool had content validity.

Procedure of the study

1. Identification of the problem in conventional method by administering pre-test to the both groups.
2. Preparation of Computer Assisted Learning.
3. Preparation of Activities.
4. Validation of the Computer Assisted Learning in the classroom practices.
5. Preparation of Achievement test.
6. Administering Pretests for the both control group and Experimental group.
7. Post tests for the both groups.
8. Treatment is given for both control group and Experimental group
9. Administering post-test.
10. Data Analysis.
11. Findings and Interpretations.
12. Educational Implications.

Procuring Data

The researcher got permission from the management of the school and administered pretest to the students with the help of a teacher. The question paper was given to the individual student and evaluated learning obstacles of the students were identified by the pretest. The causes of low achievement by unsuitable methods were found out. Computer Assisted Learning was used in the classroom for eliminating the problems of learning grammar for one week. The posttest was administered and the effectiveness of the Computer Assisted Learning was found out. Treatment was given for the both groups.

Statistical technique

ttest was computed for the study.

RESULT**Hypothesis :1**

Students of standard VIII have problems in learning Tamil Grammar through conventional methods.

In the post-test Experimental group, students of standard VIII scored 70% of marks but in the pre-test Experimental group students scored 30% of marks in the conventional method. Hence Students of standard VIII have problems in learning Tamil Grammar through conventional methods.

Hypothesis : 2

There is no significant difference in achievement mean score between the pre test of control group and the post test of control group.

Stages	N	Mean	S.D	df	t-value	Result
Pretest control group	30	10.48	2.85	58	0.38	Insignificant at 0.05 level
Post test control group	30	10.78	3.12			

The table showing achievement mean scores between pre test of control group and posttest of Control group.

The calculated" t' value is (0.38) less than table value (2.00). Hence null hypothesis is accepted at 0.05 levels. Hence there is no significant difference between the pre test of control group and post test of control group in achievement mean scores of the students in learning Tamil Grammar through conventional method at standard VIII.

Hypothesis:3

There is no significant difference in achievement mean score between the pre test of Experimental group and post test of Experimental group

Stages	N	Mean	S.D	df	t-value	Result
Pretest experimental group	30	12.70	3.21	58	5.99	Significant at 0.05 level S
Post test experimental group	30	17.70	3.25			

The table showing achievement mean scores between pretest of Experimental group and posttest of Experimental group.

The calculated 't' value is (5.99) greater than table value (2.00). Hence null hypothesis is rejected at 0.05 level. Hence there is significant difference between the pre test of Experimental group and post test experimental group in achievement mean scores of the students in learning Tamil Grammar through CALL at standard VIII.

Hypothesis :4

There is no significant difference in achievement mean score between the post test of controlled group and the second post test of treatment given to controlled group

Stages	N	Mean	S.D	df	t-value	Result
Post test Control group	30	10.78	3.12	58	6.771	Significant at 0.05 level S
Post test Contol group after treatment	30	16.35	3.25			

The table showing achievement mean scores between Post test of control group and second post test of treatment given controlled group

The calculated 't' value is (6.77) greater than table value (2.00). Hence null hypothesis is rejected at 0.05 level. Hence there is significant difference between the pre test of Post test of control group and post test of treatment given to the control group in achievement mean scores of the students in learning Tamil Grammar.

Hypothesis :5

Computer Assisted Learning is more effective than conventional methods in learning Tamil Grammar

Achievement mean scores of the learners in post-test of control group is 10.78 and the achievement mean scores of the learners post test of control group after treatment of using CALL is 16.35. Score of the post test of Experimental group (17.70.) is greater than Pre test of Experimental group (12.70). Above both assure that acquiring the skill of learning Tamil Grammar by using CALL is more effective than conventional methods.

FINDINGS

1. In the post-test, students scored 70% of marks and the trainees scored 30% of marks in the conventional method.
2. There is no significant difference between the Pret test of control group and post test of control group in achievement mean scores .
3. There is significant difference in achievement mean score between the pre test of Experimental group and post test of Experimental group
4. There is significant difference in achievement mean score between the post test of controlled group and the second post test of treatment given controlled group
5. Computer Assisted Language Learning is more effective than existing methods in learning Tamil grammar

EDUCATIONAL IMPLICATIONS

1. CALL in Tamil grammar can be extended to primary level, secondary level and higher secondary level.
2. It can be encouraged to implement to use in adult education
3. It may be activated in teachers education also.
4. It can be used in the classroom of Tamil as a second language learners.

References

- **Bailey R &. Dugard Co (2007)** Lights, camera, action! Digital video in the languages classroom (CILT New Pathfinder No.6), London: CILT.
- **Bertin J-C (2001)** "CALL material structure and learner competence". In Chambers A. & Davies G. (eds.) Information and Communications Technology: a European perspective, Lisse: Swets & Zeitlinger.
- **Bkkertmn D. (2000)** "Can (and should) academic linguists become multimedia authors?" In Fremdsprachenlernen mit Multimedia [. .], Triangle 17, 30-31 janvier 1998, Paris: ENS Editions (for Goethe-Institut, ENS Fontenay/Saint-Cloud, The British Council).
- **Chapelle Co (1997)** "CALL in the Year 2000: still in search of research paradigms?" Language Learning & Technology 1, 1: 19-43
- **Chapelle Co (1998)** "Multimedia CALL: lessons to be learned from research on instructed SLA", Language Learning & Technology 2, 1: 22-34
- **Lee, Cynthia (2013)** Immediate Web-Based Essay Critiquing System Feedback and Teacher Follow-Up Feedback on Young Second Language Learners' Writings: An Experimental Study in a Hong Kong Secondary School (EJ1005234) Computer Assisted Language Learning, v26 n1 p39-60 2013

Using Augmented Reality as a Visual Aid to Teach Tamil Language in Malaysian Tamil Primary Schools

Kumaraguru S/O Poobalan, MSC Student, University Malaysia Pahang Gambang

Abstract: Augmented Reality is a technology which combines computer vision and virtual reality. This technology works in a way that, when certain printed image is brought under a certain vision based software, a 2D or 3D image can be generated in the devices. Hence this technology has greatly potential to be used as a visual aid in education specially in teaching Tamil language in class room environment in Malaysian Primary Tamil Schools. Besides that, this method will allow a much more interactive and interesting teaching and learning method. Several problems involving the learning and teaching can be solving using this method like maintaining the interest level of learning of students in classroom. Student will not need to require following the traditional aspect of learning where a student is needed to open a book to learning, here they can scan a printed image to learning. In addition, it also provides a new teaching and learning method in teaching and learning Tamil language especially in schools. The main emphasis on this technology is computer vision focusing on image processing, 3D modeling and vision manipulation in cameras in computers and similar devices. It gives great details into detection and recognition of object using image processing techniques. This paper delves into possibility to use augmentation reality for Tamil language education in Malaysian Tamil schools via any mobile devices like laptop, iPad, iPhone and any other similar devices.

Keywords: Augmented Reality, Tamil, Education, Teaching, Learning

Introduction

Augmented Reality (AR), is a useful technology which deal with the mixed of reality and virtual reality. Generally AR is the visual results which are generated by computer by merging with a real view to create an augmented display. According to Roched (2011) has stated that Augmented Reality (AR) has existed in sci-fi movies for decades and until recently has been a thing of the future. This can supported since AR has been an interesting and developing technology in recent year, however it was focus more toward advertising and entertainment. Due the recent development in mobile computing devices has enable AR to be used in education as well. Hence there is a great chance that this technology can be used as a visual aid to teach Tamil language in Malaysian Tamil School. Klopfer and Squire (2008) stated that recent push from desktop computers and laptop computers to handheld computers has certain obvious advantages in terms of cost and maintenance, the educational potential of this new platform have been sparsely explored. Indicating that augmented reality has great potential to be developed to toward the cause. Besides that, augmented reality can be consider the level of teaching Tamil language in school due to the fact that the technology has the nature ability to attract people attention due the unique results generated by it. In term of education, this technology has great potential to further improve the teaching and learning level of Tamil language in Malaysian Tamil School. Hence this paper provides a simple prototype on how the technology can be used in such way.

Problem Statement

This paper aims to solve the following problems;

1. Maintaining the interest level of learning in student while studying in classroom
2. Provides a new teaching and learning method in teaching and learning Tamil language especially in Tamil schools.
3. Implement a new scan to read method to replace the old methods used in Tamil schools in Malaysia.

Implementation

Augmented Reality allows virtual images to be imposed over live video of the reality world. This process uses marker which is describe below. A general AR application works as follow;

1. A camera works record videos of reality world and send into computer
2. Then an application will search of any markers
3. If the markers are found, then a pre-drawn graphic model will be generated over the marker.

Implementation of this paper is divided into 3 parts. These parts are details described below;

AR tag markers

Augmented tag markers are printed images with black or white details. The markers are used in collaboration with edge detection to detect any edge pixels which serves as basic for the marker detection process. The figure 1.0 below shows a sample marker which can be used in augmented reality application. It is to be noted that, it is totally different from QR markers which works in almost similar way but requires internet connection to display information. These AR markers do not require any internet connection to display the visual results but required a vision image processing software to work. Therefore setting up the markers is the foremost process in creating an augmented reality application.

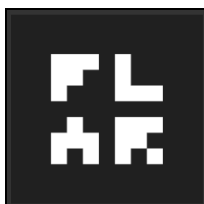


Figure 1.0

Image processing

The image processing technique used in creating the application is edge detection. This process involve a detail analyze of the black and white details in the marker. The edge detection which used in this application was Canny Edge Detection which it involves 5 different steps. These steps include smoothing which blurs the image to remove unwanted noise or data. The next is the find the gradient since the edges should be marked to identify the magnitudes. This step is followed by suppression of non-maximum where the only local maxima should be marked as edges. Next, the potential edges are determines by thresholding using the concept of double thresholding. The final step will be to remove edges which are not connected with each other.

Creating the visual image

The image that required to be generated can be in any format either 2D or 3D. But this paper only 2D image will be used. The image will be created using Paint software in Windows. But if a 3D image is needed softwares like lightwave or 3D max which can be used to generator the image. The last step is to combine the image processing technique and the visual to produce the image.

Results

This section highlights how the proposed application can be used as a Visual Aid to Teach Tamil Language in Malaysian Tamil Primary Schools The application is able to produce any Tamil language characters or any type of learning mode of the language. A webcam captures the video stream of the markers which is places on the location which the images are produced. In this case, the marker is glued to wall to obtain a sample result. However a prebuild marker on the wall or any surface can be used. The results are as followed.

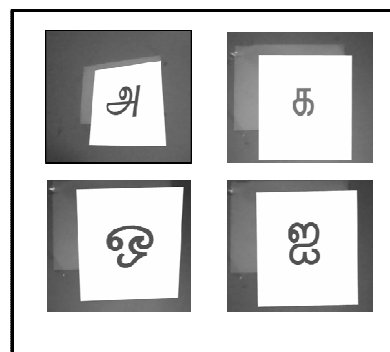
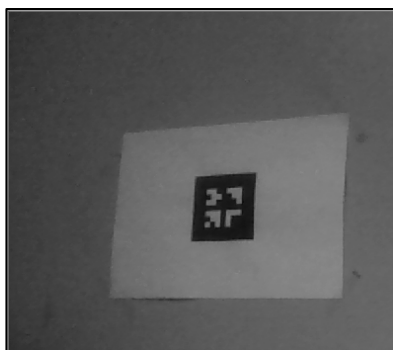


Figure 2.0 shows the marker

Figure 3.0 shows the results

Future Improvement

The current application uses only 2D image, perhaps in the future implementation can lead to 3D images, video or even animation. Besides that, it is also possible to use advanced vision system techniques like motion detection to further manipulate the image from reality world. This will enable to modify the visual result generated on the marker.

Conclusion

As conclusion the usage of augmented reality in teaching Tamil language can be very useful and beneficial to the social and student. Besides that, the technology has the potential to be future progression of the education system in Malaysian Tamil schools. As a visual aid the technology is able to give better understanding the language also acts a much better interaction mode of learning the language. Teachers will be able to maintain the level of interest of class as well delivering an effective teaching and learning process to the student. Besides that, students will no longer need to depend on books or normal visual aids to learn when this technology is being used.

References

1. Eric Klopfer, Kurt Squire. (2008). The Development of an Augmented Reality Platform for Environmental Simulations. *Environmental Detectives*. 1 (1), 50.
2. Kyle Roche (2011). *Pro iOS 5 Augmented Reality*. USA: Apress. 356.

Bibliography

1. John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8 (6):679–698, Nov. 1986.
2. Lu, S.C.-Y., Shpitalni, M., & Gadh, R. (1999). Virtual and augmented reality technologies for product realization. *CIRP Annals*, 48 (2), 471–495.
3. Maiga Chang, Wu-Yuin Hwang, Ming-Puu Chen, Wolfgang Müller. (2011). *Edutainment Technologies. Educational Games and Virtual Reality/Augmented Reality Applications*. 6th International Conference on E-learning and Games, Edutainment 2011, Taipei, Taiwan, September 2011. 6 (4), 99.
4. M. Maidi et al., A performance study for camera pose estimation using visual marker based tracking, *Machine Vision and Applications*, Springer-Verlag, 2008.
5. R. Azuma et al., Recent advances in augmented reality, *IEEE Computer Graphics and Applications*, 21 (6), pp. 34-47, 2001
6. Yuan-Ping Luh, Jeng-Bang Wang, Jin-Wan Chang, Shun-Ya Chang, Chih-Hsing Chu. (2012). Augmented reality-based design customization of footwear for children.13.

Dubbing and Subtitling Free open course wares in Tamil

Ragunaath Rathnam, ragu.dpi@gmail.com

The idea of dubbing and subtitling free open course wares in Tamil, emerges from Khan academy website, the lack of contribution when compared to the rest of languages, despite the fact, Tamil is one of the largest spoken language across countries. The idea is to reach the existing quality science education to our young minds in our language.

Khan Academy

Salman Khan, who reinvented the free and open Education through Video. Started with Yahoo doodle to teach Mathematics for his cousin Nadia, later he uploaded it on youtube for the benefit of everyone, Khan's teaching were practical, relaxed and in a humorous way, which helped more young minds all over the world. Khan quit his hedge fund analyst job in 2009 and dedicated his work towards free education under the name Khan Academy. www.khanacademy.org

Course Content

Mr. Khan started with his doodle and now their library contains more than 4200 videos, which focus on mathematics, physics, chemistry, biology and extends to history, economics and so on.

Course Languages

Lessons have been dubbed in more than 22 languages with its own lingual sub-domain and the URL's of each video are mapped to original video. Among Indic languages, Telugu and Bengali owns their sub-domain, carrying their language's code.

- <http://bn.khanacademy.org> (Bengali - BN)
- <http://te.khanacademy.org> (Telugu – TE)
- <http://de.khanacademy.org> (German – DE)

MIT Open Course wares

Massachusetts Institute of Technology, top ranking technical university which offers open course ware from their site, MIT's open course ware has both Text lessons and Audio Video lectures, ranging from Computer Science, Aeronautics, Physics, Anthropology, Media Arts to Environmental Science for free, these are also translated in other languages.vAudio / Video Lectures

- <http://ocw.mit.edu/courses/audio-video-courses>

Tamil Contribution

Tamil contribution to Khan academy is relatively very low, the Tamil youtube page of Khan Academy contains only 18 mathematics lessons, which were dubbed in Tamil.

- <http://www.youtube.com/user/KhanAcademyTamil>

Conclusion

Edunation Malaysia is a non-profit volunteer initiative to ensure quality education to every Malaysians for free. Likewise, It is important to address the need of contributors to dub most of the lessons in Tamil. which will benefit the young minds of Tamil across the globe and make a presence of Tamil site with in Khan Academy.

- ta.khanacademy

About Edunation (from their website)

Edunation is a volunteer driven non-profit educational initiative and we are here to ensure that every Malaysian student has all the educational resources he needs to succeed academically in school.

For the past year we have been hard at work in putting the Malaysian school syllabus online by mapping the Khan Academy videos and creating our own videos.

- www.edunation.my

References

<https://www.sites.google.com/a/khanacademy.org/forged-for-translators/translations-update>

Bridging the Digital Divide for the Differently-Abled (Tamil) Students in Higher Education

J. Jerald Inico, Assistant Professor, Department of Computer Science
Dr. T. Edwin Prabakaran, Associate Professor, PG & Research Department of Statistics,
Loyola College, Chennai
jeraldinico@gmail.com , teprabakaran@yahoo.com

Abstract

Information and Communication Technology, popularly known as ICT, brought us many new avenues to create, connect, communicate, collaborate and do many creative works. Educators are using ICT for demonstrating their students various aspects of a concept. The students themselves update their knowledge and determine what they want to learn from time to time. The Parents find enough guidelines to guide their wards to choose right subject at the right time. The members of faculty are searching for the alternate theory or alternate solution for the similar kind of problem. The managements of the educational institutions too are looking for information about new courses to be offered that fetch immediate job to the new generation of students. Hence, all the stake holders of education are highly relying on ICTs to grow further.

The ICT consists of a computer system, internet connectivity and other associated assistive devices with necessary software system to create, connect, communicate and to collaborate. The computer system may range from small PDAs, Palmtops, Laptops, Personal Computers, Mini computers, Super Computers, up to MainFrame Computers. Today even the smart phones or other modern gadgets have the similar capacity of a computer which help the user to connect to the internet. There is no end to innovation and numerous numbers of innovative computer hardware and software products are produced every day. The free form style of internet encourages everyone to publish their articles, stories, photos, videos, etc, and it brought numerous volume of data available in the web. The search engines are employed to fetch the right data wanted by the users. The social networking websites have also created a new way of communicating and collaborating with the users. The modernized interfaces are easy to learn and help everyone to interact with the web. Not only the experts but also the kids can operate computer systems with a little bit of experience. Though the ICT has given us a lot of sophisticated hardwares, softwares and data, they have also created a digital divide among us. The differently-abled, persons with no sight or partial sight, are really under privileged, as they can not use mouse with ease.

This research aims to analyze the utility of ICT by the differently-abled (Tamil Literature) students and to promote necessary assistive devices/ facilities in Tamil language to enable them to use web on a par with others.

Keywords

Information and Communication Technology, Internet, Persons with Visual Disability, Web Accessibility, Web Content Accessibility, Web Personalization

Introduction

Every one is enjoying the fruits of ICT as it opens many new avenues to create, connect, communicate, and to collaborate world wide. Almost all the stake holders of education are utilizing it for their academic purpose also. A student with vision impairment mostly depend on someone who can read texts that they want or require someone who can write what they tell, namely Scribes. They also can not use mouse and other sensitive devices which others can use. This creates a Digital Divide. There are few accessibility softwares called screen readers, which may read the contents in the screen and they have emerged recently as a great help in utilizing the resources in the web. The screen readers including JAWS, NVDA, VOICE OVER, ZOOM TEXT, WINDOWS EYES, CHROME VOX, and SYSTEM ACCESS are the popular among them. Apart from the screen readers, a student with low sight can utilize the features such as Screen Enlarger, Increased Browser Text, Browser Zoom and High Contrast, introduced recently. The screen readers for the mobile devices were also introduced to help them, utilize Nuance Talks, Mobile Speak, Voice Over, Talk back for Android, Mobile Accessibility for Android, IDEAL, Orator for BlackBerry, Galaxy Talk, etc. Many such facilities are available for English language but are not available for Tamil language. The emergence of new hand held devices namely Tablets and other touch screen devices have also created some challenges only for the persons with vision impairment, which can be termed as a Digital Divide. This paper primarily aims to identify the level of digital divide by conducting a preliminary survey among the differently-abled Tamil Literature

students in Higher Education. The problem identification is the first step towards finding the solution.

The Objective of the Study

ICT consists of many components including Information, Computer and Internet. The devices used vary from user to user, depending on the features, comfort, availability, affordability, support, etc. Since the differently-abled can not see or realize any graphical items that others may refer in ICT, they need some assistive devices or softwares which help them to understand what information is available in front of them. The assistive devices are mostly from the Optical Character Recognition, Speech Recognition, Web Accessibility and Web Content Mining. The main objective of this study is to identify whether or not those products of ICT are available, accessible, affordable and usable by the differently-abled (Persons with no sight and low sight) using those assistive devices and to identify the problems faced by them.

The Motivation of the study

The motivation behind this study is very unique. Many NGOs are closely working with differently-abled students by arranging them scribes who are working professionals, for reading and writing on certain days of the week. To avail these facilities, they may have to travel from one place to another, sometimes to far away places where they may not reach in time, may lose their opportunity or miss their scribe and to their portions too. Due to non availability of standard scribes, many such students are disappointed. The only viable alternative to all the above problems is ICT. ICT is really a boon in the life of every differently-abled person in this World. Text to Speech, Speech to Text and the Web Accessibility are the main research areas where the difficulties of differently-abled can be mitigated. There are some problems in the usage of ICT that create digital divide among the Differently-Abled that can be eliminated by giving proper attention at the right time. Since they can not use graphical oriented menu system that is triggered by the mouse, they have to use only Tab Keys and the Arrow keys to navigate the page. There is a chance of missing a menu item that leads to repeated pressing of Tab key in a cyclic way and fail to go to the required position and it may eat the productive time of them. They also have to remember all the short Cut Keys to select a menu item which also contribute to the development of Digital Divide. This is the right time to initiate any suitable measure to reduce those Digital Divide and to achieve 100% of utilization of web resources by the differently-abled students.

The Scope and coverage of the study

The Universities in India, usually offer courses that are recognized by the University Grants Commission of India (UGC) for the undergraduate, postgraduate, Master of Philosophy and Doctor of Philosophy programmes. The Colleges in India offer courses that are recognized by their affiliating universities. Those who have successfully completed their higher secondary level of school education are permitted to enroll in higher education. The differently-abled are also enrolling themselves for higher education either in a College or in a University. They are now-a-days very successful in Tamil Literature, English Literature, History, Sociology, Business Administration and Commerce.

It is obvious that the differently-abled students are scattered all around the country but this study covers the students who are presently engaged in their study at either a University or a College in Chennai particularly those pursuing their B. A., B. Ed., M. A., M. Ed., M. Phil., and PhD degree in Tamil Literature and who have consented to record their opinion on a particular date and time during a visit to their institution and to their hostels.

This study primarily tries to measure the level of digital divide created by ICT among the Persons with no sight and Persons with less sight, starting from the awareness about ICT, usage of assistive softwares, Tamil web resources etc.

The Methodology used in the Study

In order to conduct preliminary survey, model questions were selected from the popular web site WEBAIM₂ that conducts online survey for screen reader accessibility. Thirty questions were prepared to measure the Digital Divide among the Differently-Abled (Tamil Literature) students. The questionnaire to be filled was analyzed thoroughly to check for the awareness, availability, usability condition of the web resources before actual conduct of the survey. Some additional columns were given to record their personal data including name, roll number, course studying, date of birth, place of birth, father's occupation, mother's occupation, school last studied, demography, gender and the nature of disability to make it more reliable.

The questions present in the questionnaire were: Do you own a Personal Computer / Laptop with internet facility? Do you use internet regularly? Do you refer to any website for your study? How many Tamil websites are available for you to refer

your subjects matter? Do you use a screen reader due to a disability? Please rate your screen reader proficiency. Please rate your proficiency using the Internet. On which of the following devices, do you use a screen reader? Which operating system do you use in your desktop/laptop? Which of the following is your primary desktop/laptop screen reader? What is the main reason for using your primary screen reader? Which of the following visual features, if any, do you use? Which browser do you use most often with your screen reader? Have you ever customized your screen reader settings? Do you feel that free or low-cost screen readers (such as NVDA or VoiceOver) are viable alternatives to commercial screen readers? Do you use a screen reader on a mobile phone or mobile handheld device? Which of the following is your primary mobile platform? Which of the following mobile screen readers do you commonly use? Which of the following best describes your feelings regarding the accessibility of web content? In general, how accessible are social media web sites to you? How often do you navigate by landmarks in your screen reader? When trying to find information on a lengthy web page, which of the following do you use mostly? When a "skip to main content" or "skip navigation" link is available on a page, how often do you use it? CAPTCHAs are images presenting text used to verify that a user is a human user. How difficult are CAPTCHAs to you for usage? Do you visit Tamil websites using mobile phone? If yes, how do you rate the accessibility of Tamil websites in your mobile phone? For what purpose, do you visit Tamil web sites mostly? How will you know about any new Tamil web sites mostly? Do you exchange Tamil greetings in your email? If yes, which font do you use mostly? Please specify. Do you use any assistive device for your study? If yes, which device do you use mostly? Please specify.

The questions in the questionnaire were clearly explained to the student respondents to avoid any error in data collection. While observing data, the one to one approach, the direct personal interview method was adopted, to avoid influences of their neighbor. The Digital Divide can be measured through the questions like the use of screen readers, proficiency level of using screen readers, the ability to change the screen reader settings, proficiency level of using internet, usage of visual features, accessibility of social media web sites, the ability to identify CAPTCHAs, accessibility of tamil web sites, typing tamil greetings, use of screen readers in Mobile, etc. The IBM SPSS software package is used to analyze the opinion taken. The data entry was checked thoroughly and analyzed to understand the needs of the differently-abled and are narrated in the following section:

The Statistics

Twelve city colleges were considered for conducting the survey, as listed below with details:

Name of the College	Gender	Number of Students participated											
		Disability type		Course studying								Total	
				I B A	II BA	III BA	I MA	II MA	B. Ed	M. Ed	M. Phil		PhD
		No sight	Less sight										
Apollo College of Education	Female		1						1				1
	Male												
Chellamal College for women	Female	1	2		3								3
	Male												
Ethiraj College for Women	Female	1	1					2					2
	Male												
Govt Arts &	Female												

Name of the College	Gender	Number of Students participated											
		Disability type		Course studying									Total
				I B A	II BA	III BA	I MA	II MA	B. Ed	M. Ed	M. Phil	PhD	
Science College, Nandanam	Male	3	1	2	1		1						4
Loyola College	Female												
	Male	11	10	6	7	6		2					21
Madras Christian College	Female												
	Male		1									1	1
Pachayappa’s College	Female						1	3			1		5
	Male	4	1										5
Presidency College	Female		3		1	2							3
	Male	8	10	4	4	6		3			1		18
Quaid-E-Milleth College for Women	Female	1	2			2		1					3
	Male												
Queen Mary’s College	Female	14	13	17	4	6							27
	Male												
University of Madras	Female												
	Male	6	5				3	5		1	1	1	11
World Tamil Research Centre	Female	1									1		1
	Male												
Total		50	50										100

It is really fortunate to have the responses from the equal number of respondents, both the persons with no sight and the persons with less sight.

Summary table on the Digital Divide between Person with no sight and less sight

Measure of Digital Divide	Responses from Person with no Sight		Responses from Person with less Sight	
Own a Computer/ Laptop with Internet	21% YES	29% No	24% YES	26% No
Use internet regularly	22% YES	28% No	21% YES	29% No
Refer Websites for Academic Purposes	21% YES	29% No	24% YES	26% No
Tamil Web sites for Academic Purposes	0 20%, 1-5 22%, 6-10 3%, 11-20 1%, above 20 4%		0 19%, 1-5 22%, 6-10 3%, 11-20 3%, above 20 3%	
Use of screen reader	35% YES	15% NO	45% YES	5% NO
Screen Reader usage - proficiency level	7% High, 20% Moderate, 23% Low		8% High, 24% Moderate, 18% Low	
Internet usage- proficiency level	6% High, 16% Moderate, 28% Low		2% High, 25% Moderate, 23% Low	
The device where those screen readers are used mostly	13% None, 15% Desktop, 19% Laptop, 3% Mobile,		6% None, 14% Desktop, 25% Laptop, 5% Mobile	
The Operating System used	1% iOS, 49% Windows		2% Linux, 48% Windows	
The Screen Reader Used (Multiple Responses obtained)	33% JAWS, 28% NVDA		34% JAWS, 1% Voice Over, 24% NVDA, 2% Zoom Text	
The reason to choose a particular screen reader	14% Comfort, 7% Features, 1% Cost, 22% Availability, 6% Support		14% Comfort, 5% Features, 2% Cost, 20% Availability, 9% Support	
Visual features used by the persons with less sight	50% none		19% Screen Enlarger, 7% Increased browser text, 11% Browser Zoom, 13% High Contrast	
The browser used with the screen reader	41% Internet Explorer, 6% FireFox, 2% Safari, 1% Netscape		37% Internet Explorer, 10% FireFox, 2% Chrome, 1% Opera, 1% Netscape	
Their ability to change the settings of the screen reader	18% YES	32% NO	30% YES	20% NO
Is low cost screen readers are alternate to the commercial softwares	19% YES, 18% No, 13% I don't know		26% YES, 15% No, 9% I don't know	
The usage of screen readers in Mobile devices	12% YES	38% NO	18% YES	32% NO

The Primary Mobile Platform	2% Android, 3% Windows Phone, 45% Nokia		2% iPhone, 2% Android, 1% Windows Phone, 43% Nokia, 2% Blackberry	
The Mobile screen reader	31% None, 4% Voice Over, 5% Nuance Talk, 8% Mobile Speak, 1% Talkback for Android, 2% Orator		27% None, 5% Voice Over, 7% Nuance Talk, 7% Mobile Speak, 2% Talkback for Android, 1% IDEAL, 1% Orator	
Accessibility of Web Content	7% VV Good, 8% V Good, 10% Good, 15% Bad, 10% Very Bad		2% VV Good, 12% V Good, 14% Good, 11% Bad, 11% Very Bad	
Accessibility of Social Media Web sites	5% Very Much Accessible, 4% Some What Accessible, 3% Somewhat Inaccessible, 38% Very Inaccessible		1% Very Much Accessible, 12% Some What Accessible, 4% Somewhat Inaccessible, 33% Very Inaccessible	
The use of Landmarks	2% Very Often, 7% Often, 3% Sometimes, 38% Never		6% Very Often, 8% Often, 8% Sometimes, 28% Never	
The most usage pattern	37% Headings, 3% Find, 3% Links, 2% Landmark, 5% entire page		29% Headings, 9% Find, 5% Links, 2% Landmarks, 5% entire page	
Usage of “Skip to main content” option or “Skip Navigation”	8% Very Often, 4% Often, 14% Sometimes, 24% Never		9% Very Often, 4% Often, 19% Sometimes, 18% Never	
Tolerate level of CAPTCHA’s while authentication	41% VV difficult, 4% V Difficult, 3% Difficult, 2% Not Difficult		32% VV difficult, 12% V Difficult, 4% Difficult, 2% Not Difficult	
Accessing Tamil Webs in mobile phone	4% YES	46% NO	10% YES	40% NO
Accessibility of Tamil Web sites in Phone	46% No Access, 3% Access Partially, 1% Access little bit		40% No Access, 2% Access Completely, 4% Access Partially, 4% Access little bit	
The purpose of visiting Tamil Webs	24% Entertainment, 10% Academic, 10% Downloading, 6% News		15% Entertainment, 20% Academic, 6% Downloading, 9% News	
The propagators	8% Professors, 1% Lab Instructors, 37% Friends, 1% Parents, 3% Ads		10% Professors, 4% Lab Instructors, 28% Friends, 3% Parents, 5% Ads	
Tamil Greetings in Email	9% YES	41% NO	10% YES	40% NO
Assistive Devices for study	45% YES	5% NO	41% YES	9% NO

The Findings from the survey

This study reveals that 86% of the Differently-Abled students are using some kind of assistive devices for their study which is very much encouraging. 45% of them own a desktop/ laptop with internet facility, 43% of them are regularly using internet, 45% of them are referring internet for their academic purposes.

The Digital Divide measured from the above table is described below: The use of screen readers is 80%. The proficiency level of using screen readers are 15% High, 44% Moderate, 41% Low. The ability to change the screen reader settings are 48% Only. The proficiency level of using internet are 8% High, 41% Moderate, 51% Low. The usage of visual features are 50% Not Possible, 19% use Screen Enlarger, 7% use Increased browser text, 11% use browser zoom, and 13% use high contrast. The accessibility of social media web sites are revealed, 71% Very Inaccessible, 7% Somewhat inaccessible, 16% Somewhat Accessible, 6% Very much Accessible. The usability status of CAPTCHAs are 73% Very very difficult, 16% Very difficult, 7% Difficult, and 4% No difficult. The accessibility status of of tamil web sites are 14% Yes, 86% No. The use of tamil greetings in email are 19% Yes and 86% No. The use of screen readers in Mobile are 30% Yes and 70% No.

Apart from the survey, a personal interaction is also made to realize the status of Digital Divide among the Differently-Abled students. Most of them have agreed that ICT is the gift to reduce the divide that make them live on a par with others. They appreciated web sites like Madurai Project, Tamil Virtual University, etc, from where they get contents for their academic purposes. A few of them informed that the group mails including “Valluvan Parvai”, “Inaiya Thendral”, etc, created a new way of exchanging their idea, but most of them, are not aware of these group emails or some of them are not participating due to the non availability of a Computer with internet facility or the restrictions in the hostel as well as in the Colleges keep them away from their social networking groups. Most of them said that NVDA and eSpeak, the screen reading softwares, offering them an opportunity to learn from Tamil web sites but are not satisfying their expectations that is also creating some amount of Digital Divide. In order to attract the users, many web sites are using Graphical tools that are not accessible by the person with no sight. Too many frames, image maps, Java scripts, etc in a web page creates navigational problems. Lack of assistive softwares for non-unicode fonts made them inaccessible. Hence, the study proved that the ratio of Digital divide increased for the Differently-Abled to utilize the resources available now.

A few suggestions to mitigate disability to use ICT

The study reveals that the Person with Visual disabilities can not use many of the modern facilities available. In order to overcome those disability, the whole Tamil Computing community must follow the guidelines given by the World Wide Web (W3) Consortium like the Web Accessibility Guidelines₃ and the Web Content Accessibility Guidelines₄ to make web resources are accessible by everyone. The Differently-Abled are definitely depending on TAB keys, Short Cut Keys and the Arrow keys for navigation, the number of frames in a web page can be limited, the list of short cut keys and other navigational information can be added in the help menu that may reduce Digital Divide.

The Device Personalization₁ is a technique by which the type of device can be identified so that the web shall automatically generate exclusive content that is suitable for the device or selects the readymade web page that is suitable to view in that particular device. The Tamil Web sites may adopt such a technique to eliminate digital divide and hence, the resources can be fully utilized.

Future Direction

This survey is the first of its kind to measure the Digital Divide through availability, accessibility, usability of Tamil Web Resources for the differently-abled or the real users, undergoing higher education in Chennai, Tamilnadu, India. This survey can also be observed for the students in other parts of Tamilnadu or in India with a little modification in the questionnaire. It can also be modified to accommodate new facilities in Tamil Language that emerge as of now. The same survey can be extended to the members of the differently-abled faculty of Tamil literature to find their usage pattern of Tamil Web sites.

Acknowledgement

The authors are thankful to the management of Loyola College, Chennai, India for providing all the necessary facilities to conduct the survey smoothly. They are grateful to the Computer Society of India, Chennai Chapter for the financial assistance to present this paper in the 12th International Tamil Internet Conference. They also acknowledge the help rendered by the Principals, Head Of the Departments and the esteemed members of faculty of Tamil Department in Chennai City Colleges.

Conclusion

The study proved that there are many differently-abled (Tamil Literature) students in higher education, who rely on ICT for their academic purposes. It also revealed that there are difficulties in using the resources currently available that paved the way for the Digital Divide among differently-abled. The degree of Digital Divide may vary from person to person but can be eliminated by giving proper attention to the new and existing resources and offering periodical training to the users. It must be an ongoing process that the technical team should interact regularly with the students and make some remedy to meet the demands of the real users.

References

- Jerald Inico, J. and Edwin Prabakaran, T. (2012). “ Computing Intelligence to enhance Web Personalization”, “INTELLIGENT SYSTEMS” @ Computer Society of India COMMUNICATIONS Dec 2012, 10-11, 18.
- <http://webaim.org/projects/screenreadersurvey4/>, Screen Reader User Survey #4 Results, seen on June 10, 2013.
- <http://www.w3.org/WAI/eval/preliminary.html>, Easy Checks - A First Review of Web Accessibility updated on \$Date: 2013-07-09 12:29:50 \$
- <http://www.w3.org/TR/WCAG20/>, Web Content Accessibility Guidelines (WCAG) 2.0, seen on July 01, 2013.

தமிழ்மொழி கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தின் பயன்பாடு - மலேசியத் தமிழாசிரியர்களிடம் ஓர் ஆய்வு

முனைவர் சாமிக்கண்ணு ஜெபமணி ஈசாக்கு சாமுவேல்
விரிவுரையாளர், சுல்தான் இட்ரிசு கல்வியியல் பல்கலைக்கழகம்

முன்னுரை

மலேசியாவில் 532 தமிழ்ப்பள்ளிகள் செயல்படுகின்றன. இப்பள்ளிகளில் தமிழ்மொழிப் பயிற்று மொழியாக அங்கிகாரம் பெற்றுள்ளது. இப்பள்ளிகளில் மலாய், ஆங்கில மொழிப் பாடங்களைத் தவிர்த்து மற்ற பாடங்கள் அனைத்தும் தமிழ்மொழியிலேயே கற்பிக்கப்படுகின்றன. தமிழ்மொழிக் கற்றல் கற்பித்தலுக்குக் கலைத்திட்டமும் பாடநூல்களும் மலேசியக் கல்வி அமைச்சால் தயார் செய்யப்பட்டு வழங்கப்பட்டுள்ளன. பாட நூல்களைத் தவிர்த்து, தமிழ்மொழிக் கற்றல் கற்பித்தலுக்குத் தேவையான பாடத் துணைப்பொருட்கள் இணையத் தளத்தில் அதிகமாகக் காணப்படுகின்றன. இருப்பினும், தமிழாசிரியர்கள் தமிழ்மொழிக் கற்றல் கற்பித்தலுக்குத் தமிழ்மொழிப் பாடநூலையே அதிகமாகச் சார்ந்திருக்கின்றனர்.

அரசாங்கம் பெரும்பாலான தமிழ்ப்பள்ளிகளில் கணினி அறைகளை அமைத்துக் கொடுத்துள்ளது. பள்ளி நிர்வாகம் இக்கணினி அறைகளில் இணைய வசதிகளை ஏற்பாடு செய்து கொடுத்துள்ளது. தமிழாசிரியர்கள் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளை இணையத்தைப் பயன்படுத்தி மேற்கொண்டால் மாணவர்கள் சிறப்பாகக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் ஈடுபடுவதற்கு வாய்ப்பாக அமையும். தமிழாசிரியர்கள் இணையத்தைத் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளுக்கு எவ்வாறு பயன்படுத்துகின்றனர், அவற்றைப் பயன்படுத்துவதில் எவ்வாறான சிக்கல்களை எதிர் நோக்குகின்றனர் என்பதை ஆய்வு செய்வதே இக்கட்டுரையின் நோக்கமாகும். இவ்வாய்வு தஞ்சாவூர் மாலிம் வட்டாரத்தில் இடம்பெற்றுள்ள மூன்று தமிழ்ப்பள்ளிகளில் பணியாற்றுகின்ற 45 தமிழாசிரியர்களிடம் மேற்கொள்ளப்பட்டுள்ளது. மேலும் இம்மூன்று தமிழ்ப்பள்ளித் தலைமையாசிரியர்களிடமும் நேர்க்காணல் நடத்தப்பட்டது.

தர அடிப்படையிலான கலைத்திட்டம்

தொடக்கத் தமிழ்ப்பள்ளிகளுக்கான தர அடிப்படையிலான தமிழ்மொழிக் கலைத்திட்டத்தில் கேட்டல், பேச்சு, வாசிப்பு, எழுத்து ஆகிய அடிப்படை மொழித் திறன்களோடு செய்யுளும் மொழியணியும், இலக்கணம் ஆகிய கூறுகளும் இணைக்கப்பட்டுள்ளன. இத்திறன்கள் உள்ளடக்கத் தரம், கற்றல் தரம் என்று இரண்டு வகையாகப் பிரிக்கப்பட்டுள்ளன. கேட்டல் பேச்சு திறனில் 11 உள்ளடக்கத் தரமும் வாசிப்புத் திறனில் 8 உள்ளடக்கத் தரமும் எழுத்துத் திறனில் 10 உள்ளடக்கத் தரமும் செய்யுளும் மொழியணியும் திறனில் 13 உள்ளடக்கத் தரமும் இலக்கணத் திறனில் 9 உள்ளடக்கத் தரமும் இடம் பெற்றுள்ளன. இவ்வுள்ளடக்கத் தரங்கள் அனைத்தும் ஆண்டு வாரியாகப் பல்வேறு கற்றல் தரங்களை உள்ளடக்கியதாகப் பிரிக்கப்பட்டுள்ளன. இவ்வனைத்துக் கற்றல் தரங்களையும் இணையத்தின் வழி பல்வேறு நடவடிக்கைகளைப் பயன்படுத்திக் கற்பிக்க இயலும்..

ஆய்வு முடிவு

தர அடிப்படையிலான கலைத்திட்டத்தில் இடம்பெற்றுள்ள அனைத்துக் கற்றல் தரத்தையும் இணையத்தைப் பயன்படுத்திக் கற்பிப்பதற்கு வாய்ப்பு இருப்பினும், ஆசிரியர்கள் தங்கள் தமிழ்மொழிக் கற்பித்தலுக்கு இணையத்தை எந்த அளவுக்குப் பயன்படுத்துகின்றனர் என்று அறிய ஓர் ஆய்வு 45 தமிழாசிரியர்களிடம் மேற்கொள்ளப்பட்டது. ஆய்வின் முடிவுகள் அட்டவணை 1-இல் கொடுக்கப்பட்டுள்ளது.

கருத்தறி வினாக்கள் ஆம் இல்லை மொத்தம்

எண். % எண். %

1. தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்தலாம்.

43 96 2 4 45

2. தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்த பயிற்சி பெற்றுள்ளேன்

33 73 12 27 45

3. என்னுடைய தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்துகிறேன்

33 73 12 27 45

4. இணையத்தைப் பயன்படுத்தி கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொண்டால் மாணவர்கள் ஆர்வத்துடன் பயில்வார்கள் 45 100 0 0 45

5. கேட்டல் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்.

33 73 12 27 45

6. பேச்சுத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்

33 73 12 27 45

7. வாசிப்புத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்

39 87 6 13 45

8. எழுத்துத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்

36 80 9 20 45

9. இலக்கணத்தைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்

31 69 14 31 45

10. செய்யுளையும் மொழியணியையும் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்தலாம்

36 80 9 20 45

11. உங்களிடம் அகண்ட அலைவரிசை (broadband) உள்ளதா?

19 42 26 58 45

12. உங்களிடம் அகண்ட அலைவரிசை இருந்தால் அதனைப் பயன்படுத்தி தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைக்கு இணையத்தைப் பயன்படுத்துகிறீர்களா?

7 37 12 63 19

அட்டவணை 1 கருத்தறி வினா பகுப்பாய்வு

தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தின் பயன்பாடு

ஆய்வு மேற்கொள்ளப்பட்ட 45 ஆசிரியர்களில் 43 (96%) ஆசிரியர்கள் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்தலாமென தெரிவித்துள்ளனர். ஆனாலும் இவர்களில் 33 (73%) ஆசிரியர்கள் மட்டுமே தங்கள் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்துகின்றனர். இவர்களில் அநேகர் இணையத்தில் காணப்படும் பயிற்றுத் துணைப் பொருட்களை ஓய்வு நேரத்தில் பதிவிறக்கம் செய்து பின் அவற்றை வகுப்பறை கற்றல் கற்பித்தல் நடவடிக்கைகளில் பயன்படுத்துவதாகத் தெரிவித்துள்ளனர். இவ்வாய்வின் வழி, வகுப்பு நேரத்தில் இணையத்தைப் பயன்படுத்திக் கற்பிக்கும் நிலை பெரும்பாலான ஆசிரியரிடமும் இல்லை என்பது தெரியவருகிறது. வகுப்பு நேரத்தில் இணையத்தைப் பயன்படுத்திக் கற்பித்தால்தான் மாணவர்களை இணையத்தைப் பயன்படுத்தும் நடவடிக்கைகளில் ஈடுபடுத்த முடியும்.

ஆய்வு மேற்கொண்ட மூன்று பள்ளிகளிலும் இணையச் சேவை இருக்கின்றது. மேலும், ஒரு பள்ளியில் கணினி அறை இருக்கின்றது. அவ்வறையில் 18 பயன்படுத்தக்கூடிய கணினிகள் இருக்கின்றன. ஆகவே இணையத்தைப் பயன்படுத்தி தர அடிப்படையிலான கலைத்திட்டத்தில் இடம்பெற்றுள்ள அனைத்துக் கற்றல் தரத்தையும் கற்பிப்பதில் இம்மூன்று பள்ளிகளிலும் சிக்கல் இல்லை.

கேட்டல் - பேச்சுத் திறன்

கேட்டல் திறனைக் கற்பிப்பதற்கு இணையத்தளத்தில் போதுமான துணைப் பொருட்கள் இடம்பெற்றுள்ளன. உதாரணத்திற்கு, நான்காம் ஆண்டு பாடத்திட்டத்தில், 1.4 செவிமடுத்தவற்றைக் கூறுவர்; அதற்கேற்பத் துலங்குவர் எனும் உள்ளடக்கத் தரத்தில் இடம்பெற்றுள்ள 1.4.7 செவிமடுத்தவற்றைக் கோவையாகக் கூறுவர் என்ற கற்றல் தரத்தைக் கற்பிப்பதற்கு ஆசிரியர் இணையத்தளத்தில் இடம்பெற்றுள்ள உரை, விளம்பரம், செய்தி, கதை போன்ற ஏதேனும் ஒன்றைத் தேர்ந்தெடுத்து அதனை மாணவர்களைக் கேட்க செய்து பின் செவிமடுத்தவற்றைக் கோவையாகக் கூறச் செய்யலாம். இந்நடவடிக்கையை மேற்கொள்வதற்குக் கணினி அறை தேவையில்லை. வகுப்பில் ஆசிரியர் அகண்ட அலைவரிசையைப் பயன்படுத்தி இணையத்தளத்திலிருந்து மேற்கூறியவற்றில்

ஏதேனும் ஒன்றைத் தேர்ந்தெடுத்துப் பயன்படுத்தலாம். ஆனாலும், இவ்வாய்வில் 12 (27%) ஆசிரியர்கள் கேட்டல் - பேச்சுத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்த முடியாதென தெரிவித்துள்ளனர். இவர்களுக்குத் தகுந்த பயிற்சி வழங்கப்பட வேண்டும்.

வாசிப்புத் திறன்

கேட்டல் - பேச்சுத் திறனைப் போலவே, வாசிப்புத் திறனைக் கற்பிப்பதற்கும் இணையத்தளத்தில் போதுமான துணைப்பொருட்கள் இடம்பெற்றுள்ளன. உதாரணத்திற்கு, நான்காம் ஆண்டு பாடத்திட்டத்தில், 2.8 வாசித்துத் தகவல்களைச் சேகரிப்பர் எனும் உள்ளடக்கத் தரத்தில் இடம்பெற்றுள்ள 2.8.3 பனுவல்களை வாசித்துத் தகவல்களைச் சேகரிப்பர் என்ற கற்றல் தரத்தைக் கற்பிப்பதற்கு ஆசிரியர் அகப்பக்கத்தில் இடம்பெற்றுள்ள ஏதேனும் பனுவலைத் தேர்ந்தெடுத்து வாசிக்கச் செய்து, அப்பனுவலில் இடம்பெற்றுள்ள தகவலைச் சேகரிக்கப் பணிக்கலாம் அல்லது கொடுக்கப்பட்டுள்ள வினாக்களுக்கு விடையளிக்கக் கூறலாம். இதனால், மாணவர்கள் ஆர்வத்துடன் கற்றல் கற்பித்தல் நடவடிக்கைகளில் கலந்துகொள்வார்கள். ஆனாலும், இவ்வாய்வில் 6 (13%) ஆசிரியர்கள் வாசிப்புத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்த முடியாதென தெரிவித்துள்ளனர்.

எழுத்துத் திறன்

எழுத்துத் திறனைக் கற்பிப்பதற்கும் இணையத்தளத்தில் போதுமான வசதிகள் இருக்கின்றன. உதாரணத்திற்கு நான்காம் ஆண்டு பாடத்திட்டத்தில், 3.10 பல்வகை வடிவங்களைக் கொண்ட எழுத்துப் படிவங்களைப் படைப்பர் எனும் உள்ளடக்கத் தரத்தில் இடம்பெற்றுள்ள 3.10.9, 80 சொற்களில் நட்புக் கடிதம் எழுதுவர் என்ற கற்றல் தரத்தைக் கற்பிப்பதற்கு ஆசிரியர் மாணவர்களைக் கொடுக்கப்பட்ட தலைப்பில் நட்புக் கடிதத்தைக் கணினியில் தட்டச்சுச் செய்து சக மாணவருக்கு மின்னஞ்சல் அனுப்பப் பணிக்கலாம். சக மாணவர் மின்னஞ்சலைத் திறந்து கடிதத்தை வாசித்துத் தன் நண்பனுக்குப் பதில் கடிதம் தட்டச்சுச் செய்து மின்னஞ்சல் அனுப்பலாம். இந்நடவடிக்கை வழி மாணவர்கள் கடிதம் எழுதும் நடவடிக்கையை இணையம் வாயிலாகவே மேற்கொள்ள வாய்ப்பாக அமைகிறது. ஆனாலும், இவ்வாய்வில் 9 (20%) ஆசிரியர்கள் எழுத்துத் திறனைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்த முடியாதென தெரிவித்துள்ளனர்.

செய்யுளும் மொழியணியும்

செய்யுளையும் மொழியணியையும் கற்பிப்பதற்கும் இணையத்தளத்தைத் தாராளமாகப் பயன்படுத்தலாம். உதாரணத்திற்கு, நான்காம் ஆண்டு பாடத்திட்டத்தில், 4.13 பழமொழிகளின் பொருளை அறிந்து சரியாகப் பயன்படுத்துவர் எனும் உள்ளடக்கத் தரத்தில் இடம்பெற்றுள்ள 4.13.4 நான்காம் ஆண்டுகான பழமொழிகளின் பொருளை அறிந்து சரியாகப் பயன்படுத்துவர் என்ற கற்றல் தரத்தில் இடம்பெற்றுள்ள “சிறு துரும்பும் பல் குத்த உதவும்” என்ற பழமொழியைக் கற்பிப்பதற்கு ஆசிரியர் இணையத்தளத்தில் இடம்பெற்றுள்ள சிறுவர் கதைப் பகுதிக்குச் சென்று அப்பகுதியிலுள்ள “சிங்கமும் சுண்டெலியும்” எனும் கதையைத் தேர்ந்தெடுத்து ஒளிப்பரப்பி மாணவர்களுக்குப் பழமொழியின் பொருளைக் கதை வழி உணர்த்தி பழமொழியின் பொருளை அறிந்து சரியாகப் பயன்படுத்தக் கற்பிக்கலாம். இந்நடவடிக்கையை மேற்கொள்வதற்குக் கணினி அறை தேவையில்லை. வகுப்பில் ஆசிரியர் அகண்ட அலைவரிசையைப் பயன்படுத்தி இணையத்தளத்திலிருந்து பதிவிறக்கம் செய்து பயன்படுத்தலாம். ஆனாலும், இவ்வாய்வில் 9 (20%) ஆசிரியர்கள் செய்யுளையும் மொழியணியையும் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்த முடியாதென தெரிவித்துள்ளனர்.

இலக்கணம்

இலக்கணம் கற்பிப்பதற்கும் இணையத்தளத்தில் நிறைய மென்பொருள்கள் உள்ளன. ஆசிரியர்கள் இவற்றை அடையாளங்கண்டு தேவைக்கு ஏற்ப பயன்படுத்தலாம். உதாரணத்திற்கு உயிர் எழுத்துகளைக் கற்பிப்பதற்கு இணையத்தளத்தில் ‘அம்மா இங்கே வா வா’ என்ற ஒரு பாடல் காட்சியுடன் இடம் பெற்றுள்ளது. ஆசிரியர்கள் இக்காணொலியைப் பதிவிறக்கம் செய்து உயிரெழுத்துகளைக் கற்பிக்கும் போது பயன்படுத்தலாம். இதுபோலவே இலக்கணக் கூறுகளைக் கற்பிப்பதற்கு நிறைய பாடத் துணைப்பொருட்கள் இணையத்தில் இடம்பெற்றுள்ளன. இவற்றை ஆசிரியர்கள் முழுமையாகப் பயன்படுத்திக்கொள்ள வேண்டும். ஆனாலும், இவ்வாய்வில் 14 (31%) ஆசிரியர்கள் இலக்கணத்தைக் கற்பிப்பதற்கு இணையத்தைப் பயன்படுத்த முடியாதென தெரிவித்துள்ளனர்.

தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்துவதில் ஏற்படும் சிக்கல்கள்

தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்துவதில் ஆசிரியர்கள் பல்வேறு சிக்கல்களை எதிர் நோக்குகின்றனர். முதலாவதாக, தமிழாசிரியர்களுக்குப் போதுமான பயிற்சிகள் வழங்கப்படவில்லை. ஆய்வு மேற்கொள்ளப்பட்ட 45 ஆசிரியர்களில் 12 (27%) ஆசிரியர்கள் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்த எந்தவிதமான பயிற்சியும் பெறவில்லை. மேலும், தலைமையாசிரியர்களிடம் மேற்கொள்ளப்பட்ட நேர்காணல் வழி ஆய்வு மேற்கொள்ளப்பட்ட மூன்று பள்ளி

நிர்வாகமும் தங்களுடைய தமிழாசிரியர்களுக்கு இணையம் தொடர்பான எந்தப் பயிற்சியையும் ஏற்பாடு செய்யவில்லை என்பது தெரியவருகிறது. தொடர்ந்து, ஆய்வு மேற்கொள்ளப்பட்ட மூன்று தமிழ்ப்பள்ளிகளிலும் போதுமான நீர்மபடிம உருகாட்டி (LCD) இல்லாததால் ஆசிரியர்கள் இணையத்தைப் பயன்படுத்தித் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொள்வதில் சிக்கலை எதிர்நோக்குகின்றனர் என்பதும் தெரியவருகிறது.

இதனைத் தவிர்த்து, நேரப் பற்றாக்குறையின் காரணமாகவும் ஆசிரியர்கள் இணையத்தைப் பயன்படுத்தி தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொள்வதைத் தவிர்கின்றனர். மாணவர்களைக் கணினி அறைக்கு அழைத்துச் சென்று, கணினியை முடுக்குவதில் பாட வேளையில் பாதி நேரம் விரையமாவதால் இணையத்தைப் பயன்படுத்துவதில்லையென்று ஆசிரியர்கள் கருத்துத் தெரிவித்துள்ளனர். தொடர்ந்து ஆய்வு மேற்கொள்ளப்பட்ட இரண்டு பள்ளிகளில் கணினியறை இல்லை. இப்பள்ளிகளில் பணியாற்றும் ஆசிரியர்கள் பள்ளியில் போதுமான வசதியில்லை என்பதால் இணையத்தைப் பயன்படுத்திக் கற்பிப்பதில் சிக்கல் ஏற்படுகிறது என கருத்துத் தெரிவித்துள்ளனர். ஆய்வு மேற்கொள்ளப்பட்ட 45 ஆசிரியர்களில் 19 (42%) ஆசிரியர்கள் பிராட்பேண்டு வைத்துள்ளனர். இவர்கள் முயன்றால் தங்களுடைய பிராட்பேண்டைப் பயன்படுத்தி இணையம் வழி தங்கள் கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொள்ளலாம்.

முடிவுரை

ஆய்வில் இடம்பெற்றுள்ள அனைத்து ஆசிரியர்களும் இணையத்தைப் பயன்படுத்தி கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொண்டால் மாணவர்கள் ஆர்வத்துடன் பயில்வார்கள் என்ற கருத்தை ஏற்றுக்கொண்டுள்ளனர். ஆகவே, ஆசிரியர்கள் தங்கள் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடவடிக்கைகளில் இணையத்தைப் பயன்படுத்தி மாணவர்களின் கற்றல் திறனை மேம்படுத்தினால் தமிழ்மொழியின் அடைவு நிலை உயரும் என்பது திண்ணம்.

மேற்கோள் நூல்கள்

- தர அடிப்படையிலான கலைத்திட்டம். (2010). கலைத்திட்ட மேம்பாட்டுப் பிரிவு, புத்ர ஜெயா: கல்வி அமைச்சு.
- ரவிந்திரன் தனபால் & சரஸ்வதி கண்ணன். (2010). தமிழ்மொழிப் பாட நூல், ஆண்டு ஒன்று, **Petaling Jaya: Multi Educational Book Enterprise.**
- ரவிந்திரன் தனபால் & சரஸ்வதி கண்ணன். (2011). தமிழ்மொழிப் பாட நூல், ஆண்டு ஒன்று, **Petaling Jaya: Multi Educational Book Enterprise.**
- சரஸ்வதி கண்ணன் & சந்திரன் வீரமுத்து. (2012) தமிழ்மொழிப் பாட நூல், ஆண்டு மூன்று, **Petaling Jaya: Multi Educational Book Enterprise.**

Computer-based and Paper-based Vocabulary Games for Tamil Learners

Paramasivam Muthusamy¹, Kanthimathi Letchumanan² and Tan Bee Hoon³

¹*Faculty of Modern Languages and Communication, UPM, param@fbmk.upm.edu.my*

²*SMK Seri Kembangan, kanthi65@hotmail.com*

³*Faculty of Modern Languages and Communication, UPM, tanbh@fbmk.upm.edu.my*

ABSTRACT

This study aimed to investigate if language games via computer or paper helps to expand the English vocabulary size of Form Four students especially at the 2000 word-level. Two intact classes were used as the experimental groups, and they played computer-based and paper-based games as learning activities. The study was conducted in three stages. First, the respondents sat for the Productive Vocabulary Levels Test (PVLТ) as a pre-treatment test. Then the experimental group 1 played the computer-based games and the experimental group 2 paper-based games for seven weeks and sat for the first post-treatment parallel PVLТ. Then both groups switched treatments and played the respective games for another seven weeks before being assessed by the second parallel PVLТ (post-treatment test 2). The pre and post-treatment tests scores were analysed statistically by t-test. The results show a significant difference for both modes of games, but the computer-based games had a higher mean difference than the paper-based games. Therefore, the study provided evidence that the computer-based games had better impact on students' vocabulary enhancement than the paper-based games. Thus, computer-based vocabulary games can also be used and practiced among Tamil Learners at both primary and secondary schools in enhancing Tamil vocabulary for learners.

Keyword: vocabulary, computer-based games, paper-based games, productive vocabulary levels test

Introduction

Words are tools we use to communicate with others, as well as to explore and analyze the world around us. Most people who have learned a language and tried to use it for their daily tasks are aware of how insufficient vocabulary retards their communication (Raimes, 1985) and this is usually the case of less proficient writers who are unable to express their ideas due to insufficient vocabulary. As a result, their messages are not delivered. The most obvious effect of an under-developed vocabulary will be seen in the productive skills of speaking and writing (Read, 2000), and learners will also struggle to extract meaning from reading passages. Therefore, vocabulary learning or acquisition is increasingly viewed as crucial to language learning.

Research Objective

The objective of this study is to investigate if learners are able to enhance vocabulary through language games using two different modes via computer and paper based games. Students' vocabulary enhancement was measured by the Productive Vocabulary Levels Test (PVLТ). The computer-based games used in this study are those that are free downloadable from the internet due to the realization that not all learners have Internet facilities at home and the MySchoolNet is not always accessible to learners. Besides the computer-based games, this study also uses paper-based games which are provided on gamesheets to learners. Both the computer and paper-based games include crossword puzzles, word maze, mystery games, identifying jumbled words, matching exercises and solving riddles.

Research Design

This was a quasi experiment that used a counterbalanced design to compare the effects of two modes of language games. Two intact classes were used as experimental groups as both groups were exposed to the treatments. The experimental group 1 received treatment A (Computer-based Games) first followed by Treatment B (Paper-based Games), while the experimental group 2 were exposed to Treatment B first followed by Treatment A. At the end of the experiment, both the experimental groups had been exposed to both treatments of learning vocabulary: computer and paper-based games. Each experimental group consisted of thirty-five subjects.

Research Procedure

The study was conducted in three stages. At first a Pre-treatment Test on vocabulary was given to all the respondents to determine their current state of vocabulary knowledge. The subjects were given the Productive Vocabulary Levels Test (PVLТ) from the 2000 word-level based on West's (1953), General Service List (GSL). In the test, the initial letters were provided and respondents were required to complete the words based on the context of the sentences. After the Pre-treatment Test, the experimental group 1 played the computer-based games and the experimental group 2 paper-based games. Both the groups were exposed to the treatments for seven weeks. After the seventh week, the first Post-treatment was conducted which comprised a parallel Productive Vocabulary Levels Test. After the first Post-treatment, the experimental group 1 and group 2 switched treatments, whereby the experimental group 1 played the paper-based games and the experimental group 2 played the computer-based games. Both the groups underwent the treatments for another seven weeks. After the fourteenth week,

another parallel assessment (Post-treatment Test 2) on vocabulary was administered to both the groups. The results of these assessments were used to compare the effects of both treatments on the respondents' vocabulary performance.

Results and Discussion

The vocabulary level of the respondents at three different stages was tested. The pre-treatment test result shows the mean vocabulary test score of Group 1 ($M=11.11$, $SD=2.742$) and Group 2 ($M=11.71$, $SD=2.607$). It indicates that there is no significant difference between the two groups at the beginning of the treatment. Hence, both the experimental groups were deemed homogenous in their vocabulary level.

After pre-treatment, Group 1 played computer-based games and Group 2 played paper-based games for seven weeks. After Group 1 played computer-based games, there was a significant difference in the result of the PVLTL ($MD=3.60$, $SD=1.98$, $t(35)=10.71$, $p=.000$) as shown in Table 1. The mean score is greater ($M=14.71$) than Pre-treatment Test ($M=11.11$). A similar comparison was made with the Post-treatment Test 2 (see Table 2) after they played paper-based games where there was a significant difference in the mean scores ($MD=0.83$, $SD=2.24$, $t(35)=2.19$, $p=.036$). The mean score of Post-treatment Test 2 shows that their mean score decreased slightly ($M=13.89$) when compared to their Post-treatment Test 1. Thus, the result shows that although there is a significant difference for Group 1 after both treatments, they show better improvement in their PVLTL after they played computer-based games (Post-treatment Test 1) since their mean score is higher ($M=14.71$) than paper-based games (Post-treatment Test 2).

Table 1: Changes in Productive Vocabulary Levels Test of Group 1 after Playing Computer- based Games (n=35)

Table	Pre Treatment Test Mean	Post Treatment Test 1 Mean	Mean Difference	t-value	SD	p-value $p<.05$	2
	11.11	14.71	3.60	10.71	1.98	.000	

Changes in Productive Vocabulary Levels Test of Group 1 after Playing Paper-based Games (n=35)

Post Treatment Test 1 Mean	Post Treatment Test 2 Mean	Mean Difference	t-value	SD	p-value $p<.05$
14.71	13.89	0.83	2.19	2.24	.036

Likewise, when Group 2 who played paper-based games first, their PVLTL shows that there was significant difference in their PVLTL ($MD=1.66$, $SD=2.58$, $t(35)=3.80$, $p=.001$) as shown in Table 3. The PVLTL of Group 2 shows that their mean score is greater ($M=13.37$) when compared with their Pre-treatment Test ($M=11.71$). A similar comparison were made with the Post-treatment Test 2 after they played the computer-based games where there was also a significant difference in the scores ($MD=3.23$, $SD=2.42$, $t(35)=7.89$, $p=.000$). Between the two treatments, the result shows that Group 2 has better gain in their PVLTL after they have played the computer-based games as shown by their mean score ($M=16.60$) to paper-based games ($M=13.37$). Thus, the results show that Group 2 also has shown an improvement in their PVLTL at each mode of vocabulary games and have gained in their productive use of vocabulary (see Table 4). Therefore, the findings suggest that both the groups had better improvement in their productive use of vocabulary after playing the computer-based games.

Table 3: Changes in Productive Vocabulary Levels Test of Group 2 after Playing Paper-based Games (n=35)

Table	Pre Treatment Test Mean	Post Treatment Test 1 Mean	Mean Difference	t-value	SD	p-value $p<.05$	4:
	11.71	13.37	1.66	3.80	2.58	.001	

Changes in Productive Vocabulary Levels Test of Group 2 after Playing Computer-based Games (n=35)

Post Treatment Test 1 Mean	Post Treatment Test 2 Mean	Mean Difference	t-value	SD	p-value $p<.05$
13.37	16.60	3.23	7.89	2.42	.000

Conclusion

The Productive Vocabulary Levels Test was conducted to assess the respondents' ability to use a target word correctly from the 2000 word-level. The result shows that both the experimental groups had attained active productive vocabulary use at the 2000 word-level shown by the mean differences between pre-treatment test scores and the post-treatment test scores after they played computer-based and paper-based games (see Table 1 - Table 4). Besides, the study also shows that both the experimental groups have shown better improvement in their productive use of vocabulary after playing computer-based games. This difference on the mode of games could be due to the nature of computer-based games which were more interactive allowing the respondents to use multisensory elements, text, sound, pictures and animation thus providing meaningful contexts to facilitate comprehension (Butler-Pascoe & Wilbur, 2003). Thus, it may be concluded that language games were able to increase the vocabulary size of the respondents. This is because the nature of repetitiveness in the games allowed the respondents to give informational feedback regarding the games. Feedback evokes the correct behavior, thoughts and actions of the players (Kapp, 2012).

References

- Butler-Pascoe, M. E., & Wilbur, K.M. (2003). *Technology and teaching English language learners*. Boston: Pearson Education.
- Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly*, 19 (2), 229-225.
- Read, J. (2000). *Assessing vocabulary*. United Kingdom: Cambridge University Press.
- Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education*. San Francisco: Pfeiffer
- West, M. (1953). *A general service list of English words*. London: Longman, Green and Company.

மொழி விளையாட்டுகளும் சொற்கள் வளப்படுதலும்

முனைவர் அ.பரிமளகாந்தம்,

இணைப்பேராசிரியர், அகராதியியல் துறை, பி.எஸ். தெலுங்குப் பல்கலைக்கழகம், ஐதராபாத் – 500 004.

முன்னுரை

மொழி, சமுதாயப் பிராணியான மனிதனின் கருத்துக்களை, சிந்தனைகளை தன்னுடன் வாழும் தன்னை ஒத்த மனிதனுக்கு தெரியப்படுத்த உதவும் ஒரு அருமையான தொடர்பு சாதனமாகும். அத்தோடு தன்னையொத்த மனிதனுடன் நல்லுறவுடன் வாழ வைக்கும் சாதனமாகவும் விளங்குகின்றது. நவீனமயமான, உலகமயமான இன்றைய சமுதாயத்தில் கருத்துப் பரிமாற்றத்திற்குக் காரணமான மொழி கற்றலின், கற்பித்தலின் முக்கியத்துவம் அதிகமாகிவிட்டது. தொடக்க காலத்தில் மொழியைக் கற்பிக்க குருகுலங்களும், அதன்பின்பு பள்ளிகளும், ஆசிரியர்களும் இன்றியமையாதவைகளாகியன. தற்போது தொழில் நுட்பத்தில் ஏற்பட்ட மறுமலர்ச்சியால் ஆசிரியர்களோ பாடசாலைகளோ இல்லாமல் மொழி கற்கும் வசதியும் வாய்ப்பும் ஏற்பட்டுள்ளது, அத்தகைய வசதிகளில் ஒன்று கணினி. கணினியின் ஒரு பாகமான இன்டர்நெட் மூலம் மொழி கற்க வசதி செய்யப்பட்டுள்ளது.

கட்டுரையின் நோக்கம்

இக்கட்டுரையின் நோக்கம் கணினி இன்டர்நெட் மூலம் மொழியைக் கற்பதற்கு, குறிப்பாகச் சொற்களைக் கற்பதற்கு உள்ள முறைகளில் ஒன்றான விளையாட்டு முறை பற்றி விவரிப்பதேயாகும்

உடலுக்கு ஆரோக்கியத்தையும், மனத்திற்கு உற்சாகத்தையும் ஏற்படுத்தும் விளையாட்டு, மொழி கற்பித்தலிலும் முக்கியப் பங்கு வகிக்கிறது. மொழியின் அனைத்து பாகங்களையும் விளையாட்டின் மூலம் கற்றுக் கொள்ளலாம் என்றாலும் குறிப்பாக மரபுத் தொடர், எதிர்ப்பதச்சொல், தொகைச்சொல், பலசொல் ஒரு பொருள், பல பொருள் ஒரு சொல் போன்றவைகளைக் கற்றுக் கொள்ள சில விளையாட்டுகளை இக் கட்டுரை குறிப்பிட்டு, அதன் செயல்பாட்டை விவரிக்கிறது. இந்த முறைக்கு, மூலமாக இன்நெட்டிலுள்ள ஆங்கிலச் சொற்களை வளர்த்துக் கொள்ள ஏற்படுத்தியுள்ள சில விளையாட்டு முறைகளை எடுத்துக்காட்டாகக் கொண்டு, தமிழ் மொழிக்கு அவ்விளையாட்டு முறைகளைப் பயன்படுத்த எடுத்துக் கொண்ட முயற்சியே இக்கட்டுரையின் உருவாக்கம்.,

விளையாட்டு

பொதுவாக விளையாட்டு பிள்ளைகளின் உடலை நலத்துடனும் மனத்தை உற்சாகத்துடனும் வைத்திருக்க உதவும். எனவே தான் பள்ளிகளில் பிள்ளைகளுக்கு விளையாட நேரத்தை ஒதுக்கிறார்கள். பிள்ளைகளுக்கு அதாவது மாணவர்களுக்கு ஓய்வு நேரம் கிடைத்தால் அநாவசியான விசயங்களைப் பேசுவதற்காகவும், புறங்கூறுதல் அல்லது அல்ப விசயங்களுக்காக சண்டை போட்டுக் கொள்ளவும் பயன்படுத்தக் கூடும். இந்தச் செயல்கள் பயனற்றவை மட்டுமல்லாமல் பள்ளிகளுக்கும் பெற்றோர்களுக்கும் தலைவலியை ஏற்படுத்தும். அத்துடன் மாணவர்களின் குணநலன்களையும் பாதிக்கும். ஆகையால் ஓய்வு நேரங்களில் சில பயனுள்ள விளையாட்டுகளை விளையாடச் செய்தால் அவை மனதுக்கு மகிழ்ச்சியையும் அறிவையும் வளர்த்துக் கொள்ளவும் உதவுவதாக அமையும். விளையாட்டுகள் பலவிதம். உடலுக்கு ஆரோக்கியத்தையும் மனத்திற்கு மகிழ்ச்சியையும் உற்சாகப் படுத்தக் கூடிய விளையாட்டுக்கள் கால்பந்து, கூடைப் பந்து, கபடி, ஓட்டப் பந்தயம், கில்லி, கிரிகெட், டென்னிஸ் போன்றவையாகும். அறிவை வளர்க்கும் விளையாட்டுக்கள் மொழி விளையாட்டு, செஸ் போன்றவையாகும்.

மொழி விளையாட்டு

அத்தகைய விளையாட்டுகளில் முக்கியமானது மொழி கற்றுக் கொள்ளும் விளையாட்டு. இவ் விளையாட்டுகள் மாணவரின் மொழியை அதாவது பேசுதல், கேட்டல், படித்தல், எழுதுதல் போன்ற திறமைகளையும் சொற்களையும், வாக்கியங்களையும், இலக்கணத்தையும், வளர்த்துக் கொள்ள மட்டுமல்லாமல் மனத்தை உற்சாகமாக வைத்து யோசிக்கும் திறனையும் படைப்புத் திறனை வளர்த்துக் கொள்ளவும் உதவும். அத்துடன் மேற்கூறிய மற்ற அநாவசியமான விசயங்களைச் செய்யாமல் தடுக்கவும் உதவும். விஞ்ஞானமும் தொழில் நுட்பமும் வளர்ச்சியடையும் முன் பிள்ளைகள் தங்கள் தோழர்களுடன் மொழி விளையாட்டை விளையாடினார்கள். ஆனால் விஞ்ஞானமும் தொழில் நுட்பமும் வளர்ச்சியடைந்த இக்காலத்தில் தோழர்கள் இல்லாமலேயே விளையாடும் வசதி ஏற்பட்டு இருக்கிறது.

இண்டர்நெட் மூலம் மொழி விளையாட்டு

விஞ்ஞானமும் தொழில் நுட்பமும் வளர்ச்சியுற்ற இக்காலத்தில் கணினி வழியாக இண்டர்நெட்டில் மொழியைக் கற்றுக் கொள்ள குறிப்பாக சொற்களை வளப்படுத்திக் கொள்ள நல்ல வாய்ப்பும் வசதியும் ஏற்பட்டு இருக்கிறது. இந்த விளையாட்டுகளை விளையாட தோழர்கள் தேவையில்லை. ஆசிரியர் தேவையில்லை. தேவையான விதி முறைகளுடன் விளையாட்டுகள் உள்ளன. ஆகையால் இண்டர்நெட்டைத் திறந்து விளையாட்டுப் பகுதியில் நுழைந்தவுடன் பிள்ளைகளுக்கு அந்த விளையாட்டுகளை விளையாடவேண்டும் என்ற ஆர்வம் தன்னிச்சையாக ஏற்பட்டு விடும்.

இண்டர்நெட்டில் ஆங்கிலச் சொற்களை வளப்படுத்திக் கொள்ள பல விளையாட்டுகள் கொடுக்கப்பட்டுள்ளன. அவை ஒப்புமை விளையாட்டு, எதிர்பதச் சொல் விளையாட்டு, தொகைச் சொல் விளையாட்டு, சூழ்நிலை விளையாட்டு, சுருக்கமொழி விளையாட்டு, ஆங்கில மொழி விளையாட்டு, அயல் மொழி விளையாட்டு, ஹிக்பிக் விளையாட்டு, ஒலி ஒன்றி விளையாட்டு, மரபுத் தொடர் சொல் விளையாட்டு, இலத்தீன் மொழி விளையாட்டு, இலக்கிய மொழி விளையாட்டு, மொழி பகுதி விளையாட்டு, ஒலி/தவனி விளையாட்டு, முன் ஒட்டு விளையாட்டு, வேர்ச் சொல் விளையாட்டு, சேட் விளையாட்டு, எழுத்து விளையாட்டு, பின் ஒட்டு விளையாட்டு, அசைச் சொல் விளையாட்டு, ஒரு பொருள் பல சொல் விளையாட்டு, டைப்பிங் விளையாட்டு, சொல் விளையாட்டு போன்ற பல விளையாட்டுகள் கொடுக்கப்பட்டுள்ளன.

இவற்றில் கட்டுரையாளர் ஒப்புமை விளையாட்டு, எதிர்பதச் சொல் விளையாட்டு, தொகைச் சொல் விளையாட்டு, மரபுத் தொடர் சொல் விளையாட்டு, ஒரு பொருள் பல சொல், எழுத்து விளையாட்டு போன்ற விளையாட்டுகளை மட்டும் தேர்ந்தெடுத்து அவற்றைத் தமிழ் மொழியை வளப் படுத்த இக் கட்டுரையில் பயன்படுத்தி உள்ளார்.

விளையாடும் முறை

முதலில் இண்டர்நெட்டில் ஆங்கில மொழி வளத்திற்கு எவ்வாறு விளையாட்டு அமைக்கப்பட்டுள்ளது, அதை எவ்வாறு விளையாடுவது என்பதை விவரித்து, அதன் அடிப்படையில் தமிழுக்கு ஒரு சில விளையாட்டுகள் அமைக்கப்பட்டுள்ளது.

ஆங்கிலத்தில் தமிழ் சொற்களை வளப்படுத்த தேர்ந்தெடுத்துக் கொண்ட ஒவ்வொரு விளையாட்டும் ஆரம்ப நிலையிலிருந்து இண்டர்மீடியட் நிலை வரை ஆறு பகுதிகளாக பிரிக்கப்பட்டுள்ளது. ஒவ்வொரு விளையாட்டும் பல வகைகளை உடையது. உதாரணமாக ஒப்புமை விளையாட்டில் ஒப்புமை பல பொருள் ஒரு சொல், ஒப்புமை (parody is to satire as rogue is to scoundrel) இன்னும் பகுதி-முழுமை ஒப்புமையுடன் (part and whole analogy) சேர்த்துப் (Eiffel Tower .is to France as Great Pyramid is to Egypt) பல வகைகள் உள்ளன. ஒப்புமை விளையாட்டு 3rd to 5th grade ஒப்புமை விளையாட்டு, 6th to 8th grade ஒப்புமை விளையாட்டு இன்னும் 9th to 12th grade ஒப்புமை விளையாட்டு. மேலும் 3rd to 5th grade ஒப்புமை விளையாட்டில் இரண்டு பகுதிகள் உள்ளன, அவை ஆரம்ப நிலையும் அதற்கு மேற்பட்ட நிலையுமாகும், 6th to 8th grade ஒப்புமை விளையாட்டில் ஆரம்ப நிலை விளையாட்டும் நடு நிலை பள்ளி விளையாட்டு என்ற இருபகுதிகளும், 9th to 12th grade ஒப்புமை விளையாட்டில் உயர்நிலைப் பள்ளி ஒப்பமை பொருத்தமும் உயர்நிலைப் பள்ளி ஒப்புமை என்ற இரண்டு பிரிவும் உள்ளன. ஒவ்வொரு பிரிவிலும் எட்டிலிருந்து பத்து சொற்கள் இருக்கும். இதற்கு மாணவர் சரியான ஒப்புமையைக் கண்டு பிடிக்க வேண்டும். இந்த ஒப்புமைகளில் ஆரம்ப பள்ளி ஒப்புமை விளையாட்டு நகைச்சுவையாகவும், நடு நிலைப்பள்ளி ஒப்புமை விளையாட்டு analogical reasoningகாவும், மேல் நிலைப் பள்ளி ஒப்புமை விளையாட்டு ஒப்புமை பிரச்சனையை தீர்ப்பதாகவும் அமைந்துள்ளது. உதாரணமாக ஆரம்ப நிலை ஒப்புமை விளையாட்டை ஆங்கிலத்தில் பார்க்கலாம்.

ஆங்கிலத்தை இண்டர்நெட் ஒப்புமை விளையாட்டு அமைப்பு கீழ்க்கண்ட விதமாக இருக்கிறது.

Which one

Instructions

start

Back

start

என்று இருக்கிறது. எப்படி விளையாடவேண்டும் என்று தெரிந்து கொள்ள Instructionsஐக் கிளிக் செய்தால் Select the answer that matches best என்று வரும். பின்பு Back என்பதைக் கிளிக் செய்து முதல் இடத்திற்குத் திரும்பி வந்து start

பட்டனைக் கிளிக் செய்தால் நாம் விளையாட வேண்டிய கீழ் கண்ட அமைப்பு வரும். இந்த அமைப்பில் மாணவன் சரியான ஒப்புமையைக் கிளிக் செய்யவேண்டும்.

Choose the answer that best completes the analogy

Water is to boil as

Snow is to freezes

Temperature is to cold

Ice is to melt

Rain is to evaporates

Check

Next

(The correct answer is water is to boil as ice is to melt)

பிறகு 'next' என்ற வார்த்தையைக் கிளிக் செய்ய விளையாட்டுத் தொடரும்

Goldfish is to aquarium as

Barn is to house

horse is to barn

Form is to silo

garage is to room

Check

Next

(goldfish is to aquarium as horse is to barn)

Milk is to cow as

Goat is to lamb

Egg is to chicken

Wing is to feather

Fruit is to orange

Check

Next

(milk is to cow as egg is to chicken)

இதே ஒப்புமைகளைத் தமிழில் மொழி பெயர்த்து கீழ்க்கண்ட விதமாகக் கொடுக்கலாம்.

தண்ணி கொதிக்கும் அதுபோல

பனி உறையும்

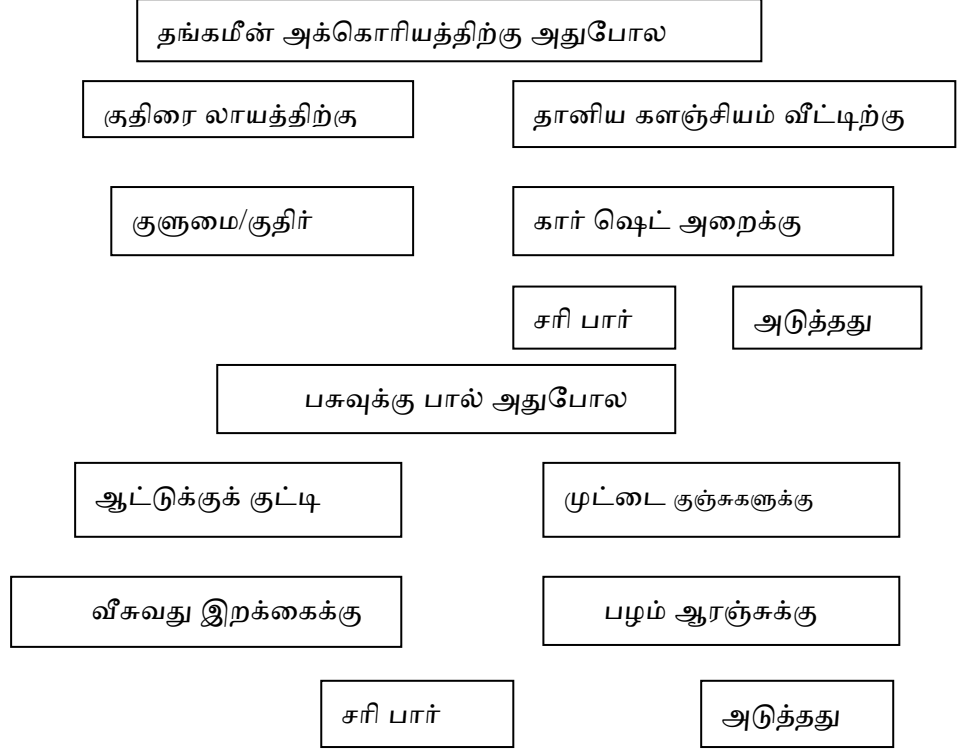
உஷ்ணநிலை குளிராகும்

ஐஸ் உருகும்

மழை ஆவியாகும்

சரி பார்

அடுத்தது



தமிழ் கற்றுக் கொள்ள நினைக்கும் மாணவர்களுக்கு இந்த விளையாட்டு ஒரு பொழுது போக்காகவும் மொழி கற்றுக் கொள்ளும் முறையாகவும் அமையும் என்பதில் சந்தேகமில்லை. ஒரு சுற்றுக்கு எட்டோ பத்தோ சொற்களைக் கொடுக்கலாம்.

இவ்வாறே கட்டுரைக்கு கட்டுரையாளர் எடுத்துக் கொண்ட மற்ற விளையாட்டுக்களும் அமைந்துள்ளன. இவ்வாறு ஆங்கிலத்தில் சொல்லப் பட்ட முறைகளை முழுமையாக இல்லாவிட்டாலும் கூடுமான வரை தமிழுக்குப் பயன் படுத்த முயற்சி செய்து இருக்கிறார்.

மரபுத் தொடர்கள்

மரபுத் தொடர்களைத் பொருத்துக, சரியான விடையைத் தேர்ந்தெடுக்கவும் என்ற முறைகளில் கற்பிக்கலாம்

இதன் அடிப்படையில், தமிழ் மொழியில், கீழ்க்கண்ட மரபுத் தொடர்களை அதற்கு எதிரில் கொடுக்கப் பட்டுள்ள வரைவிலக்கணத்துடன் பொருத்துமாறு சொல்லப்பட்டிருக்கிறது.

அ வரிசையில் உள்ள மரபுத் தொடர்களை ஆ வரிசையில் உள்ள வரைவிலக்கணத்துடன் பொருத்துக,

அ
கம்பி எண்ணுவிது
கயிறு திரிப்பது
கதைசொல்லுவது
கால் கட்டுப் போடுவது
காதில் பூச் சுற்றுவது
காது குத்துவது

ஆ
பொய் சொல்வது
இல்லாத ஒன்றைச் சொல்வது
ஜெயிலுக்குப் போவது
ஏமாற்றுவது
கற்பனையாகச் சொல்வது
திருமணம் செய்வது

சரி பார்	உதவி	மறுபடியும் விளையாடு
----------	------	---------------------

சிகப்பு விளக்குப் பகுதி
மஞ்சள் பத்திரிக்கை
கைநாட்டு
கள்ளநோட்டு
கறுப்புப் பணம்
நீலப்படம்
மாமியார் வீடு
மஞ்சள் காணி

கணக்கில் வராத பணம்
அரசு உத்தரவில்லாமல் அச்சடிக்கப்பட்ட பணம்
ஆபாசப்படம்
விபசாரம் செய்யும் பகுதி
சிறை
தாய் வீட்டுச் சீதனம்
படிக்காதவன், பாமரன்
பிரபலங்களைப்பற்றிய கிசு கிசு வரும் பத்திரிக்கை

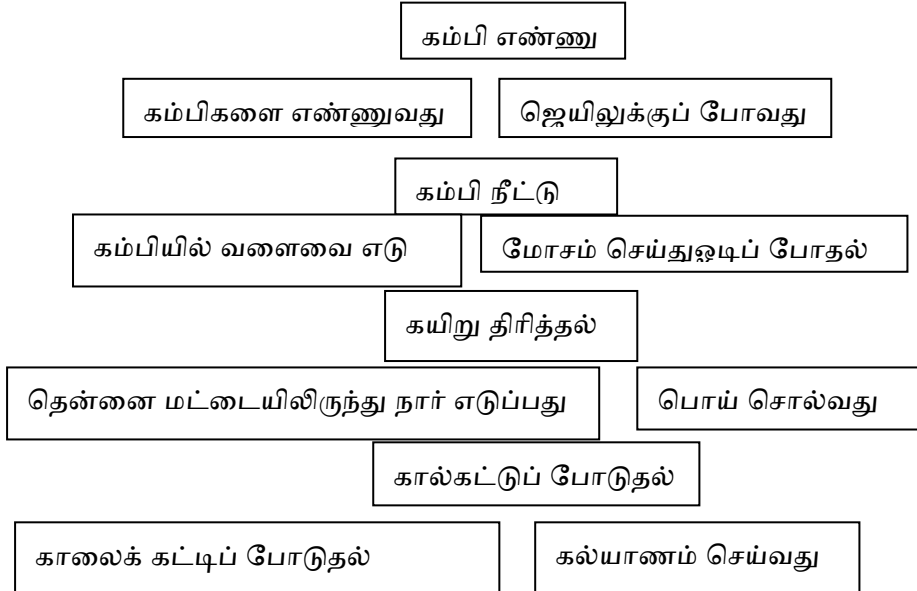
சரி பார்

உதவி

மறுபடியும் விளையாடு

இந்த விளையாட்டை தரம் பிரித்து முதல் படி இரண்டாவது படி மூன்றாவது படி என்று கற்பவர்களின் திறத்தைப் பொறுத்து எளிமையான மரபுத் தொடர்களில் இருந்து கடினமான மரபுத் தொடர்கள் வரை அமைக்கலாம். இவ்வாறு கொடுப்பதன் மூலம் நேரடியாகக் கற்றுக் கொள்ளாத பொருள்களைத் தமிழை இரண்டாவது மொழியாகப் படிக்கும் மாணவர்களுக்கு கற்றுக் கொடுக்கலாம். இந்த முறையில் இண்டர்நெட்டில் தமிழ் மொழியில் உள்ள மரபுத் தொடர்களை வரிசைப் படுத்திக் கொடுப்பதின் மூலம் மாணவர்களின் சொற்கோவை உயர வாய்ப்பு ஏற்படும். ஆசிரியர் இல்லாமல் சொற்களை வளப்படுத்தும் முறையாகவும் ஆகிவிடுகிறது.

மரபுத் தொடர்களை கீழ்கண்ட விதமாகக் கொடுத்து சரியான விடையை தேர்ந்தெடுக்கும் படி சொல்லலாம். சரியான விடையைத் தேர்ந்தெடுக்கவும்



தொகைச் சொற்கள்

அதே போல தொகைச் சொற்களை உருவாக்கவும் வளப்படுத்த படங்கள், பொருத்துக, தொகைச் சொல்லைப் பிரித்தல் போன்ற முறைகளை இண்டர்நெட்டில் உருவாக்கலாம்.

தொகைச் சொற்களை உருவாக்க படங்களை உப.யோகிக்கலாம்



+



=



+



=



+



=

தொகைச் சொற்களை பொருத்துக முறையைப் பயன்படுத்தியும் கற்பிக்கலாம்.

பொருத்துக

Column A

column B

சிவப்பு

பத்திரிகை

கறுப்பு

விளக்கு

மஞ்சள்

படம்

நீலம்

குதிரை

வெள்ளை

துணி

தொகைச் சொற்களை மாணவர்களிடம் வளப்படுத்த மொழியில் உள்ள தொகைச் சொற்களைப் பிரித்து காட்டச் சொல்லி பின் வரும் விளையாட்டு முறையைப் பயன்படுத்தலாம். உதாரணமாக புளிமூட்டையை, புளி+மூட்டை என்று பிரிக்கலாம்.

தொகைச் சொற்களைப் பிரிக்கச் சொல்லுதல்

மாட்டுக்கொட்டம்

கிளிகூண்டு

உப்புமூட்டை

அரிசிசாக்கு

கூண்டுகிளி

புளியம் பழம்

கம்மஞ் சோறு

வேப்பங் கொழுந்து

சாப்பாட்டுக் கூடை

செஞ்சாந்து

கருங்குழம்பு

வெண்தாமரை

இரு சொற்கள் சேர்த்து தொகைச் சொற்களை ஏற்படுத்தும் படி சொல்லலாம்

சுடு + தண்ணி = சுடுதண்ணி

சுடு + சோறு = சுடுசோறு

வண்டி + மாடு = வண்டிமாடு

வாடகை + வீடு = வாடகைவீடு

கவசம்+ குண்டலம்= கவசகுண்டலம்
 காடு+ யானை= காட்டு யானை
 வீடு+ பூனை =வீட்டுப் பூனை
 தண்ணீர்+ தொட்டி= தண்ணித்தொட்டி
 கம்பு+ சோறு= கம்மஞ்சோறு
 புளி+ கொம்பு= புளியங்கொம்பு
 பூ+ தொட்டி= பூந்தொட்டி
 பூ+பந்து= பூப் பந்து
 கறுப்பு+ குதிரை= கறுப்புக் குதிரை
 வெள்ளை+ புறா= வெள்ளைப் புறா

மேற்கூறிய பிரித்தெழுது சேர்த்தெழுதுக்குரிய உதாரணங்களில் பல வகையான புணர்ச்சி விதிகள் உள்ளன. இச் சொற்களைப் பிரித்துச் சேர்த்து எழுதுவதன் மூலம் அப் புணர்ச்சி விதிகளை சரியான முறையில் கற்றுக் கொள்ளவும் இந்த விளையாட்டு உதவுகிறது.

எதிர்பதச்சொற்கள்

எதிர்பதச் சொற்கள் பெயரடைவோ, வினையடையாகவோ, த, சிறிதளவு பெயர்களும் எதிர்பதச் சொற்களாக விளங்கும் .உதாரணமாக உயரம், மெதுவாக, வேகம் போன்றவைகளுக்கு எதிர்பதச் சொற்கள் இருக்கின்றன . ஆனால் மேசை என்பதற்கு எதிர்பதம் இல்லை .எதிர்பதச் சொல்லை பொருத்துக போன்ற விளையாட்டுகள் மூலம் கற்றுக் கொள்ளலாம் .எதிர்பதச் சொல் விளையாட்டுகள் ஆரம்ப, நடுநிலை உயர்நிலை பள்ளிகளுக்கு நிறைய இருக்கின்றன.

ஆங்கிலத்தில் இருப்பது போல தமிழிலும் பலவிதமான எதிர்பதச் சொற்கள் இருக்கின்றன .கீழ் கண்ட விதமாக கணினியின் இண்டர்நெட் மூலம் கற்றுக் கொடுக்கலாம் .

பொருத்துக

புகழ்ச்சி	குள்ளம்
உயரம்	மேடு
உயர்வு	தட்டை
பள்ளம்	தாழ்ச்சி
மலை	ஒல்லி
உருண்டை	மடு
குண்டு	தாழ்வு

சரியான சொல்லுடன் சேர்

அநாகரிகம்

நாகரிகமின்மை நாகரிகம்

அநாவசியம்

அவசியம் அவசியமின்மை

அவநம்பிக்கை

நம்பிக்கையின்மை நம்பிக்கை

அவமரியாதை

மரியாதையின்மை மரியாதை

பல சொல் ஒரு பொருள்

விலங்குகள்

A

கவி

சுனகம்

வராகம்

B

நாய்

குரங்கு

மூஞ்சூறு

ரிஷபம்	பன்றி
மேஷம்	ஆடு
சிங்கம்	அரிமா
மூசிகம்	காளை
அஸ்வம்	யானை
வாரணம்	புரவி

எழுத்து விளையாட்டு

சிறு பிள்ளைகளுக்கும், தமிழை இரண்டாம் மொழியாகக் கற்பவர்களுக்கும் எழுத்து சோதனைக் கருவி சரியான எழுத்துக்கள் மிகவும் தேவை. கல்வி கற்கும் போது மட்டுமல்லாமல் வேலை பார்க்கும் போதும் சரியாக எழுதுவது மிகவும். சரியாக எழுதுவது படிப்பதையும் எழுதுவதையும் மெருகுப்படுத்துவதுடன் சொற்களையும் அறிவுணர்திறனையும் வளப்படுத்த உதவுகிறது. அத்தகைய எழுத்தை வளப்படுத்த இண்டர்நெட் எழுத்து விளையாட்டு உதவுகிறது. சரியான எழுத்து எழுத பயிற்சி பெறுவதற்கு ஆன்லைன் விளையாட்டு மிகவும் சிறந்த முறை. பலவிதமான முறைகள் சரியான எழுதுப் பயிற்சி பெற இருக்கின்றன. அவற்றில் ஒன்று எழுத்துக்களைக் கலைத்துப் போட்டு சரியாகச் சேர்க்கச் சொல்லுவது ஒருவகை.

எழுத்துக்களைக் கலைத்துப் போடுதல் என்பது எழுத்தை வளப்படுத்துவதுடன் சொல் வளத்திற்கும் உதவும் ஒரு வகையான விளையாட்டு. உதாரணமாக

லி பு
ர டி க
வே ரி கோ தை ழு க
யா நீ னை ர்
க ம் ட் ஓ ட
ரை ரி வ தி கு
மை கா எ ட் ரு டு

இதைப் போல இன்னும் வேறு விளையாட்டுக்களும் தவறில்லாமல் தமிழில் எழுத்துக்களை எழுத தயாரிக்கலாம். இவற்றை ஆங்கில மொழியைப் போல இண்டர்நெட்டில் போட்டால் மாணவர்கள் கற்றுக் கொள்ள வசதியாக இருக்கும்.

முடிவுரை

மேற்கூறிய விளையாட்டுகளை விளையாடுவதன் மூலம் தமிழ் கற்க விரும்பும் மாணவர்களுக்கு தமிழில் சொற்கள் வளப்படுவதுடன் சொற்களைப் பிழையில்லாமல் எழுதவும், புணர்ச்சி விதிகளைப் போன்றவைகளை கற்பிக்கவும் முடியும். மொழி வளையாட்டால் பொது அறிவு வளர்ச்சி அடையவும்.

The Potential Role of YouTube in the process of 21st century approach to Classical Tamil Literature (2012 – 2013 Report)

Dr.M.Semmal MBBS, D.L.O, M.Sc, M.D, M.Phil
Managing Director, Manavai Mustafa Scientific Tamil Foundation, Chennai
Associate Professor, Physiology, Sri Ramachandra University, Chennai

Exploring the scientific acumen embedded in the scripts of the ancient world in general and that of the Ancient Tamils in particular is an extremely important task left to be accomplished by the linguistic scholars and academicians of the modern world. With a huge cache of literary works the arena of classical Tamil literature is a potential area for scientific inquiry and exploration particularly from the Neurobiological point of view. The universe of ancient Tamil literary works is a huge arena with multiple literary works penned at a timeline which extends many centuries even before the period accepted to be the classical Tamil timeline. Even after losing many valuable works, we the Tamils are left with a considerably admirable volume of literary works.

Due to various reasons which are beyond the preview of this paper, awareness about the scientific acumen of the ancient Tamils is meager in the western world, as proper representation for it has been missing for various reasons. In this era of technical marvel, where the web has melted all the solid human barriers reinforced for centuries by human territorial tendencies and bias, it is time for the Tamils to reach the greater world in the technically most advanced format and show the world, what they have in store for the modern regenerations to ponder and decipher and this necessitates us to use the state of the art technology available to us rather than via the usual outlets of the past via printed booklets and lectures alone, we need to accept that these techniques are already passing through a difficult and diminishing phase of acceptance among the human species even in the developing world.

The stubbornness and closed outlook of the technically weaker segments of a language to embrace modern methods of human expression, which springs out of the territorial behavior is only causing a detrimental impact on the overall importance and attention drawn by the classical Tamil Literary works in the modern world. Here is a scenario where an outlook can actually harm a concept. Acquisition of technical skills must be made as a compulsory expectation rather than an option for the Tamil language scholars who have the time and linguistic expertise to deal with the core of classical Tamil Literary works; this will open the gates for them to outreach to the technically empowered Tamils who may be non linguistic experts hailing from various scientific disciplines and this will create a scenario for Classical Tamil Literature to be elaborated in multi disciplinary direction in the future.

Philocine – the Philosophical Aspect of Medicine is a vibrant and nascent subspecialty in Medical Education. It plays with ease the role of a bridge between what appears to be two distant arenas, [1] Crafted by the ancient world, the classical Tamil Literary works penned between the periods of 300 B.C - 300 A.D and [2] the concept of streaming free video sites (like the You Tube) made possible by the technical advancements of the modern era with a gaining popularity and reach during the past decade. Fusing the two is ought to create an exploding impact on the reach of the Classical Tamil Literary works and it would ignite a scenario undreamt of in the past.

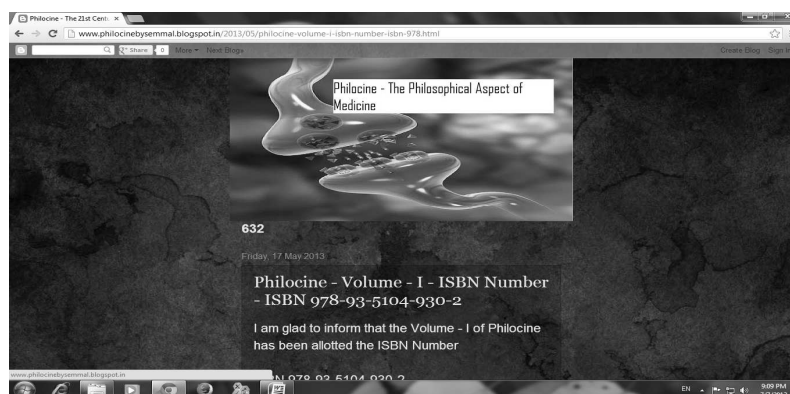


Figure – I – Free Blogspot can embed the videos away from its hosting mother sites



Figure–II–Simple digital cameras were used for the Learning module to create the educational videos.

The Manavai Mustafa Scientific Tamil Foundation [MMSTF] made a pilot observation using the **Philocine** learning module that integrates the two concepts, the classical Tamil literature and the Modern science using free streaming video hosting sites (like the you tube), a brief account of the experiences of the learning module will be recalled here.

Critical scientific enquiry was conducted on the first Kural “அகர முதல எழுத்தெல்லாம் ஆதி; பகவன் முதற்றே உலகு” using the Philocine learning module to decipher the various scientific interpretations springing from the Kural numbered 1 in the long order of 1330, the findings were presented in the usual method of conversion into a oral research paper and was presented in the World Classical Tamil conference in Coimbatore, India during June 2010. The post conference presentation work on the above mentioned paper deserves a closer look by the technical world, a part of the entire scientific document which expands to nearly 850 pages was converted into a more palatable and absorbing format, as a drama script with seven imaginary characters; 5 higher secondary school students; a doctor and a Tamil Language scholar.

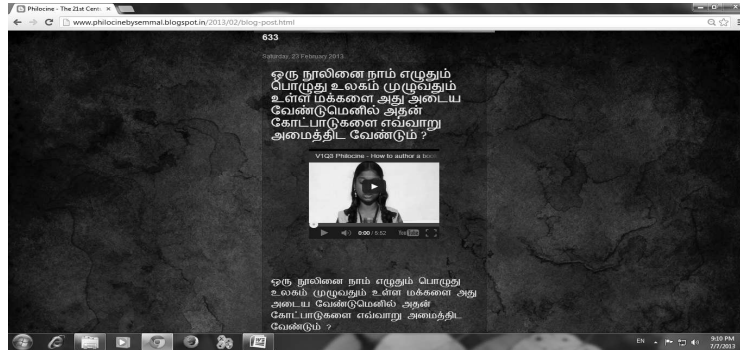


Figure – III – Activation of Mirror Neurons is the Neurobiological Phenomenon used for creating a vibrant generation of Tamils

All the characters were crafted by the author for a specific purpose; the characters were designed to be reflective of the various possible domains of thought process of present day higher secondary school students, one character was inclined towards the traditional roots, one was convinced to abandon the past for no obvious reason, one was revolutionary, one inclined to spirituality and one convinced to atheism. This arrangement was made to make a clear and open multidimensional scientific investigation into the core of the literary work with liberal room for the usual emotions springing in the adolescent age.

The scientific material was transferred onto the script in the drama with its above mentioned framework and was aligned into a discussion format; the script was then handed over as individual question and answer style to a selected group of 50 student volunteers of class Eleven of Lakshmi Matriculation Higher Secondary School, Manapparai.

The students video graphed the questions as separate video files and finally they were converted into individual educational videos each containing 2 parts, the student understanding and the authors (Dr.M.Semmal) 21st century interpretation of the Kural.



Figure– IV - content is more important, for it will draw its audience

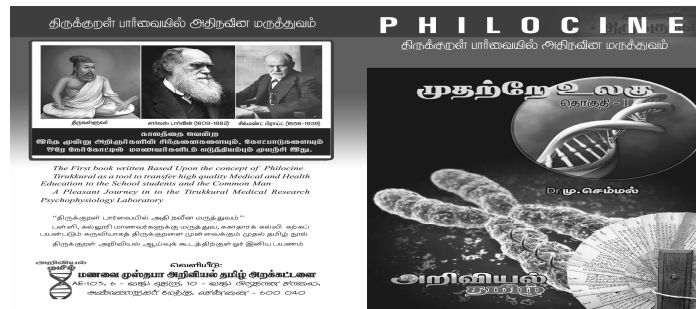
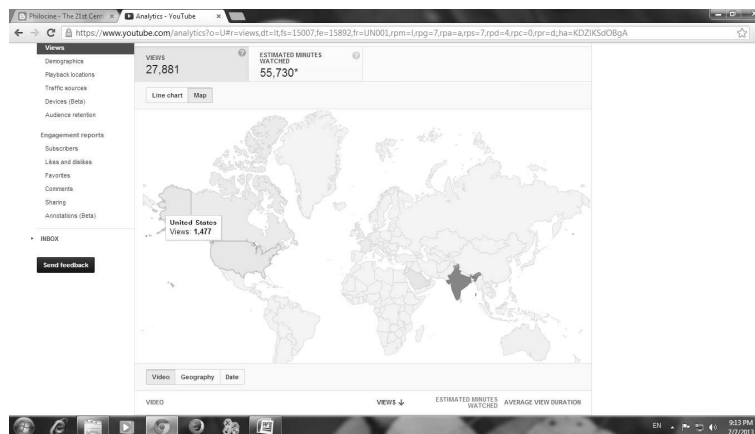


Figure – V – Printed Book with ISBN Number supplemented by

Free educational videos in the internet

The videos were supplementary to the printed book 'Mudhatrae Ulagu' containing 112 pages and bearing an ISSN Number (ISBN 978–93–5104–930–2). The videos were uploaded in the free educational digital Scientific Tamil Channel “Ariviyal Tamil Mandram” in the You Tube, then the videos were embedded in a dedicated blog www.philocinebysemmal.blogspot.in, thus this module is able to create the world’s first ever research work in Classical Tamil Literature accomplished by a medical doctor and a group of (50) higher secondary school students, published as an ISBN numbered book and as supplementary educational videos. An elegant fusion of 2000 year old literary work and a less than 20 year old technology.



Figure– VI – The Analytical data will motivate the students for more vigorous involvement

The mission was accomplished by the Manavai Mustafa Scientific Tamil Foundation (MMSTF), in the future each and every single Classical Tamil Literary work will be screened for the scientific acumen embedded in it and the findings will be in published in 3 separate formats –

- [1] As a research paper
- [2] As a printed ISBN numbered Book
- [3] As a collection of free educational videos in free video hosting sites.

ஆண்ட்ராய்டு இயக்கியில் தமிழ்ப்பள்ளியுடன்

எளிய தொடர்பு

Megavarnan Jagadesan, SJK(T) Ladang Strathmashe, 36100 Bagan Datoh, Perak, Malaysia

ஆய்வுச் சுருக்கம்

இன்றைய நவீன உலகத்திற்கேற்ப கற்றல் கற்பித்தலுக்கான உபகாரணப் பொருள்களும் மாற்றி வருகின்றன. மலேசிய கல்வி அமைச்சு ஒவ்வொரு பள்ளியையும் தகவல் தொழில்நுட்ப ரீதியில் முன்னேற்றி வருகிறது. சில தோட்டப்புற தமிழ்ப்பள்ளிகளில் இத்திட்டங்களில் கடைசி இடத்தையே வகிக்கின்றன. அவ்வகையில் ஸ்திராத்தமாஷி தோட்டத் தமிழ்ப்பள்ளி, தொலைபேசி, தொலைநகல் வசதி கூட இல்லாத ஒரு தமிழ்ப்பள்ளி. கைத்தொலைபேசி சேவையும் சரி வர கிட்டைக்காத ஒரு செம்மண் சாலைத் தோட்டத்தில் உள்ள தமிழ்ப்பள்ளி. பிறர் இந்த பள்ளியுடன் எளிதில் தொடர்பு கொள்ள ஆண்ட்ராய்டு விவேக கைப்பேசிக்கான ஒரு மென்பொருளை இலவசமான முறையில் தயாரித்து பயன்பாட்டில் வைத்துள்ளது. இந்த மென்பொருளை இணைய வசதியுடன் கணினிகளிலும் பயன்படுத்தலாம். இரண்டு இலவச இணைய மென்பொருள் கட்டுமான தளத்தின் துணையுடன் இந்த மென்பொருள் நிறுவப்பட்டது.

ஆய்வுப் பின்னணி

ஆண்ட்ராய்டு என்பது லினக்ஸ் கெர்னலில் இயங்கக்கூடிய செல்லிடப்பேசிகாக அதுவும் முதன்மையாக தொடுத்திரையுடன் கூடிய செல்லிடபேசிக்காக உருவாக்க பட்ட ஒரு இயங்குதளம் ஆகும். தொடக்கத்தில் இது அண்ட்ராய்டு இன்க்., கால் உருவாக்கப்பட்டு பிறகு இந்த நிறுவனம் கூகுளால் வாங்கப்பட்டது. மேலும் அண்மையில் ஓப்பன் ஹேண்ட்செட் அலையன்ஸால் இந்த நிறுவனம் வாங்கப்பட்டது.^[5] கையாளப்பட்டக் குறியீட்டை ஜாவா நிரலாக்க மொழியில் எழுதுவதற்கு உருவாக்குநர்களுக்கு இது இடமளிக்கிறது. மேலும் கூகுளால் உருவாக்கப்பட்ட ஜாவா லைப்ரரீஸ் மூலம் சாதனத்தை கட்டுப்படுத்துகிறது. கூட்டுநிறுவனங்களான 47 வன்பொருள், மென்பொருள் மற்றும் தொலைத் தொடர்பு நிறுவனங்களின் செல்லிடப்பேசி சாதனங்களுக்கான கட்டுப்பாடற்றத் தரத்தை ஊக்கப்படுத்துவதற்கு ஈடுபாடுகொண்ட ஓப்பன் ஹேண்ட்செட் அலையன்ஸ் நிறுவப்பட்டு, 5 நவம்பர் 2007 அன்று அண்ட்ராய்டின் விநியோகம் வெளிக்கொணரப்பட்டது. பெரும்பாலான அண்ட்ராய்டு குறியீட்டை அப்பாச்சி உரிமத்தின் கீழ் கூகுள் வெளியிட்டது, அப்பாச்சி என்பது ஒரு இலவச மென்பொருள் மற்றும் திறந்த மூல உரிமம் ஆகும்.

ஆய்வின் நோக்கம்

உட்புற பகுதியில் உள்ள தமிழ்ப்பள்ளிகளின் தகவல்களை எப்படி எளிமையாக பெருவது என்பதைப் பற்றியே இந்த ஆய்வின் நோக்கம். மேலும் இன்றைய நவீன தொலைத்தொடர்பு வசதிகளைப் பயன்படுத்தி எப்படி கையாளலாம் என்பதையும் இந்த ஆய்வு கண்டுபிடிக்க வழி செய்கிறது.

ஆய்விற்குரிய சிக்கல்

தோட்டப் புறத்திலுள்ள / உட்புற தமிழ்ப்பள்ளிகளுக்கு பெரும்பாலும் மாணவர்கள், ஆசிரியர்கள் தவிர வேறு யாரும் வருவதில்லை. பெற்றோர்கள் வருவதும் மிக அரிது. அதிகாரிகள் இது போன்ற பள்ளிகளுக்கு வருவது மிகவும் அரிது. இந்த புறநகர் அல்லது உட்புற பள்ளிகளில் என்ன நடக்கின்றன என்பது யாவரும் அறிவதில்லை. நகர்புற பள்ளிகளில் எது நடந்தாலும் மக்கள் சுலபமாக அறிந்துக் கொள்ளலாம். தொலைபேசி, தொலைநகல் போன்ற தொடர்பு சாதன வசதி இல்லாத இந்த பள்ளிகளில் வனெலி பொருத்தப்பட்ட இணைய வசதி மட்டுமே தொடர்பு சாதனம். இது போன்ற தமிழ்ப்பள்ளிகளுடன் எப்படி தொடர்பு கொள்வது என்பதே இந்த ஆய்வுக்கான சிக்கல்.

செயலாக்கம்

இணையம் வாயிலாக பல இலவச இணைய மென்பொருள் கட்டுமானத்தளம் தேடப்பட்டது. பின் சில கட்டுமானத்தளம் பரிசோதிக்கப்பட்டது. அதில் எளிமையாகவும், விரைவாகவும் ஈர்ப்புடையதாகவும் உள்ள ஒரு கட்டுமானத்தளம் தேர்வு செய்யப்பட்டது. அது தான் <http://www.theappbuilder.com/>. அந்த கட்டுமானத்தளத்தின் வழி மென்பொருள் நிறுவப்பட்டது. மென்பொருள் கட்டுமானம் முடிந்த பின், மென்பொருள் பயன்பாட்டிற்குத் தயாராக்கப்பட்டது. இணையம் வசதியைக் கொண்டு, பள்ளியின் பெயர், படங்கள், பள்ளியின் புதிய தகவல்கள், காணொளிகள் மற்றும் பள்ளியின் முக்கிய தொடர் முறைகளையும் அறியலாம். இது அனைத்து இணைய வசதி

கொண்ட கருவிகளிலும் இயங்கும். உதாரணத்திற்கு <http://myapp.is/SJKTLadangStrathmashie> எனும் இணைய முகவரியைப் பயன்படுத்தி இந்த மென்பொருளைப் பயன்படுத்தலாம்.

அதனையடுத்து <http://www.appsgeyser.com/> எனும் முகவரிக்குச் சென்று ஒரு கணக்கை திறந்து, <http://myapp.is/SJKTLadangStrathmashie> எனும் முகவரியை இணைக்கவும். பின் அது ஆண்ட்ராய்டு இயங்குதளத்திற்கான .apk எனும் கோப்பை தரவிறக்கம் செய்யும். அந்த கோப்பைப் ஆண்ட்ராய்டு கருவிகளில் நிறுவி, இணையத் தொடர்பின் வழி பள்ளியின் தகவல்களை உள்ளங்கையில் அறியலாம். வரைபடம் 1 இந்த செய்முறையைச் சுருக்கமாக காட்டுகிறது.



முடிவுரை

இந்த ஆண்ட்ராய்டு இயங்குதளத்திற்கான இலவச இணைய மென்பொருளைப் பயன்படுத்தி எளிதில் தொடர்பு கொள்ள முடிகின்றது. உட்புற அல்லது தோட்டப் புற தமிழ்ப் பள்ளிகள் இதனைப் பயன்படுத்துவதன் வழி, தங்கள் பள்ளியைப் பற்றி தகவல்களையோ அல்லது நிகழ்வுகளையோ எளிதில் மக்களிடம் கொண்டு சேர்க்க முடிகின்றது.

மேற்கோள்

<http://www.appsgeyser.com/>
<http://appinventor.mit.edu/explore/teach.html>
<http://www.theappbuilder.com/>
<http://developer.android.com/training/basics/firstapp/creating-project.html>
http://ta.wikipedia.org/wiki/ஆண்ட்ராய்டு_இயங்குதளம்

Role of Information Technology in Teaching and Learning of Classical Literature

Saravanan Raja

CAS in Linguistics, Annamalai University, Tamil Nadu, India.

Abstract

Tamil language is one of the longest surviving classical languages in the world. Tamil language and literature has existed for over 2000 years. Cilappatikaram is one of the five great epics, the others being Manimekalai, Civakacintamani, Valaiyapati and Kundalakesi. The poet prince Ilango Adikal, popularly believed to have been a Jain monk has created this epic. He is reputed to be the brother of Senguttuvan from Chera dynasty. As a literary work, it is held in high regard by the Tamils. The nature of the epic is narrative and has a moralistic undertone. It contains three chapters and 5270 lines of poetry. The epic revolves among Kannagi, who having lost her husband to a miscarriage of justice at the court of the Pandiya king, wreaks her revenge on his kingdom. The present paper mainly focuses on teaching and learning of classical literature in general and in particular, to teaching and learning of Cilappatikaram through computers. In the recent past, information technology has found a major place in all spheres of education. Information technology helps in promoting opportunities of knowledge sharing throughout the world. These can help the teachers and learners having up - to - date information and knowledge. Information technologies are assumed to play a constructive role in education to make the teaching and learning process more productive through collaboration in an information rich society. In this current scenario, development of innovative teaching methods for teaching of Tamil classical literature like Cilappatikaram is the need of the day. The use of internet for downloading the original text in Tamil along with the translated version in English will be much helpful for the learners when compared to the conventional classroom teaching. Besides, a combination of computer audio and video discs related to Cilappatikaram will enhance the learning process. Well designed computer mediated instruction is more likely to impart effective teaching and learning of Tamil classical literature which is discussed in detail in this paper.

Introduction

At present, knowledge may be regarded as power and it comes from having information. Information encompasses and relies upon the use of different communication channels or technologies called information technologies, for its effectiveness and equal access. Information technologies may extend knowledge beyond the geographical boundaries of a state or country providing relevant information to the relevant people round the clock. Information Technology “is any computer-based tool that people use to work with information and support the information and information processing needs of an organization” (Haag, 1998; pp.17. 518). It includes computers and its related technologies; WWW, Internet and Videoconferencing etc. Information technology can be used to promote the opportunities of knowledge dissemination. Accurate and right information is necessary for effective teaching and learning; and information technology (Haag, 1998; p.10) is a “set of tools that can help provide the right people with the right information at the right time.” Information technologies help in promoting opportunities of knowledge sharing throughout the world. These can help the teachers and students having up-to-date information and knowledge. Students are able to work in collaborative and interactive learning environments effectively communicating, sharing information and exchanging ideas and learning experiences with all in the environment.

Information rich societies are developed and dominating and they are controlling the information throughout the world. Information encompasses and relies on the use of different channels of communication, presently called information and communication technologies (Hussain, 2005) and would be incorporating better pedagogical methods to cope with such emerging situations. These have changed the scenario of education particularly, pedagogy and instruction making teaching learning process more productive, creating collaborative, learner centered and interactive global learning environments. Therefore, information technologies are assumed to play a constructive role in education to make the teaching and learning process more productive through collaboration in an information rich society.

Information technologies can promote the opportunities of restructuring the teaching learning process. These can transform teaching and learning by offering alternatives to the teacher provided information, access to virtually unlimited resources and opportunities for real world communication, collaboration and competition. Traditional lectures and demonstrations can become web-based multimedia learning experiences for learners. Use of information technologies in classroom situation particularly interactive technologies however; ensure attention and active involvement of students. Well-designed computer-mediated instruction is more likely to engage individuals for effective learning than simple lectures and book reading in a classroom. Information technologies are the result of knowledge explosion. These include hardware and software technologies and facilitate teaching learning process. Using information technologies learners are now able to participate in learning communities throughout the world. They are independent and free in choice of their programmes of study and access

to the resources. They may learn collaboratively, share information, exchange their learning experiences and work through cooperative activities in virtual learning communities. Information technologies facilitate teaching learning process in more productive fashion.

Visualization is any technique for creating images, diagrams or animations to communicate a message. Visualization through visual imagery has been an effective way to communicate both abstract and concrete ideas since the dawn of man. Typical of a visualization application is the field of computer graphics. The invention of computer graphics may be the most important development in visualization and the development of animation has also helped advance visualization. It is a process of transforming information into a visual form that enables user to observe the information. On the other side, it uses techniques of computer graphics and imaging. Successful visualization can reduce the time to get information, make sense, and enhance creative thinking. Visualization uses various aspects such as texts, graphics, images, video, audio and animation. Teaching of Cilappatikaram to students by using these techniques has been discussed in this paper.

Cilappatikaram is one of the five great epics, the others being Manimekalai, Civakacintamani, Valaiyapati and Kundalakesi. The poet prince Ilango Adikal, popularly believed to have been a Jain monk has created this epic. He is reputed to be the brother of Senguttuvan from Chera dynasty. As a literary work, it is held in high regard by the Tamils. The nature of the epic is narrative and has a moralistic undertone. It contains three chapters and 5270 lines of poetry. The epic revolves around Kannagi, who having lost her husband to a miscarriage of justice at the court of the Pandiya king, wreaks her revenge on his kingdom. The present paper mainly focuses on teaching and learning of classical literature in general and in particular, to teaching and learning of Cilappatikaram through computers. Cilappatikaram contains three chapters and the Canto five of the first chapter Pukarkantam has been taken for this study.

In Canto five, the festival of Indra (*Indra vizha*) has been explained in detail. To the delight of the earth and the admiration of heaven, the festival of Indra was celebrated by the people of Puhar. One thousand and eight kings bore on their heads gold pots filled with cool and holy water, rendered fragrant by floating pollen of flowers of the Cauvery taken from where it merges the sea, and performed the bathing ceremony of the Lord of Gods, Indra. The festival of Indra comprises of several events which will be explained in detail with the help of various visual objects like graphics, images, video, audio clips and animation along with the relevant texts. The festival of Indra has been depicted in the form of images as given below in one of the websites.

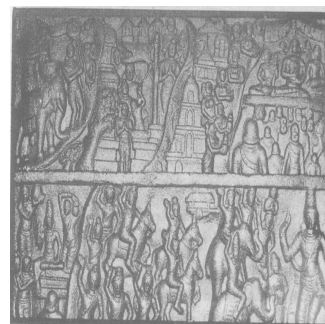
Poompuhar – Pattinapaakkam



Poompuhar –Nalangadi



Indra Festival in Poompuhar



Apart from these images, the events related with this festival can be further explained in detail with the other forms of visualization like audio clips, animated clippings, etc.

Conclusion

Information technologies are the result of knowledge explosion. These include hardware and software technologies and facilitate teaching learning process in a more productive fashion. Information technologies are restructuring teaching learning process to meet the International standards. In this current scenario, development of innovative teaching methods for teaching of Tamil classical literature like Cilappatikaram, one of the five great epics assumes great importance.

References

- Aggarwal, A. 2000. Web-Based Learning and Teaching Technologies: Opportunities and Challenges. London; Idea Group Publishing.
- Haag, Cummings and Dawkins (1998). Management Information Systems for the Information age, McGraw Hill USA.
- Menges, R. J. 1994. Teaching in the age of Electronic Information. In Wilbert J. McKeachie (eds.), Teaching Tips: 9th Ed. U.S.A.; D.C Heath and Company.
- Pillai, R.S. 1989. Cilappatikaram, Thanjavur: Tamil University.
- Venkataswamy Nattar. 2011. Cilappatikaram: muulamum uraiyum, Chennai: Ramaiya Publication.

காணொளி பேசும் தமிழ்

(Tamil Teaching via Windows Movie Maker)

சங்கீதா இராமகிருஷ்ணன்

துவான்கு பைனான் ஆசிரியர் கல்விக் கழகம், பிளாங்கு

முன்னுரை

இந்தத் தகவல் தொழில்நுட்ப உலகத்தில் கணினி நமது கையடக்கமாகி விட்டது. ஆகவே, இந்நிலையில் உலகத்தோடு அவ்வ துறை வதறிவு என்பதற்கேற்ப ஓர் ஆசிரியர் உலகியல் அறிவோடு தனது கற்றல் கற்பித்தலை நடத்த வேண்டும், (திருக்குறள், எண்: 426). உலகின் மூத்த மொழியான தமிழ் மொழியையும் கணினி மூலம் கற்பிக்க முடியும் என்பதையே இவ்வாய்வு ஆராய்கிறது.

ஆரேபியன் லாடாக் (Arabian Ladak, 1992) கூற்றுப்படி மாற்றம் ஒன்றே மாற்றம் அடையாதது. மாற்றமே நிலையானது. ஆகவே, தமிழ் மொழி கற்பிக்கும் முறைகளையும் கற்கும் முறைகளையும் மாற்றினால் தமிழ் மொழியும் சுவைக்கும். நம்மில் பெரும்பாலான ஆசிரியர்கள் புதிய வழிமுறைகளைக் கையாளுவதற்கு அச்சம் கொண்டிருப்பதனாலும் புது வழியை முயற்சிக்காததாலும் மாணவர்களிடையே தமிழ் மொழி வெறுக்கப்படுவதற்கான காரணம் எனலாம்.

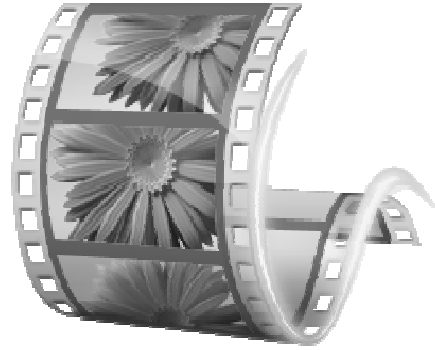
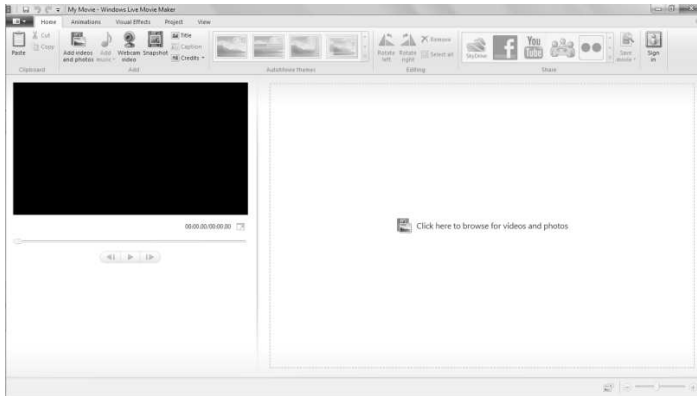
செம்மொழியான தமிழ் மொழி இலக்கண இலக்கிய வளம் பெற்றாதாகும். தமிழ் மொழியின் இவ்விரண்டு கண்களையும் கணினி வாயிலாக மாணவர்கள் மனத்தில் எளிதில் பதிய வைக்கலாம். கணினியின் வரைகலை (கிராஃபிக்ஸ்) ஆசிரியரின் கற்றல் கற்பித்தலை எளிமையாக்குவது மட்டுமல்லாமல் இனிமையாக்குகிறது.

காணொளி

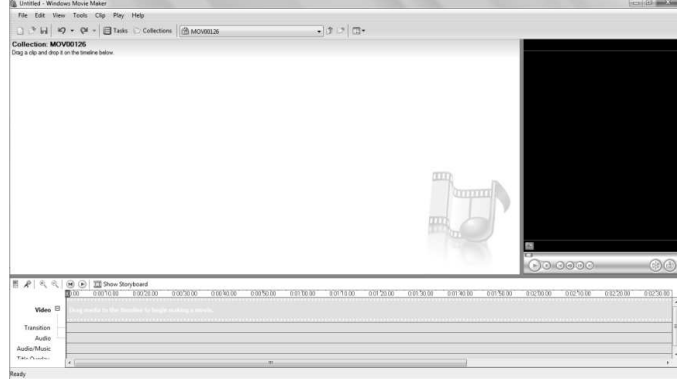
காணொளியை நாம் காண் + ஒளி எனப் பிரித்திடலாம். காணொளி என்பது திரையில் உருவங்கள் அசைந்து நகர்வதைப் போல காட்டும் படம். (<http://wiktionary.org>) ஒடுவது, நடப்பது போன்ற நிகழ்வுகளை நேரில் பார்ப்பதுபோலவே, ஒரு திரையில் காட்டும் அசைப்படம் ஆகும். ஒரு நொடியில் நிகழும் ஒரு காட்சியை நொடிக்கு, 24 முறை அல்லது அதற்கும் அதிகமான தடவையாக ஒளிப்படமாக எடுத்து, பின்னர் ஒன்றன் பின் ஒன்றாக, அப்படங்களை ஒரே திரையில், ஒருங்கே விழுமாறு செய்தால், நிகழ்வுகள் நேரே நடப்பது போன்றே தோற்றம் அளிக்கும்.

இத்தகைய அசையும் படங்களின் வரிசை, காணொளி என்றழைக்கப் படுகிறது. திரைப்படம் என்பதும் இதே அடிப்படையான கருத்திலேயே இயங்குகின்றது.

இந்தத் தொழில்நுட்பக் காலத்தில் காணொளியைத் தயாரிக்கப் பல மென்பொருள் வந்த வண்ணமாக உள்ளது. ஆயினும், நான் என் ஆய்விற்கு தேர்ந்தெடுக்கப்பட்டது, பரவலாக எல்லா கணினிகளிலும் காணப்படுவது, windows live movie maker மற்றும் windows movie maker ஆகும்.



படம் 1: windows live movie maker



படம் 2: windows movie maker 2.6

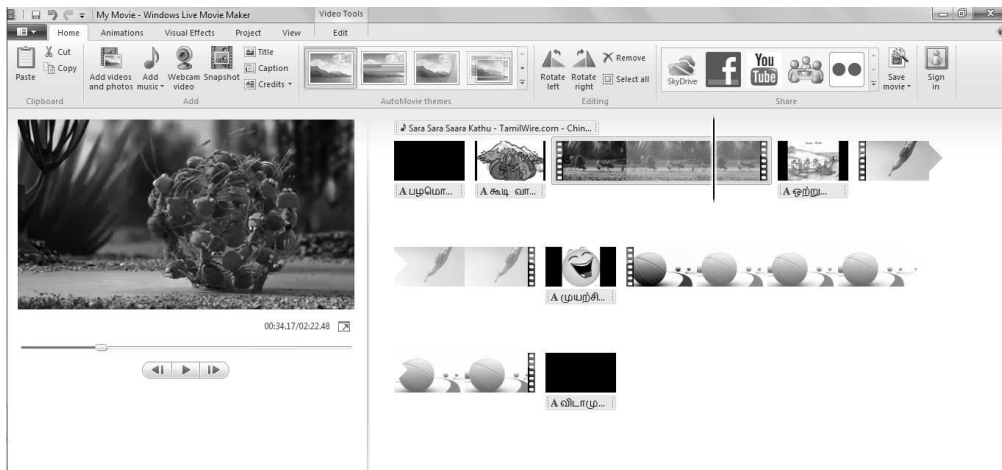
windows live movie maker அல்லது windows movie maker-ல் தயார் செய்யப்படும் காணொளியை எவ்வாறு நாம் தமிழ் மொழி கற்றல் கற்பித்தலுக்கு உபயோகிக்கலாம்? காணொளி மூலம் தமிழ் மொழியை எவ்வாறு கற்பிக்கலாம்? இவ்விரண்டு கேள்விக்கான பதில்தான் இவ்வாய்வு.

காணொளி பேசும் தமிழ்

ஒரு தமிழாசிரியர் காணொளி உத்திமுறை கொண்டு மொழியை மாணவர்களுக்குக் கற்பிக்கலாம். தமிழ்மொழியில் இலக்கணம், இலக்கியம் ஆகிய இவ்விரண்டிற்கும் இந்த உத்திமுறை பொருந்தும். பாடல், கதை, செய்யுள் ஆகியவை அடங்கும் இலக்கியத்தை இக்காணொளி முறையில் கற்பிக்கலாம். அதாவது முதலில் ஒரு தமிழாசிரியர் தான் கற்பிக்க விரும்பும் பாடத்தை, அதன் கருத்துகளைத் தொகுத்து காணொளியாகப் படைத்திடலாம். காட்டாக, பாடலுக்கு ஏற்ற படங்களும், கதைக்கு ஏற்ற சித்திரங்களும் மற்றும் செய்யுளின் கருத்தைத் தொகுத்தும் காணொளி செய்திடலாம்.

முதலில், இலக்கியத்தில் பழமொழியைக் காண்போம். பழமொழிகளைக் கற்பிக்க விழையும் ஆசிரியர் மாணவர்களின் தரத்திற்கு ஏற்ற சில பழமொழிகளைத் தெரிவு செய்ய வேண்டும்.

பின்னர், தெரிவு செய்யப்பட்ட பழமொழியை விளக்கும் அல்லது அதன் பொருளை விளக்கும் படங்கள் அல்லது கதையைத் தேர்ந்தெடுக்க வேண்டும். பின்னர், அதற்கேற்ற இசை, வரைகலையுடன் காணொளி ஒன்றை windows live movie maker @ windows movie maker-ல் தயார் செய்திட வேண்டும். காட்டாக, கூடி வாழ்ந்தால் கோடி நன்மை என்ற பழமொழியைக் காண்போம்.



படம் 3: பழமொழிக்கான காணொளி

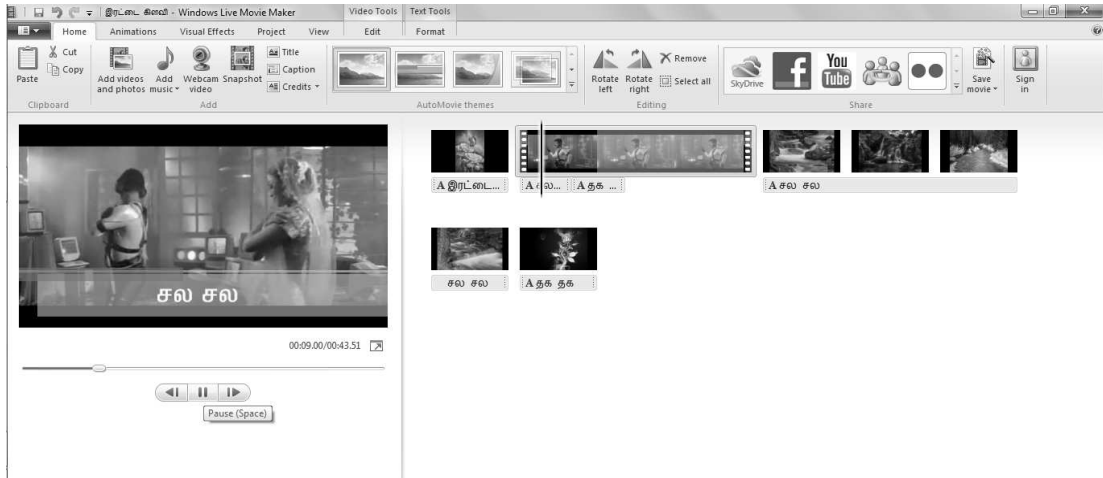
தயார் செய்த காணொளியை மாணவர்களுக்குப் படைத்து அதற்கேற்ற பயிற்சிகளை வழங்கிடும் பொழுது மாணவர்கள் மகிழ்ச்சியாகவும் ஆர்வத்துடனும் ஈடுபடுவது மட்டுமல்லாமல் மாணவர்கள் கற்பித்த பாடங்களைக் கிரகித்து கொண்டனரா என்பதையும் உறுதி செய்ய முடிகிறது. (<http://www.thirteen.org>)

காணொளியை இலக்கியத்திற்கு மட்டுமே பயன்படுத்த முடியும் என்பது பரவலாகப் பேசப்படும் வரும் கருத்து. ஆனால், இலக்கியத்தைத் தவிர்த்துக் காணொளியை இலக்கணத்திற்கும் உபயோகிக்கலாம், (பேராசிரியர் ந.தெய்வசுந்திரம், 10-ஆம் உலகத் தமிழாசிரியர் மாநாடு). காணொளியின் துணைகொண்டு ஓர் ஆசிரியர் இலக்கணத்தை எளிதாகக் கற்பிக்கலாம்.

இலக்கணத்தில் எழுத்ததிகாரம் முதல் சொல்லதிகாரம் வரை அனைத்திற்கும் இக்காணொளி முறையை உபயோகிக்கலாம். எழுத்ததிகாரத்தில், எழுத்து உச்சரிப்பு முதல் சொல்லதிகாரத்தில் உரிச்சொல் வரை ஆசிரியரின் அறிவாற்றலுக்கும் சிந்தனைத் திறனுக்கும் ஏற்பக் காணொளியைத் தயார் செய்து மாணவர்களுக்கு ஒரு சிறப்பான கற்றல் கற்பித்தலை நடத்திடலாம்.

காட்டாக, இலக்கணத்தில் இரட்டைக்கிளவியை எடுத்துக் கொள்வோம். ஆசிரியர் முதலில், மாணவர்களின் தரத்திற்கு ஏற்பத் தான் கற்பிக்க விழையும் இரட்டைக்கிளவிகளைத் தேர்ந்தெடுத்தல் வேண்டும். பின்னர், தேர்ந்தெடுக்கப்பட்ட இரட்டைக்கிளவிற்கு ஏற்பப் படங்கள் அல்லது கதையைத் தேர்ந்தெடுத்து அதற்கேற்ற இசையையும் வரைகலையுடனும் காணொளி ஒன்றைத் தயார்ச் செய்திடல் வேண்டும்.

ஆசிரியர் இரட்டைக்கிளவி குறிக்கும் ஒலியையும் தனது காணொளியில் இணைத்திடலாம். அதாவது, சல சல எனும் இரட்டைக்கிளவிக்கு நீர் ஓசையையும் கல கல வெனும் இரட்டைக்கிளவிக்குச் சிரிப்பொலியையும் சேர்த்திடலாம். இவ்வாறு செய்திடும் பொழுது, மாணவர்கள் இரட்டைக்கிளவிகளைப் படிக்கும் பொழுதே அதனின் பொருளையும் உணருகின்றனர்.



படம் 4: இரட்டைக்கிளவிக்கான காணொளி

ஆகவே, காணொளி இலக்கியத்திற்கும் மட்டுமல்லால் இலக்கணத்திற்கும் உபயோகிக்கலாம். காணொளியின் வழி ஆசிரியரின் கற்பிக்கும் வேலை எளிமையாகுவது மட்டுமல்ல, மாணவர்கள் இலக்கணம் மற்றும் இலக்கியத்தை எளிதில் புரிந்து கொள்வதற்கும் காணொளி உற்ற தோழனாக அமைகிறது.

காணொளி உத்திமுறை எவ்வாறு வேறுபட்டுள்ளது?

பெரும்பான்மையான தமிழாசிரியர்கள் நூல்களைக் கொண்டே தமிழ் மொழி கற்றல் கற்பித்தலை நடத்துகின்றனர். ஆகவே, இந்நிலையில் மாணவர்களின் செவி மட்டுமே செயல்படுகிறது. ஆனால், ஒருவர் காணொளியைப் பார்த்திடும் பொழுது அவரின் கண்கள் மற்றும் செவி ஆகியவைச் செயல்படுகின்றன, (Brown J.W., Lewis R.B, Harclerod F.F, 1985). ஏனெனில், அசையும் படங்கள், இசை, ஒலி, ஒளி, வரைகலை என்பதின் கூட்டே காணொளி ஆகும். மேலும், “I hear, I forget, I see, I remember, I do I understand” எனும் Confucius-ன் கூற்றுப்படி செவிப்புலனோடு கண்களும் செயல்பட்டால் மாணவர்கள் பாடங்களைப் புரிந்து கொள்வதில் சிரமம் எதிர்கொள்ள மாட்டார்கள். இஃது, மாணவர்களின் ஆர்வத்தையும் அதிகரிக்கிறது.

மேலும், மூளையில் இருக்கும் ஒக்கிபிடல் லோப் (occipital lobe) தான் நாம் காணும் காணொளியை நமக்கு (interpret) விளக்குகிறது. இது பின்மூளையின் நான்கு பகுதிகளுள் பெரிய பரப்பளவைக் கொண்டதாகும். (Paul Grobstein, Bryn Mawr College). ஆதலால், காணொளியின் வழி நாம் காண்பதும் கேட்பதும் நம் மனத்தில் எளிதாகப் பதிந்து விடுகிறது.

மேலும், இக்காணொளி உத்திமுறை நாடிக்கற்றலுக்கும் உதவுகிறது. அதாவது, ஆசிரியர் அல்லாமல் மாணவர்களே சுயமாகக் காணொளியின் வழி பாடங்களைக் கற்றிடலாம். கற்ற பாடங்களை ஆசிரியரின் துணை அல்லாமல் அவற்றை புரிந்து கொள்ளவும் இக்காணொளி வழிவகுக்கிறது, (Forster's, 1972). காணொளி என்பதால் மாணவர்கள் இவ்வகையான நாடிக்கற்றலில் ஆர்வத்துடன் ஈடுபடுகின்றனர்.

தமிழ் மொழியைக் கற்பிக்க உபயோகிக்கும் பல அணுமுறைகளில் காணொளி உத்திமுறை மாணவர்கள் மத்தியில் பெரும் வரவேற்பைப் பெற்றுள்ளது. காணொளி மூலம் கற்பிக்கப்படும் பாடங்கள் அனைத்தும் மாணவர்கள் மனத்தில் பசுமரத்தாணி போல பதிந்து விடுகிறது.

முடிவுரை

ஆகவே, இந்தக் கட்டுரை தொடக்கப்பள்ளி, இடைநிலைப்பள்ளி, உயர்கல்விக்கூடங்கள் அனைத்திலும் தமிழ் மொழி கற்பிக்கும் ஆசிரியர்கள் அனைவருக்கும் உதவியாக அமையும் என்று நம்புகிறேன். இக்கட்டுரையில் இலக்கணத்தில் இரட்டைக்கிளவிக்கும் இலக்கியத்தில் பழமொழிக்கும் மட்டுமே காணொளி தயார் செய்து ஆராயப்பட்டுள்ளது. தமிழ் மொழி ஆசிரியர்கள் இம்முறையைக் கொண்டு வேறு கூறுகளையும் கற்பிக்கலாம் ஆகவே, தமிழ் மொழி கற்றல் கற்பித்தலில் கணினி மூலம் மாற்றம் செய்திடுவோம்; தமிழ் மொழியைத் தரணி எங்கும் பரப்பிடுவோம்.

மேற்கோள் நூல்கள்

டாக்டர் சொ. பரமசிவம், (2008). நற்றமிழ் இலக்கணம், பட்டுப் பதிப்பகம், சென்னை.

பேராசிரியர் வி.கணபதி, (2010). நற்றமிழ் கற்பிக்கும் முறைகள், சாந்தா பப்ளிஷர்ஸ், சென்னை.

அ.கி. பரந்தாமனார், (2009). நல்ல தமிழ் எழுத வேண்டுமா ?, அல்லி நிலையம், சென்னை.

காணொளி என்றால் என்ன?

<http://ta.wiktionary.org/wiki/%E0%AE%95%E0%AE%BE%E0%AE%A3%E0%AF%8A%E0%AE%B3%E0%AE%BF> (Accessed on July 2013)

Penggunaan video dalam Pengajaran dan Pembelajaran. <http://www.scribd.com/doc/25832190/Penggunaan-Video-Dalam-P-P> (Accessed on July 2013)

Structure of brain, <http://serendip.brynmawr.edu/bb/kinser/Structure1.html> (Accessed on July 2013).

கிராமப்புற ஆசிரிய கல்விப் பட்ட மாணவர்களின் அறிவியல் கற்பித்தல் திறனை மேல்புல கற்பித்தல் வடிவமைப்பைப் பயன்படுத்திய மின்பொருள் மூலமாக மேம்படுத்துதல்

முனைவர். ச. அமுதா, உதவிப் பேராசிரியர்

கல்விநுட்பவியல் துறை, பாரதிதாசன் பல்கலைக்கழகம், திருச்சிராப்பள்ளி - 620 023.

ஆய்வுச்சுருக்கம்

நமது அன்றாட வாழ்வில் அறிவியல் பெரிதும் பங்கேற்றாலும் அவை மாணவர்களைச் சரியான முறையில் சென்றடைகிறதா என்றால் இல்லை என்பதே நிதர்சனமான உண்மை. மாணவர்கள் பரிட்சை நோக்கிலேயே கற்பதால் கருத்துக்களை நீண்ட நாட்கள் நினைவில்கொள்ள முடிவதில்லை. இந்த நிலை மாறவேண்டுமானால் ஆசிரியர்கள் முதலில் மாறவேண்டும். ஆகவே பின்நாளில் ஆசிரியராகப் போகும் ஆசிரிய மாணவர்களை மையமாக வைத்து செய்யப்பட்டதே இந்த ஆராய்ச்சி ஆகும். எந்த ஒரு புதுமையும் நகர்ப்புறங்களை சார்ந்தவர்களையே சென்றடைகிறது. அந்த நிலையும் மாறவேண்டும் என்ற நோக்கில் கிராமப்புற ஆசிரிய மாணவர்களை இந்த ஆராய்ச்சிக்கு தேர்ந்தெடுக்கப்பட்டது. முதலில் வாய்மொழிமூலம் நடத்தினால் சரிவரப் புரிந்துகொள்ள முடியாது என்ற பாடங்களாகிய செல் பிரிதல், மகரந்தச் சேர்க்கை, சவ்வுடுபரவல் ஆகிய பாடங்களில் மின்பொருள் தயாரிக்கப்பட்டது. பிறகுஅதற்கானஅடிப்படைவிபரம்எடுப்பதற்கானகேள்வித்தாள்கள்உருவாக்கப்பட்டன.

- அறிவியல் கற்பித்தல் திறனுக்கான அளவுகோல்
- மேல்புல அறிவிற்கான விழிப்புணர்வு வினா நிரல்
- மேல்புல கற்பித்தல் வடிவமைப்பு
- தகவல் தொழில்நுட்பம் மற்றும் பல்ஊடக வினா நிரல்

இந்த நான்கு வினா நிரல்களும் உருவாக்கப்பட்டு தகுந்த வல்லுநர்களிடம் மதிப்பீடு செய்யப்பட்டு 54 ஆசிரிய மாணவர்களுக்கு வழங்கப்பட்டது. 54 ஆசிரிய மாணவர்கள் 3 குழுவாக முறையே கட்டுப்பாட்டுக் குழு, சோதனைக் குழு 1, சோதனைக் குழு 2 என்று பிரிக்கப்பட்டது. சோதனை 21 நாட்கள் நடத்தப்பட்டது. முடிவில் எண்ணிக்கை சார்ந்த, தரம் சார்ந்த பகுப்பாய்வு, பலம், பலவீனங்கள், வாய்ப்புகள், சவால்கள் பற்றிய பகுப்பாய்வும் நடத்தப்பட்டது. மின்பொருள் கல்வி நுட்பவியல், தாவரவியல் துறைகளைச் சார்ந்த வல்லுநர்களிடம் கொடுத்து கருத்துக்கள் சேகரிக்கப்பட்டது. பகுப்பாய்வின் முடிவில் மேல்புல அறிவிற்கான ஆற்றுப்படுத்தல் பயிற்சி கொடுக்கப்பட்ட குழு மற்ற இரண்டு குழுக்களையும்விட அதிக மதிப்பெண் பெற்று, தயாரிக்கப்பட்ட மின்பொருள் அறிவியல் கற்பித்தலை மேம்படுத்துகிறது என்பது நிரூபணமாகிறது.

முன்னுரை

கல்வி என்பது ஒருவன் தன்னை மேம்படுத்திக்கொள்ள உதவும் ஒரு தலைசிறந்த கருவி ஆகும். அதனை நல்லமுறையில் கற்றுக்கொடுக்கும் கடமை ஆசிரியரைச் சாரும். ஆசிரியக்கல்வி என்பது அறிவு மற்றும் திறமை சார்ந்த தொழிற்கல்வியாகும். அறிவியல் என்பது கற்பனைசார்ந்த செய்தி அல்ல. உள்ளதை உள்ளபடியே தெள்ளத்தெளிவாக புரிந்துகொள்ளுதலும் அதனை அன்றாட வாழ்வில் பயன்படுத்துதலுமே அறிவியல் கற்றலின் நோக்கமாகும். அறிவியல் மிக நுட்பமான செயலாக்கம் (Process), வேதியியல் மாற்றங்கள் மற்றும் உடல் இயங்கியல் (Physiological) பற்றியதாகும். பொதுவாக ஒருவனின் முந்தைய கற்றல் அனுபவங்கள் புதிய கற்றலைப் புரிந்துகொள்ள வழிவகுக்கும். ஆனால் அறிவியல் பாடத்தை அவரவர் அறிவிற்கேற்பவோ முந்தைய கற்றல் அனுபவங்களைப் பொறுத்தோ புரிந்துகொள்ள முடியாது. வகுப்பிலுள்ள அனைவருமே ஒரேமாதிரியான சரியான கற்றல் அனுபவத்தைப் பெறவேண்டும். அறிவியல் கற்பித்தல் என்பது உள்ளூர் மக்களின் சமுதாயத் தேவையை வெளிப்படுத்தக்கூடிய உண்மையான, ஆர்வமூட்டக்கூடிய, பயன்பாடுடையதாக இருக்கவேண்டும் (கோத்தாரி, 1966).

அதற்கான வழி மின்பொருளைப் பயன்படுத்திய கற்பித்தல் முறை ஆகும். மின்பொருளைப் பயன்படுத்திய கற்றல், கற்பித்தல் அனுபவங்களைப் பெற தகவல் தொடர்பு தொழில்நுட்பத்தைப் பற்றிய அறிவு தேவைப்படுகிறது. மேலும் உளவியல் அடிப்படையில் பார்க்கும்போது மேல்புல அறிவு என்னும் கருத்தை அடக்கிய கற்பித்தல் வடிவமைப்பைப் பயன்படுத்தி கற்பிக்கும்போது அறிவியல் எளிமையாக மாணவர்களைச் சென்றடைகிறது.

மேல்புல அறிவு என்பது உயர்மட்ட சிந்தனையாகும் (Higher order thinking). இதன்மூலம் ஒருவன் தனது சிந்தனையைப் பயன்படுத்தி செய்கின்ற செயலை சுயதிட்டமிடல், சுயமேற்பார்வையிடல் மற்றும் சுயமதிப்பீடு செய்தலுமாகும். இவ்வாறு செய்வதின்மூலம் செயலின் தன்மை மேம்படுகிறது (அமுதா, 2010). மேல்புல அறிவைப் பயன்படுத்தி தயாரிக்கப்பட்ட மின்பொருள் வடிவமைப்பு கற்பித்தல் திறனை மேம்படுத்துகிறது (அமுதா & இராமகணேஷ் 2011).

ஆய்வின் நோக்கங்கள்

கீழ்க்கண்டவை ஆய்வின் நோக்கங்களாக உருவாக்கப்பட்டுள்ளது.

- கிராமப்புற ஆசிரிய மாணவர்களின் கற்பித்தல் திறனை மேம்படுத்த
- மேல்புல கற்பித்தல் வடிவமைப்பை உருவாக்கவும் மதிப்பீடு செய்யவும்.
- மேல்புல கற்பித்தல் வடிவமைப்பின் அடிப்படையில் மின்பொருள் உருவாக்கவும், மதிப்பீடுசெய்யவும்.

கீழ்க்கண்ட மாறிலிகள் மின்பொருள்மீது ஏற்படுத்தும் தாக்கத்தை அறியவும்.

- அறிவியல் கற்பித்தல் திறன்
- மேல்புல அறிவிற்கான விழிப்புணர்வு
- மேல்புல கற்பித்தல் வடிவமைப்பு
- தகவல் தொழில்நுட்பம் மற்றும் பல்ஊடக வினா நிரல்

முடிவுகளின் அடிப்படையில் புதிய கோட்பாடுகளை உருவாக்க பரிந்துரை செய்தல்.

கருதுகோள்கள்

ஆய்வின் நோக்கங்கள் அடிப்படையில் உருவாக்கப்பட்ட கருதுகோள்கள் பின்வருமாறு

1. ஆசிரிய கல்வி மாணவர்களின் அறிவியல் பயிற்றுவித்தல் சோதனையில் கட்டுப்பாட்டுக்குழு, சோதனைக்குழு ஒன்று மற்றும் சோதனைக்குழு இரண்டிற்குமிடையே நுழைவு நடத்தையில் எந்த வேறுபாடும் இல்லை.

2. ஆசிரிய கல்வி மாணவர்களின் அறிவியல் பயிற்றுவித்தல் சோதனையில் கட்டுப்பாட்டுக் குழு, சோதனைக்குழு ஒன்று மற்றும் சோதனைக்குழு இரண்டிற்குமிடையே இறுதி நடத்தையில் எந்த வேறுபாடும் இல்லை.

ஆய்வுமுறையியல்

இவ்வாய்வு சோதனைமுறை, எண்ணிக்கை சார்ந்த, முறை மற்றும் தரம் சார்ந்த பகுப்பாய்வு போன்றவற்றின் அடிப்படையில் செய்யப்பட்டது.

ஆய்விற்காக நான்கு வினா நிரல்கள் ஆய்வாளரால் தயாரிக்கப்பட்டன. அவை முறையே

- அறிவியல் கற்பித்தல் திறனுக்கான அளவுகோல் - Science Teaching Competence Scale (STCS)
- மேல்புல அறிவிற்கான விழிப்புணர்வு வினாநிரல் - Metacognitive awareness Inventory for science student-teachers (MAISST)
- மேல்புல கற்பித்தல் வடிவமைப்பு - Metacognitive Instructional design Questionnaire for student-teachers (MIDQST)

- தகவல் தொழில்நுட்பம் மற்றும் பல்ஊடக வினா நிரல் - Checklist on student-teachers' knowledge towards ICT and Multimedia components. (CLKICTMC)

அறிவியல் கற்பித்தல் திறனுக்கான அளவுகோல் (Science Teaching Competence Scale - STCS)

இந்த தரஅளவுகோல் 5 பரிமாணக் கூறுகளையும் 35 உருப்படிகளையும் கொண்டது. ஒவ்வொரு வினாவும் நான்கு சாத்தியக்கூறுகளைக் கொண்டது. சரியான பதிலுக்கு 1 மதிப்பெண் வழங்கப்பட்டுள்ளது.

மேல்புல அறிவிற்கான விழிப்புணர்வு வினா நிரல் (Metacognitive awareness Inventory for science student-teachers - MAISST)

இந்தவினாநிரல் 9 பரிமாணங்களையும் 48 உருப்படிகளையும் உள்ளடக்கியது. 5 புள்ளிதரஅளவுகோல்பயன்படுத்தப்பட்டுள்ளது. (எ.கா) எப்பொழுதும்செய்யமாட்டேன் (0 மதிப்பெண்) அவ்வப்போதுசெய்வேன் (1 மதிப்பெண்) சிலசமயங்களில்செய்வேன் (2 மதிப்பெண்) வழக்கமாகச்செய்வேன் (3 மதிப்பெண்) எப்போதும்செய்வேன் (4 மதிப்பெண்) என்பதாகும்.

மேல்புல கற்பித்தல் வடிவமைப்பு (Metacognitive Instructional design Questionnaire for student-teachers - MIDQST)

இந்த நிரல், ஐந்து புள்ளி தரஅளவுகோலைக் கொண்ட 20 வினாக்களை உடையது. அ - எப்போதும் இல்லை. ஆ - அரிதாக. இ - சில சமயங்களில். ஈ - பல சமயங்களில். உ - எப்போதும் என்பதாகும். இவை முறையே 0, 1, 2, 3, 4 என்ற மதிப்பெண்களைக் கொண்டது.

தகவல் தொழில்நுட்பம் மற்றும் பல்ஊடக வினாநிரல் (ICT and Multimedia components - CLKICTMC)

இந்தநிரல் ஆம், இல்லை என்ற கருத்துக்களை தெரிவிப்பதாக உள்ளது. 44 உருப்படிகள் தகவல் தொழில்நுட்பம் மற்றும் பல்ஊடகத்தைப் பற்றியதாகவும் உள்ளது. இவை முறையே 1, 0 என்ற மதிப்பெண்களைக் கொண்டது.

அட்டவணை 1: மூன்றுகுழுக்களின்நுழைவுநடத்தையின்சராசரி

குழு	STC	MAI	MID	CLKICTMC
கட்டுப்பாட்டுக்குழு	10.28	16.72	11.72	27.33
சோதனைக்குழுஐ	9.78	20.56	14.11	26.89
சோதனைக்குழுஐஐ	11.39	20.67	14.56	24.94

குழு	STC	MAI	MID	CLKICTMC
கட்டுப்பாட்டுக்குழு	14.0	20.5	14.83	31.00
சோதனைக்குழுஐ	17.11	23.22	16.33	31.22
சோதனைக்குழுஐஐ	21.0	24.5	17.06	31.94

பகுப்பாய்வின்விபரங்களின்பின்வருமாறு

ஆசிரிய மாணவர்களின் அறிவியல் கற்பித்தல் திறனைப் பொருத்தவரை கட்டுப்பாட்டுக் குழுவைவிட சோதனைக்குழு அதிக சராசரியைப் பெற்றுள்ளது. சோதனைக்குழு ஒன்றைவிட சோதனைக்குழு இரண்டு அதிகமான சராசரியைப் பெற்றுள்ளது. இதிலிருந்து அறிவியல் பயிற்றுவித்தலை மின்பொருள்மூலம் மட்டும் கற்றுக் கொடுக்காமல் அதனைப்பற்றிய ஆற்றுப்படுத்தல் பயிற்சி கொடுக்கும்போது இன்னும் அதிகமாகச் சென்றடைகிறது.

தரம் சார்ந்த பகுப்பாய்வு (Qualitative Analysis)

தாவரவியல் மற்றும் தகவல் தொழில்நுட்பத் துறையைச் சார்ந்த வல்லுநர்களின் பின்னூட்டத் தகவல்கள் கீழே தொகுத்தளிக்கப்பட்டுள்ளது.

- ஆய்வின் நோக்கங்கள் சரியானதாகவும் அடையக்கூடிய வகையிலும் உள்ளது.
- மின்பொருள் உருவாக்கத்திற்கான எழுத்தாக்கம் (script) நன்றாக உள்ளது.
- கலைச்சொற்களும் அதன் அர்த்தங்களும் சிறப்பாக தரப்பட்டுள்ளது.
- மின்பொருள் தயாரிப்பதற்கான வழிமுறைகள் எளிமையாக தரப்பட்டுள்ளது. இதனைப் பயன்படுத்தி ஆசிரிய மாணாக்கர்கள் தானாக மின்கற்றல் பொருளை உருவாக்க முடியும்.
- மேல்புல அறிவு ஆசிரிய மாணாக்கர்களின் பயிற்றுவித்தல் திறனை மேம்படுத்த உதவுவதாக அமைந்துள்ளது.
- மின்கற்றலில் அசையக்கூடிய வீடியோக்கள் அமையப்பெற வேண்டும் என்றும் கருத்து தெரிவித்துள்ளனர்.

SWOT பகுப்பாய்வின் மூலம் பெறப்பட்ட தகவல்கள்

- ஒப்படைப்பு தொழில்நுட்பம் (Delivery) நன்றாக உள்ளது
- மின்கற்றல் உருவாக்க தேவையான தகவல்களைப் பெறக்கூடிய வலைதளங்கள் தரப்பட்டுள்ளது
- மேல்புல அறிவு பற்றிய விழிப்புணர்வு மாணவ ஆசிரியர்களுக்கு பயனளிக்கக் கூடியது

பலவீனங்கள்

- தாய்மொழியில் மின்பொருள் இல்லாமை
- குறிப்பிட்ட தலைப்புகளுக்கு மட்டும் மின்பொருள் உருவாக்கப்பட்டுள்ளது
- தகவல் தொழில்நுட்ப அறிவு இருந்தால் மட்டுமே பயன்படுத்த முடியும்

வாய்ப்புகள்

- கற்பித்தலுக்கு மின்பொருள் ஒரு சிறந்த துணைமூலமாகமுகம்
- மேல்புல அறிவை பற்றிய விழிப்புணர்வு எற்படுத்துவதன்மூலம் திட்டமிடல், மேற்பார்வையிடல், ஆய்வுசெய்தல் (Evaluate) போன்றவற்றின் மூலம் ஆசிரியர்கள் தங்களை மேம்படுத்திக்கொள்ள முடியும்
- ஊடகத்தின்மூன் நின்று பேசக்கூடிய தயக்கம் நீங்கும்

சவால்கள்

- மின்பொருள் உருவாக்கம் காலத்தை எடுத்துக்கொள்ளக் கூடியது
- மின்பொருள் உருவாக்கப் பயன்படும் மென்பொருள், மற்றும் வன்பொருள் அதிக பணம்
- மின்பொருளுக்கான எழுத்தாக்கம் (Script) கடினமான செயலாகும்
- ஆற்றுப்படுத்தல் பயிற்சி சிறந்த வல்லுநர்கள் மூலம் தரப்பட வேண்டும்.

பரிந்துரைகள்

மேற்கண்ட புள்ளிவிபர அடிப்படையில் கீழ்க்கண்ட பரிந்துரைகள் செய்யப்படுகின்றன.

- மின்கற்றலை உருவாக்க தேவையான அடிப்படை வசதிகளை ஒவ்வொரு கல்வி நிறுவனத்திலும் செய்துதர வேண்டும்.
- மின்பொருளை சிறந்த முறையில் பயன்படுத்த ஆற்றுப்படுத்தல் பயிற்சி அளிக்கப்படவேண்டும்.
- மேல்புல அறிவு, அறிவியல் கற்பித்தல் திறனை அதிகரிக்கச் செய்வதால் ஆசிரியக்கல்விக் கலைத்திட்டத்தில் இதனை இணைக்கவேண்டும்.

முடிவுரை

இன்றைய தொழில்நுட்ப வளர்ச்சியின் காரணமாக உலகமே ஒரு சிறிய கிராமமாக சித்தரிக்கப்படுகிறது. மின்கற்றல்மூலம் உலகிலுள்ள யாரையும் எப்பொழுது வேண்டுமானாலும் தொடர்பு கொள்ளமுடியும், தகவல் மற்றும் கருத்துக்களை பரிமாறிக்கொள்ள முடியும். அறிவியல் கற்பித்தல் என்பது மக்களின் சமுதாயத் தேவையை வெளிப்படுத்தக்கூடிய உண்மையான, ஆர்வமூட்டக்கூடிய, பயன்பாடுடையதாக இருக்கவேண்டும் என்ற கோத்தாரியின் கூற்றுப்படி; அறிவியல் கற்பித்தல் மற்றும் கற்றலின் இடைவெளி மின்கற்றல் மூலம் குறைக்கப்படுகிறது என்பது இவ்வாய்வின்மூலத்தெளிவாக புரிந்துகொள்ள முடிகிறது.

உசாத்துணைகள்

- Amutha, S., & Ramganes, E. (2010). Application of Metacognition in the Instructional Design and e-content development, e-Resources in Higher education, International Conference Proceedings, Department of Educational Technology, Bharathidasan University, Tiruchirappalli. 121-124.
- Amutha, S., & Ramganes, E. (2011). Perception of Student teachers of science at secondary level about Metacognitive instructional Design, Experiments in Education, Vol 38 (2), 19-24.
- Kothari Commission. (1960, 1966) retrieved from <http://www.education.nic.in/cd50years/g/t/HB/0THB0J01.htm>.
- National Commission on Excellence in Education. retrieved from <http://www2.ed.gov/pubs/NatAtRisk/appenda.html>.
- Education Commission, (1966) Retrieved from <http://www.education.nic.in/cd50years/q/3B/7C/3B7C0701.htm>.

தமிழ் ஆசிரியர்களிடையே பலகைக் கணினிகளின் வழி தமிழ் வாசிப்பதிலும் சமூகக் கட்டமைப்பு வலைகளில் தமிழ் பதிப்பதிலும் உள்ள சிந்தனைப் போக்கு: ஒரு பார்வை.

இல.வாசுதேவன், மலேசியா. மின் அஞ்சல்: vas_devan@hotmail.com

1.0 முன்னுரை:

உலகம் ஊடகத் துறையில் உன்னதத்தைத் தொட்டுவிட்டது; மாந்தர் அதன் ஊடே வினையாட வேண்டிய சூழல் ஏற்பட்டுவிட்டது. தகவல் தொடர்புத் தொழில்நுட்பத்தின் வளமையோ உலகமயமாதலின் கட்டமைப்பில் விரிந்து பல எல்லைகளைக் கடந்தும் விட்டது. பயனர்களின் தகவல்களும் பல்கிப்பெருகி விட்டதால் மாற்றுக்கொள்கலனாக, இன்று மேகக் கணிமை தோன்றிவிட்டது. இந்நிலை சாத்தியமானதில் இணையத்தின் பங்களிப்பு அளப்பறியது. தமிழ்ப் பயனர்களுக்கு இன்று மாற்று நவீன கையடக்கக் கருவியாக விளங்குவது பலகைக்கணினிகள் என்றால் மிகையிலலை.

கல்வியுலகில் குறிப்பாகக் கற்றல் கற்பித்தலில் செல்பேசிகளின் சிறிய திரைகளைக் காட்டிலும் ஐபேட் (iPad) போன்ற சற்றே பெரிய திரைகளே பிரபலம் என்பது ஆசிரியர்களின் பொதுவான கருத்து. இவ்வகை பலகைக்கணினிகளின் பயன்பாடு தமிழ் ஆசிரியர்களிடையே எவ்வளவு வளர்ச்சியும் பிரபலமும் அடைந்துள்ளது? அதோடு ஐ.ஓ.எஸ், ஆண்டிராய்டு மற்றும் விண்டோஸ் இயங்கு தளங்களின் வழி தமிழ் வாசிப்பதிலும் சமூகக் கட்டமைப்பு வலைகளான வலைப்பூ (Blog), ஃபேஸ்புக் (Facebook), டிவீட்டர் (Twitter) எனும் குறுஞ்செய்தியோடைகளில் தமிழ் பதிப்பதிலும் அவர்களின் சிந்தனைப் போக்கும் ஆர்வமும் எப்படி இருக்கிறது? என ஆய்வின் வழி கட்டுரையாளர் தெரிந்து கொள்ள விழைகிறார்.

எனவே, அதன் அடிப்படையில் மலேசியாவின் சில மாநிலங்களில் தேர்ந்தெடுக்கப்பட்ட தமிழ் ஆசிரியர்களிடம் ஆய்வு மேற்கொள்ளப்பட்டு தரவுகளின் ஆய்வு முடிவுகள் கட்டுரையில் இடம்பெற்றுள்ளன.

2.0 கல்வியுலகு : மடிக்கணினியிலிருந்து பலகைக்கணிக்கு மாறும் புதிய செல்நெறி

கல்வியுலகில் புதிய தொழில்நுட்பங்களின் வரிசையில் கையடக்கக்கணினிகளின் வரவில் தற்போது பலகைக்கணினியே பிரபலமடைந்து வருகிறது. இந்த நவீன இலத்திரன் தொழில்நுட்ப வரவு தொடக்கப்பள்ளி முதல் உயர்க்கல்விக்கூடமான பல்கலைக்கழகம் என்று மட்டுமல்லாமல், அதையும் தாண்டி வெவ்வேறு துறைகளிலும் மிளர்கிறது. அதன் முகாமை பயனர்களாக, ஆசிரியரும் மாணவரும் திகழ்கிறார்கள்.

கல்வியுலகையானாலும் 'மாற்றி யோசி' சிந்தனை உலகெங்கும் வலுவடைந்து வருகிறது. 'ஸ்டீவ் ஜோப்ஸ் பள்ளிகள்' (Steve Jobs School) என்ற கொள்கை பெயரில் பதினொரு பள்ளிகள் ஆகஸ்டு, 2013 – நெதர்லாந்து நாட்டில் தொடங்கிவிருக்கிறது. முற்றிலும் 'ஐபேட்' (iPad) பலகைக்கணினியைக் கொண்டே மின்வெளி பாடத்திட்டங்கள் மூலம் கற்றல் நடைபெறவிருக்கிறது. ஆசிரியர்கள் இங்கு 'கற்றல் பயிற்றுனர்' (Learning Coaches) என்றே செயல்படுவர். கரும்பலகைகள் இன்றி, வெண்கட்டி இன்றி, பாரம்பரிய வகுப்பறைகள் இன்றி, வழமையான கற்றல் கற்பித்தல் இன்றி, பாடநூல்கள் இன்றி, பள்ளிச்சூழல் உருமாற்றம் பெற்று வருகிறது. இப்படி நிறையவே மாற்றி யோசித்து இருக்கிறார்கள். 4 வயது முதல் 12 வயது மாணவர்கள் பயிலும் இவ்வகை பள்ளிகளில், ஆறு வாரங்களுக்கு ஒருமுறை பெற்றோர் – கற்றல் பயிற்றுனர் சந்திப்பில் மாணவர்களின் அடைவு நிலைக்கு ஏற்ப அடுத்த நிலைக்கான மாணவர் கற்றல் திறன்கள் முடிவுசெய்யப்படும்.

பொதுவாக பலகைக்கணினிகளின் (Tab or Tablet) குறிப்பிடத்தக்கப் பயன்பாடுகள் யாதெனில்,

- * மிக விரைவாக இலத்திரன் குறிப்பெடுக்கும் வசதி உள்ளது
- * நான்காட்டி மேலாண்மை, அதிகப்படியான வேலைச்சூழலுக்கு உதவுகிறது
- * சொற்பொழிவு, படைப்புக்குத் தேவையான குரல் மற்றும் காணொளி பதிவை எளிமையாக்குகிறது
- * அதிகப்படியான இலத்திரன் படங்களை நிறைவாய் காண்பிக்கிறது
- * மின்நூல் வாசிக்க ஏதுவாகிறது; இனி அனைத்து பாடநூல்களையும் எடுத்துவர தேவையில்லை.
- * ஒருங்கிணைக்கப்பட்ட வைஃபை (WiFi) மற்றும் 3G தொழில்நுட்பம் இணையத்தில் தொடர்பு கொண்டு ஆய்வுக்கும் கூட்டுக்கல்விக்கும் உதவுகிறது.

அடிப்படையில் மடிக்கணினியைவிட பலகைக்கணினிகளின் எடை குறைவு. எனவே, அதை இலகுவாக எங்கு வேண்டுமானாலும் ஆசிரியர்களும் மாணவர்களும் எடுத்துச் செல்லலாம். இத்தகைய தாக்கங்களுக்கும் மாற்றங்களுக்கும் தமிழ் ஆசிரியர்கள் தயாராகி வருகிறார்களா? இலத்திரன் இடைவெளியைக் (Digital divide) குறைக்கிறார்களா? பலகைக்கணினிவழி சமூகக் கட்டமைப்பு வலை (Social Networks), இதர தளங்களில் தமிழ்

வாசிப்பதிலும் இணையத்தில் பதிப்பதிலும் ஆசிரியர்களின் சிந்தனைப் போக்கு எவ்வாறு இருக்கிறது என்பன போன்ற கேள்விகளுக்கு விடை காணும் வகையில் இவ்வாய்வு மேற்கொள்ளப்பட்டது.

3.0 பலகைக்கணினிகளின் வழி தமிழ் வாசிப்பதிலும் சமூகக் கட்டமைப்பு வலை போன்ற நவீன தகவல் தொடர்பு நொழில்நுட்ப ஊடகங்களில் தமிழ்ப்பள்ளி ஆசிரியர்களின் ஈடுபாடு:

சமூகக் கட்டமைப்பு வலையிலிருந்துதான் கல்வி கட்டமைப்பு வலை பரிணமிக்கிறது. இவ்விரு கட்டமைப்புகளும் அடிப்படையில் ஒன்றெனினும் அதன் நோக்கங்கள் வெவ்வேறானவை. கல்வி கட்டமைப்பு வலையானது கல்வித் துறைச் சார்ந்ததாகும். அவ்வாறு கல்வி கட்டமைப்பு வலைவழி வளப்படுத்தப்படும் குழுவும், வலை கட்டமைப்பு உருவாக்கத்திற்குப் பிறகு இலக்கு நோக்கி பரிணமிக்கப்படுதல் அவசியமாகிறது. அதன் வளர்ச்சிக்கும் விளைபயன்மிக்க மாற்றங்களுக்கும் போதிய கால அவகாசமும் தொடர்ந்த ஒருங்கிணைக்கப்பட்ட கூட்டுமுயற்சியும் தேவைப்படுகிறது.

இன்றைய காலச்சூழலில், தமிழ் ஆசிரியர்கள் பலகைக்கணினியின்வழி தமிழ் வலைப்பூ (Blog), முகநூல் (FaceBook), டிவீட்டர் (Twitter) என்னும் குறுஞ்செய்தியோடைகளில் தமிழில் வாசிப்பதற்கும், மறுமொழி இடுவதற்கும் ஆர்வம் காட்டி வருகிறார்கள். ஆசிரியர்களின் ஈடுபாட்டினை ஓர் ஆய்வு மூலம் பெறுவதற்கு கட்டுரையாளர் முனைந்தார்.

4.0 ஆய்வின் நோக்கம்:

மலேசிய தமிழாசிரியர்களிடையே பலகைக்கணினியின் பயன்பாடு, அதன்வழி இணையத்தில் தமிழ் வாசிப்பும் பதிப்பும் அவர்களின் சிந்தனைப் போக்கும் எவ்வாறு உள்ளது என்பதை அறியவும் இந்த ஆய்வு மேற்கொள்ளப்பட்டது.

4.1 ஆய்வுக்குட்பட்ட ஆசிரியர்கள்:

இந்த ஆய்வில் ஜோகூர், மலாக்கா, சிலாங்கூர், பேராக், கூட்டரசுப் பிரதேசம், பினாங்கு, கெடா என மலேசியாவின் ஏழு மாநிலத்தைச் சார்ந்த ஆசிரியர்கள் பங்கு பெற்றனர்.

4.2 ஆய்வு முறைமை

பலகைக்கணினி பயன்படுத்தி தமிழ் வலைதளங்கள், வலைப்பூ அல்லது ஃபேஸ்புக் போன்ற சமூக வலை கட்டமைப்புகளில் தமிழ் வாசிப்பு மற்றும் தமிழில் பதிப்பதிலும் ஆசிரியர்களின் சிந்தனைப்போக்கை அறிய ஆய்வுக்காக 150 ஆய்வு வினாத்தாள்கள் வழங்கப்பட்டது. அவற்றுள் 90 ஆய்வு வினாத்தாள்களே முழுமைபெற்ற தரவுகளாக ஆய்வுக்கு கிடைக்கப்பெற்றது.

5.0 ஆய்வு வினாத்தாள் / தரவுகள்

ஆய்வுக்குட்பட்ட 90 தரவுகளில் ஆண் ஆசிரியர்கள் 42 பேர், பெண் ஆசிரியர்கள் 48 பேர். அவர்களுக்கு ஆய்வுக்கான வினாத்தாள் வழங்கப்பட்டது. அவை கீழ்க்காணும் துணைக் கேள்விகளை ஒட்டி இருந்தன.

- பலகைக்கணினி (TABS) பயன்பாடு
- பலகைக்கணினியில் இணையம் பயன்பாடு:
- தமிழில்@மாற்றுவழியில் (Romanized Tamizh) நான் பயன்படுத்தும் சமூகக் வலைக்கட்டமைப்பு (Social Networks)
- பலகைக்கணினியில் (TABS/iPAD) பயன்பாட்டு நேரம்
- பலகைக்கணினிவழி (TABS/iPAD) தமிழ் மின்னூல்கள் (Tamil eBook) வாசிப்பு
- பலகைக்கணினிவழி (TABS/iPAD) தமிழில் சமூக வலைக்கட்டமைப்புகளில் பதிப்பு
- 3-வது தலைமுறை செல்பேசி (3G/ smart phone Screen) திரை & பலகைக்கணினி திரை (TABS/ Tablet PC /iPAD Screen) வழி இணைய தமிழ் வாசிப்பு
- தமிழ் கற்பித்தலில் பலகைக்கணினி (TABLET PC/iPADs) & 3ஜி செல்பேசி பயன்பாடு
- பலகைக்கணினி வழி இணையம் சமூகக் கட்டமைப்பு வலையானது தமிழாசிரியர்கள்

6.0 பகுத்தாய்வு

இந்த ஆய்வுக்கான தரவுகள் அனைத்தும் 'SPSS Version 20.0' மென்பொருளைக் கொண்டு பகுத்தாயப்பட்டன.

அட்டவணை : 1 ஆசிரியர்களின் கற்றல் கற்பித்தல் அனுபவம்

	அனுபவம்	எண்ணிக்கை	விழுக்காடு
1.	< 2 ஆண்டுகள்	4	4.44
2.	2 - 5 ஆண்டுகள்	15	16.67

3	6 - 10 ஆண்டுகள்	16	17.78
4	11 - 15 ஆண்டுகள்	31	34.44
5	> 16 ஆண்டுகள்	24	26.67
	மொத்தம்	90	100.0

அட்டவணை 1-இல் குறிப்பிட்டுள்ளபடி, 4.44 விழுக்காடு ஆசிரியர்கள் 2 ஆண்டுக்கும் குறைவான கற்றல் கற்பித்தல் அனுபவம் பெற்றுள்ளனர். 2 இருந்து 5 ஆண்டுகள் அனுபவம் பெற்றவர்கள் 15 பேர் (16.67%). ஒளவுக்குட்பட்ட ஆசிரியர்களில் 16 பேர் (17.78%) 6 இருந்து 10 அண்டுகள் வரையில் அனுபவத்தைப் பெற்றுள்ளனர். 11 இருந்து 15 ஆண்டுகள் அனுபவம் பெற்றவர்கள் 31 பேர் (34.44%). மேலும், 16 ஆண்டுகளுக்கு மேல் ஆசிரியராக அனுபவம் பெற்றவர்கள் 24 பேர் (26.67%) ஆவர்.

அட்டவணை:2 : பலகைக்கணினி வழி சமூக / கல்வி கட்டமைப்பு வலை பயன்பாடு

பலகைக்கணினி பயன்பாடு	முற்றும் மறுத்தல்	மறுத்தல்	உறுதியின்மை	ஏற்றல்	முற்றும் ஏற்றல்
1. தமிழில் ஒருங்குறி	12 (13.33 %)	0 (0 %)	0 (0 %)	45 (50.00 %)	33 (36.67 %)
2. தமிழில் மின்னஞ்சல்	10 (11.1 %)	0 (0 %)	2 (2.22 %)	46 (51.11%)	32 (35.56 %)
3. ச.க.வலை நிர்வகிக்கிறேன்	13 (14.44 %)	22 (24.44 %)	0 (0 %)	25 (27.78 %)	30 (33.33%)
4. வலைப்பூ (Blog)	23 (25.56 %)	10 (11.11 %)	0 (0 %)	15 (16.67 %)	42 (46.67 %)
5. டிவீட்டர் (Twitter)	40 (44.44 %)	12 (13.33%)	0 (0 %)	23 (25.56 %)	15 (16.67 %)
6. ஃபேஸ்புக் (FaceBook)	10 (11.1 %)	19 (21.11 %)	0 (0 %)	28 (31.11 %)	33 (36.67 %)

அட்டவணை 2-இல் பலகைக்கணினி வழி தமிழில் ஒருங்குறி பயன்படுத்துவோர் 78 பேர் (86.67%), பயன்படுத்தாதவர்கள் 12 பெர் (13.33 %) ஆவர். ஆசிரியர்களில் தமிழில் மின்னஞ்சல் பயன்படுத்துவோர் 78 பேர் அல்லது 86.67 விழுக்காட்டினர். தமிழில் பயன்படுத்தாதவர்கள் 10 பேர் அல்லது 11.1 விழுக்காட்டினர். உறுதியின்மை என்றவர்கள் 2 (2.22 %) பேர். ஆசிரியர்களில் சமூக கட்டமைப்பு வலை நிர்வகிப்பவர்கள் 55 (61.11 %) பேர். நிர்வகிக்காதவர்கள் 35 (37.3 %) பேர். சமூக கட்டமைப்பு வலைகளான வலைப்பூ, டிவீட்டர் மற்றும் பேஸ்புக் ஆகிய மூன்றிலும் முறையே 57 (63.33%) பேர், 38 (42.22%) பேர், 61 (67.78%) பேர் பயன்படுத்துகிறார்கள். பயன்படுத்தாதவர்கள் முறையே வலைப்பூவில் 33 (36.67 %) பேர், டிவீட்டரில் 52 (57.78 %) பேர் மற்றும் ஃபேஸ்புக்கில் 29 (32.22 %) பேர். சமூக கட்டமைப்பு வலை பயன்பாட்டில் உறுதியின்மை என்றவர்களில் வலைப்பூ 40 (23.7 %) பேர், டிவீட்டர் 53 (31.4 %)பேர், மற்றும் ஃபேஸ்புக் 40 (23.7 %) பேர் ஆவர்.

அட்டவணை:3 : பலகைக்கணினி பயன்பாடும் தமிழில் மின்னூல் வாசிப்பும்

பலகைக்கணினி பயன்பாடும் & தமிழில் மின்னூல் வாசிப்பும்	முற்றும் மறுத்தல்	மறுத்தல்	உறுதியின்மை	ஏற்றல்	முற்றும் ஏற்றல்
1. ஒவ்வொரு நாளுமும்	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)	31 (36.67 %)
2. வாரத்தில் 1 நாள்	0 (0 %)	0 (0 %)	0 (0 %)	12 (13.33%)	0 (0 %)
3. வாரத்தில் 2 - நாட்கள்	0 (0 %)	0 (0 %)	0 (0 %)	15 (27.78 %)	0 (0%)
4. வாரத்தில் 4 - 5 நாட்கள்	0 (0 %)	0 (0 %)	0 (0 %)	20 (22.2 %)	0 (0 %)
5. வாரக் கடைசி மட்டும்	0 (0 %)	0 (0 %)	0 (0 %)	12 (13.33 %)	0 (0 %)
6. தமிழில் மின்னூல் வாசிப்பு	10 (11.1 %)	26 (28.89 %)	0 (0 %)	34 (37.78 %)	20 (22.22 %)
7. புத்தாக்கச்சிந்தனை	12	14	14	32	18

ஏற்படுகிறது	(11.1%)	(15.56 %)	(15.56 %)	(35.56 %)	(20.00 %)
-------------	---------	-----------	-----------	-----------	-----------

அட்டவணை 3-இல் தமிழ் ஆசிரியர்கள் பலகைக்கணினி பயன்படுத்தும் காலத்தைக் காட்டுகிறது. அவ்வாறு பார்க்கையில், ஒவ்வொரு நாளும் பயன்படுத்துபவர்கள் 31 பேர் (36.67 %), வாரத்தில் ஒரு நாள் மட்டும் பயன்படுத்துவோர் எண்ணிக்கை 12 பேர் (13.33%), வாரத்தில் 2 அல்லது 3 நாட்கள் பயன்படுத்துவோர் 15 பேர் (27.78 %), வாரத்தில் 4 அல்லது 5 நாட்கள் பயன்படுத்துவோர் 20 பேர் (22.2 %) ஆவர். அவர்களில் வார இறுதியில் மட்டும் பயன்படுத்துவோர் 12 பேர் (13.33 %). பலகைக்கணினியின் வழி மின்நூல்கள் வாசிப்பவர்கள் 54 பேர் (60.00 %) ஆவர். பலகைக்கணினி வழி வாசிப்பு அவர்களின் புத்தாக்கச் சிந்தனையைத் தூண்டுவதாகக் குறிப்பிட்டவர்கள் மொத்தம் 50 பேர் (55.56 %), இல்லை என்று மறுத்தவர்கள் 36 பேர் (40.00%). இருப்பினும் உறுதியில்லை என்று தெரிவித்தவர்கள் 14 பேர் (15.56%) ஆவர்.

7.0 ஆசிரியர்கள் பலகைக்கணினி வழி இணையத்தில் சமூக / கல்வி கட்டமைப்பு வலை பயன்பாட்டில் எதிர்கொள்ளும் சவால்கள்:

ஆசிரியர்கள் பலகைக்கணினி வழி இணையத்தில் தமிழ் சமூக / கல்வி கட்டமைப்பு வலை தொழில் நுட்பம் பயன்படுத்தும்போது எதிர்கொள்வதாகக் குறிப்பிட்ட சவால்கள் இங்கு தொகுத்து வழங்கப்படுகின்றன (அட்டவணை : 4).

அட்டவணை : 4 ஆசிரியர்கள் எதிர்கொள்ளும் சவால்கள்

எண்	சவால்கள்
1.	பலகைக்கணினிவழி பயன்படுத்தும் முறையை முழுமையாக தெரிந்திருக்கவில்லை.
2.	ஆசிரியர்களிடையே தொழில் நுட்ப விழிப்புணர்வு குறைவு.
3.	பள்ளிச் சூழல் பயன்படுத்த தடையாக உள்ளது.
4.	ஆண்டிராய்டு அல்லது ஐ.ஓ.எஸ் இயங்கு தளங்கள் பற்றிய தகவல் அறியாமை
5.	அடிப்படையில் கணினி, இணையம் பயன்படுத்தும் ஆர்வம் குறைவு.
6.	ஆசிரியர்களின் வேளைப் பளு (நிர்வாகப் பணிகள், பள்ளி நடவடிக்கைகள்) பெருந்தடை
7.	தமிழ் ஒருங்குறி பயன்படுத்த தகவல் தெரியாமை.
8.	தமிழில் மின்னூல் வாசிப்பது குறைவு
9.	இணையவழி கல்வி பற்றிய விழிப்புணர்வின்மை.
10..	சமூக கட்டமைப்பு வலை தொடர்பான தகவல்கள் அறியாமை.
	ஆசிரியர்களிடையே சமூக / கல்வி கட்டமைப்பு வலைகளில் தமிழ் கற்றல் கற்பித்தல் தொடர்பான கலந்துரையாடல்கள் மிகக் குறைவு.
	இலவச குறுச்செயலிகள் (Free Apps) உருவாக்கம் பற்றிய போதிய தகவல் தெரியாமை

8.0 ஆசிரியர்களுக்கான பரிந்துரைகள்:

ஆய்வில் கலந்து கொண்ட 90 ஆசிரியர்கள் ஆசிரியர்கள் எதிர்நோக்கும் தொழில்நுட்பம் சார்ந்த பிரச்சனைகளுக்குத் தீர்வு காண பின்வரும் பரிந்துரைகளை மேற்கொள்ளலாம்.

ஆசிரியர்களுக்கான சில பரிந்துரைகள்

பரிந்துரைகள்:
<ul style="list-style-type: none"> ➤ பலகைக்கணினி வழி சமூக / கல்வி கட்டமைப்பு வலை தொடர்பான பயிலரங்கில் ஆர்வமுடன் கலந்து கொள்ளல் ➤ ஆசிரியர்கள் தொழில் நுட்ப அறிவு வளர்த்துக் கொள்ள தாமாகவே முயற்சி மேற்கொள்ளல். ➤ ஆண்டிராய்ட் அல்லது ஐ.ஓ.எஸ் இயங்கு தளங்கள், குறுஞ்செயலிகள் பற்றிய தகவல்கள் அறிந்துகொள்ளல். ➤ ஆசிரியர்களிடையே சமூக/கல்வி கட்டமைப்பு வலைகளால் ஏற்படும் நன்மைகளை எடுத்துரைத்து ஆர்வத்தைக் கூட்டுதல். ➤ இணையக் கல்வியின் அவசியத்தை தமிழாசிரியர்களின் கவனத்திற்கு கொண்டு செல்லுதல். ➤ தமிழ் ஒருங்குறி பயன்படுத்தும் வழிமுறைகளை ஆசிரியர்களுக்குப் பயிற்சியின் வழி கற்றுக் கொடுத்தல். ➤ பலகைக்கணினி வழி நடைபெற்றுவரும் கல்வியுலக மாற்றம் குறித்த விழிப்புணர்வு பெறுதல். ➤ இணையவழி தமிழ் மின்னூல்கள் வாசிக்க, பதிக்க ஆர்வத்தைக் கூட்ட வேண்டும். ➤ கையடக்கக்கணினி வழி சமூக கட்டமைப்பு வலை தொடர்பான கருத்தரங்குகள், பட்டறைகள் போன்றவற்றில் கலந்து கொள்ளல். ➤ நவீன கையடக்கக்கணினிகளின் வழி தமிழாசிரியர்களிடையே ஆக்கரமான கலந்துரையாடலுக்கும்,

9.0 தொகுப்பாய்வு:

ஆய்வில் பங்கெடுத்த 90 பேரில் 78 பேருக்கு இணையத்தில் தமிழ் ஒருங்குறி பயன்படுத்தத் தெரிந்திருப்பது மகிழ்ச்சியளிக்கிறது. அவர்கள் தமிழிலும் மின்னஞ்சல் அனுப்பும் முறை அறிந்திருக்கிறார்கள். பலகைக்கணினி பயன்படுத்தி தமிழில் மின்னூல்கள் வாசிப்போர் எண்ணிக்கை 60 விழுக்காட்டினர் (54 பேர்) என்பதும் வரவேற்கக்கூடிய தகவலாகும். அந்நடவடிக்கை அவர்களின் புத்தாக்கச் சிந்தனையை அதிகரிக்கிறது என்று குறிப்பிட்டவர்கள் 50 பேர் அதாவது 55.56 விழுக்காட்டினர். இந்த ஆய்வின் வழி மலேசிய தமிழ் ஆசிரியர்களிடையே கையடக்கக்கணினியான பலகைக்கணினியின் பயன்பாடு அதிகரித்து வருவதை அறிகிறோம். அதே வேளையில், பயனாளர் எனும் அடிப்படையில் தமிழ் ஆசிரியர்கள் தகவல் தொழில்நுட்பம் சார்ந்த தகவல்களை ஓரளவிற்கு அறிந்தே வைத்திருக்கிறார்கள் என்பது தெரியவருகிறது. இருப்பினும் இந்த பலகைக்கணினி பயன்பாட்டின் செல்நெறி இன்னும் பரவலாக தமிழ் ஆசிரியர்களிடையே புலக்கத்தை அதிகரிக்க வேண்டும் என்பதே ஆய்வாளரின் அவா ஆகும்.

10.0 முடிவுரை

கல்வியியல் என்றால் முதன்மை பொறுப்பாளர்களாகத் திகழ்பவர்கள் ஆசிரியர்தாம். தமிழ்மொழி கற்பித்தலிலும் துறை சார்ந்த தகவல் தொழில்நுட்ப மாற்றங்களுக்கு ஈடுகொடுப்பவர்களாக இருப்பவர்களும் தமிழ் ஆசிரியர்களே. இன்றைய தகவல் தொழில்நுட்ப செல்நெறி இரண்டு பிரமாண்டமான பெரிய சவால்களை எதிர்கொள்கிறது. முதலாவது நாம் ஏற்கிறோமோ இல்லையோ, நம் வாழ்வின் ஓர் அங்கமாக 'இலத்திரன் வாழ்க்கை' (digital life) அமைந்து விட்டது. இலத்திரன் ஒப்பம், நம் ஏற்றுக்கொண்ட ஒப்பம். அன்றாட வாழ்வின் அத்தனை தகவல்களும் தற்போது இலத்திரன் மேகக்கொள்கலனில் (digital cloud) சேமிக்கப்படுகிறது. இதன் கட்டுக்கடங்காத தகவல் சேமிப்புதான் இரண்டாவது பெரிய சவால். 'பெருந்தகவல்' (big data) என்று கூறுப்படுகிறது. எனவே, தமிழ் ஆசிரியர்கள் பலகைக்கணினிகளின் வெறும் பயனர்களாக மட்டும் இருக்கலாகாது. தகவல் தொழில்நுட்பம் அறிந்த, கொஞ்சம் நிரல்மொழி தெரிந்த, கற்றல் கற்பித்தலுக்கான தமிழ் 'குறுஞ்செயலிகள்' (Apps) உருவாக்கும் திறன் கொண்டவர்களாக தங்களை மறு உருமாற்றத்திற்கு வித்திட வேண்டும். தமிழாசிரியர்களின் கல்விச் சிந்தனை மாற்றங்களினால்தான் ஆக்கப்பூர்வமான கற்றல் கற்பித்தலை மாணவர்களுக்கு வழங்க முடியும். எனவே, தமிழாசிரியர்கள் கையடக்கக்கணினியான பலகைக்கணினி வழி மின்னூல்கில் கல்வியியல் தொடர்பான வலைகளில் மின் தமிழில் புலங்கவும் பயன்விளைவு மிக்க மாற்றங்களுக்கு வித்திடுபவர்களாகத் திகழ வேண்டும்.

துணை செய்த நூல்கள் / ஆவணங்கள்:

- Boyd, D.M. & Ellison, N. B. (2008). *Social Network Sites: Definition, History, and Scholarship*. Michigan State University. MI, USA
- Clark, C. & Gruba, P. (2010). The use of social networking sites for foreign language learning: An autoethnographic study of *Livemocha*. In C.H. Steel, M.J. Keppell, P. Gerbic & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney 2010* (pp.164-173).
- Geisert, P.G & Futnell, M.K (1995). *Teachers, Computers, and Curriculum: Microcomputers in the Classroom*. Allyn & Bacon., A Simon & Schuster Company, NeedhamHeights, Massachusetts.
- Rajendran, Nagappan (2008). *Malaysiavil Tamil Kalviyum Katral Karpittalum*, Uma Publications. Kuala Lumpur.

துணை செய்த வலைதளங்கள்:

1. <http://www.widbit.com>
2. <http://ascilite.org.au/conferences/sydney10/procs/Cclark-full.pdf>
3. <http://stevejobsschoolsneek.nl/home/>
4. <http://thinkingmachine.pbworks.com/w/page/22187720/Think-Social-Networking-for-Education>
5. http://www.maclife.com/article/news/dutch_steve_jobs_schools_use_ipads_replace_traditional_schools

தமிழ்மொழிக் கற்றல் கற்பித்தலில் இணையத்தின் பயன்பாடும் பங்களிப்பும்

சுப.நற்குணன், மலேசியா.

suba.nargunan@gmail.com

கட்டுரைச் சுருக்கம்:-

இந்தக் கட்டுரையானது, தமிழ்மொழிக் கற்றல் கற்பித்தலுக்கு இணையத்தை எவ்வாறு பயன்கொள்ள முடியும் என்பதையும்; இணையம் தொடர்பான ஏந்துகளைத் (வசதிகள்) தமிழ்மொழி கற்றல் கற்பித்தலுக்கு எவ்வாறெல்லாம் பயன்படுத்த முடியும் என்பதையும் ஆராய்கின்றது. இணையத்தின் வரவானது கற்றல் கற்பித்தலில் மிகப்பெரிய உருமாற்றங்களை ஏற்படுத்தியுள்ளது. மரபு வழிசார்ந்த கற்றல் கற்பித்தல் அணுகுமுறைகள் இருந்த வகுப்பறைகளில், தற்போது நவீனமான முறைமைகள் இடம்பிடித்துள்ளன. இணையம் வழியான கல்வி முறை ஆசிரியர்களின் கற்பித்தலிலும் மாணவர்களின் கற்றலிலும் புதிய வகைப் பரிணாமங்களை உண்டாக்கிக் கொடுத்திருக்கின்றன. எளிமை, விரைவு, விரிவு, விளைபயன், ஈர்ப்பு, மனமகிழ்வு முதலான தன்மைகளைக் கொண்டிருப்பதால், இணையம் வழியான கல்வி முறை இன்றையக் காலத்திற்கு மிகவும் ஏற்றதாகவும் தவிர்க்க இயலாத ஒன்றாகவும் ஆகிவிட்டதை இந்தக் கட்டுரை நிறுவுகின்றது. தவிர, கற்றல் கற்பித்தலுக்குப் பொருந்தும் வகையில் பல்வகையான இணைய ஏந்துகள் (வசதிகள்), செயலிகள், மென்பொருள்கள், வலைத்தளங்கள் முதலியவை நிறைய வந்துவிட்டதையும் இக்கட்டுரை எடுத்துக்காட்டுகின்றது. இவேந்துகளைப் பயன்படுத்தி தமிழ்மொழியைக் கற்கவும் கற்பிக்கவும் கூடிய வழிவகைகள் அறிந்து செயல்படுத்த வேண்டும். மொத்தத்தில், தமிழ்மொழிக் கற்றல் கற்பித்தலுக்கு இணையத்தையும் இணைய ஏந்துகளையும் முறைப்படுத்தித் திட்டமிட்டுச் சரியாகப் பயன்படுத்தினால் நன்மை விளையும் என்பதில் ஐயமில்லை.

1.0 முன்னுரை

தகவல் தொழில்நுட்பத்தின் மிக விரைவான வளர்ச்சி கல்வித் துறையிலும் மிகப் பாரிய தாக்கத்தை ஏற்படுத்தியுள்ளது. கணினியையும் இணையத்தையும் பயன்படுத்தாத கல்வி நிறுவனங்களே இல்லை என்கின்ற அளவுக்குத் தொடக்கப் பள்ளி தொடங்கி உயர்க்கல்விக் கழகங்கள் வரையில் எல்லா இடங்களிலும் கணினியும் இணையமும் ஆதிக்கம் செலுத்தி வருகின்றன. இத்தகையச் சூழலில், கற்றல் கற்பித்தலில் இணையமும், இணைய ஏந்துகளும் முக்கிய இடத்தைப் பிடித்துவிட்டன எனலாம். கோலின்சு (1996) என்பார் கூற்றின்படி, 1996 / 1997 ஆம் ஆண்டுகளிலேயே கற்றல் கற்பித்தலில் இணையம் ஊடுருவி வளர்ந்துவிட்டது என்கிறார்.

இதன் அடிப்படையில், இணையம் வழியான கற்றல் கற்பித்தல் முறைமை இன்று மிகவும் புகழ்பெற்றதாகவும் எல்லாராலும் விரும்பப்படுவதாகவும் இருக்கின்றது. ஆகவேதான், இணையத்தைப் பயன்படுத்தி கற்றல் கற்பித்தலை வழிநடத்துவதற்கு உலகம் முழுவதிலும் உள்ள எல்லாக் கல்வி நிறுவனங்களும், தனியார், தன்னார்வ நிறுவனங்களும் முண்டியடித்துக்கொண்டு முயற்சிகளை முன்னெடுக்கின்றன.

இதன் அடிப்படையில், இணையத்தின் வழியாக அல்லது செயலிகள், மென்பொருள்கள், வலைத்தளங்கள், மின்னிதழ்கள், ஒலி ஒளி மின்னாடகங்கள், மடற்குழுக்கள், சமூக வலைத்தளங்கள் முதலான இணைய ஏந்துகள் வழியாகத் தமிழ்மொழியைக் கற்கவும் கற்பிக்கவும் வேண்டிய தேவைகள் ஏற்பட்டுவிட்டன. ஆகவே, காலத்திற்கு ஏற்ற வகையில் புதிய அணுகுமுறையில் தமிழ்மொழிக் கற்றல் கற்பித்தலை மேம்படுத்த வேண்டும். அதற்குரிய விரிவான வாய்ப்புகள் இன்று உருவாகி இருப்பதை ஊக்கத்துடன் வரவேற்று விரிவாகப் பயன்கொள்ள வேண்டும்.

2.0 இணையம் வழிக் கற்றல்

இணையம் வழிக் கற்றல் (e-learning) இரண்டு வகையில் செயல்படுகின்றது. ஒன்று இணையம் வழி கற்பித்தல் (web based instruction, WBI). மற்றொன்று, கல்வி வலைத்தளங்கள் வழி கற்றல் (educational web sites, EWS). இவ்விரு வகை கற்பித்தலும் கற்றலும் ஆங்கிலத்தில் மட்டுமே செயற்பட முடியும் என்ற நிலைமை மாறி, இன்று தமிழிலும் மிக இலகுவாகச் செயல்படுத்த முடிகின்றது.

இணையம் வழி கற்பித்தலில் (WBI), இணையம் ஒரு பயிற்றுத்துணைப் பொருளாக அல்லது கற்றல் ஊடகமாகப் பயன்படுகின்றது. அதாவது, ஆசிரியர் மாணவருக்குக் கற்பிக்க இணையத்தை ஓர் ஊடகமாகப் பயன்படுத்திக்கொள்வதாகும். இதன்வழி ஆசிரியர் மாணவருக்கு அல்லது ஒரு மாணவர் இன்னொரு மாணவருக்கு இணையம் வழியாக ஒன்றைப் பயிற்றுவிக்கவும் தகவல்களைப் பரிமாறிக்கொள்ளவும் முடிகின்றது. இவ்வகை கற்பித்தலுக்கு மின்னஞ்சல் (email), இணைய உரையாடல் (chat), நிகழ்ப்பட கருத்தரங்கு (video conference), மடற்குழு (news group), கோப்பு பரிமாற்று நெறிமுறை (File Transfer Protocol, FTP), சமூக வலைத்தளம் (social network) முதலான இணைய ஏந்துகள் மிகவும் உறுதுணையாக உள்ளன. மேலும் ஆயிரக்கணக்கிலான தமிழ் வலைத்தளங்களும் வலைப்பதிவுகளும் கற்பித்தலுக்கு உதவக்கூடிய தரவுகள், செய்திகள், தகவல்கள், கட்டுரைகள், கதைகள், கவிதைகள் எனப் பல்வேறு வகையான கற்றல் கற்பித்தல் மூலங்களை (Teaching and Learning Source) வழங்குகின்றன.

அடுத்து, கல்வி வலைத்தளங்கள் வழி கற்றல் (EWS) அணுகுமுறையைப் பற்றி கண்ணோட்டமிடுவோம். கல்வி (education) தொடர்பான இணையத்தளங்களைப் பயன்படுத்தி கற்றுக்கொள்வதே இந்த அணுகுமுறையாகும். இவ்வகையிலான இணையத்தளங்கள் தற்போது தமிழில் அதிகமாகச் செயல்படுகின்றன. பல்கலைக்கழகங்கள், கல்வி நிறுவனங்கள், தன்னார்வ அமைப்புகள் முதலியவை பற்பல இணையத்தளங்களை நடத்திவருகின்றன. கல்வி இணையத்தளங்கள் தவிர மின்நூலகம், மின்நூல்கள், கலைக்களஞ்சியங்கள், மின்னகராதிகள் ஆகியனவும் இவ்வகையான கற்றலுக்கு மிகவும் உறுதுணையாக உள்ளன.

ஆகவே, ஆசிரியர்களும் மாணவர்களும் இவ்வகை நவீனக் கால கற்றல் கற்பித்தலை நடைமுறைபடுத்தும் வழிவகைகளை அறிந்து சிறந்த முறையில் பயன்பெற வேண்டும்.

3.0 தமிழ்மொழிக் கற்றல் கற்பித்தலில் இணைய ஏந்துகளின் பயன்பாடு

தமிழ்மொழிக் கற்றல் கற்பித்தலுக்கு உதவும் இணைய ஏந்துகள் தற்போது பல்கிப் பெருகி வருகின்றன. மழலையர் கல்வி தொடங்கி முனைவர் படிப்பு வரையில் தமிழில் படிப்பதற்குரிய வாய்ப்புகளை இன்றைய இணைய உலகம் ஏற்படுத்திக் கொடுத்திருக்கிறது. கேட்டல், பேச்சு, வாசிப்பு, எழுத்து முதலான அடிப்படைப் படிப்பு முதல் ஆராய்ச்சிப் படிப்பு வரையில் இணையத்தில் படிக்கக்கூடிய நிலைமை உருவாகிக்கொண்டு வருகின்றது. இலக்கணத்தையும் இலக்கணத்தையும் புதிய ஊடகத்தின் வழியாக புதிய அணுகுமுறைகளில் கற்பதற்கும் கற்பிப்பதற்கும் வாய்ப்புகள் பெருகி வருகின்றன.

3.1 தமிழ்க் கல்வி வலைத்தளங்கள்

தமிழ்மொழியைக் கற்கவும் கற்பிக்கவும் பல இணையத் தளங்கள் உள்ளன. தமிழை முதல்மொழியாகக் கற்பதற்கும், இரண்டாம் மொழியாகக் கற்பதற்கும் இணையத்தளங்கள் வந்துவிட்டன. அதுமட்டுமல்லாது, ஆங்கிலத்தின் வழியாகத் தமிழ் கற்றல், மலாய்மொழி வழியாகத் தமிழ் கற்றல் என உலகத்தின் பெருமொழிகள் பலவற்றின் வழியாகத் தமிழைப் படித்துக்கொள்வதற்குரிய சூழல்கள் உருவாகியுள்ளன. முறையாகச் செய்யப்பட்ட கலைத்திட்டங்கள், பாடங்கள் ஆகியவற்றை இணையத்திலேயே படிப்பதன் வழியாகத் தமிழ்மொழியைக் கற்றுக்கொள்ள முடியும்.

தமிழ் இணையப் பல்கலைக்கழகம் தமிழ்மொழிக் கற்றலுக்கு மிகச்சிறந்த முன்னோடியாகத் திகழ்வதை மறுக்க இயலாது. அதுபோல, உலகின் பல நாடுகளில் செயல்படும் தமிழ்க் கல்விக் கழகங்கள் அல்லது கல்விக் கழகத் தமிழ் இருக்கைகள் இணையம் மூலமாகத் தமிழ்க் கல்வியை முன்னெடுத்து வருகின்றன. அவ்வாறான வலைத்தளங்கள் சில பின்வருமாறு:-

- அ) www.tamil.net
- ஆ) www.tamil.org
- இ) www.southasia.upenn.edu/tamil
- ஈ) www.languageshome.com/English-Tamil.htm
- உ) www.tamildigest.com/
- எ) www.tamilvu.org/

3.2 கலைத்திட்டம் - பாடநூல்

தமிழ்நாடு, இலங்கை, மலேசியா, சிங்கப்பூர் முதலான நாடுகளில் தமிழ்மொழிக் கல்வி முறையாகப் பயிற்றுவிக்கப்படுகிறது. அதற்காக, தனியான கலைத்திட்டமும் பாடநூல்களும் பயன்படுத்தப்படுகின்றன. கனடா, அசுத்திரேலியா, அமெரிக்கா முதலிய இன்னும் பிற நாடுகளில் தமிழ்மொழிக் கல்வி தனியார் நிறுவனங்கள்

வழியாகவும் தன்னார்வ அடிப்படையிலும் கற்பிக்கப்படுகின்றது. மேற்சொன்ன அவ்வந்த நாடுகளில் பயன்படும் கலைத்திட்டமும் பாடநூல்களும் இணையத்தில் ஏற்றப்பட்டு பகிர்ந்துகொள்ளப்படுகின்றன. இந்தப் பகிர்ந்தலின் வழியாக, தமிழ்மொழிக் கற்றல் கற்பித்தலில் புதிய வளர்ச்சிகளை எட்டுவதற்குரிய வாய்ப்புகள் நிரம்ப உள்ளன.

அ) <http://www.moe.gov.my/bpk>

ஆ) <http://www.etatamilschool.org/shtml/curriculum.shtml>

இ) <http://www.textbooksonline.tn.nic.in/>

3.3 வலைத்தளங்கள் - வலைப்பதிவுகள்

தமிழ் வலைத்தளங்கள் இன்று ஆயிரக்கணக்கில் இருக்கின்றன. இவற்றுள் மிகப் பெரும்பான்மையானவை கேளிக்கை, பொழுதுபோக்கு, மனமகிழ்ச்சி ஆகியவற்றை நோக்கமாகக் கொண்டவை என்பதை மறுக்க முடியாது. ஆயினும், தமிழ்க் கல்வியை வளர்த்தெடுக்கும் நோக்கத்தோடு செயல்படும் வலைத்தளங்களும் வலைப்பதிவுகளும் ஏராளமாகவே இருக்கின்றன. இவை கற்றல் கற்பித்தலுக்கு உதவக்கூடிய கற்றல் மூலங்களை (learning source) வழங்குகின்றன. அதுமட்டுமல்லாது, கற்றல் கற்பித்தல் தொடர்பான வலைத்தளங்களையும் வலைப்பதிவுகளையும் உருவாக்கி நிருவகிக்கவும் இவை பெரும் வாய்ப்பினை வழங்கிக்கொண்டிருக்கின்றன.

அ) www.tamilheritage.org

ஆ) www.tamil.net/projectmadhurai

இ) http://hongkongtamil.blogspot.com/2010/03/blog-post_6167.html

ஈ) <http://spmilakiam.blogspot.com/>

உ) <http://thirutamil.blogspot.com>

3.4 கலைக்களஞ்சியம் - மின்னகராதி

ஆங்கிலத்தில் பெரும் வரவேற்பைப் பெற்றுள்ள கலைக்களஞ்சியம், மின்னகராதி போன்றவை தற்போது தமிழிலேயே இணையத்தில் கிடைக்கின்றன. இவற்றின் வழியாக, ஆசிரியர்களும் மாணவர்களும் தங்கள் கற்பித்தலையும் கற்றலையும் பொருளுள்ளதாகவும் பயனுள்ளதாகவும் ஆக்கிக்கொள்ள முடியும். அதிகம் விலைகொடுத்து பெரிய பெரிய கலைக்களஞ்சியம், அகரமுதலி ஆகியவற்றை வாங்கிச் சேர்த்துவைக்க வேண்டிய வேலையோ அல்லது எங்கு சென்றாலும் அவற்றைச் சுமந்துசெல்ல வேண்டிய அவசியமோ இல்லை. இருக்கும் இடத்திலேயே விரல்நுனியில் விவரங்களைப் பெற்றுக்கொள்ளக்கூடிய வாய்ப்பினை இணையம் ஏற்படுத்திக் கொடுத்திருக்கிறது.

அ) <http://ta.wikipedia.org>

ஆ) <http://www.tamilvu.org/library/kalaikalangiyam/lkk00/html/lkk00hom.htm>

இ) <http://ta.wiktionary.org>

ஈ) <http://www.thozhilnutpam.com>

உ) <http://www.tamildict.com>

3.5 மின்னூலகம்

தற்போது தமிழ் மின்னூலகங்கள் இணையத்தில் நிறைய வந்துவிட்டன. இவற்றின் வழியாக ஆயிரக்கணக்கிலான நூல்கள் காணக்கிடைக்கின்றன. பழைய புதிய நூல்களோடு கிடைத்தற்கரிய தமிழ் நூல்களை மின்வடிவில், மின்னூலாகப் படிப்பதற்குரிய வாய்ப்பினை இந்த மின்னூலகங்கள் ஏற்படுத்திக் கொடுத்துள்ளன. அதுமட்டுமல்லாது, பழங்கால ஓலைச்சுவடிகளையும் இணையத்தில் படிப்பதற்குரிய வாய்ப்புகள் இன்று சாத்தியமாகியுள்ளன.

அ) <http://tamilibrary.org/>

ஆ) <http://www.chennaiLibrary.com/index.html>

இ) <http://www.tamilvu.org/library/libcontnt.htm>

3.6 தமிழ் ஒருங்குறி

தற்போது நடப்பில் இருக்கும் தமிழ் ஒருங்குறி எழுத்துரு முறை தமிழ்மொழி வளர்ச்சிக்குப் பெரும் பங்காற்றுகின்றது என்றால் பொய்யில்லை. தேடு பொறிகளில் (search engine) தமிழிலேயே தட்டச்சி தேடுகின்ற நிலைமை இன்று

சாத்தியமாகியுள்ளது. இதனைத் தமிழ்மொழிக் கற்றல் கற்பித்தலிக்கு ஆக்கமான முறையில் பயன்படுத்த முடியும். குறிப்பிட்ட ஒன்றைப் பற்றிய விவரங்களைத் தமிழிலேயே தேடிப்பெற்று பயனடைய முடிகிறது.

அ) <http://www.google.com/webhp?hl=ta>

ஆ) http://ezilnila.com/nila/unicode_writer.htm

3.7 அச்ச வடிவத்தை ஒலி வடிவமாக மாற்றுதல்

இணையத்தில் தமிழ்மொழியில் நிகழ்ந்துள்ள புதிய வளர்ச்சிகளாக சிலவற்றைக் குறிப்பிடலாம். முதலாவது, எழுத்துருவை ஒலியாக்குதல். இரண்டாவது, பேச்சொலியை எழுத்துருவுக்கு மாற்றுதல். மூன்றாவது, பிறமொழிகளைத் தமிழுக்கும் தமிழைப் பிறமொழிகளுக்கும் மொழிபெயர்க்கும் தானியங்கி செயலி. இவ்வாறான, புதிய வளர்ச்சிகளைத் தமிழ்மொழிக் கற்றல் கற்பித்தலுக்குப் பயன்படுத்திக்கொள்ள வழிவகைகளைக் காண வேண்டும்.

அ) http://mile.ee.iisc.ernet.in:8080/tts_demo/

ஆ) <http://www.google.com/transliterate/Tamil>

3.8 சமூக வலைத்தளங்கள்

மின்னஞ்சல், மடற்குழு போன்றவை வழியாகவும் முகநூல், டுவீட்டர் போன்ற சமூக வலைத்தளங்கள் வழியாகவும் கற்றல் கற்பித்தல் தொடர்பான செய்திகள், பாடக்குறிப்புகள், பாடப்பொருள்கள், படங்கள், அசைவுப்படங்கள், நிகழ்ப்படங்கள் ஆகியவற்றை ஆசிரியர்களும் மாணவர்களும் பரிமாறிக்கொள்ள முடியும். இதற்கு அதிகப்படியான நேரமோ அல்லது பணமோ தேவையில்லை. குறைவான நேரத்தில் மிகக் குறைந்த செலவில் மிகப்பெரிய பலன்களை எல்லாரும் அடைய முடியும். அதுமட்டுமல்லாது, உலகத்தின் எந்த மூலையில் இருந்தாலும் எந்த நேரமாக இருந்தாலும் இணையத்தின் வாயிலாகக் கற்றல் கற்பித்தலை மேற்கொள்ளும் வாய்ப்பும் இருக்கின்றது.

அ) <http://en-gb.facebook.com/people/Tamil-Aasiriyam/>

ஆ) <http://mail.google.com>

இ) <http://mail.yahoo.com>

3.9 இணையம் ஒரு பல்லாடகம்

இணையம் இன்று ஒரு பல்லாடகக் கருவியாக பரிணமித்து இருக்கின்றது. எழுத்துரு, படங்கள், அசைவுப்படங்கள், நிகழ்ப்படங்கள், ஒலி ஒளி சேர்க்கைகள், இணைய உரையாடல், நிகழ்ப்படக் கருத்தரங்கு, வானொலி, தொலைக்காட்சி எனப் பன்முகப் பயன்பாடுகளைக் கொண்டதாக விளங்குகின்றது. இவை அனைத்தும் தற்கால நவீன முறையிலான கற்றல் கற்பித்தலுக்கு மிகவும் ஏற்றவை மட்டுமல்ல, எளிமையாகப் பயன்படுத்தக்கூடியவை. இவ்வாறு பல்லாடகமாகச் செயல்படும் இணையத்தில் உள்ளவற்றை கணினியில் சேமித்து வைத்தோ அல்லது பதிவிறக்கம் செய்தோ தேவையான நேரத்தில் பயன்படுத்திக்கொள்ள வாய்ப்பாக இருக்கின்றது. அதேபோல், நம்மிடம் உள்ளவற்றை இணையத்தில் பதிவேற்றம் செய்து உலகம் முழுவதும் உள்ள கணினிப் பயனாளர்களோடு பகிர்ந்துகொள்ளவும் முடிகிறது. இப்படியான, அரிய வாய்ப்பினைத் தமிழ்மொழிக் கற்றல் கற்பித்தலுக்கும் பயன்படுத்திக்கொள்ள வேண்டியது நம்முடைய ஆற்றலும் திறமும் ஆகும்.

அ) www.youtube.com

ஆ) www.tamilradios.com

இ) <http://www.4shared.com/>

ஈ) <http://www.tubetamil.com/>

உ) <http://www.formatoz.com/>

4.0 இணைய வழிக் கற்றல் கற்பித்தலின் மேன்மைகள்

இணையம் வழியாகத் தமிழ்மொழிக் கற்றல் கற்பித்தல் நடைபெறுவதால் குறிப்பிடத்தக்க நன்மைகள் விளைகின்றன. அவற்றைப் பின்வருமாறு பட்டியலிடலாம்:-

1) மாணவர்கள் பாடங்களை எளிதாகப் புரிந்துகொள்ள முடிகிறது.

- 2) ஒன்றைப் பற்றிய மேலதிக தகவலை அறிந்துகொள்ளும் வகையில் தொடுப்புகளைக் (hypelinks) கொண்டிருக்கிறது.
- 3) வெவ்வேறு ஆற்றலும் விருப்பமும் கொண்ட மாணவர்கள் தாங்கள் விரும்பும் வகையில் கற்பதற்கு வாய்ப்பினை வழங்கிகின்றது. எழுத்து (text), ஒலி (sound), காட்சி (visual), அசைவுப்படம் (graphics), நிகழ்ப்படம் (video), உடலியக்கம் (psychomotor), இருவழித் தொடர்பு (interactive) எனப் பலதரப்பட்ட வகையில் கற்பதற்குரிய சூழல் இருக்கின்றது.
- 4) மாணவர்களின் கற்றல் திறனையும் தன்னம்பிக்கையையும் வளர்க்கிறது.
- 5) தனியாகக் கற்பதற்குரிய (individualise learning) வாய்ப்பு கிடைக்கிறது.
- 6) மாணவரை இலக்காகக் கொண்ட கற்றல் கற்பித்தல் சிறப்பாக நடைபெறுகிறது.
- 7) மனமகிழ்ச்சியுடன் கற்பதோடு வெல்விளி (சவால்) நிறைந்த கற்றல் சூழலை உருவாக்கிக் கொடுக்கிறது.
- 8) மாணவர்களின் ஆக்கச் சிந்தனையையும் ஆய்வுச் சிந்தனையையும் வளர்த்துக்கொள்ள உதவுகிறது.
- 9) குறிப்பிட்ட காலம், இடம், சூழல் என எந்தவித கட்டுப்பாடும் இல்லாமல் கற்பதற்குரிய வாய்ப்பு கிடைக்கிறது.
- 10) கிடைப்பதற்கு அரிய தகவல்களை விரைவாகவும் விரிவாகவும் பெற முடிகின்றது.
- 11) குறைந்த செலவில் விலைமதிப்பில்லாத் தகவல்களையும் தரவுகளையும் நொடிப்பொழுதில் மிக எளிதாகப் பெற முடிகிறது.
- 12) கற்றலில் ஏற்படும் சிக்கல்களுக்குரிய தீர்வுகளைப் பல முனைகளிலிருந்தும் மூலங்களிலிருந்தும் உடனடியாகப் பெற முடிகிறது.
- 13) உலகத்தின் எந்தப் பகுதியில் இருந்தாலும் ஒருவரோடு ஒருவர் தொடர்பு கொள்ளவும், இணைய உரையாடல் கருத்துப் பரிமாற்றம் செய்யவும், நிகழ்ப்பட கலந்துரையாடல் நடத்தவும், மின்னஞ்சல் வழி தகவல்களைப் பரிமாறவும் வாய்ப்புகள் திறந்து கிடக்கின்றன.
- 14) தகவல்களைத் திரட்டவும், சேமிக்கவும், புதிய தகவல்களை இற்றைப்படுத்தவும் (update), தேவையற்ற விவரங்களை நீக்கவும், விரும்பியபடி மாற்றங்களைச் செய்துகொள்ளவும் முடிகிறது.

5.0 முடிவுரை

மொழிக் கற்றல் கற்பித்தல் எளிமையான ஒன்றல்ல. அதுவும் இணையத்தில் தமிழ்மொழியைக் கற்பதும் கற்பிப்பதும் மிகக் கடுமையான ஒன்றாகும். இருந்தபோதிலும், ஒலியியல், எழுத்தியல், வரிவடிவம், இலக்கணம், இலக்கியம் என விரிந்து கிடக்கும் தமிழ்மொழியை இணையத்தின் துணைகொண்டு கற்கவும் கற்பிக்கவும் கூடிய வாய்ப்புகள் இன்று நடைமுறைக்கு வந்திருக்கின்றன. தற்போதைய சூழலில் இருக்கின்ற வாய்ப்புகளையும் ஏந்துகளையும் முழுமையாகப் பயன்படுத்தி தமிழ்மொழிக் கற்றல் கற்பித்தலை முன்னெடுக்கும் முயற்சிகள் நடைபெற வேண்டும். மரபு வழியான கற்றல் கற்பித்தலுக்கு இடையில் இணையம் சார்ந்த நவீன முறையிலான கற்றல் கற்பித்தலுக்கும் இடங்கொடுக்க வேண்டும். இணையம் வழி கற்பித்தலானாலும் சரி அல்லது இணையத்தளங்களைப் பயன்படுத்தி கற்பதானாலும் சரி, இவ்விரண்டினையும் விரிவுபடுத்தி வளர்த்தெடுக்கும் வழிகளை ஆராய வேண்டும்.

தமிழ்மொழிக் கற்றல் கற்பித்தலைப் புதிய இலக்கு நோக்கி நகர்த்த வேண்டிய பொறுப்பும் கடமையும் கல்வியாளர்களுக்கும் கணிஞர்களுக்கும் இருக்கின்றது. ஆகவே, உலக உருண்டையில் தமிழும் தமிழ்க் கல்வியும் நிலைபெற வேண்டுமானால், இணையம் வழியாகவும் இணைய ஏந்துகளின் வழியாகவும் தமிழைக் கற்கவும் கற்பிக்கவும் வேண்டியது காலத்தின் கட்டாயமாகும். இதனை உணர்ந்து, தமிழ்மொழியைக் கற்பிக்கும் ஆசிரியர்களும் தமிழைக் கற்கும் அல்லது கற்க விரும்பும் மாணவர்களும் இணையத்தையும் இணைய ஏந்துகளையும் பயன்படுத்த முன்வர வேண்டும்.

இனிவரும் காலம் இணையத்தோடு இரண்டறக் கலந்துவிடப்போகின்றது. அதற்கேற்றால்போல, தமிழ்மொழிக் கற்றல் கற்பித்தலும் இணையத்தை உள்வாங்கிக்கொள்ள வேண்டும். அப்பொழுதான், தமிழும் தமிழ்க்கல்வியும் நீடுநிலவ முடியும்; தமிழ்மொழி தனது தொன்மையின் தொடர்ச்சியைத் தொலைத்துவிடாமல் நிலைத்து வாழும்.

மேற்கோள்கள்

குழந்தைவேல் பன்னீர்செல்வம்.சு. (2009). இணையம்வழி மொழிக் கற்றல் கற்பித்தலில் புதிய அணுகுமுறைகள், தமிழ் இணைய மாநட்டு மலர், செருமானியம்.

ரபி சிங். எம்.ஜே. (2010). மின்னணு வழியில் தமிழ்மொழி கற்பித்தல் மற்றும் கற்றல், தமிழ் இணைய மாநாட்டு மலர், கோவை

நக்கீரன்.பி.ஆர். (2010). தமிழ் இணையப் பல்கலைக்கழக மென்பொருள்கள் – ஒரு கண்ணோட்டம். தமிழ் இணைய மாநாட்டு மலர், கோவை

இளஞ்செழியன்.வே & இளந்தமிழ்.சி.ம. (2011). தமிழில் தகவல் தொழில்நுட்பத்தைக் கற்பித்தல்: வாய்ப்புகளும் சிக்கல்களும். தமிழ் இணைய மாநாட்டு மலர், பென்சில்வேனியா.

பெரியண்ணன்.கோ. (2011). இணையம் மற்றும் கணினி வழி தமிழ் கற்றல் கற்பித்தல். தமிழ் இணைய மாநாட்டு மலர், பென்சில்வேனியா.

Dr. Seetha Lakshmi. (2011). Facebook and Tamil Language in Singapore's Teacher Education . தமிழ் இணைய மாநாட்டு மலர், பென்சில்வேனியா.

Dr.Sajap Maswan. (1998). Kelebihan Penggunaan Internet dan Laman Web dalam Pengajaran dan Pembelajaran. Institut Perguruan Tuanku Bainun. Pulau Pinang, Malaysia.

Collins, B. (1996). The Internet As An Educational Innovation: Lesson From Experience With Computer Implementation. Educational Technology, 36 (6), pg. 21- 30

கணினி சார்ந்த கற்றல்கற்பித்தலில் தமிழ்மொழிப் பயன்பாடும் நன்மைகளும்

(மாணவர்களின் பார்வையில் -பரதன், தனலட்சுமி, அக்ஷயா, ரஷிதா)
எழுத்தாக்கம்: திருமதி.கல்பனா. மூத்ததமிழ் ஆசிரியை. சீலிங் உயர்நிலைப்பள்ளி. சிங்கப்பூர்.

முன்னுரை:

‘மாற்றம் என்பது மானிடத் தத்துவம்; மாறாதிருக்க யான் வன விலங்கு அல்லேன்’ என்று பாடிச்சென்றார் கண்ணதாசன். மனிதன் மட்டுமா காலத்துக்கேற்ப மாற்றத்தை விரும்பி ஏற்கிறான்? மனிதனால் உருவாக்கப்பட்ட கணினியும் நேற்றொரு தோற்றம் இன்னொரு மாற்றம் பார்த்தால் பார்வைக்கும் தெரிகிறது என்ற வேகத்தில் அல்லவா காலத்திற்கேற்ற மாற்றங்களைக் கண்டு வருகிறது. கற்றல் கற்பித்தலில் கணினியின் பயன்பாடும் உத்திகளும் காலத்திற்கேற்றவாறு பல்வேறு வகையான மாற்றங்களைக் கண்டு வருகின்றன என்பதும் மறுக்க முடியாத உண்மை. கணினி சார்ந்த கற்றல் கற்பித்தலில் மொழிப்பயன்பாடும் நன்மைகளும் என்ற இத்தலைப்பில் மாணவர்களாகிய எங்கள் அனுபவங்களைப் பகிர்ந்துகொள்ள வேண்டும் என்ற நோக்கத்தில் அரும்பியதே இந்த அனுபவக் கட்டுரை.

கணினி சார்ந்த கற்றல் கற்பித்தலின் படிநிலைகள்:

கணினி சார்ந்த கற்றல் கற்பித்தலில் அடிப்படைச் செயற்பாடுகள் (Basic Operations) இணையத் தேடுதலும் வழிக்கற்றல் (Learning with Searches), பனுவல் வழிக்கற்றல் (Learning with Text), பல்லாடகங்கள் வழி கற்றல் (Learning with Multimedia), கருத்துப்பரிமாற்றக் கருவிகள் வழிக்கற்றல் (Learning with Communication Tools), தகவல் திரட்டுகருவிகள்வழிக் கற்றல் (Learning with Data Collection Tools) போன்ற அடிப்படை அடைவுநிலைகளை உயர்நிலைப் பள்ளிக்கல்வியை முடிப்பதற்குள் நாங்கள் பெற்றிருக்க வேண்டும். மேலும் நிகழ்நிலையிலான இருவழிக் கருத்துப் பரிமாற்றத்திற்கான கருவிகளைப் பயன்படுத்தவும் தெரிந்திருக்க வேண்டும். இதற்கான பயிற்சிகளும் எங்களுக்குப் பள்ளிகளில் அளிக்கப்படுகின்றன

கணினி சார்ந்த கற்றல்கற்பித்தலில் மொழிப் பயன்பாடு:

‘ஒரு சமுதாயம் இன்றைய பணிகளை இன்றைய கருவியைக் கொண்டுதான் செய்ய வேண்டும். இன்றைய பணியை நேற்றைய கருவியைக் கொண்டு செய்யும் சமுதாயத்தின் நாணய வாழ்வு நலியும். இது தவிர்க்க முடியாதது’ என்கிறார் வா.செ. குழந்தைசாமி. அவரது கூற்றிற்கேற்ப எங்கள் நாட்டில், தொடக்கப்பள்ளிமுதல் பல்கலைக்கழகம்வரை கற்றல் கற்பித்தலில் தகவல்தொழில் நுட்பத்தின் பயன்பாடு பெரிதும் போற்றப்பட்டு வருவது கண்கூடு. மாணவர்களாகிய எங்களுடைய மொழித்திறன்களை வளர்க்கவும் தமிழ்மொழிப் பயன்பாட்டை அதிகரிக்கவும், வாழ்க்கைச் சவால்களைச் சமாளிக்கவும், அவற்றை எதிர்கொள்ளவும், புத்தம் புதிய மென்பொருட்களைப் பயன்படுத்திப் படைப்பாற்றலை வளர்க்கவும், மொழி கற்றலில் ஆர்வத்தைப் பெருக்கவும், பயன்முனைப்பு மிக்க முறையில் கற்றல் கற்பித்தல் நடவடிக்கைகளை மேற்கொள்ளவும் தகவல் தொடர்புத் தொழில்நுட்பம் ஒரு கருவியாக எங்கள் பாடத்திட்டத்தில் வலியுறுத்தப்பட்டுள்ளது.



தடைக்கற்களும் படிக்கற்களும்:

‘சென்றிடுவீர் எட்டுத்திக்கும் கலைச்செல்வங்கள் யாவும் கொணர்ந்திங்கு சேர்ப்பீர்’ என்று பாடினார் பாரதி. அவரது கூற்றிற்கேற்ப ஆங்கிலம், சீனம், மலாய் போன்ற மொழிகளில் பயன்படுத்தப்படும் புதிய வளங்களைத் தமிழிலும் பயன்படுத்தி வருகிறோம். மொழி வளர்ச்சிக்கு அடிப்படையே பேசுதல் நடவடிக்கையாகும். கணினி சார்ந்த நிலையில் அந்நடவடிக்கையைக் மேற்கொள்ளும்போது வகுப்பில் ஈடுபாடு மிக்க கற்றலை மேற்கொள்ள முடிகிறது. எடுத்துக்காட்டாக ஒகாரு.காம் (vocaroo.com)என்ற

இணையப்பக்கத்தைப் பயன்படுத்திப் பேசுதல், வாசித்தல் நடவடிக்கைகளை மேற்கொண்டு வருகிறோம். எங்களுடைய வாசித்தலையும் பேசுதலையும் பதிவு செய்த பின்னர், அதைப்பதிவேற்றம் செய்கிறோம். இதன்மூலம் ஒலிப் பனுவலையும், உரையாடல் பகுதியையும் மீண்டும் மீண்டும் கேட்டு எங்களை நாங்களே சுயமதிப்பீடு செய்துகொள்ள முடிகிறது. சக மாணவர்கள் ஒருவரையொருவர் மதிப்பீடு செய்யவும் முடிகிறது. ஆசிரியர்கள் மாணவர்களின் திறன்களை எளிதில் மதிப்பீடு செய்யவும் முடிகிறது.

யுனிகோட் எழுத்துருக்களை வேகமாகத் தட்டச்சு செய்யமுடியாத நிலையில் இருக்கும் மாணவர்களுக்குக் குகன்மொழிபெயர்ப்புக் கருவியைப் பயன்படுத்தி ஆங்கிலத்தில் தட்டச்சு செய்து அவற்றைத் தமிழ் வரிவடிவங்களாகப் படிக்கும் உத்திகளும் கற்பிக்கப்பட்டன. அவற்றைப் பயன்படுத்தும்போது மாணவர்களுக்கு மொழிப்பாடத்தில் சோர்வோ அலுப்போ ஏற்படுவதில்லை. மேலும் இருமொழிகளிலும் நிகழும் மொழிபெயர்ப்பு முறைகளைப் பற்றியும் தெரிந்து கொள்ள முடிந்தது. Adobe CS6 என்ற கட்டமைப்பிலிருக்கும் InDesign, Word for newsletter design போன்ற வளங்களைப் பயன்படுத்திப் புத்தாக்கத்துடன் எழுதுதல் நடவடிக்கையாக அறிவிப்புக் கடிதங்களைத் (Newsletter) தயாரிக்கவும் முடிந்தது. இத்தகு நடவடிக்கைகளின்மூலம் நாங்கள் வாசித்தல், பேசுதல், கேட்டல் சிந்தித்தல், எழுதுதல் போன்ற ஒருங்கிணைந்த திறன்களை வளர்த்துக் கொள்ளவும் முடிகிறது.

எதிர்காலப் படைப்பாளிகள்!

கணினிக் கூடத்தில் பயிற்சியைப் பெறப்போகிறோம் என்றாலே சர்க்கரைப் பந்தலில் தேன்மாரி பொழிவது போலத்தான் எங்களுக்கு இருக்கும். மாணவர்களின் மொழிப்பற்றையும், கற்றலில் ஈடுபாட்டையும் ஒன்றிணைக்கும் வகையில் இணையத் தளங்கள், வலைப்பூக்கள், யூ-டியூப்பில் இடம்பெறும் குறும்படங்கள், காட்சிகள் போன்றவற்றைப் பயன்படுத்தி ஆசிரியர்கள் கற்றல் கற்பித்தலை மேற்கொள்கிறார்கள். அதே நேரத்தில் செய்முறைப்பயிற்சியாகத் தமிழில் வலைப்பூக்களை உருவாக்குதல், மின் புத்தகங்களை உருவாக்குதல், குறும்படங்களைத் தயாரித்தல் போன்றவற்றிலும் நாங்கள் ஈடுபடுத்தப்பட்டோம்.



பள்ளிகளில் நிகழும் நிகழ்ச்சிகள், கற்றல் பயணங்கள் அவற்றில் இடம்பெறவேண்டிய செய்திகள், படங்கள் அவற்றைப் பற்றிய பின்னாட்டங்கள் போன்றவற்றை ஒருங்கிணைக்கும் வகையில் தமிழில் வலைப்பூக்களை உருவாக்கும் பயிற்சியும் குழுநிலையில் எங்களுக்கு அளிக்கப்பட்டது. அதன்பின்னர் அவரவர் திறமைகளை வெளிப்படுத்தும் வகையில் தனிநிலையில் வலைப்பூக்களை உருவாக்கினோம். அவற்றைப் பற்றிச் சகமாணவர்கள் ஆரோக்கியமான முறையில் மதிப்பீடு செய்யும்போது நாங்கள் பெற்ற இன்பமே அலாதி யானது.

அதற்கு அடுத்த நிலையில் நாங்கள் குறும்படங்களைத் தயாரித்தலுக்கான பட்டறையில் குழுவாகக் கலந்துகொண்டு பயிற்சி பெற்றோம். திரு.சி.குணசேகரன் அவர்கள் தமிழில் இப்பயிற்சிப் பட்டறையை நடத்தியதோடு செய்முறைப் பயிற்சிகளையும் அளித்தார். அதன் அடிப்படையில் கருப்பொருள், உரையாடல், நடப்பு, பின்னணி இசை, ஒளிப்பதிவு போன்றவற்றை ஒருங்கிணைத்துப் பல்வேறு வகைப்பட்ட மின்னிலக்க வளங்களைப் பயன்படுத்திக் குறும்படத்தைத் தயாரிக்கும் முயற்சியில் ஓரளவு திறன் பெற்றோம். இவ்வனுபவம் எங்களுக்கு ஊக்கத்தையும் உற்சாகத்தையும் அளித்ததால் கல்வி அமைச்சு நடத்தும் அனைத்துப் பள்ளிகளுக்கான குறும்படத் தயாரிப்புப் போட்டியில் கலந்துகொண்டோம். வெற்றி தோல்வியைப்பற்றிக் கவலைப்படாமல் போட்டிக்கு நாங்கள் தயாரித்த குறும்படத்தை அனுப்பியுள்ளோம். இதுபோன்ற படைப்புகளைத் தயாரிக்க ஆசிரியர் எங்களை ஊக்குப்படுத்துவதன்மூலம் எங்களுடைய சிந்தனைத் திறனும், படைப்பாக்கத்திறனும் வளர்க்கப்படுகின்றன. எங்களாலும் சுயமாகப் படைப்புகளை உருவாக்கமுடியும் என்ற தன்னம்பிக்கையையும், நாங்களும் படைப்பாளிகள் என்ற பெருமிதத்தையும் பெறமுடிகிறது.

மாணவ எழுத்தாளர்கள்:

தகவல் தொழில்நுட்பக் கல்வியின் மூலமாகவும் 21ஆம் நூற்றாண்டுத் திறன்களை வளர்ப்பதைக் கல்வியமைச்சு பெரிதும் வலியுறுத்தி வருகிறது. ஊடகப் பயன்பாட்டையும் தகவல் பரிமாற்றத் திறனையும் ஒருசேரப்பெறும் வகையில் மின் இதழ்களை உருவாக்கும் முயற்சியில் நாங்கள் ஈடுபட்டோம். அதற்கு ஆதார நடவடிக்கையாகக் கூடிக்கற்றல் முறையில் (Collaborative Learning) பள்ளியில் நடத்தப்பட்ட கணினி பட்டறையில் கலந்துகொண்டோம். எங்கள் பள்ளியில் அப்பட்டறையைத் திரு.சி.குணசேகரன் அவர்கள் நடத்தினார். மின் இதழ்களைத் தயாரிப்பதற்கான அடிப்படை





விளக்கங்களையும், உத்திகளையும் அவர் வழங்கினார். டேல்ஸ்ஸ்பிரிங் என்ற இணையவளத்தைப் பயன்படுத்தி மின் இதழ்களைத் தயாரிக்கும் பயிற்சிகளையும் அளித்தார். பின்னர் நாங்களாகவே மின்னிதழ்களை உருவாக்கும் பணியில் ஈடுபட்டோம்.

சந்தேகங்கள் எழுந்தபோது கணினி உரையாடல் மூலம் அவருடன் நாங்கள் தொடர்பு கொண்டு ஐயங்களைத் தீர்த்துக்கொண்டோம். அதன் பின்னர் ஆசிரியரின் வழிகாட்டுதலோடு (Self-directed Learning) சுயமாகவே படைப்புகளை உருவாக்கும் முயற்சிகளில் ஈடுபட்டோம். எங்களுடைய தனித்துவம் வெளிப்படும் வகையில் மிகச் சிறிய அளவிலான படைப்புகளை உருவாக்கினோம். எங்கள்

ஆசிரியரும் அவ்வப்போது ஆக்ககரமான பின்னூட்டத்தை வழங்கினார். நாங்கள் உருவாக்கிய இதழ்களைப் பற்றிய மதிப்பீட்டைச் சக மாணவர்களின் வாயிலாகவும் பெற முடிந்தது. நாங்கள் கணினிக்கூடத்தில் குழுவாகவும் தனிநிலையிலும் பயிற்சிகள் செய்யும்போது ஒருவருக்குத் தெரிந்த உத்திகளை மற்றவர்களுக்குக் கற்பிக்கும் வாய்ப்பும் கிடைத்தது. தன்னம்பிக்கையோடு சகமாணவர்களுக்கு வழிகாட்டவும், நாங்கள் கற்றுக்கொண்ட உத்திகளை நன்றாக நினைவில் வைத்துக் கொள்ளவும் முடிகிறது. ஆசிரியர்களும் ஊடகப் பயன்பாட்டை மேம்படுத்திக் கற்பிக்கவும், எங்கள் திறமைகளுக்கேற்றவாறு வேறுபடுத்திக் கற்பிக்கவும், கண்காணிக்கவும், மதிப்புரை வழங்கி ஊக்கப்படுத்தவும் கணினி சார்ந்த கற்றல் கற்பித்தல் துணை நிற்கிறது.

முடிவுரை: கணினி சார்ந்த கற்றல் கற்பித்தலால் புத்தம் புதிய துறைகளைப் பற்றியும், உலக நிகழ்வுகளைப் பற்றியும் எளிதில் அறிந்து கொள்ள முடிகிறது. தமிழ் ஆர்வலர்களின் மின்னிதழ்களைப் படித்துப் பாராட்ட முடிகிறது. மின்னிதழ்களில் வலியுறுத்தப் பெறும் பண்பாட்டு விழுமியங்களைக் கண்டறிந்து போற்றவும் முடிகிறது. எங்களுடைய படைப்பை மற்றவர் படைப்புகளோடு ஒப்பிட்டுப் பார்க்கவும், சுயமதிப்பை நிலைநாட்டும் வகையில் எங்கள் படைப்பை மேம்படுத்திக் கொள்ளவும் முடிகிறது. 'அங்கங்கே கலைகள் தேரும் அறிவன்' என்று திருவிளையாடற்புராணம் மாணவர்களைப் பற்றிக் கூறுகிறது. அதுபோன்று எங்களுக்குத் தேவைப்படும் இணையவளங்களை, புதிய மென்பொருட்களைத் தேடிக்கண்டறிந்து, மொழிப்பயன்பாட்டை அதிகரித்துக் கொள்ளும் வாய்ப்புகளை நாங்கள் பெற்றுள்ளோம். ஆகவே எங்களைப் போன்ற மாணவர்களின் சின்னஞ் சிறுபடைப்புகளை ஊக்கப்படுத்தித் தமிழ் வளர்ச்சிப் பணியில் எங்களையும் ஈடுபடுத்தினால் எங்கும் தமிழ் எதிலும் தமிழ் என்ற நிலை உருவாகும்.



துணை நின்ற இணைய வளங்கள்:

<http://www.adobe.com/sea/>

<http://www.apple.com/education/>

www.nearpod.com

www.google.com

<http://tamilmurasu.com.sg/>

www.ida.gov.sg

www.mda.gov.sg

<http://www.nac.gov.sg/>

<http://www.moe.edu.sg/>

www.silingsec.moe.edu.sg/

PDF வடிவத்தில் பல்லாடகப் பனுவல் தயாரித்தல், இதன் மூலம் கற்றல், கற்பித்தல், ஒரு பொருளை விளம்பரப்படுத்தி விளக்குதல்

காசி வீராசாமி - தமிழாசிரியர் (விக்டோரியா தொடக்கக்கல்லூரி- சிங்கப்பூர்)

Kasinathan_veerasamy@moe.edu.sg

சாத்தப்பன் மெய்யப்பன் - தமிழாசிரியர் (விட்லி உயர்நிலைப் பள்ளி-சிங்கப்பூர்)

sathappan_meyyappan@moe.edu.sg

நோக்கம்-

மாணவர்களுக்குத் தற்போது காகித வடிவிலும் கணினி மூலமாகவும் பாடங்களும் பயிற்சிகளும் கொடுக்கப்படுகின்றன. அதில் ஏற்படும் இடர்ப்பாடுகளைக் களைவதற்காக `PDF வடிவத்தில் பல்லாடகப் பாடம் கற்பித்தல்` என்ற எளிய முறையைப் பயன் படுத்தினோம். இதனைத் தயாரிக்கும் முறையும், அதனால் விளைந்த பயன்களையும் விளக்குவதே இக்கட்டுரையின் நோக்கமாகும்.

இந்த “PDF வடிவத்தில் உள்ள பல்லாடகப் படைப்பு”, தமிழில் உரையாடி, வியாபாரம் செய்யும் வணிகர்கள் தங்களது பொருளைத் தமிழ் பேசும் நுகர்வோரிடம் கொண்டு செல்லப் பெரிதும் உதவும்.

Ios, Android சாதனங்களில் PDF வடிவத்தில் பல்லாடகத்தமிழ்:

கணினி, இணையத்தளம் இவற்றிலுள்ள செய்திகளை மாணவர்கள் கீழறிக்கம் செய்வதற்கு இடப்பற்றாக்குறை, நேரப்பற்றாக்கறை, மற்றும் சிரமம் ஏற்படுவதை நீக்க அல்லது குறைக்க இம்முறை பயன்படுகிறது. இதன் மூலம் வாசித்தல், பாடல்கள் வழி இலக்கியச்சுவை உணர்தல், கருத்தறிதல், கட்டுரை முதலிய பாடங்களை மாணவர்களுக்குச் சுவைபடக் கற்பிக்கலாம். இது மாணவர்களின் வாசிப்புத் திறனை வளர்க்க உதவுகின்றது. கடினமான சில சொற்களை எவ்வாறு உச்சரிக்க வேண்டும் என்பதை மாணவர்கள் வீட்டில் இருந்த படியே, தானே, பலமுறை கேட்டு, பார்த்துத் தெரிந்து கொள்ள முடியும்.

வாசிப்புப்பனுவல், பாடல் வரிகள் முதலியவற்றை எழுத்து வடிவிலும் ஒலி, ஒளி வடிவிலும் (காணொளி) பார்க்கும் கேட்கும் மாணவர்கள் ஆர்வத்துடன் பாடங்களைக் கற்பர். இதனை ஒரு மாணவர் பலமுறை பயன்படுத்தலாம். எங்கும், எதிலும் இணைய இணைப்பில்லாத கணினி, நவீன கைத்தொலைபேசி போன்றவற்றில் திறந்து படிக்கலாம், பார்க்கலாம்.

வகுப்பறையில் கணினி மூலம் ஆசிரியர் இவற்றைக் கற்பிக்கும் போது மாணவர்கள் ஒரு முறை தான் கற்க முடியும், பார்க்க முடியும். ஆனால் PDF வடிவத்தில் ஒளி ஒலியுடன் மின்னஞ்சலில் அனுப்பும் போது அவர்கள் அதனைச்சேமித்து வைத்துக்கொண்டு பலமுறை அவற்றைப் பார்த்துப் படித்தறிய முடியும். வாகனங்களில் பயணம் செய்யும் போதும் கையடக்கத்தொலைபேசி மூலம் பார்த்து, படிக்க முடியும்.

வியாபாரிகளுக்குப் பெரிதும் உதவும்:

ஒரு வியாபாரி தான் தயாரித்த பொருளைப் பிறருக்கு அறிமுகப்படுத்த, அல்லது நுகர்வோருக்கு அறிமுகப்படுத்த இம்முறையைப் பயன்படுத்தலாம்.

உதாரணமாக ஒரு நூலாசிரியர் தனது படைப்பிலுள்ள ஒரு முக்கியப்பகுதியை PDF படிவமாக மாற்றி அப்பக்கத்தில் அவரது விளக்க உரையை அல்லது வேண்டுகோளை ஆடியோ வடிவத்தில் வைத்து அனுப்பலாம். மேலும் புத்தக வெளியீட்டு விழாவின் வீடியோவையோ, நூலைப்பாராட்டிப் பேசிய அறிஞரின் வீடியோ, ஆடியோவையோ அதில் வைத்து மின்னஞ்சலில் நண்பர்களுக்கு அனுப்பலாம். மேலும் நூலில் இடம்பெற்றுள்ள பனுவலின் பொருளை, அல்லது விளக்கத்தை அதில் எழுத்தாளர் மறைத்துவைக்கலாம். தேவையானவர்கள் மட்டும் அந்த இடத்தில் மவுசை வைத்து அழுத்தினால் விளக்கம் வெளியே வரும். மவுசை வைத்து அழுத்தா விட்டால் விளக்கம் வெளியே வராது. இதனைப் படிக்க, பார்க்க எந்தத் தமிழ் எழுத்துருவும் வேண்டாம்

இக்கோப்பை கணினியில் திறந்து பார்த்துப் படிக்கலாம், ஆடியோ, வீடியோவை கேட்கலாம், பார்க்கலாம். அதற்குச் சாதாரண ஆடோபி ரீடர் (Adobe Reader- latest version) போதுமானது. வேறு எந்த மென்பொருள் உதவியும் வேண்டாம்.

ios, Android சாதனங்களில் செம்மையாக இயங்கும்:

அனைத்து ios, Android கையடக்கத்தொலைபேசியிலும், கையடக்கக்கணினியிலும் (ipad & notepad) தெரியும், திறக்கும், செயல்படும்.

ஆனால் இந்த இரண்டு சாதனங்களிலும் **EZ-PDF Multimedia Reader** என்ற மென்பொருளைக் கீழறிக்கம் செய்து பொருத்திக் கொண்டால் போதும். இதன் விலை \$ 3 (மூன்று வெள்ளி) மட்டுமே. இம்மென்பொருள் இருந்தால் வீடியோ, ஆடியோ சிறப்பாக இயங்கும். இல்லை என்றால் பனுவலை மட்டும் படிக்கலாம்.

இதன் மூலம் மாணவர்களுக்கு நாங்கள் பாடங்கள் தயாரித்து அனுப்பி மாணவர்களை மயக்குகின்றோம். தமிழில் நிகழும் அற்புதங்களைக் கண்டு தமிழால் எல்லாம் நிகழும் என்பதை இளைய சமுதாயம் ஏற்று வரவேற்று மகிழ்கின்றது.

தமிழ் வணிகர்கள் தங்களது தொழில் தொடர்பான விளம்பரங்களைத் தமிழில் தயாரித்து, அதனைக் கையடக்கத்தொலைபேசி மூலம் பிறருக்கு அனுப்பி, தனது தொழிலைப் பெருக்கலாம். வியாபாரத்தைப்பெருக்கலாம்.

இப்பல்லுடகப் பாடத்தைத் தயாரிக்க சிறப்பான மென்பொருள் எதுவும் தேவையில்லை. மாணவர்கள் பார்க்கவும், படிக்கவும் வேறு எந்த மென்பொருளும் தேவையில்லை. **ADOBE PROFESSTIONAL WRITER** மட்டும் இந்தப் பல்லுடப்பகுதியைத் தயாரிப்பவருக்கு இருக்க வேண்டும். பனுவல் மட்டுமே வேண்டும். அப்பனுவல் எந்த வடிவத்தில் இருந்தாலும் அதனை அப்படியே **PDF படிவமாக** மாற்றி அதில் பல்லுடகங்களை வைத்து (இணைத்து), அளவையும் மின்னஞ்சலில் செல்லும் வகையில் குறைத்து, மின்னஞ்சலில் அனுப்பி விடலாம். இப்பல்லுடகப் பாடத்தைப் பயன்படுத்தி மாணவர்கள் பயன் அடைந்து உள்ளனர்.

பயன்பாடும் செயல்முறையும்

PDF வடிவத்தில் பல்லுடகப்பாடம் கற்பித்தல் முறை கற்றல் கற்பித்தலில் பலவழிகளில் பயன்படுகின்றது. (1) இணையம்வழிக்கற்றல் (E-learning) (2) மொழிக்கூறுகளை எளிமையாகப் புரிந்துகொள்ளுதல், (3) வாய்மொழிப் பயிற்சிக்கு உதவுதல், (4) விளம்பரத்துறைக்குப் பயன்படுத்துதல்.

இணையம்வழிக்கற்றல் (E-learning)

தற்போதைய சூழ்நிலையில் பல பள்ளிகளில் வீட்டுப்பாடங்களை இணையம் வழிசெய்யுமாறே பெரிதும் வலியுறுத்துகின்றனர். அதற்கு PDF வடிவத்தில் பல்லுடகப்பாடம் கற்பித்தல் முறை மிகச் சிறந்த வாய்ப்பாக மாணவர்களுக்கு அமைகிறது. இம்முறையில் மாணவர்கள், ஆசிரியர்கள் தயாரித்து PDF வடிவத்தில் அனுப்பும் பாடங்களை எவ்வித இடர்ப்பாடும் இல்லாமல் வீட்டில் பார்க்க முடியும். இம்முறையில் பனுவல்கள் வரிவடிவிலும் ஒலி, ஒளி வடிவிலும் காணொளியாக இருப்பதால் மாணவர்கள் மிகுந்த ஆர்வத்துடன் எளிமையாகப் பகுதிகளைப் புரிந்து கொண்டு வினாக்களுக்க விடை தருவர்.

மொழிக்கூறுகளை அறிமுகப்படுத்துதல்-

சிங்கப்பூர், மலேசியா போன்ற நாடுகளில் தமிழ்மொழி இரண்டாம் மொழியாகக் கற்பிக்கப்படுகிறது. மேலும் சுமார் இரண்டுகோடி தமிழ் மக்கள் உலகம்முழுவதும் புலம்பெயர்ந்து வாழ்கின்றனர். அவர்களுக்குப் பள்ளிகளில் தனிப்பாடமாகத் தமிழ்மொழி பயிற்றுவிப்பதில்லை. பெற்றோர்கள் ஊடகங்களின் மூலமாகத்தான் தமிழைத் தங்கள் பிள்ளைகளுக்குக் கற்றுக்கொடுக்கிறார்கள் என்பதை விட அறிமுகப்படுத்துகிறார்கள் என்று கூறலாம். அத்தகைய பெற்றோர்களுக்கு இம்முறை ஒரு வரப்பிரசாதமாகும். இன்றைய இளையர்கள் கையடக்கத் தொலைபேசி இல்லாமல் வெளியில் உலா வருவேத இல்லை. அவர்களது கைகளில் கையடக்கத் தொலைபேசி பல வழிகளில் பல நிலைகளில் பயன்பாட்டில் உள்ளது. அதன் மூலமாகத் தமிழை உலாவர வைத்தால் எளிதில் இளையோரைச் சென்றடையும். இரண்டாம் மொழியாகத் தமிழ் கற்கும் மாணவர்கள் மொழிக் கூறுகளை, இலக்கணப் பகுதிகளைப் புரிந்து கொள்வதில் பல்வேறு சிரமங்களை எதிர்நோக்குகின்றனர். கற்பிக்கும் ஆசிரியர்களும் இலக்கணப் பகுதிகளைச் சுவைபடக் கற்பிப்பதில் இடர்ப்படுகின்றனர். இந்தச் சிக்கலை PDF வடிவத்தில் பல்லுடகப்பாடம் கற்பிக்கும் முறை வெகுவாகக் குறைக்கின்றன.

இன்றையஇளையர்களில் இசைக்கு மயங்காதவர்களே இல்லை எனலாம். அவர்களை மயக்கும் இசையோடு மொழியையும் சேர்த்துக் கொடுத்தால், `காணொளி` வடிவில் கையடக்கத் தொலைபேசியில் வரி வடிவத்தில் கண்ணில் ஊடுருவி, மனதில் தைத்து, கருத்தில்பதியும். கண்ணில் ஊடுருவும் என்பதைவிட, ஊடுருவியுள்ளது என்றே ஆணித்தரமாகக் கூறலாம்.

இதில் இலக்கியநயமிக்க சில பாடல்வரிகளும் அதன் ஒலி, ஒளி வடிவங்களும் இணைக்கப்பட்டிருக்கும். அதனைப் படிக்கும், கேட்கும், பார்க்கும் மாணவர்கள் பாடல் வரிகளில் புதைந்துள்ள இலக்கணக் கூறுகளையும் புதிய சொற்களையும் புரிந்துகொள்வர். மாணவர்களுக்குத் தெரியாத, புரியாத பகுதிகளுக்கு ஆசிரியர் விளக்கம் எழுதி அப்பகுதிகளுக்கு அருகில் சேமித்துவைத்தால், மாணவர்கள் கையடக்கத் தொலைபேசியைப் பயன்படுத்தும் போது தங்கள் கைவிரல்களையும் கணினியைப் பயன்படுத்தும் போது மவுசையும் (சுட்டியையும்) அருகில் வைத்து அருஞ்சொற்களுக்கு உரிய சரியான பொருளைத் தாமே அறிந்துகொள்வர். இம்முறையால் மாணவர்கள் ஆசிரியர் உதவியின்றி எவ்விடத்திலும் எந்நேரத்திலும் பாடல்களில் இடம்பெற்றுள்ள அணிகள், அடுக்குத்தொடர்கள், இரட்டைக்கிளவிகள், மரபுத்தொடர் இணைமொழிகள், சொல்வளம், பொருள்வளம் முதலிய மொழிக்கூறுகளை எளிதில் அறிந்து கற்றுக்கொள்வர்.

3. வாய்மொழிப் பயிற்சி

இன்றைய உலகம் மாணவர்களின் வாய்மொழித் திறனுக்கும் பேச்சுத் தமிழின் பயன்பாட்டிற்கும் அதிக முக்கியத்துவம் கொடுத்துவருகிறது. அவ்வகையில் வாய்மொழிப் பயிற்சி கொடுப்பதற்கும் இந்த புதிய முறை மிகவும் பயனுள்ளதாக அமைகிறது.

ஆசிரியர் வாய்மொழிக்குரிய படங்களையும் சிறுகுறிப்புகளையும் தயாரித்து PDF வடிவத்தில் மாணவர்களுக்கு மின்னஞ்சல் மூலம் அனுப்பலாம். அப்பகுதியை மாணவர்கள் தங்கள் வீட்டுக் கணினியிலோ, பள்ளியிலுள்ள கணினிக்கூடக் கணினியிலோ திறந்து பார்த்து தேவையான கதையை உருவாக்கி அவற்றைத் தங்கள் குரலிலேயே பேசிப் பதிவு செய்து திரும்ப ஆசிரியருக்கு அனுப்பினால் அவர் அதனைக்கேட்டு நிறைகுறைகளை மாணவர்களுக்கு மின்னஞ்சல் மூலம் தெரியப்படுத்துவார். இவ்வகையில் இது ஒரு இருவழிக் கருத்துப் பரிமாற்றக் கல்விமுறையாகவும் இணையம் வழிக்கற்றலாகவும் (elearning)அமைகிறது.

4. தொழில்முறைப் பயன்பாடு

PDF வடிவத்தில் பல்லாடகப் பனுவல் தயாரித்தல் முறையைக் கற்றல் கற்பித்தலுக்கு மட்டுமன்றி, விளம்பரத்துறைக்கும் பயன்படுத்தலாம். இன்றைய நவீன தொழில்நுட்ப உலகில் தாளின் பயன்பாடு மிகவும் குறைந்து வருகிறது. தாளின் பயன்பாட்டிற்காக மரங்கள் அதிகம் அழிக்கப்படுகின்றன. இதனால் சுற்றுச்சூழல் பாதிப்பு ஏற்படுகிறது. இதனாலேயே இன்று பெரும்பாலான வணிக நிறுவனங்கள் டிஜிட்டல் முறையையே தங்கள் விளம்பரத்திற்குப் பயன்படுத்துகின்றன. தொலைக்காட்சி, இணையம், திரைப்படங்கள் போன்ற தகவல் சாதனங்கள் வழி இவற்றை விளம்பரம் செய்ய அதிகம் செலவாகும். மேலும் பழைய முறைப்படி விளம்பரங்களை அச்சிட்டு விநியோகம் செய்வதிலும் சில சிக்கல்கள் உள்ளன. அவற்றைப் பயனீட்டாளர்களிடம் கொண்டு போய்ச் சேர்ப்பதற்குத் தேவையான ஆட்கள் கிடைப்பதில்லை. அவ்வாறு சேர்த்தாலும் மக்கள் அவற்றை ஆர்வத்துடன் படிக்காமல் குப்பைத் தொட்டியில் எறிந்துவிடுவதைக் காணலாம். பயனீட்டாளர்களின் இத்தகைய மனப்போக்கால் வணிகநிறுவனங்களுக்குச் செலவுதான் ஏற்படுகிறதே ஒழிய எதிர்பார்த்த பயன் கிடைப்பதில்லை.

ஆனால், IOS, அண்ட்ராய்டு தளங்கள் உள்ள iPhone, Ipad, Samsung Galaxy போன்ற சாதனங்களில் தமிழில் விளம்பரங்கள் தயாரித்து அனுப்பும் போது பல்லாடகங்கள் முறையில் விளம்பரங்கள் அமைவதால் பயனீட்டாளர்கள் அதிக ஆர்வத்துடன் விளம்பரங்களைப் பார்க்கும், கேட்கும், படிக்கும்வாய்ப்பு ஏற்படுகிறது. MMS, E-mail முறைகளில் விளம்பரங்கள் அனுப்பப்படுவதால் பயனீட்டாளர்கள் குறைந்தது ஒரு முறையாவது அவற்றைப்பார்க்கும் வாய்ப்பு ஏற்படுகிறது. இதனால் மிகக்குறைந்த செலவில் விரைவாக விளம்பரப்படுத்த முடிகிறது. இவ்விளம்பரங்களைச் சேமித்து வைத்துக்கொண்டு எத்தனை முறை வேண்டுமானாலும் எங்கு வேண்டுமானாலும் பயன்படுத்தலாம். ஒரு முறை கீழிறக்கம் (download) செய்து சேமித்துவிட்டால் எப்போதும் பயன்படுத்தலாம். வேண்டாத போது அழித்துவிடலாம். நம் நண்பர்களுக்கு அதனை அனுப்பி (Forward) வைத்துக் கருத்துகளைப் பகிர்ந்துகொள்ள முடியும்.

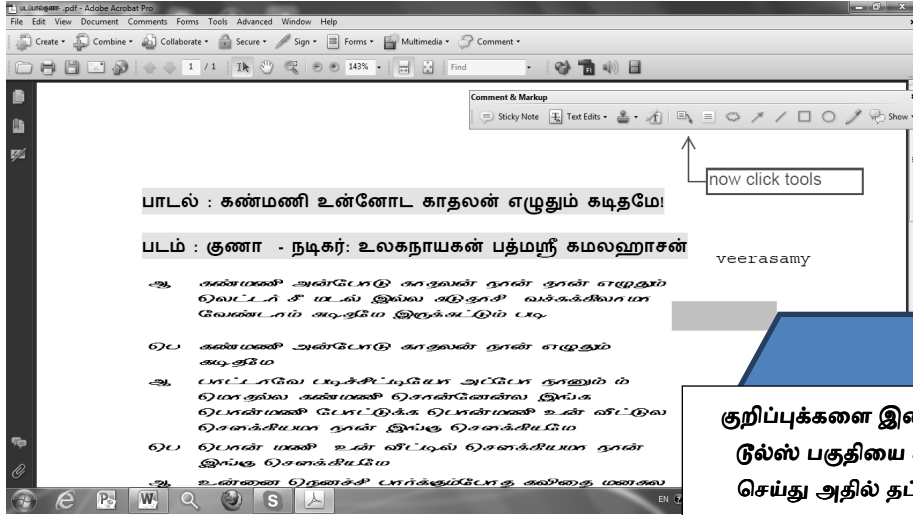
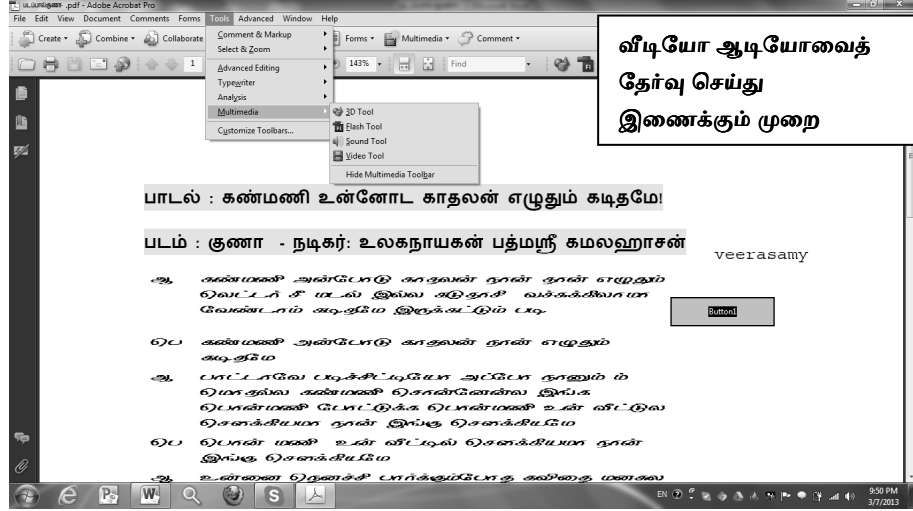
இதனால் விளம்பரங்கள் எளிதில் பலரைச் சென்றடையும். அவர்கள் அதனைக் கணினியில் சேமித்து வைத்துப் பார்க்க, படிக்க, அனுபவிக்க எந்த விதச் சிறப்பு மென்பொருளும் தேவையில்லை. `Adobe Acrobat Reader 9` என்ற மென்பொருள் மட்டும் இருந்தால் போதுமானது. இது இலவசமாகவே கிடைப்பதால் அனைவரும் கீழிறக்கம் செய்துவைத்துக் கொள்ளலாம்.

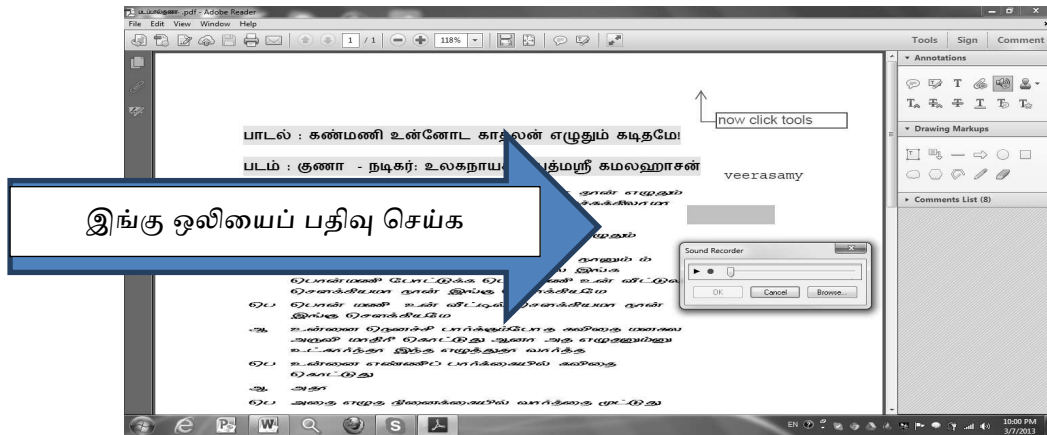
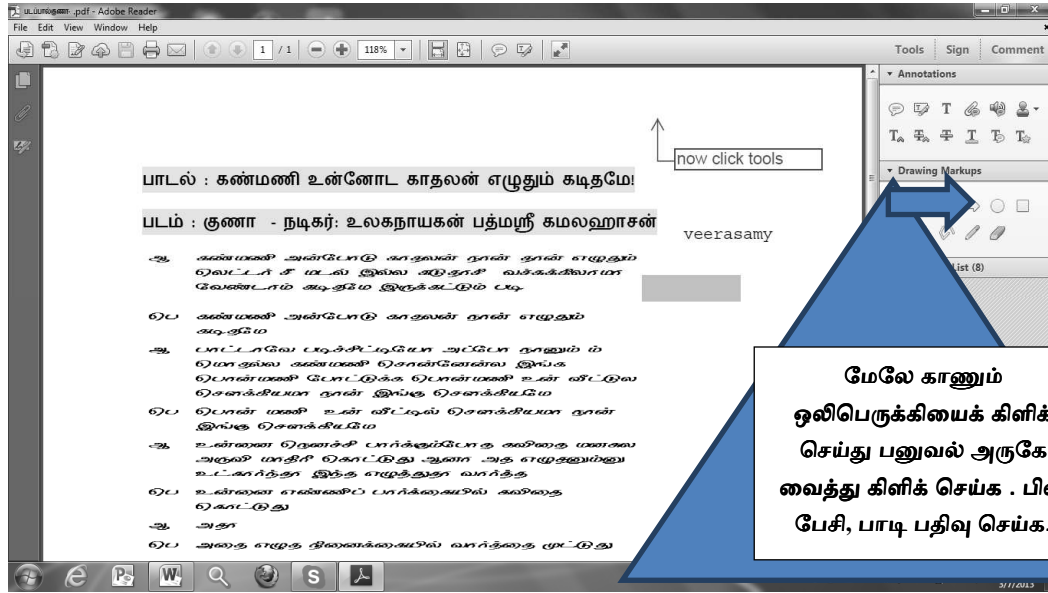
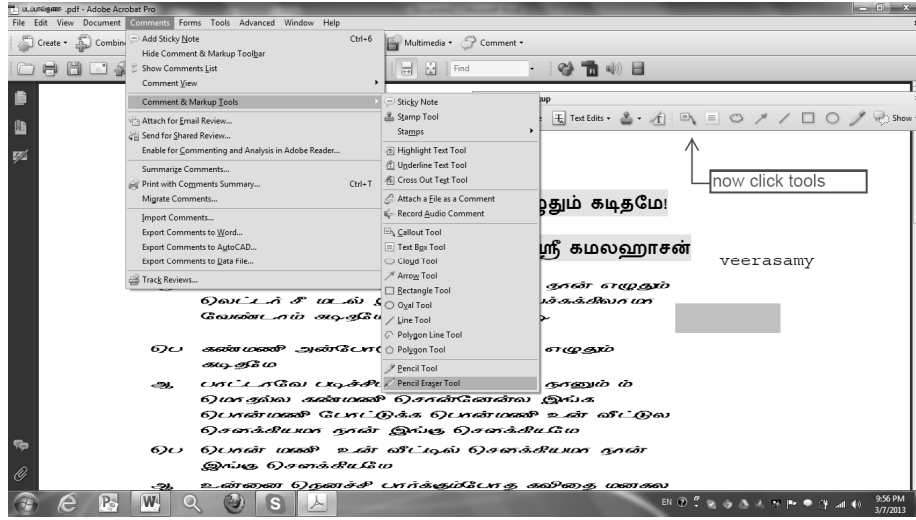
செய்முறை

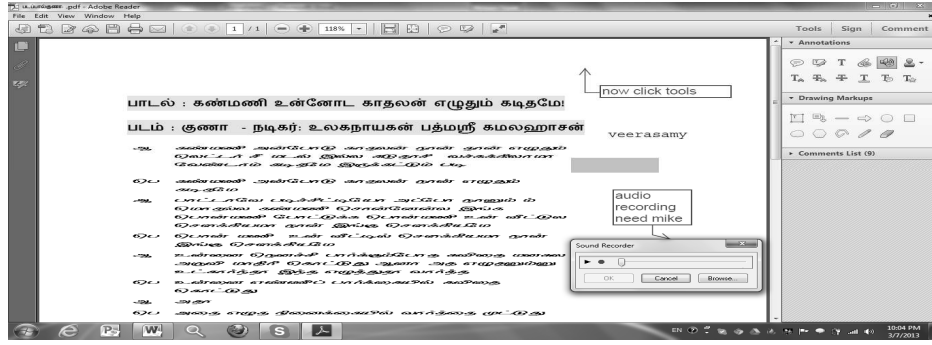
நாம் முதலில் தயாரிக்க வேண்டிய ஒன்றை முறையாக word, power point, publisher, போன்ற மென்பொருள்களைப் பயன்படுத்தி பனுவலையும். படம் போன்றவைகளையும் வைத்துத் தயாரித்து, பிறகு அதனை PDF Format இல் மாற்றிச் சேமிக்க வேண்டும். அதற்குப் பிறகு தான் பல்லாடகங்களை இணைக்க வேண்டும்.

வீடியோ அனைத்தும் mp4 format இல் தயாரித்து வைத்துக்கொண்டு, கீழ்க்கண்ட முறையில் சேர்க்க வேண்டும். பாடல் mp3 format இல், wav format- இல் இருக்க வேண்டும்.

இந்தப் பல்லாடகங்களைக் கீழ்க்கண்ட முறைகளைக் கையாண்டு தான் இணைக்க வேண்டும். ADOBE READER TOOLSஐ- திறந்து பார்த்தால் அனைத்து வழிகளும் தெரியும்.







குறிப்பு : ஒவ்வொரு மாற்றத்திற்குப் பிறகும், இணைப்புக்குப்பிறகும், கோப்பைச் சேமிக்க வேண்டும்.இது ஒரு மின்னிலக்க விளம்பரம்.-

Digital Advertisment - (INTERACTIVE MOBILE EXPERIENCE)



முடிவுரை.

இம்முறையை உலகத்தமிழ் இணைய வல்லுநர்கள் அறிந்தால் தான் தமிழுலகம் அறியும், இதனைப் பின்பற்றும். இதனைத் தமிழாசிரியர்களும் பின்பற்றி, பயன்படுத்தி தமிழ்க் கற்பித்தலைத் தொழில் நுட்பத்தோடு இணைத்து, இனிமையாக்கலாம். வணிகர்கள் அறிந்தால் வியாபாரம் அதிகரிக்கும்.

நம் தமிழ் இனிமையானது ; அதனை இனிமையான முறையில் போற்றுவோம்.

பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பப் பயன்பாட்டின் வாயிலாக மூன்றாம் ஆண்டு மாணவர்களிடையே திருக்குறள் ப யன்பாட்டினை மேம்படுத்துதல் - ஓர் ஆய்வு

சிவகுமாரி த/பெ சாமியப்பன்

விரிவுரையாளர், ஆசிரியர் கல்விக் கழகம், ராஜா மெலேவார் வளாகம்
சிரம்பான், நெகிரி செம்பிலான்

முன்னுரை

கணினி மற்றும் தகவல் தொடர்பு தொழில்நுட்பத் துறை சார்ந்த பல்லாடகத் துணைக்கருவிகளின் வருகைகல்வியுலகில் விவேகமான கற்றல் கற்பித்தல் சூழலுக்கு இட்டுச்சென்றுள்ளது. தொழில்நுட்பம் நாளுக்கு நாள்புத்தாக்கச் சிந்தனையின் அடிப்படையில் வளர்ச்சியடைந்து வருகிறது. ஒரு காலத்தில் போதிய தகவல் இல்லை என்ற நிலை; இன்றோ அளவிற்கு அதிகமான தகவல்கள் பெருகிவிட்டன. இலத்திரனியல் தொழில்நுட்பம், கணினிக்களுக்கிடையிலான வலைப்பின்னல் மற்றும் 'வயர்லெஸ்' தொழில்நுட்பம் போன்றவை கல்வித்துறையில் புதிய இலக்கினை எட்டியுள்ளன.

பல்லாடகம் எனப்படுவது கணினியைப் பயன்படுத்தி எழுத்து, வரைகலை, இசை, வீடியோ மற்றும் இயங்குபடம்ஆகிய ஊடகங்களின் ஒருங்கிணைந்த படைப்பின் மூலம் மாணவர்களை கற்றலில் செல்நெறிப்படுத்தவும், உருவாக்கவும், இருவழித்தொடர்புபடுத்தவும் தகுந்த விசைக்கருவிகளைக் கையாள்வதேயாகும். மேலும்பல்லாடகம் என்பதை ஒன்றுக்கு மேற்பட்ட ஊடகங்களைப் பயன்படுத்தி தகவலை வெளிப்படுத்தும் தொழில்நுட்பம் என்றும் கூறலாம். தொழில்நுட்ப மாற்றங்களுக்கு ஆசிரியர்கள் குறிப்பாகத் தமிழ்பள்ளி ஆசிரியர்கள் ஈடுபாடுகாட்டுகின்றனரா?. தமிழ்பள்ளி ஆசிரியர்கள் ஈடுபாடு காட்ட ஆரம்பித்துள்ளனர். இருப்பினும் கணினியின் பயன்பாடுமேலும் அதிகமாகக் காணப்பட வேண்டும். வாழ்க்கையை பண்படுத்தும் இலக்கியத் திறனைஆசிரியர்கள்மாணவர்களுக்கு இளமையிலேயே வளர்த்து விட வேண்டும்.

1.1 பின்னணி

இலக்கியக் கூறுகளுள் ஒன்றாகத் திகழும் திருக்குறள் அதனைக் கற்போரின் அறிவை வளர்ப்பது மட்டும்அல்லாமல், வாழ்க்கையையும் நெறிப்படுத்துகிறது. மாணவர்கள் தமிழ்மொழியில் பல சிரமங்களைஎதிநோக்குவதற்கும், இலக்கியப் பகுதிகளில் தேவையான அடைவுநிலைகளைப் பெறத் தடுமாறுவதற்கும்முக்கியக் காரணங்களில் ஒன்றாக அமைவது தமிழ் கற்றல் கற்பித்தலில் பல்லாடகத் தகவல் தொடர்புதொழில்நுட்பப் பயன்பாடு நிறைவாக நடைபெறுவது கிடையாது. பாடத்திட்டம் மாணவர்களிடையே இலக்கியக்கூறுகளை வலியுறுத்த வேண்டியதன் அவசியத்தை எடுத்துரைத்தபோதிலும், நடைமுறைமுறையில் ஆசிரியர்கள்இதைச் செயல்படுத்த பல சிரமங்களை எதிர்கொள்ள வேண்டியுள்ளது. இலக்கியக் கூறுகளுள் ஒன்றாகத் திகழும்திருக்குறள் தொடக்கப் பள்ளியில் பயில மாணவர்களுக்கு வாய்ப்பு வழங்கப்பட்டுள்ளது. இருப்பினும் அதிகமானமாணவர்கள் தேர்வுக்காக மட்டுமே மனனம் செய்கின்றனரே தவிர அதனின் முழு இலக்கியச் சுவையைஉணராமலேயே இருக்கின்றனர். இதற்குக் காரணம் கற்றல் கற்பித்தலின்போது தமிழ் ஆசிரியர்கள் கையாளும்கற்பித்தல் உத்திகள், அணுகுமுறைகள், மாணவர்களின் தரத்திற்கேற்றவாறும், ஈர்க்கும் வகையிலும்அமைவதில்லை.

தமிழர்களின் பண்பாட்டையும் பாரம்பரியத்தையும் காலத்தால் அழியாத பண்பாட்டு மேன்மைகளையும்எடுத்துக்காட்டி தமிழனின் பெருமைகளை உலக அரங்கில் உலாவரச் செய்த உன்னத நூல் திருக்குறள். திருக்குறள்மிகச்சிறந்த இலக்கியப் பெட்டகமாகவும் என்றென்றைக்கும் பொருந்தக்கூடிய அரிய தத்துவங்களையும்அறிவுரைகளையும் ஈரடிக்குள் செதுக்கிவைத்த சிந்தனைக் களஞ்சியமாகவும் திகழ்கிறது. இரண்டாயிரம்ஆண்டுகளுக்கு முன்னர் வாழ்ந்த தமிழனின் அரசியல், பண்பாடு, காதல் போன்ற அம்சங்களைக் காட்டுகின்ற காலக்கண்ணாடியாகவும் திகழ்வது திருக்குறள். மேலும் இதுவரை வந்த தமிழ் நூல்களிலேயே அதிகமான மொழிகளில்மொழி பெயர்க்கப்பட்டு அகில உலகத்தையும் கவர்ந்தது திருக்குறளாகும்.

வான்புகழ் வள்ளுவன் விட்டுச் சென்ற “திருக்குறள்” தமிழ் இனத்திற்கும் தமிழ் மொழிக்கும் கிடைத்த அரியவரப்பிரசாதமாகும். தம் இனத்திற்கும் மொழிக்கும் காலம் உள்ளவரைக்கும் பெருமை சேர்க்கும் “குறளை” தமிழ்மறை என்றும் அறிஞர்கள் அழைப்பதுண்டு. இதற்குக் காரணம் அதில் தமிழ், தமிழன், தமிழ்நாடு என்ற ஒருகுறிப்பிட்ட வட்டத்தைக் குறிக்கின்ற சொற்கள் அடங்காமால் மனித இனத்தையே அது

ஒட்டுமொத்தமாகக்கூறுவதாகும். “திருக்குறள்” இன்று உலகின் பலமொழிகளில் மொழி ஆக்கம் செய்யப்பட்டு பல்லினத்தாரும் அதனை வாழ்க்கைக்கான வழிகாட்டியாகப் பயன்படுத்தி வருகின்றனர். இத்தகையச் சிறப்பு வாய்ந்த திருக்குறளை தமிழ்ப் பள்ளி மாணவர்கள் ஆர்வத்துடன் கற்றுத்தேற தமிழ்மொழி ஆசிரியர்கள் கற்றல் கற்பித்தல் நடவடிக்கையில் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பப் பயன்பாட்டினை மேம்படுத்த வேண்டும்.

எல்லா மாணவர்களும் மொழித்திறனில் ஒரே விதமான ஆற்றல் கொண்டவர்களாக இருப்பார்கள் என்றும், மொழிகற்றல் அடைவுநிலைகள் எல்லா மாணவர்களுக்கும் ஒரேவிதமாக இருக்கும் என்றும் எண்ணுவது என்பது தவறான அனுமானங்களாகும். இதற்கு ஒரே வழி ஆய்வாளர், மாணவர்களை ஈர்க்கும் வகையில் புதியதொரு அணுகுமுறையைப் பயன்படுத்திக் கற்றல் கற்பித்தலில் நடவடிக்கையைச் செயலாக்க வேண்டும் என்பதை உணர்ந்தார். இதனைக் கருத்தில் கொண்டுதான் பல்லாடகக் கற்றலின் வழி மூன்றாம் ஆண்டு மாணவர்களிடையே திருக்குறள் பயன்பாட்டினை மேம்படுத்த முடியும் என்று இந்த ஆய்வினை மேற்கொண்டார். இந்த ஆய்வில்குறிப்பாக எழுத்து, இசை அல்லது குரல் பதிவு, திரைப்படக்காட்சிகள் மற்றும் இயங்குபடம் போன்ற ஊடகங்கள் மாணவர்களின் திருக்குறள் பயன்பாட்டினை வளப்படுத்தப் பயன்படுத்தப்பட்டன.

1.2 ஆய்வுக் குரிய சிக்கல்

இந்த ஆய்வின் சிக்கலானது மாணவர்களிடையே திருக்குறள் பயன்பாடு மிகவும் குறைவாக உள்ளதே ஆகும். மாணவர்கள் இலக்கியம் என்றாலே ஈடுபாடு இழந்து காணப்படுவதை ஆய்வாளர் கண்டறிந்தார். மாணவர்கள்திருக்குறளை தேர்வுக்காக மனனம் செய்வதையே முக்கியமாகக் கருதுகின்றனர். மனனம் செய்யும் மாணவர்கள்திருக்குறளின் இலக்கியச் சுவையை உய்த்துணராமல் இருக்கும் சூழ்நிலையைக் காணமுடிந்தது. குறிப்பாகக் கட்டுரை எழுதும் வேளை சூழலுக்கு ஏற்றவாறு பொருளுணர்ந்து பயன்படுத்துவதில் சிரமத்தை எதிர்நோக்கினர். மேலும் அவர்களின் உரையாடல்கள் மற்றும் அன்றாட நடைமுறை வாழ்க்கையிலும் திருக்குறள் பயன்பாடு மிகவும் குறைவாகவே காணப்பட்டது. இதற்குக் காரணம் கற்றல் கற்பித்தலின்போது திருக்குறளின் பொருளை முழுமையாகப் புரிந்து கொள்வதில்லை. மேலும் திருக்குறளில் பயன்படுத்தப்பட்டிருக்கும் சொற்கள் புரிந்து கொள்ளக்கூடியதாக இருப்பதனால் மாணவர்கள் தேர்வின் அடிப்படையில் பொருளறியாமல் மனனம் செய்யும் சூழ்நிலையைக் காணமுடிந்தது. இதனால் மாணவர்கள் தேர்வுக்காக மட்டுமே கற்பது போன்ற சூழ்நிலைக்காணப்படுகின்றது. திருக்குறள் கற்றல் கற்பித்தலின்போது ஆசிரியர்கள் பல்லாடகக் கற்றலின் வழி குறிப்பாக எழுத்து, இசை அல்லது குரல் பதிவு, திரைப்படக்காட்சிகள் மற்றும் இயங்குபடம் போன்ற ஊடகங்களின் வழி மாணவர்களின் திருக்குறள் பயன்பாட்டினை வளப்படுத்த முடியும் என்ற கருத்து இந்த ஆய்வின் வழிகண்டறியப்பட்டது.

1.3 ஆய்வு நோக்கம்

- கற்றல் கற்பித்தல் நடவடிக்கையில் மூன்றாம் ஆண்டு மாணவர்களிடையே திருக்குறள் பயன்பாடு அடைவுநிலையைக் கண்டறிதல்.
- மாணவர்கள் திருக்குறள் கற்பதில் எதிர்கொள்ளும் சிக்கலின் காரணத்தைக் கண்டறிதல்.
- பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பப் பயன்பாட்டின் வாயிலாக மாணவர்களிடையே திருக்குறள் பயன்பாட்டினை மேம்படுத்துதல்.

1.4 ஆய்வு வினா

- ஆசிரியரின் கற்றல் கற்பித்தல் நடவடிக்கையின்போது மூன்றாம் ஆண்டு மாணவர்களிடையே திருக்குறள் பயன்பாடு அடைவுநிலை எந்த அளவில் உள்ளது?
- மாணவர்கள் திருக்குறளைக் கற்பதில் எத்தகையச் சிக்கலை எதிர்நோக்குகின்றனர் மற்றும் அதற்கானக் காரணங்கள் யாவை?
- கற்றல் கற்பித்தலில் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்ப அணுகுமுறை மாணவர்களின் திருக்குறள் பயன்பாட்டினை எவ்வாறு மேம்படுத்துகிறது?

1.5 ஆய்வு எல்லை

இந்த ஆய்வு சிரம்பான் வட்டாரத்தில் அமைந்த தமிழ்ப்பள்ளியில் பயிலும் மூன்றாம் ஆண்டு மாணவர்களிடையே நடத்தப்பட்டது. கற்றல் கற்பித்தலின்போது மாணவர்களின் திருக்குறள் பயன்பாட்டினை எவ்வாறு வளப்படுத்த முடியும் என்பதே இவ்வாய்வின் நோக்கமாகும். இவ்வாய்வின் முடிவானது இக்கட்டுரைகளின் மாதிரி தேர்ந்தெடுக்கப்பட்டப் பள்ளிக்கே உரித்தான ஒன்று. இவ்வாய்வின் மூலம் கிடைக்கும் முடிவுகளைப் பொதுமையாக்கிக் கொள்ள முடியாது.

2.1 ஆய்வு வடிவமைப்பு

ஆய்வு நடத்துவதற்கு முன் ஆசிரியரின் கற்றல் கற்பித்தலின் போது நுண்ணோக்கல் வழி (observation) மாணவர்கள்திருக்குறள் பயன்பாட்டில் எத்தகைய அடைவநிலையை அடைந்துள்ளனர் என்ற தகவல்கள் திரட்டப்பட்டன. மேலும் அன்றாடக் கற்றல் கற்பித்தலில் மாணவர்களின் திருக்குறள் பயன்பாடு சில கூறுகளின் அடிப்படையில் உற்று நோக்கலின் வழியும் தகவல்கள் திரட்டப்பட்டன. அடுத்து, திருக்குறள் கற்றல் கற்பித்தலின்போது மாணவர்கள் எதிர்நோக்கும் சிக்கல்களை அடையாளம் காண தர மதிப்பீட்டு கோளுடைய வினா நிரல் பயன்படுத்தப்பட்டது. மேலும் நடவடிக்கையின் அமலாக்கத்திற்கு முன் முன்னறிச் சோதனை நடத்தப்பட்டது. செயல்திட்டத்தின் பிறகு மாணவர்களிடையே திருக்குறள் பயன்பாட்டினைக் கண்டறிய பின்னறி சோதனை மற்றும் உற்றுநோக்கலும் நடத்தப்பட்டது. நடவடிக்கையின் அமலாக்கத்திற்கு முன் மாணவர்களின் அன்றாட உரையாடல்களிலும், கட்டுரைகளிலும் திருக்குறளின் பயன்பாடு பின்தங்கிய நிலையில் இருந்தது கண்டறியப்பட்டது.

2.2 ஆய்வுக்கு உட்படுவோர்

இந்த ஆய்விற்கு உட்பட்டவர்கள் 20 மாணவர்கள் ஆவர். இந்த வகுப்பில் மொத்தம் 8 ஆண் மாணவர்களும் 12 பெண்மாணவர்களும் உள்ளனர். தமிழ்பள்ளி ஆசிரியர்கள் திருக்குறள் கற்றல் கற்பித்தலின்போது பல்லாடகக் கற்றலின்வழி குறிப்பாக எழுத்து, இசை அல்லது குரல் பதிவு, திரைப்படக்காட்சிகள் மற்றும் இயங்குபடம் போன்ற ஊடகங்களைப் பயன்படுத்தினால் மாணவர்களின் திருக்குறள் பயன்பாட்டினை வளப்படுத்த முடியும் என்ற கருத்து இந்த ஆய்வின்வழி கண்டறியப்பட்டது.

2.3 ஆய்வுக் கருவிகள்

இந்த ஆய்விற்கான தரவுகளைத் திரட்டுவதற்கு ஆய்வுக் கருவிகளாக நுண்ணோக்கல், வினா நிரல் பட்டியல், நேர்க்காணல் ஆகியவைகள் கையாளப்பட்டன. இந்த ஆய்வின் வழி, மாணவர்களை ஈர்க்கும் வகையில் ஏற்படையநவீன அணுகுமுறைகளை கற்றல் கற்பித்தலில் பயன்படுத்தினால் மூன்றாம் ஆண்டு மாணவர்களிடையே திருக்குறள் பயன்பாட்டினை மேம்படுத்த முடியும் என்ற உண்மை கண்டறியப்பட்டது.

2.4 ஆய்வு அமைப்புமுறை

செயல்திட்டத்தின் அமலாக்கம்

அட்டவணை 1 : செயல்திட்ட நடவடிக்கைகள்

வாரம் நடவடிக்கைகள்

1 -முன்னறித் தேர்வு

ஒளிபரப்பப்படும் கதையில் இடம்பெற்ற திருக்குறள் மற்றும் அதன் பொருளைக் கூறப் பணித்தல்.

தனியாள் முறையில் கூறப் பணித்தல்

2 ஒளிபரப்பப்பட்ட சிறுவர் நாடகங்களில் பயன்படுத்தியத் திருக்குறள் மற்றும் அதன் பொருளைக் குழுமுறையில் கலந்துரையாடிக் கூறப் பணித்தல்

3 ஒளிபரப்பப்பட்ட திரைப்படப்படல்களில் பயன்படுத்தியத் திருக்குறளை அடையாளம் காண்பணித்து அதன்பொருளை கூறப் பணித்தல்.

தனியாள் முறையில் கூறப்பணித்தல்

இசை, குரல் பதிவு (audio) நடத்தப்பட்டது

4 ஒளிபரப்பப்பட்ட சினிமா உரையாடலுக்கேற்ற திருக்குறளை கூறப்பணித்தல்.

குழுமுறை போட்டி நடத்தப்பட்டது

5 ஒளிபரப்பப்பட்ட சிறுவர் கதையில் இடம்பெற்ற திருக்குறளைக்

கலந்துரையாடி கதைமாந்தர் பாகமேற்று நடிக்கப் பணித்தல்

-இயங்குபடம்

6 கணினியில் பதிவு செய்து ஒளிபரப்பப்பட்ட உரையாடலைச் செவிமடுத்து , பொருத்தமானத் திருக்குறளைக் கூறப் பணித்தல்.

- 7 ஒளிபரப்பப்பட்ட சிறுவர் பாடலில் இடம்பெற்ற திருக்குறளைக் கலந்துரையாடித் தனியாள் முறையில் கருத்துகளைக் கூறப் பணித்தல்
-இயங்குபடம்
- 8 கணியைப் பயன்படுத்தி திருக்குறள் மொழி விளையாட்டு
- 9 மூன்றாம் ஆண்டுக்கான திருக்குறளைத் தொடர்படங்களைப் பயன்படுத்தி அதற்கான பொருளைக் கூறப்பணித்தல் (கணினியைப் பயன்படுத்தி எழுத்து (text), குரல் பதிவு (audio)
-சூழல் அமைத்தல்
- 10 முன்னறித் தேர்வில் பார்த்த அதே கதை பின்னறித் தேர்விலும் ஒளிபரப்பப்பட்டது.
திருக்குறளைப் பயன்படுத்தி கதை/சூழல் கூறப் பணித்தல்
குரல் பதிவு (audio) நடத்தப்பட்டது

3.1 தரவுகள் பகுப்பாய்வு

இந்த ஆய்வில் நடத்தப்பட்ட நடவடிக்கைகள் திருக்குறள் பொருளை அடையாளம் காணுதல், திருக்குறளைப்பொருளுணர்ந்து கூறுதல், சூழலுக்கேற்றத் திருக்குறளையும், திருக்குறளுக்கேற்றச் சூழலையும் கூறுதல், சுயமாகஉருவாக்கும் கதைகள்/ உரையாடல்களில் திருக்குறளைப் பயன்படுத்துதல், சிறுவர் பாடல்கள் மற்றும் கதைகளில்இடம்பெற்றத் திருக்குறளை அடையாளம் கண்டு பொருள் கூறுதல், கணினியைப் பயன்படுத்தி மொழிவிளையாட்டுபோன்ற கூறுகளின் அடிப்படையில் அமையப்பெற்றன. முன்னறிச் சோதனை, பின்னறிச் சோதனை இவற்றின்முடிவுகளைக் கொண்டு அட்டவணைகள் உருவாக்கப்பட்டன.

அனைத்து நடவடிக்கைகளும் திருக்குறளின் பயன்பாட்டினை மாணவர்களிடையே மேன்மைபடுத்தும் வகையில்பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பத்தை பயன்படுத்தி நடத்தப்பட்டது. ஆய்வு நடவடிக்கைகளுக்கு முன்திருக்குறள் கற்பதை சுமையாக நினைத்த மாணவர்கள் ஆய்வு நடவடிக்கைகளுக்கு பின் திருக்குறள் கற்றலைசுகமானதாக கருதினர். இதனை மாணவர்களிடையே திருக்குறள் பயன்பாடு கற்றல் கற்பித்தலிலும், அன்றாடநடைமுறை வாழ்க்கையில் பயன்படுத்துவதன் வழியும் காண முடிந்தது.

அட்டவணை 2 மாணவர்களின் திருக்குறள் பயன்பாட்டின் அடைவுநிலை

மாணவர்கள் முன்னறிச் சோதனை /100 பின்னறித் சோதனை /100 வேறுபாடு

(ஏற்றம்) சதவிகிதம் (%)

(ஏற்றம்)

1	35	77	42	54.5
2	35	76	41	53.9
3	32	68	36	52.9
4	33	69	36	52.1
5	34	70	36	51.4
6	30	65	35	53.8
7	30	61	31	50.8
8	42	82	40	48.7
9	43	82	39	47.5
10	40	77	37	48
11	38	75	37	49.3
12	44	85	41	48.2
13	30	64	34	53.1

14	33	66	33	50
15	35	67	32	47.7
16	30	60	30	50
17	45	86	41	47.6
18	34	73	39	53.4
19	32	60	28	46.6
20	32	64	32	50

சராசரி 35.35 71.35 36 50.4

மேற்கண்ட முன்னறிச் சோதனை மற்றும் பின்னறிச் சோதனை முடிவினைக் கீழ்க்கண்டவாறு சராசரியின் அடிப்படையில் கணக்கிடப்பட்டுப் பகுப்பாயப்பட்டுள்ளது.

அட்டவணை 3 முன்னறி மற்றும் பின்னறிச் சோதனையின் பகுப்பாய்வு

புள்ளிகள் முன்னறிச் சோதனை

(மாணவர் எண்ணிக்கை) பின்னறிச் சோதனை

(மாணவர் எண்ணிக்கை)

0 - 39 15

(75%) 0

(0%)

40 - 60 5

(25%) 2

(10%)

61 - 79 0

(0%) 14

(70%)

80 - 100 0

(0%) 4

(20%)

மேற்கண்ட அட்டவணையைப் பகுப்பாய்வு செய்யும் பொழுது பின்னறிச் சோதனைகளின் முடிவுகளைப் பார்த்தோமானால் எல்லா மாணவர்களின் அடைவுநிலையும் சற்று வளர்ச்சிநிலை அடைந்துள்ளதை காணமுடிகிறது. முன்னறிச் சோதனையில் மொத்தம் 15 அல்லது 75% மாணவர்கள் குறைவான புள்ளிகளைப் பெற்றுள்ளனர். மேலும் 5 அல்லது 25% மாணவர்கள் நடுத்தரமானப் புள்ளிகளையேப் பெற்றிருந்தனர். இம்முன்னறிச் சோதனை ஆய்விற்கு உட்படுத்தப்பட்ட 20 மாணவர்களில் எந்த ஒரு மாணவரும் சிறப்பு தேர்ச்சிப்பெறவில்லை. அதாவது கற்றல் கற்பித்தல் நடவடிக்கையின்போது மாணவர்களிடையே திருக்குறள் பயன்பாடுமிகவும் குறைந்த நிலையில் காணப்பட்டது. இருப்பினும் பின்னறிச் சோதனையில் 70% மாணவர்கள் சிறப்பானத்தேர்வும் 20% விழுக்காடு மாணவர்கள் மிகவும் சிறப்பானத் தேர்ச்சியினையும் பெற்றுள்ளனர். அதாவது மாணவர்களின் முறையான திருக்குறள் பயன்பாடு அதிகமாக காணப்பட்டன. இப்பின்னறிச் சோதனையில் எந்தமாணவரும் குறைந்தப் புள்ளிகள் பெறவில்லை என்பது குறிப்பிடத்தக்கது. முன்னறிச் சோதனையில் சிறப்பு தேர்ச்சி நிலை இல்லாது, பின்னறிச் சோதனையின் போது சிறப்பு தேர்ச்சி விகிதம் 20% விழுக்காடாக உயர்ந்திருப்பதுநல்ல முன்னேற்றத்தைக் காட்டுகிறது. திருக்குறள் கற்றல் கற்பித்தல் நடவடிக்கையில் பல்லாடகத் தகவல்தொடர்பு தொழில்நுட்பப் பயன்பாடு மாணவர்களிடையே சிறந்த முன்னேற்றம் காணப்பட்டுள்ளதை அறியமுடிகின்றது.

4.1 ஆய்வு கண்டு பிடிப்புகளும் முடிவும்

மாணவர்களின் அடைவுநிலைக்கேற்றவாறும், தரத்திற்கேற்றவாறும், மனமகிழ்வு ஊட்டும் வகையில் பல்லாடகத்தகவல் தொடர்பு தொழில்நுட்பத்தைப் பயன்படுத்தி கற்றல் கற்பித்தல் முறையே நடைபெற்றால் மாணவர்களின் திருக்குறள் பயன்பாட்டினை வளப்படுத்தி வெற்றி காணலாம் என்பதே இந்த ஆய்வின் முடிவு.

கற்றல் கற்பித்தலில் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பத்தைப் பயன்படுத்துவதன் வழி மாணவர்களிடமிகவும் ஆர்வத்துடனும், மகிழ்வுடனும் காணப்படுகின்றனர். ஆய்விற்குப் பிறகு மாணவர்களால் குறளுக்கு ஏற்றபொருளையும், சூழலுக்கு ஏற்ற திருக்குறளையும் சுயமாக தயாரிக்க முடிந்தது. மாணவர்களின் கட்டுரை பயிற்சிபுத்தகங்களைக் கண்ணோட்டமிடும்போது கட்டுரையில் திருக்குறளின் பயன்பாடு அதிகமாகக் காணப்பட்டது. ஆசிரியரின் கற்றல் கற்பித்தலில் மாணவர்களின் ஈடுபாடு அதிகமாகக் காணப்பட்டது. ஆய்வாளரின் கற்றல் கற்பித்தல் நடவடிக்கைகள் யாவும் மாணவர்கள் மையமாகவே இருந்தது. மாணவர்களே தகவல்களை சேகரித்து குழு நடவடிக்கைகளில் செயல்படுவதால் மாணவர்களின் பங்கேற்பு அதிகமாகக் காணப்பட்டது. இந்நடவடிக்கைகள் மாணவர்களின் சிந்தனை ஆற்றலை வளப்படுத்தும் வகையில் அமைந்திருந்தது. குழு நடவடிக்கையின்போது ஆர்வத்துடன் தயக்கமின்றி தங்களுடைய கருத்துகளைப் பரிமாறிக்கொள்கின்றனர். மாணவர்கள் இணையத்தின்வழி திருக்குறள் அடங்கிய கதை, பாடல், உரையாடல், நாடகம், திரைப்படக் காட்சிகள், மொழி விளையாட்டு போன்ற நடவடிக்கைகளில் ஈடுபடும்பொழுது மாணவர்களின் சுயகற்றல் ஆர்வமும் காணப்படுகிறது. ஆய்வுக்குப் பின்மாணவர்களிடையே திருக்குறள் பயன்பாடு மிகவும் அதிகமாகவும், ஆர்வத்துடனும் காணப்பட்டது. மேலும்பள்ளிகளில் நடத்தப்படும் திருக்குறள் போட்டி, கதைக்கூறும் போட்டி, பேச்சுப் போட்டிகளில் பொருத்தமானப் திருக்குறள்களைப் பயன்படுத்தினர்.

கற்றல் கற்பித்தல் நடவடிக்கையின்போது பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பத்தைப் பயன்படுத்துவதன்வழி மாணவர்களிடையே திருக்குறள் பயன்பாட்டினை மிகவும் சிறப்பாக வளப்படுத்த முடியும் என்பதே இந்த ஆய்வின் கண்டு பிடிப்பாகும்.

மாணவர்கள் திருக்குறள் கற்பதில் எதிர்கொள்ளும் சிக்கலின் காரணத்தை ஆராயும்பொழுது கிடைக்கப்பெற்றதகவல்கள் கீழ்க்கண்டவாறு வகுக்கப்பட்டுள்ளன:

தமிழ்மொழி கற்றல் கற்பித்தலின்போது தமிழ்மொழி ஆசிரியர்களிடம் பல்லாடகப் படைப்பான குறுந்தட்டு போன்றதுணைக் கருவிகளின் பயன்பாடு மிகக் குறைவு

போதிய தமிழ் குறுந்தட்டுகள் கிடைப்பதில்லை.

பள்ளியில் கணினி வசதியிருந்தும் பயன்படுத்தாத சூழ்நிலை

தேர்வுக்கு மட்டும் முக்கியத்துவம் அளிக்கும் சூழ்நிலை

தொழில் நுட்ப மாறுதலுக்கு ஆசிரியர்கள் தங்களைத் தயார்படுத்திக் கொள்ளாமை

பாடப் புத்தகங்களிலுள்ள உள்ளடக்கங்களே கற்றல் கற்பித்தல் நடவடிக்கைகளுக்கு போதுமானது என்ற ஆசிரியரின் சிந்தனை.

பழைய கற்றல் கற்பித்தல் (பாடப் புத்தகத்தை மட்டும் பயன்படுத்துதல்) சிறப்பு என்ற பொதுவான மனப்போக்கு

மேற்கண்ட செயல்பாடுகள் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பப் பயன்பாட்டிற்கு தடையாக அமைந்துமாணவர்கள் திருக்குறள் கற்றல் கற்பித்தல் நடவடிக்கையில் ஆர்வத்துடனும், மனமகிழ்ச்சியுடனும் ஈடுபடமுடியாத சிக்கலுக்கு வழிவகுத்தது.

பரிந்துரைகள்

- 1 பல்லாடகத் தகவல் தொடர்பு தொழில் நுட்பத்தின் வழி கற்றல் மிகுந்த பலனை அளிப்பதால் இம்முறையை ஆசிரியர்கள் முறையாகப் பயன்படுத்துவது சிறப்பை அளிக்கும்.
- 2 தமிழாசிரியர்கள் அடிப்படைப் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பத் திறன் பெற்றிருப்பவர்களாகத் திகழவேண்டும்.
- 3 மாணவர்களின் அடைவு நிலைக்கு ஏற்றவாறும், சிந்தனையை வளப்படுத்தும் வகையிலும் பல்லாடகத்தகவல் தொடர்பு தொழில் நுட்பத்தைக் கற்றல் கற்பித்தலில் பயன்படுத்தி பயிற்சிகள் அளித்து மாணவர்களின் ஆர்வத்தை வளர்க்க வேண்டும்.

- 4 மாணவர்களின் தமிழ்மொழிப் பாட கற்றல் கற்பித்தலின்போது ஆசிரியர் கணினியைப் பயன்படுத்தி எழுத்து, வரைகலை, இசை, வீடியோ மற்றும் இயங்குபடம், திரைப்படக் காட்சிகள் ஒளிப்பதிவு வழி மாணவர்களின் மொழிவளத்தை வளப்படுத்துவதோடு கற்பனை வளத்தையும் தூண்ட வேண்டும்..
- 5 ஆசிரியர்கள் கணினியைப் பயன்படுத்தி எழுத்து, வரைகலை, இசை, வீடியோ மற்றும் இயங்குபடம், திரைப்படக் காட்சிகள் ஒளிப்பதிவு வழி கற்றலை நடத்துவதற்கு முயற்சி செய்யவேண்டும்
- 6 பள்ளி ஆசிரியர்கள் இணைந்து எளிய முறையில் பல்லாடகக் குறுந்தட்டு தயாரிப்பில் ஈடுபட வேண்டும்.மாணவர்களுக்குக் கணினி மற்றும் இணையத்தின் வழி பயிற்சிகள் அளிக்கும் வகையினைத் தயார் செய்வதற்குத்தமிழாசிரியர்கள் முயற்சி செய்யவேண்டும்.
- 7 மாணவர் நலனைக் கருத்தில் கொண்டு, தமிழாசிரியர்களுக்குப் பல்லாடகத் தொழில்நுட்பப் பயிலரங்குகளை மாவட்ட அல்லது மாநில அளவிலானத் தமிழ்மொழி பாடத்திட்ட மேம்பாட்டுக்குழு ஏற்பாடு செய்தல் வேண்டும்.
- 8 தமிழில் பல்லாடகத் தகவல் தொடர்பு தொழில்நுட்பம் சார்ந்த பட்டறை, கருத்தரங்கம், ஆய்வரங்கம் போன்றவைகளில் தமிழ்பள்ளி ஆசிரியர்கள் பங்கேற்று பயனடைய வேண்டும். இதன் வழி கற்றல் கற்பித்தலுக்குஉகந்த வலைதளங்களைப் பதிவிறக்கம் செய்து மாணவர்களை ஈர்க்கும் வகையில் பாடங்களைத் திட்டமிட்டுகற்பிக்கலாம்.

முடிவுரை

தமிழாசிரியர்களின் புத்தாக்கச் சிந்தனையில்தான் நம் மாணவர்களின் ஒளிமயமான எதிர்காலம் அடங்கியுள்ளது.சிறப்பான கடமையுணர்வு கொண்டுள்ள நம் தமிழாசிரியர்கள் காலமாற்றத்திற்கு ஏற்பவும், வளர்ந்து வரும் நவீனபல்லாடத் தொழிநுட்ப வசதிகளுக்கேற்பவும் தங்கள் கற்றல் கற்பித்தல் நடவடிக்கையில் புதிய அணுகுமுறையைக்கையாள்வதை முக்கியக் கடமையாக கொள்ள வேண்டும்.

துணை நூல் பட்டியல்

1. வேணுகோபால் இ. பா (2006) சிறப்புத்தமிழ் கற்பிக்கும் முறைகள், சாரதா பதிப்பகம் சென்னை
2. சிங்கன். வி (1973) திருக்குறளைப் படிப்பதற்குப் பத்து முறைகள், இ.வி.எஸ் எண்டர்பிரைஸ், சிங்கப்பூர்.
3. சுப்புரெட்டியார். ந (2003), தமிழ் பயிற்றும் முறை, மெய்யப்பன் பதிப்பகம் சிதம்பரம்.
4. இரபிசிங், ம.செ : இணைய வழித் தமிழ் இலக்கியம் கற்பித்தல், 5வது உலகத் தமிழாசிரியர் மாநாடு, சிங்கப்பூர்
5. சு.திண்ணப்பன் (1987), கணினியும் தமிழ் கற்பித்தலும், ஆறாவது உலகத் தமிழ் ஆராய்ச்சி மாநாடு ஆய்வுக்கட்டுரைகள், மலாயாப் பல்கலைக்கழக, கோலாலம்பூர்.
6. கருணாநிதி. மா (2008) கற்றல் கற்பித்தல் மேம்பாட்டுக்கான வழிமுறைகள், சேமமடு பதிப்பகம் கொழும்பு.

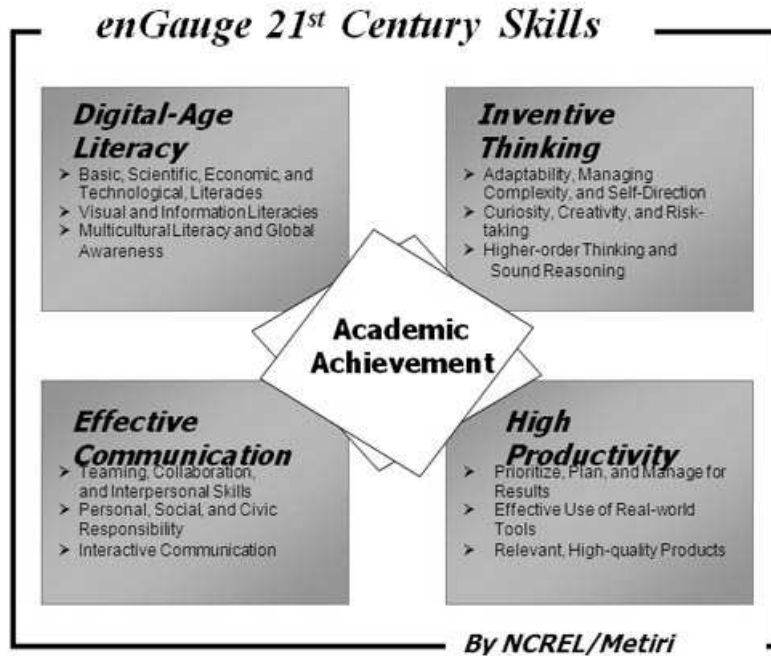
பன்முகத்தன்மை வாய்ந்த முகநூல், வாட்ஸ்ஆப் வழி ஒருங்கிணைந்த மொழித் திறன்கள் கற்றல் கற்பித்தல்

கோவிந்தசாமி சந்தானராஜ்

21-ஆம் நூற்றாண்டுக் கற்றலின் பரப்பும் நோக்கும்

21-ஆம் நூற்றாண்டுக் கற்றலில், மின்னியல் அறிவு (DIGITAL LITERACY), கற்பனையோடு கூடிய படைப்பாக்கச் சிந்தனை (INVENTIVE THINKING), பயன்முனைப்புமிக்க கருத்துப்பரிமாற்றம் (EFFECTIVE COMMUNICATION) அதிக உற்பத்தித் திறன் (HIGH PRODUCTIVITY) எனும் நான்கு பரப்புகள் முக்கியப் பங்கு வகிக்கின்றன. இன்றைய மாணவர்கள் நடைமுறைக்கேற்பக் கற்றுக்கொள்வதும் நடைமுறையோடு பொருத்திப் பார்த்துக் கற்றலில் தாம் அடைந்துள்ள நிலையைப் புரிந்துகொண்டு மேம்பட வழி வகுக்கவும் மற்றும் கற்றவற்றை நடைமுறைக்கேற்ப வெளிப்படுத்திச் செயலாற்றிப் பயன் பெறவும் 21-ஆம் நூற்றாண்டுத் திறன்கள் மிகவும் அவசியமாகின்றன. இவ்வகையில் மாணவர்கள் தங்களுடைய கற்றலை எவ்வித அச்சுறுத்தலும் இல்லாத சூழலில் நிகழ்த்திக்கொள்ளவும் மனமகிழ்ச்சியுடன் விருப்பமான முறையில் கற்றுக்கொள்ளவும் தங்களின் கற்பனைகளுக்கு உயிர்கொடுக்கும் வசதி வாய்ப்புள்ள களமாகப் பயன்படவும் சிறு முயற்சியின் மூலம் பன்மடங்குச் சிறப்பு வாய்ந்த உற்பத்தித்திறன்மிக்க படைப்புகளைப் படைக்கவும் அப்படைப்புகளைப் பன்முக நிலையில் அலசி ஆராய்ந்து மேம்படுத்தவும் தகுதி வாய்ந்தது கணினியும் கணினிவழி ஊடகமான இணையமும் இணையவசதிகொண்ட கையடக்க மின்னியல் கருவிகளும் ஆகும் என்பதில் மாற்றுக்கருத்து இல்லை. தகவல்களைத் தேடியலைந்த காலம் மறைந்து தகவல்கள் நம் கைக்குள் வந்து நிற்கும் காலமிது என்பதால் கைத்தொலைபேசி நம் கற்றல் கற்பித்தலுக்கு மிகப்பொருத்தமான கருவியாக உள்ளது என்பதில் எள்ளளவும் ஐயமில்லை. இவ்வகையில் இன்று இணையத்தில் உள்ள பன்முகத்தன்மை கொண்ட சமூக வலைத்தளமான முகநூலும் (Facebook) கைத்தொலைபேசிகளில் உள்ள வாட்ஸ்ஆப் (Whatsapp) செயலியும் (apps) நிறைவான கற்றல் கற்பித்தலுக்குப் பெருந்துணையாக உள்ளன என்பதை ஆராய்வதே என் செயல்திட்டத்தின் நோக்கமாக இருந்தது.

-21ஆம் நூற்றாண்டுத் திறன்கள்

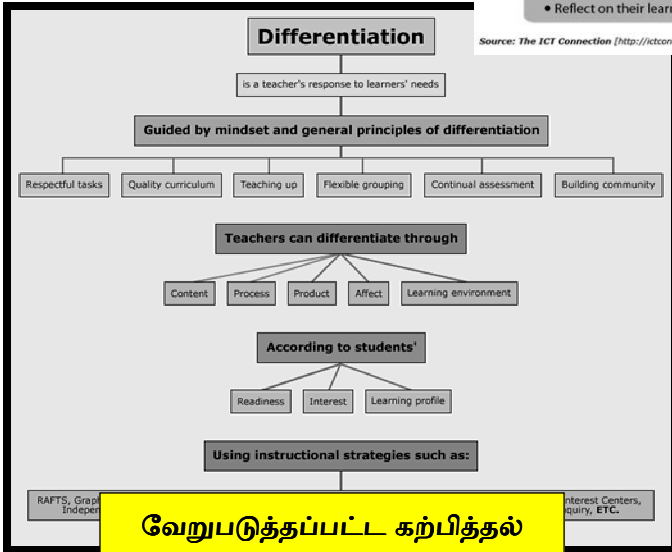
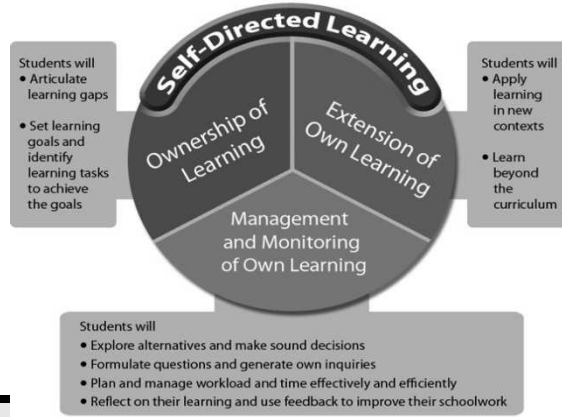


சுயவழிக் கற்றல், கூட்டாகக் கற்றல், வேறுபடுத்தப்பட்ட கற்பித்தல்

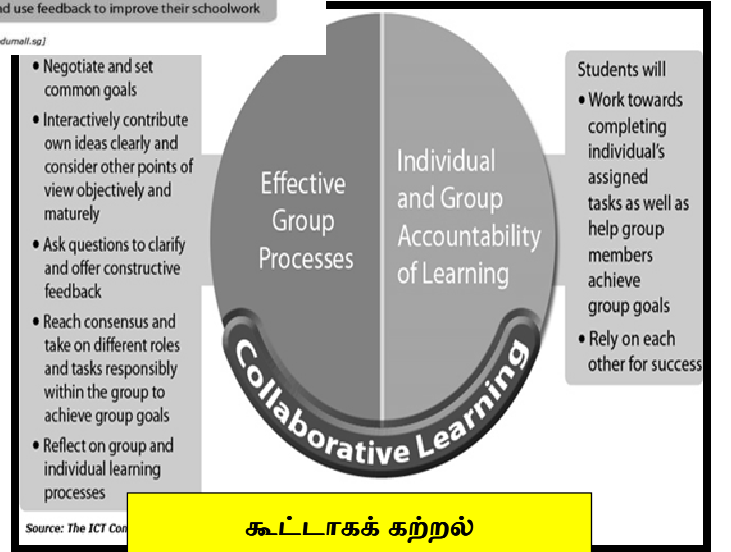
திரு மால்கம் நோல்ஸ் என்னும் அமெரிக்கக் கல்வியாளரின் வரையறைப்படி, சுயவழிக் கற்றலில் மாணவர்களை ஈடுபடுத்தத் தகுந்த ஊடகம் அவசியமென்பதால் 21-ஆம் நூற்றாண்டில் கல்வி தகவல் தொடர்புத் தொழில்நுட்பத்தை முதன்மையாகக் கொண்டு அமைந்துள்ளது என்பதை எவரும் மறுக்கவியலாது.

சுயவழிக் கற்றலில் முக்கிய மூன்று நிலைகள் உள்ளன. அவை, 1. கற்றலுக்குத் தாமே பொறுப்பேற்றல் (Ownership of learning) 2. சுய கற்றலை நிர்வகித்தலும் மற்றும் கண்காணித்தலும் (Self - management and self - monitoring) 3. வகுப்பறைக்கு அப்பாலும் கற்றல் அல்லது விரிவான கற்றல் (Extension of learning) ஆகும். இன்றைய சூழலில் மாணவர்கள் கற்றலைச் சுயமாகச் மேற்கொள்ளவும் கூட்டாக இணைந்து கற்றலை நிகழ்த்தித் தாம் புரிந்துகொண்டதைக் கொண்டு புத்தறிவினைப் படைக்குமளவிற்குத் தம்மை மேம்படுத்திக்கொள்ளவும் அவரவர்க்கு விருப்பமான வளங்கள்வழித் தேவையான அளவு தகவல்களைப் பெற்று விருப்பமானமுறையில் படைக்கவும் வாய்ப்பளிக்கப் பொருத்தமான துணைக்கருவிகள் மிகவும் அவசியம் என்று உணரப்பட்டது. அதற்கு முகநூலும் வாட்ஸ்ஆப் எனப்படும் செயலியும் மிகப்பொருத்தமானவையாக அடையாளம் காணப்பட்டன. எனவே 21-ஆம் நூற்றாண்டின் கற்றல் நோக்கத்தை நிறைவுசெய்வதை மனத்திற்கொண்டு 2012-ஆம் ஆண்டு முதல் பயா லேபா பெண்கள் உயர்நிலைப்பள்ளியில் படிக்கும் மாணவிகளுக்கு முகநூல் வழிக் கற்றுக்கொள்ளவும் 2013-ஆம் ஆண்டு முதல் முகநூலுடன் வாட்ஸ்ஆப் செயலியையும் இணைத்துக் கற்றல் நிகழ்த்துமாறும் செயல்திட்டம் வகுக்கப்பட்டுச் செயற்படுத்தப்பட்டது.

சுயவழிக் கற்றல்



வேறுபடுத்தப்பட்ட கற்பித்தல்



கூட்டாகக் கற்றல்

கற்றல் கற்பித்தலில் சமூக வலைத்தளமும் கைத்தொலைபேசிச் செயலிகளும்

இணையத்தில் இருவழிப்பரிமாற்றக் களமாகத் திகழும் சமூக வலைத்தளங்கள் உடனடிக் கருத்துப் பரிமாற்றத்தில் மாபெரும் புரட்சியைத் தோற்றுவித்தன. அவ்வகையில் முகநூல் மில்லியன்கணக்கான மக்களைத் தன்பால் ஈர்த்துப் புகழ்பெற்று விளங்குவதோடு கருத்துகளை எழுத்துப்பூர்வமாகப் பதிவுசெய்யவும், ஒலிப்பதிவாகவும் ஒளிப்பதிவாகவும் வெளிப்படுத்தும்வண்ணம் பன்முகத்தன்மை கொண்டதாக அமைந்துள்ள தன்மை கல்வி கற்றல் கற்பித்தலுக்கு ஏதுவாக உள்ளது என்பது உலக அளவில் பல கல்வியாளர்களின் கருத்தாகும். கணினியின் வாமன

அவதாரமாக இன்று கைத்தொலைபேசி உருவெடுக்கும் சூழலில் கல்வி கற்றலுக்கு உகந்தவாறு கைத்தொலைபேசியில் எண்ணற்ற வசதிகள் உருவாகியுள்ளன. அவற்றுள் முக்கியமானது செயலி மென்பொருள்கள் ஆகும். அவ்வகையில் உடனடி கருத்துப் பரிமாற்றக் களமாகத் திகழும் வாட்ஸ்ஆப் எனப்படும் செயலி மென்பொருள் பன்முகத்தன்மை கொண்டதாகத் திகழ்வதால் இன்றைய மாணவர்கள் விருப்பத்துடன் பயன்படுத்தப்படும் ஒன்றாகத் திகழ்கிறது. அத்துடன் முகநூலும் வாட்ஸ்ஆப் செயலியும் உடனடி மதிப்பீட்டுக்கு ஏதுவாக அமைந்திருப்பதால் கற்றலில் மாணவர்கள் மேம்பாடு காணவும் வழி செய்கின்றன. பன்முகத் தன்மை கொண்ட இவ்விரண்டு புதிய வசதிகளும் ஒருங்கிணைந்த மொழித்திறன் கற்றல் கற்பித்தலுக்கு மிகவும் பொருத்தமானவையாகத் திகழ்கின்றன.

தொழில்நுட்ப ஏற்பாடுகளும் கற்றல் கற்பித்தலுக்கான வழிமுறைகளும்

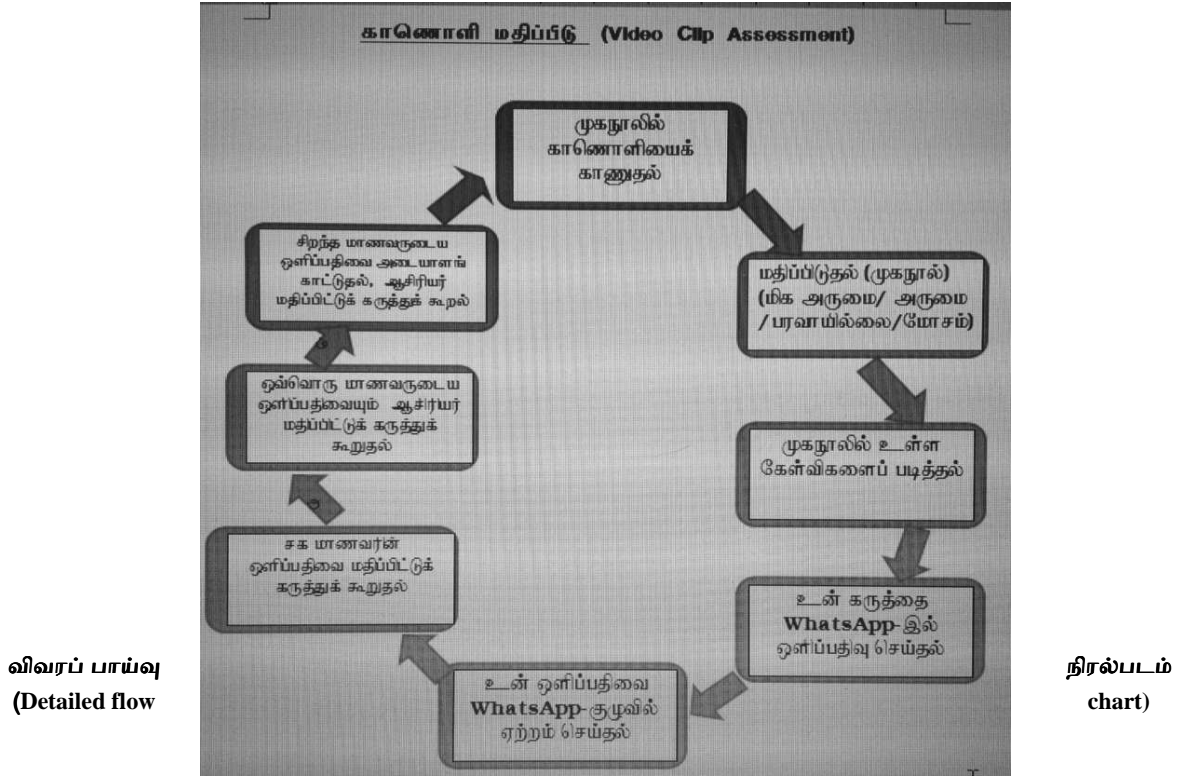
உயர்நிலை 1, 2, மாணவர்களுக்குத் தட்டச்சு செய்து எழுத்துவழியாகக் கருத்துகளைப் பதிவுசெய்வதையும், ஒலிப்பதிவு செய்தும் ஒளிப்பதிவு செய்தும் கருத்துகளைப் பதிவு செய்வதோடு சக மாணவர் மதிப்பீட்டை நிகழ்த்தி கற்றலில் மேம்படுவதையும் கருத்திற்கொண்டு முகநூலிலும் வாட்ஸ்ஆப் செயலியிலும் தகுந்த ஏற்பாடுகள் செய்து பின் பயிற்சிகள் வழங்கப்பட்டன.

1. இதற்கெனச் செயல்திட்டம் வரையப்பட்டுப் பள்ளி முதல்வரிடம் முறையான அனுமதி பெறப்பட்டது. பின்பு, தமிழ்த்துறைக்கென ஒரு முகநூல் கணக்குத் துவங்கப்பட்டு அதற்கு New Tamilmantram என்று பெயரிடப்பட்டது. மாணவர்களைக் கற்றலுக்கென சொந்தமாக முகநூல் கணக்குத் தொடங்குமாறு ஊக்குவித்து அவற்றை முறையே பெற்றோர்களிடமும் கடிதம்வழி தெரிவித்து அனுமதி பெற்றோம். ஒவ்வொரு வகுப்பிற்கென குழு உருவாக்கப்பட்டு மாணவர்கள் தத்தம் குழுவில் பாடம் தொடர்பாக ஏற்றம் செய்யப்படும் பலவகைப்பட்ட தகவல்களின் மூலம் மொழித்திறனை மேம்படுத்திக்கொள்ளும் வண்ணம் ஏற்பாடு செய்யப்பட்டது. முகநூலைப் பொருத்த அளவிற்கு பெரிய அளவில் சிக்கல்கள் இல்லாதபோதிலும் அதில் தட்டச்சு செய்வதற்குப் பயன்படுத்தப்படும் முரசு எழுத்துரு ஊள்ளீட்டுப் மென்பொருள் வழியாகச் சிக்கல் ஏற்பட்டது. 2012 வருடத்திலேயே தொடக்கநிலை 1 மற்றும் 2 மாணவர்களுக்குத் தட்டச்சுப் பயிற்சி வழங்கப்பட்டது. எனவே இணையத்தில் இலவசமாக T99 முறையில் தட்டச்சு செய்வதற்கான இணையப்பக்கங்களை மாணவர்களுக்கு அறிமுகப்படுத்தியதோடு அவற்றை எவ்வாறு பயன்படுத்த வேண்டும் என்றும் செய்துகாட்டப்பட்டது. தட்டச்சுப் பயிற்சி பெறுவதற்கு ஏதுவாக மாணவர்களுக்கான மின்னஞ்சல், கட்டுரைப் பயிற்சிகள் மின்னஞ்சல்வழி கொடுக்கப்பட்டு அவர்களைத் தட்டச்சு செய்து மின்னஞ்சல் வழியாகவே அனுப்புமாறு பயிற்சிகள் வழங்கினோம். இத்தகைய தயார்ப்படுத்துதல் எழுத்துப் பதிவினை உறுதிப்படுத்த அவசியமாகிறது. மேலும் ஒலிப்பதிவு செய்து கருத்துகளை வெளிப்படுத்த vocaroo எனப்படும் இணையப்பக்கம் சார்ந்த இலவச ஒலிப்பதிவு வசதி மாணவர்களுக்கு அறிமுகப்படுத்தப்பட்டுப் பயிற்சியளிக்கப்பட்டது. முகநூலிலேயே உள்ள ஒளிப்பதிவு வசதியைப் பயன்படுத்திக் கருத்துகளை ஒளிப்பதிவு செய்து ஏற்றம் செய்யவும் பயிற்சி அளிக்கப்பட்டது. இணையப் பக்கத்திலுள்ள கதைகள், கட்டுரைகள், you tube –இல் உள்ள காணொளிகள் எனப் பலவகைப்பட்ட தகவல்கள் மொழித்திறன் கற்பித்தலுக்குப் பயன்படுத்தப்பட்டன. இருப்பினும் முகநூலை மட்டுமே பயன்படுத்தி மாணவர்களை உற்சாகத்துடன் கற்றலில் ஈடுபடுத்துவதும் கற்பிப்பதும் சற்றுச் சிரமமாகவே இருந்தது. . அளவில் பெரிதாக உள்ள எழுத்துவழித் தகவல்களையும் ஒலி மற்றும் ஒளிவழித் தகவல்களையும் ஏற்றம் செய்வதற்கு மட்டுமின்றிக் கற்றல் கற்பித்தலை நிகழ்த்துவதற்கும் மதிப்பீடு செய்வதற்கும் உகந்ததாகத் திகழும் முகநூலைக் கைவிடுவது பொருத்தமன்று என்பது புரிந்தபோதிலும் மாணவர்களை உற்சாகத்துடன் உடனுக்குடன் கற்றலில் ஈடுபடுத்தி கற்றல் விளைவுகளை அடைவதை உறுதிசெய்ய மாற்றுவழியும் தேவைப்பட்டது. அப்போதுதான் முகநூலுடன் கைத்தொலைபேசியையும் இணைத்துக் கற்றலுக்குப் பயன்படுத்திப் பார்க்க வேண்டும் என்ற எண்ணம் தோன்றியது

2. 2013 – ஆம் ஆண்டு வாட்ஸ்ஆப் செயலியின் பன்முகத்தன்மை உடனடி கற்றலுக்கு மிகவும் உகந்ததாக இருந்ததோடு வேறுபட்ட செயல்தளங்களிலும் (Platform) செயல்படக்கூடியதாக இருந்ததால் அதைப் பயன்படுத்தத் தீர்மானித்து மாணவர்கள் பயன்படுத்தும் கைத்தொலைபேசிகளின் விபரங்கள் தொகுக்கப்பட்டன. ஐபோன் மற்றும் கைத்தொலைபேசிகளின் விசைப்பலகைகள் தமிழ் எழுத்துருக்களைத் தட்டச்சு செய்யுமாறு இல்லை. எனவே மாணவர்கள் தமிழில் தட்டச்சு செய்வதை உறுதிப்படுத்தும்வண்ணம் வாட்ஸ்ஆப் செயலியில் தமிழ் எழுத்துருக்கள் தட்டச்சு செய்யத் தேவையான தொடுதிரை மின்னியல் தமிழ் எழுத்துரு விசைப்பலகை உள்ளதா என்றும் ஆய்வு செய்யப்பட்டது. தற்போதுள்ள செயற்தளங்களில் அவ்வசதி ஏற்படுத்தப்படவில்லை. ஆனால், சீனம், மலாய், இந்தோனேஷியா ஏன் ஹிந்தி மொழியில்கூட தொடுதிரை மின்னியல் விசைப்பலகை அமைப்பை ஏற்படுத்திக்கொள்ளும்போது தமிழில் மட்டும் அமையாதது மிகவும்

வருத்தத்திற்குரியதாக உள்ளது. சமீபத்தில் கண்டறியப்பட்ட தகவலின்படி ஐபோன் செயற்தளம் 7-இல் தமிழ் தொடுதிரை விசைப்பலகை அமைப்புப் பயன்பாட்டுக்கு வருகிறது. ஆண்ட்ராய்டு செயற்தளம் பற்றி இன்னும் அறிய முயற்சி மேற்கொள்ளப்பட்டுள்ளது. அதே சமயம் விண்டோஸ் செயற்தளம் 8-இல் தமிழ் தொடுதிரை மின்னியல் எழுத்துரு பயன்பாடு சாத்தியமாகிறது என்று அறியப்பட்டது.

- மாணவர்கள் வேறுபட்ட இயங்குதளங்களைக் கொண்ட கைத்தொலைபேசிகளைப் பயன்படுத்துவதால் அவர்களை ஒருங்கிணைப்பதற்கு அவசியமான வேறுபட்ட எழுத்துரு உள்ளீட்டு மென்பொருள் பற்றியும் ஆய்வு செய்ய நேரிட்டது. ஐபோன் பயன்படுத்துபவர்கள் 'Tamil for iOS' என்ற மென்பொருள் செயலியையும் (app), ஆண்ட்ராய்டு செயற்தளம் கொண்ட கைத்தொலைபேசிகளைப் பயன்படுத்துபவர்களும் விண்டோஸ் செயற்தளக் கைத்தொலைபேசியைப் பயன்படுத்துபவர்கள் செல்லினம் எழுத்துரு செயலியையும் பயன்படுத்துவது பொருத்தம் என்பது கண்டுபிடிக்கப்பட்டது. அதே சமயம் ஆசிரியர் வேறுபட்ட செயற்தளம் கொண்ட கைத்தொலைபேசிகளைப் பயன்படுத்தி எழுத்துவழியிலான பயிற்சிகள் அளிப்பது சிரமம் எனக் கண்டறியப்பட்டது. எனவே ஆசிரியர் பயன்படுத்தும் கைத்தொலைபேசிக்குப் பொருத்தமான மென்பொருள் செயலிகளைப் பயன்படுத்தித் தகவல்களைத் தெரிவிப்பதோடு வேறுபட்ட செயற்தளம் கொண்ட கைத்தொலைபேசியைப் பயன்படுத்தும் மாணவர்களும் அறிந்துகொள்ளுமாறு அதே தகவல் திரை அச்சு (printscreen) முறையில் படமாகச் சேமிக்கப்பட்டு ஏற்றம் செய்யப்படும்.
- மாணவர்கள் முகநூலையும் வாட்ஸ்ஆப் செயலிலையும் இணைத்து வாய்மொழிப் பயிற்சியையும் மதிப்பீட்டையும் மேற்கொள்ளுமாறு பயிற்சி ஒன்று வடிவமைக்கப்பட்டு மாணவர்களுக்குத் தகவல்கள் வழங்கி செயல்விளக்கம் காட்டப்பட்டதோடு அவர்கள் அதை எவ்வாறு வீட்டிலிருந்தபடியே செய்துமுடிக்கத் தேவையான விவரப் பாய்வு நிரல்படமும் (Detailed flow chart) வழங்கப்பட்டது.



மாணவர்கள் முகநூலில் ஏற்றம் செய்யப்பட்டுள்ள காணொளியைக் கண்ணுற்று அதைப் பற்றி ஓரிரு சொற்களில் அதனுள்ளேயே கருத்தைத் தெரிவிப்பர். பின்பு அதையொட்டி முகநூலில் கொடுக்கப்பட்டுள்ள கேள்விகளுக்கான தங்கள் கருத்துகளை வாட்ஸ்ஆப் செயலியில் உள்ள ஒளிப்பதிவு வசதியைப் பயன்படுத்தி ஒளிப்பதிவு செய்து வாட்ஸ்ஆப் செயலியில் உள்ள தம் வகுப்பு மாணவர்கள் உள்ள குழுவிலேயே ஏற்றம் செய்வர். ஒவ்வொரு மாணவரின் படைப்பையும் மதிப்பீடு செய்ய வேண்டி பணி இணையாளராக ஒதுக்கப்பட்ட வேறு மாணவர் தன் இணையாளரின் படைப்பை மதிப்பீடு செய்து தன் கருத்தை ஒளிப்பதிவு

செய்து வாட்ஸ்ஆப் செயலியில் உள்ள தங்குமூவிலேயே ஏற்றம் செய்வார். ஆசிரியர் மாணவர்களின் படைப்பைப் பற்றிய தம் கருத்தை ஒளிப்பதிவு அல்லது ஒளிப்பதிவு செய்து ஏற்றம் செய்வார். அனைத்து மாணவர்களும் உணர்வோடும் சரியான உடல்மொழியோடும் பதில் அளிக்க வேண்டும் என்பதைக் கருத்திற்கொண்டு அனைவரும் கருத்துகளை ஒளிப்பதிவு செய்து ஏற்றம் செய்யுமாறு பணிக்கப்பட்டிருந்தார்கள். வாட்ஸ்ஆப் செயலி ஒளிப்பதிவு செய்ய 45 வினாடிகளுக்கு மட்டுமே அனுமதிப்பதால் விரிவாகக் கருத்துகளைக் கூறிய மாணவர்கள் தம் ஒளிப்பதிவை மின்னஞ்சல்வழியாக ஆசிரியருக்கு அனுப்பி வைத்தார்கள்.

மாணவர்கள் கற்றலில் பெற்ற வளர்ச்சி

- சுயவழிக் கற்றல், கூட்டாக இணைந்து கற்றல், வேறுபடுத்தப்பட்டமுறையில் கற்றல் போன்ற கற்றல் அணுகுமுறைகளைப் பற்றி அறிந்துகொள்ளுதல்
- தன்னுடைய கற்றல் முறையை அறிந்துகொண்டு அதற்கேற்ப வளங்களைத் தேர்ந்தெடுத்தல், படைத்தல்
- தகவல் தொழில்நுட்பத்தைப் பயன்படுத்தி மொழித் திறன் கற்றலை மேம்படுத்துதலுக்குத் தம்மைத் தயார்ப்படுத்துவதோடு விருப்பத்துடன் கற்றல்
- தகவல் தொழில் நுட்பத்தைப் பயன்படுத்தித் தமிழ் கற்பதில் ஏற்படும் நடை முறைச் சிக்கல்களால் மனம் தளராமல் மாற்று வழிகளைக் கண்டறிந்து கற்றலைத் தொடர முயலுதல்
- சுயமாகத் தகவல்களைப் பெற்றும் ஆராய்ந்தும் கூட்டாக இணைந்து கற்றலுக்கு வழிவகுத்தல்
- கற்றறிந்தவற்றைப் புதியமுறையில் படைக்க வேண்டும் எனத் தூண்டப்படுதல்
- குழுவில் ஒருவர் மற்றொருவரின் பரந்துபட்ட, வகுப்பறைக்கு அப்பாலும் நடைபெறும் கற்றலுக்கு நேரடியாகத் துணைபுரிதல்
- கற்றலுக்கான வளங்களும் கற்றலின் வெளிப்பாடும் தொடர்ந்து நிர்வகிக்கப்படுதல் மற்றும் கண்காணித்தல்
- கற்றலைச் சுய மதிப்பீடு செய்வதோடு பிறரின் கற்றலையும் மதிப்பீடு செய்தல்

நடைமுறைச் சிக்கலும் மேற்கொண்ட நடவடிக்கைகளும்

- மாணவர்கள் இணையத்தை முறையாகப் பயன்படுத்துவதற்குத் தயார்ப்படுத்துவதில் கவனச் சிதறலும் பாதுகாப்புச் சிக்கல்களும் கவனத்திற்கொள்ளப்பட்டு பெற்றோருக்கு முறையாகக் கடிதம் அனுப்பப்பட்டு அனுமதி பெறப்படல்
- அனைத்து வகுப்புகளுக்கும் ஒரே சமயத்தில் தகவல் தொழில்நுட்பப் பயன்பாட்டைத் தொடங்குவதால் ஆசிரியர்கள் போதுமான கவனம் செலுத்தமுடியாமல் போகும் என்பதைக் கருத்திற்கொண்டு ஒவ்வொரு வகுப்பாக மாணவர்களின் ஆர்வத்தைத் தூண்டியபிறகு விரிவுபடுத்துதல்
- மாணவர்கள், முரசு அஞ்சல் தமிழ் எழுத்துருவைப் பயன்படுத்துவதில் ஏற்பட்ட சிரமங்களைக் களையவும் அனைவரும் T99 விசைப்பலகை அமைப்பைப் பயன்படுத்தவும் வேண்டி இணையப் பக்கங்களில் காணப்படும் இலவச T99 விசைப்பலகைப் பயன்பாட்டைப் பற்றிய பயிலரலங்கு நடத்தப்படல் (Ex: w3.tamil.com)
- மாணவர்கள் ஒரே வகையான கைத்தொலைபேசியைப் பயன்படுத்தாததால் வெவ்வேறு வகை செயற்தளத்திற்குப் பொருத்தமான தமிழ் எழுத்துரு செயலி (Tamil Font apps) பற்றிய தகவல்களை ஆசிரியர்கள் சேகரித்து ஆய்ந்து அவற்றை இறக்கம் செய்து முறையாகப் பயன்படுத்தும் வண்ணமும் சரிவர இயங்காதபோது மாற்றுவழிகளில் தகவல் பரிமாற்றம் செய்யவும் அவர்களுக்குப் பயிற்சி அளிக்கப்படல்
- மாணவர்களைக் கொண்டு whatsapp குழு ஆரம்பிக்கப்படும்போது ஏதோ காரணத்தால் எவரேனும் விலகும்போது அவரை அடையாளங்கண்டு மீண்டும் குழுவில் சேர்ப்பது கடினம் என்பதால் ஆசிரியரே ஆரம்பத்தில் ஒவ்வொரு குழுவிலும் உள்ள மாணவர்களுக்குப் பெயரிட்டுக் குழுவை வழிநடத்துதல் (Ex: SEC 1 NT Priya)
- கருத்துகளை ஒளிப்பதிவாக ஏற்றம் செய்வதற்கு மாணவிகள் அதிகம் தயக்கம் காட்டும்போது ஆசிரியரே முன்மாதிரியாக இருந்து தம் கருத்தை ஒளிப்பதிவு செய்து ஏற்றம் செய்தல் வழி அவர்களைத் தூண்டுதல்
- மாணவர்கள் அடிக்கடி ஆங்கிலத்தில் தட்டச்சு செய்வதும் பாடத்திற்குத் தொடர்பில்லாமல் சொந்தக் கதைகளைப் பற்றிப் பேசுவதும் ஆரம்பத்தில் தொல்லையாக இருப்பினும் பொறுமையைக் கடைப்பிடித்துத் தொடர்ந்து அவர்கள் தமிழில் தட்டச்சு செய்யவும் உரையாடவும் வலியுறுத்தல்
- குழுவின் உறுதிப்பாட்டைக் காக்க வேண்டி மாணவர்கள் ஆங்கிலத்தில் வெளியிட்ட தகவல்களை எல்லாம் ஆசிரியர்களே நீக்குதல்

- மாணவர்கள் தகவல்கள் பரிமாறிக்கொள்ள காலம் அதிகம் எடுக்கும்போது ஆசிரியர்களே முன்வந்து அதனையொட்டித் தம் கருத்தை மற்றொரு மாணவர்போன்று படைத்து அவர்களை உற்சாகப்படுத்துதல்
- மாணவர்களின் படைப்பைப் பாராட்டி வாழ்த்துத் தெரிவித்தல்

எதிர்காலச் சிந்தனை

- இக்கற்றல் முறையை வட்டார அளவில் விரிவுபடுத்திப் பின்பு தேசிய அளவில் மற்ற பள்ளி ஆசிரியர்களோடு இணைந்து கற்றல் குழுமத்தை உருவாக்கி இம்முறைக் கற்றலை மேம்படுத்துதல்
- பரவலான அளவில் மாணவர்கள் தமிழைக் கைத்தொலைபேசிகளில் பயன்படுத்தும்போது இப்போது உள்ள தொழில்நுட்பக் குறைபாட்டை நீக்குவதற்கு நிறுவனங்களிடம் வேண்டுகோள் விடுத்து, மாணவர்கள் எளிதாகத் தமிழைப் பயன்படுத்த செயற்களம் மற்றும் செயலிகளின் வடிவமைப்பில் மாற்றம் கொண்டுவருதல்
- உலக அளவில் பிற நாடுகளில் உள்ள பள்ளி மாணவர்களுடன் இணைந்தும் கைத்தொலைபேசிவழிக் கற்றலை நிகழ்த்தத் தூண்டுதல்
- தமிழ் நம் மனத்தில் நிலைக்க நம் கைகளில் எங்கும் எப்போதும் தவழவைக்க வேண்டியது அவசியம் என்பதால் இம்முறைக் கற்றலைத் தொடர்ந்து செயற்படுத்துதல்

Let Us ‘Flip’ the Classroom!

Dr Seetha Lakshmi

Asian Languages & Cultures Academic Group, National Institute of Education, Singapore

Abstract:

Currently in the teaching and learning arena, a number of innovative pedagogical approaches are attracting the academics, practitioners and students at a higher education. This essay will provide more information the ‘Flipped’ classroom Approach in Tamil Teacher Education in Singapore. With new research books and pedagogical approaches on teaching and learning of Mother Tongue Language Education, the Teacher Educators find it difficult to cover their desired syllabus through lectures and tutorials. Hence, they are trying out the “Flipped” classroom approach to utilize more learning time, provide an introduction to their trainees and enable the sustained engagement in the real classrooms with pedagogical demonstration. In Tamil Language, this approach has been tested for a class and will be fully implemented in August 2013 semester for the Diploma in Education Year I class. This preliminary paper will share more about the “Flipped” classroom teaching, trainees’ inputs on this approach and the benefits in using this approach. This “Flipped” classroom approach is a joint initiative of Centre for e-Learning and Asian Languages and Cultures Academic Group of the National Institute of Education (NIE), Singapore.

Introduction

Singapore is a country which gives continuous and consistent importance to education and the development of human resources. Tamil, being one of the official languages among English, Chinese, Malay, has its own and significant place in the field of education. In fact, Tamil language outside Tamil Nadu has enhanced facilities to be developed and survives in good capacity. Here, the technology is well mixed and mingled with the Tamil teaching effects. Hence, the technology does play an important role in the Teacher Training. The Internet, Email, Web 2.0, Facebook (Seetha Lakshmi, 2011), Twitter, Video Conferencing, Wimba Live (Seetha Lakshmi & Majal Pratima, 2010), Photo story, Voice thread, Posterous, are some of the forms of technology we use in our Tamil education arena. Many of these technologies have already been practiced by our new Tamil teachers going into the service and can be given as examples through their presentations (Swarthy et al, 2012). In that sequence, we have now the “Flipped” classroom approach. This method exhibits the excellence of teaching in a disciplined way in order to improve the education of students. Let us see how the improvement of education has taken place in this approach.

In and out of the classroom

“Flipped” classroom refers to turn the traditional classroom on its head” (<http://www.knewton.com/flipped-classroom/>). Yes, it is like that. What should happen in a classroom would have happened earlier itself. Then a question will come as, what would happen in the classroom? The whole time available in the classroom will be just used for students to learn and interact with the teacher and the other students. It is teacher’s responsibility to ease the classroom environment for sustained interaction among the students which would lay the path for useful learning. This would eliminate the teacher’s worries like being afraid of leaving out anything in teaching and giving ample time to cover all with the stipulated hours of teaching time. This methodology also helps the teachers to bring forward to the students whatever they have newly learnt.

Fine, how does this classroom get established? Let us view some of the steps behind this process.

Planning

- It is about knowing what is meant by “Flipped” classroom and what are its characteristics
- How can we use this “Flipped” classroom methodology and to which subject would it suit the best?
- Interacting with the Video Content Developer of the Centre for e-Learning (CeL), NIE to choose the subject for “Flipped” classroom method.
- The lecturer sent in her formal request for the CeL to provide her with guidance and support in the “Flipped” classroom process along with the details of the Specific module and the proposed semester.
- Getting ready for the first recording after discussed with the Video Content Developer and got familiar with the “Flipped” classroom methodology

- On the day of shooting, the lecturer has to engage on discussion with two Video Content Developers on how to record the teaching session
- This has to be followed with both Video Content Developers, recording the session and copying it to harddisk before handover to the lecturer
- The lecturer, on seeing the video, can decide on where to split up the lesson into smaller parts for further editing. He / She would also plan for suitable intervals where they could have interactive sessions (MCQs) to get the understanding of students. The lecturer may also decide for the places in the video session to introduce additional teaching tools (e.g. Power Point slides, pictures, moving pictures, video and artifacts)
- Then the lecturer would have a discussion with the Video Content Developer to arrange for the next level of editing the video.
- The Video Content Developer would arrange the video into smaller parts based on the lecturer's inputs and send it to the lecturer for further review.
- The lecturer would check each and every part of the video session and if satisfied, gives signoff, for the final uploading of content.
- The Video Content Developer would then author the part of the lessons appropriately with the lesson name and the part name before uploading it to an internet page and ready for mobile devices.
- The student will go and access the internet page for viewing the lessons for learning and to learn about the contents of the lesson/lecture.
- And when the student goes into the classroom after seeing the video session, the class would start with the discussion on the lesson of that day. As the students have prior knowledge of the lesson, they are able to understand fully when the lessons are discussed in the classroom.
- Here the responsibility lies on the lecturer to prepare the lessons beforehand and keep them ready for shooting as a video session. The Video Content Developer provided his/her full cooperation by shooting the video session of teaching and later editing into appropriate parts with suitable intervals to prepare the final presentation. Hence this is a two way production for the development of an IT based pedagogy.

Apart from the above, the one and only thing which is required the foremost, is the interest in teaching and learning. In English it is termed as passion. In this "Flipped" classroom production, the Video Content Developer themselves are very much interested in working on the Tamil Pedagogy. They are also working with other subjects like English Language, Natural Science and etc. With their professional experience and interest, they have also given a number of tips to fine tune the presentation and pedagogy.

"Flipped" classroom is different from the traditional classroom. In the traditional classroom, the teaching, the related approaches, activities and evaluation would happen. But in the "Flipped" classroom, the students can access contents outside their classroom at their own pace and desired place, in their convenient timings or during travelling. They can view and listen to the video lessons in any of their favourite devices like the computer, laptop, Mac, iPod Touch, iPhone and iPad. For mobile devices - iPod Touch, iPhone and iPad, contents can be watched via NIE mVideo app (please refer to annexe). The URL is used directly without password. They are also advised to use wired internet connection or WiFi for faster playback as 3G is slower. This makes them equipped for the lessons beforehand and helps to deepen their understanding and know the objectives of the lessons clearly to be taken in the classroom on that day. Above all, their learning is not restricted to one way of communication. The students can ask questions and respond to the open questions in the classroom or the closed questions, objective questions, fill in the blanks, give their views type questionnaire to teacher either through computer devices or written in a paper to handover to the teacher in the classroom. It is their liberty to choose any of the media to respond to teacher's queries and could clarify their doubts with the teachers. This makes the learning as a responsible way of learning as the students get to understand their role in the teaching and learning process. The lecturer too has more control in monitoring the students' performance and initiates their interest whenever they sense the slowdown among them. The lecturer also has the responsibility of improving the curriculum delivery with sincere efforts and interest, in order to keep up the interests of learning of the students.

We may raise a question, of what the students will do in the "Flipped" classroom as they are already exposed to the lessons. Even we might have a doubt of classroom time of how it has been engaged. What could we do about that? When the students come to the classroom with the knowledge of the lessons, they might get engaged in deepening the understanding by means of further activities or teaching related approaches conducted by the teacher. The interaction between the students and the

teacher would broaden and deepen the knowledge about the lesson. In Tamil classrooms, the students get more opportunities to converse in Tamil. Hence they would improve in both the spoken Tamil and have confidence in speaking Tamil before anyone in the society. The skill of speaking well is one of the necessary livelihood skills of 21st century.

Teaching has the priority and more time

The teacher has more time in the “Flipped” classroom concept. In the world of knowledge, everyday new news has been delivered. The teacher who gets to read the news would like to convey it to his / her students. But in traditional way of teaching, their restricted hours would be taken up in the regular class sessions themselves. Whereas in “Flipped” classroom, they can increase their time in discussion for about two to three times and can utilize them for teaching the upcoming topics with more details and interest. They have the time to introduce new approaches and learning methodologies continuously. Thus, this new way of learning and knowledge gaining of the students expands their horizon of opportunities for greater learning and experience. The new methodology also helps the students to prepare for the future challenges.

Teachers use technology to flip their classrooms

The following are the advantages of using the “Flipped” classroom approach:

- The curriculum of the “Flipped” classroom is used to have more depth in learning.
- The learning for students becomes easy.
- It paves the way to know the lessons at home or get to know when the student is away from school and been to abroad places.
- It helps the students to continue the learning even when the teacher is not in class yet.
- It gets the ample time to Teachers to understand the problems and queries of students to give the suitable solutions.
- It leads the teachers to know about their students and enable their students to trust them for better learning.
- It allows the students to revisit their lessons from the recorded video sessions during their exams.
- It facilitates the teachers to take the role of non-detail teaching staffs and gives additional time for discussing the related things to the curriculum.
- It makes the student to take the ownership for their learning and initiates them to think about their subjects even when they are away from the classrooms.
- It helps the parents and others at home to know about the curriculum and the content being taught at the class.
- It gives the feeling as if the teacher ‘is’ present next to the student when they watch the “Flipped” classroom sessions.
- It becomes the boon to the students who have either parents who are not much educated or the busy, educated parents who might not have time to look into their curriculum or from a lower income group because it helps them to learn, revise the subjects until they understand and move forward, without losing their dignity.

Related sites on “Flipped” classroom

There have not been many academic research papers related to this topic though few of the websites are available on the internet. A school near Detroit city, in Michigan State, had followed the “Flipped” classroom methodology (<http://www.knewton.com/flipped-classroom>). For the fresh students who had joined the school, the subjects like English, Maths, Science, Social Studies were taught following this method. It was noted that the success rate of students had been increased from 20% to 40%. The improvement in the students’ attitude and the establishment of better relationship between the students and teachers had happened in that school after they started following the “Flipped” classroom method.

The classroom is for students. They are the characters of importance. But in the presence of time constraints, the teachers are completing their lessons fast and could not introduce the new thoughts since because they are not added in the existing curriculum. At this stage, the “Flipped” classroom concept will be helpful. For example, it will help to teach the Maths like subjects effectively as the students have their chances to revisit and practice their lessons again and again. There is a teacher by the name of Aaron Sams (<http://www.youtube.com/watch?v=AHYm7U0ePWY>), who teaches Biology, Botany, AP Chemistry subjects in a high school at Woodlands Park in Colorado. He says with pride, that his students are able to learn on their own due to this “Flipped” classroom method. As a teacher, his main goal is, "help students become learners who can learn for themselves and by themselves." (<http://www.techsmith.com/flipped-classroom-aaron-sams.html>). Aaron has been felicitated with the President award also, in the year 2009, for his best teaching service in the subjects Science and Maths. He

also mentions that his students have been elevated a level higher and there is almost nil number of students who score lower marks in his class. This explains significantly about the success rate guaranteed in the “Flipped” classroom method.

Another website (Tanya Roscoria, 2011) talks about the Clintondale High school where the low success rate and the poor behaviour of students lead them to change their approach to earn them positive outcomes. The Principal even explains how there were poor grades and stress in that website. So he initiated the new Camtasia Relay methodology to flip the entire school activities and achieved good results. “Redefining Learning through an Entire School Clintondale High School was going down a bad path. Passing rates were down, discipline rates were up, and the administration wasn’t sure what to do. Principal Greg Green knew drastic changes were needed to turn things around, so he “Flipped” his entire school with Camtasia Relay” (<http://www.techsmith.com/customer-stories-clintondale.html>)

Currently the “Flipped” classroom approach and related stories are highly shared by many educators (<http://www.knewton.com/flipped-classroom/>; <http://www.scoop.it/t/flipping-your-classroom>; <https://sites.google.com/site/flipperparadise/>). Ashley Tan has setup an exclusive site for “Flipped” classroom on the social media platform called Scoop.it (<http://www.scoop.it/t/flipping-your-classroom>). In that, he has curated topics on flipped classroom and also collected the useful articles of others to serve many details on “Flipped” classroom connecting with the pedagogy practices. It is significant that the site is really a good resource for “Flipped” classroom and centre for details on what the “Flipped” classroom is, what would be present in that method, what does a teacher can/ cannot do in a “Flipped” classroom and more useful guidance. Hence it can be named as one of the “Flipped” classroom related websites.

When the “Flipped” classroom is referred, a website (<http://www.knewton.com/flipped-classroom/>) says it as “A new method of teaching is turning the traditional classroom on its head”. Yes it is different from like the traditional class, but remains as a bridge to the traditional classroom. It has the traditional method of teaching also within itself. At the same time, it demonstrates to us as a new method of approach. The teaching happens in the internet. The interaction about the homework is happening in the classroom. This makes the understanding of the students deeper. It makes the lessons taught for many days in traditional class to take shorter days but at the same time gains more hours for teaching. The students need not have to come to additional classes for the incomplete lessons in this method. He or she could use the makeup class time and can understand the lessons which are placed in internet. This also helps to have understanding between the teachers and the students and the classroom remains clashes – free. As we have seen earlier, in the Detroit school class room, the lessons are prepared and placed in the internet thrice a week. The students get to access to these classes in the internet and uses the classroom for clearing the concepts which they don’t understand from the internet classes (<http://www.knewton.com/flipped-classroom/>). This helps the students to learn clearly and the lessons get by-hearted well. How can this happen? All because of the magical “Flipped” classroom methodology!

Trainees and their sharing

“Flipped” classroom is conducted here on experimental basis for the first and second year students of Diploma in Education course and the students of Post Graduate Diploma in Education course. Both of these groups just completed their courses in May 2013. Here, the lessons were delivered in internet few days prior to the actual classroom day. Also they were given three questions. The questions did not happen at the start, middle or the end but at specific intervals of the session. So, the students had to go through the classes in the session to give the answers to the questions.

The students also were come to know about the preparation for this methodology of teaching few months before itself. They knew it from their lecturers due to the preparations for the “Flipped” classroom. So naturally they were also kindled with interest to attend the “Flipped” classroom. The lecturer had prepared the sessions before handed and placed in internet. Once that was done, the internet addresses were given to the students. They also watched the sessions through iPhone, iPad, and Facebook. They could hear the video sessions while they were travelling or on the road even if they could not watch the session.

Though many trainees watched the sessions, they also faced difficulties initially. The difficulties include the issue with password access, the yearning of how to watch the session and the worries of giving the proper answers back etc. However, they were helped with Video Content Developer’s answers to overcome their initial fears to watch the sessions and respond back to the questions.

The Teacher trainees also had faced few difficulties but many of them were related to technical issues such as:

- Do not know how to get into the page
- Can't get the page even if the password is given
- Don't know anything!
- Can see in hand-phone but it is not coming in computer

Sooner we got their issues related to technical area, we have arranged to get the ways from the computer officials. Then the Teacher trainees let us know that they could see their lessons. They had helped each other and communicating the methods to watch the lessons. However, few students would like to watch the lessons together with their friends on the course day, though they had prepared their answers separately. That was appeared like a similar attitude used to be exhibited by the High school students.

They also attended the real time sessions for better understanding of the lessons being taught through internet and clarified through regular classroom sessions. They also gave their feedback on the teaching methodology. In the qualitative research conducted in this methodology, the students express the benefits of time saving and happiness to attend the classes. They also said the methodology is very useful as they can watch the sessions multiple times for clear knowledge wherein the repetition is not possible in regular classrooms due to time constraints. They also said the methodology helped for the content preparation and they also exuded a sense of freshness and interest for this type of classes. They could learn Tamil also in the internet sessions. They wondered if we could introduce the latest trends in our classroom and also were willing to take up the concept to their classrooms too. They also gained the confidence to prepare the video sessions for their students and conduct the "Flipped" classroom methodology.

After the "Flipped" class, we have asked for their feedback and the experience, they gave us their valuable thoughts and the difficulties they faced in the process.

Their positive thoughts include:

- Could focus our concentration and know the lessons
- Useful approach
- Opportunity to learn outside the class
- Opportunity to learn through hand-phone or computer
- Lessons are explained with good explanations and examples. Hence we could better understand the lessons. We can revisit always, when we don't understand any of the lessons.
- Since the lessons are divided in meaningful small parts, it gives the fulfilled learning feeling when finished the lesson.
- Could understand the existence of even more opportunities
- If the same lesson is conducted in the class, it will take many days. But in this method, we got to know more details.
- Gives the feeling that our lecturer talks to us directly
- Have to see and learn with care because there would be questions in between the lessons. We have to either write answers or prepare answers to submit in the class.

In this academic year, with the semester going to start in August 2013, there will be more opportunities to get the thoughts and feedback about the "Flipped" classroom methodology, when we are going to follow for the first year students of Diploma in Education and the second year Teacher trainees of Post Graduate Diploma in Education. We could share the experiences in the conference.

Who are going to get the most benefits?

Which is getting more importance in the "Flipped" classroom? Whether it is Students or Teachers Community? Definitely, the answer is students! If we ask pedagogy or technology, surely the pedagogy is the important aspect, because the better practice of pedagogy and teaching effects are more significant though the technology helps in the background. But we couldn't drop the value of technology too and might not ask the question if technology is not required; it would be a disaster, if we do not utilize the facilities available. The technology supports the good teacher standing behind to adopt the effective pedagogy concepts. So it is needed to have a good technical person in the educational department. He / she only can enhance the subject content several times, after the teachers hand over to them. Can we see anybody who would not take their eyes off

the screen and continuously watch for an hour? No. Hence we have to split the lesson into multiple parts so that the session would attract and sustain the students. This is possible only with the help of the technical person. In our case, our technical person showed interest and initiative to put the lessons online along with the inputs got from us. He has given good support to the team for the past one year. So with the encouragement, I could prepare the lessons for Tamil in a better way.

Pedagogy... Pedagogy... Effective and Engaged Pedagogy!

Professor Bertil Andersson (2013), President of The Nanyang Technological University (NTU) stated that the “Flipped” classroom pedagogy will be piloted soon and the infrastructure is getting ready to use this pedagogy. Professor Kam Chan Hin, Associate Provost in charge of Undergraduate education at NTU, told that, “Flipped” classroom teaching - an innovative way in which students attend lectures through online and use the class time to concentrate on the difficult work though problem solving, interaction or debates, is already being used in classrooms in small way. He also mentioned that this approach will be used across the university in the next few years (cf Sandra Devi, 2013). Special talks and courses are organized on this approach in July and August 2013 at the University. The use of prior knowledge, the step-by-step scaffolding process of bringing students towards learning, the sustained engagement, the group and pair activities, and the presence of authentic tasks where the students can use their language skills are all having the possibilities to get a place in the “Flipped” classroom. Also there are activities in between the session like answering the small questions, fill in the blanks, asking for answers to be sent through email, group presentations in classroom present in this method, so that the students would not get diverted from achieving their goals. Also, the activities are being validated to arrive at the measure of how the students have utilized the “Flipped” classroom. The teacher can archive the lessons prepared for any part of the lesson and can be used again or revisited for modifications. But it gives more time to bring in the new related matters to the students during the classroom sessions. Also the teacher gets time to interact with each of the students and explain the lessons.

Conclusion

Presently, some NIE staffs have introduced the “Flipped” classroom approach and initiative for the subjects like English Language, Maths, Natural Science, Early Childhood and Accountancy and it has been showing the progress. The way of doing a process step by step, explanation of research room concepts by depicting through pictures and content for real time understanding are some of the benefits we obtain in the “Flipped” classroom approach. It also yields good outcomes in Tamil education. It helps the students to easily practice spoken Tamil by means of hearing and listening to the contents and make way for effective understanding. It paves the route for students who come from non-Tamil speaking homes and continue to listen to their Teacher speaking Tamil in the “Flipped” classroom sessions. They can listen to it at any time and any number of times. It gives them interest to listen to Tamil and speak Tamil. It removes the inhibition of speaking Tamil and allows them to deliver their thoughts in Tamil better. The “Flipped” classroom approach gives them more training to handle the lessons to read, express their thoughts on their own and submit the suitable answers according to the questions asked.

Acknowledgements: The author wishes to express her sincere thanks to Dr Tan Min Chung Ashley (CeL, LST), Head, CeL, NIE, Mr Ngoh Kiew Meng, Henry, Video Content Developer (CeL/NIE), Ms Loh Pek Mee, Snr Video Content Developer (CeL/NIE) for their continued guidance, encouragement and support in using Information Technology in her pedagogical modules.

The author expresses her heartfelt thanks to the trainees of Diploma in Education (Aug 2011 Intake-Tamil/ALC/NIE) and Post Graduate Diploma in Education (August 2012 intake-Tamil/ALC/NIE) for their participation and inputs on the “Flipped” classroom lessons.

The author also would like to record her sincere acknowledgements to Dr Tan Min Chung Ashley (CeL, LST), Head, CeL, NIE, Mr Ngoh Kiew Meng, Henry, Video Content Developer (CeL/NIE) for their valuable inputs and advice for the development of this paper.

The author wishes to thank Sakthi and Shakthi for their kind assistance for the earlier drafts of the paper.

References:

- Andersson Bertil. (2013). From freshman to global player. ‘After 2011’s curriculum revamp, NTU now turns to transformation in teaching’. News Release. Published on 12-Apr-2013. <http://media.ntu.edu.sg/NewsReleases/Pages/newsdetail.aspx?news=faccaa48-0777-4794-ade0-a82b292356c8> website accessed on 10 July 2013.
- Sandra Devi. 2013. Want to teach at NTU? Wow students first. The Straits Times (Home B2, 13 April 2013)

Seetha Lakshmi & Majal Pratima. (2010). Introduction of Video Conferencing in Tamil Classroom: A Study. *International Journal of Educational Research, Development and Extension IJERDE*, 1 (1), 33-41,

Seetha Lakshmi (2011). Facebook and Tamil Language in Singapore's Teacher Education. In Vasu Renganathan (Ed.) *Conference Papers: Tamil Internet 2011* (pp. 76-91). Philadelphia, United States: INFITT & University of Pennsylvania.

Swathy Muralidharan, Ravathy d/o Ganasegaran, Arivinthan Ahwahday&Thiliban s/o Subramaniam. 2012. தமிழ்க் கல்விக்கான கருவிகள், உத்திகள், தொழில்நுட்பங்கள் (Equipments, Strategies and Technologies for the Tamil Education). Paper presented at the Diaspora Tamil Education Conference 2012, US: California Tamil Academy.

Tanya Roscorla. (2011). Another website Clintondale High Cuts Freshman Failure Rates with Flipped Classes. <http://www.centerdigtaled.com/classtech/Clintondale-High-Flipped-Classrooms.html> Website accessed on 8 July 2013

Websites:

<http://www.scoop.it/t/flipping-your-classroom> website accessed on 8 July 2013
<http://assets.techsmith.com/Docs/pdf-landingpages/flippedclassroom-why-flip.pdf> website accessed on 10 July 2013.

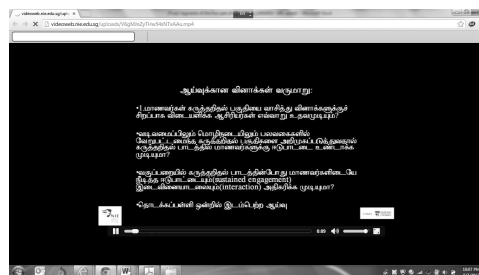
<http://www.knewton.com/flipped-classroom/> website accessed on 10 July 2013.

<http://www.techsmith.com/flipped-classroom-aaron-sams.html> website accessed on 8 July 2013.

http://www.centerdigtaled.com/news/Flipped-Institute.html?utm_source=related&utm_medium=direct&utm_campaign=Flipped-Institute website accessed on 10 July 2013.

Redefining Learning Through an Entire School. <http://www.techsmith.com/customer-stories-clintondale.html> website accessed on 8 July 2013.

Annexe: Sample Screenshots of Tamil “Flipped” classroom (1-2)



முகநூல் வழி தமிழ் கற்றல் கற்பித்தல்

ஆசிரியர் - மாணவர், மாணவர் - மாணவர், ஆசிரியர் - ஆசிரியர்

கி.விக்கினேசு

(சுல்தான் இட்ரிசு கல்வியல் பல்கலைக்கழக மாணவர்) செயின்மேரி தமிழ்ப்பள்ளி, பாரிட் புந்தார், பேரா, மலேசியா

தமிழ் கற்றல் கற்பித்தலின் புதிய அணுகுமுறை இது. அதிகமான பயனர் கொண்ட சமூக வலைத்தளம் (social network) முகநூல் ஆகும். பள்ளி நேரத்திற்குப் பிந்தியதும், 24 மணி நேர பள்ளியாகவும் இது செயல்படும். காலத்திற்கேற்றதும் கட்டாயமானதும் கூட.

முகநூல் என்பது மலேசியர்களின் வாழ்வில் இன்றியமையாத ஓர் இணைய சமூக வலைத்தளம் ஆகிவிட்டது எனலாம். அதிலும் குறிப்பாக 13 வயது முதல் 30 வயதுக்கு உட்பட்ட ஆண்களும் பெண்களும் முகநூல் பயனர்களாய் இருப்பது இயல்பு நிலையாகிவிட்டது. மலேசியச் சூழலில், இணையத் தொடர்பும் ஓர் கணினியும் ஒரு மின்னஞ்சல் முகவரியும் இருந்தாலே போதுமானது, எவரும் இம்முகநூலின் பயனர் ஆவதை தடுக்கவியலாது. எல்லைகளில்லா உலகம் இப்போது நம் கண்முன்னே!

சோசியல்பேகர்ஸ்.காம் எனும் வலைத்தளத்தின் கூற்றுப்படி, மலேசியா, உலக முகநூல் பயனர் வரிசையில் 18-வது இடத்தைப் பிடித்துள்ளது. மேலும் 13 085 000 மலேசியக் குடிமக்கள் (46.28%) முகநூல் பயனராக உள்ளனர் என்று அவ்வலைத்தளம் கூடுதல் தகவல் அளிக்கிறது. எனவே, இதில் கால்பதித்து கற்றல் கற்பித்தலில் ஒரு புதிய அணுகுமுறையைச் செயல்படுத்த இதுவே சரியான தருணம்.

தொடக்க நிலையில் முகநூல் வெறும் தொடர்பு கருவியாய் மட்டுமே செயல்பட்டது. நாளடைவில் இவ்வலைத்தளத்தில் அறிமுகப்படுத்தப்பட்ட பல்வேறு செயலிகள், இம்முகநூலின் பயன்பாட்டை வேறொரு கட்டமைப்புக்கு இட்டுச் சென்றது. உலகின் ஏதோ ஒரு மூலையில் நடக்கும் செய்திகளை விரைந்து, சுடச்சுட தருவதற்கான வசதியும் வாய்ப்பும் இதில் அமைந்துள்ளது. மேலும், அச்செய்தியைக் காணும் மற்றொரு பயனர், அது பற்றிக் கருத்திடவும், தனது முகநூல் நண்பர்களோடு பகிர்ந்து கொள்ளவும் ஏதுவாக இருக்கின்றது. இது இதுவரை இல்லாத ஒரு புதிய தொழில்நுட்பமாகக் கருதப்பட்டது.

இத்துணை ஆற்றல் மிக்க ஒன்றை கல்விக்கழகங்கள் முறையான வழியில் பயன்படுத்திக் கொள்ளுதல் நலமாகும். அதிலும் பள்ளிக்கூடங்கள் இம்முகநூல் பயன்பாட்டை தனது கற்றல் கற்பித்தலின் புதிய அணுகுமுறையாகக் கொள்ளுதல் சிறப்பை அளிக்கும் என்பதே இக்கட்டுரையின் முதன்நோக்கமாகும். பள்ளி நேரத்திற்குப் பிந்தியதும், 24 மணி நேர பள்ளியாகவும் இது செயல்படும்.

விக்கிப்பீடியா எனும் கலைக்களஞ்சிய வலைத்தளத்தின் கூற்றுப்படி, முகநூல் அதன் பயனர்களுக்கிடையே சமூகதொடர்பாடுதலை முன்னிறுத்தி, பயனரின் விருப்பங்கள், நடவடிக்கைகள் பற்றிய தகவல்களை உடனுக்குடன் பகிர்ந்து கொள்ளும் தன்மையில் செயல்படுகிறது. 2012 வாக்கில், முகநூல் நேரக்குறியீட்டினை (TIMELINE) அறிமுகப்படுத்தியுள்ளது. இந்நேரக்குறியீடானது பயனரின் தொடக்க காலம் முதல் தற்போதைய நடவடிக்கைகளைத் தன்னகத்தே குறித்துக் கொள்ளும் தன்மை வாய்ந்தது. பயனர்களும் குழுவும் அமைத்துக் கொண்டு செயல்பட முகநூலில் வசதியும் உண்டு. குழுவில், பயனர்களின் தகவல்கள் அந்தக் குழுமத்தினிடையே மட்டுமே பறிமாறிக்கொள்ளப்படும். இதனால் குழுமத்துக்கு வெளியே தகவல் கசிந்திட வாய்ப்பு இல்லை. எனவே, ஒரு பயனர் தம் கருத்தைத் தயக்கமின்றிச் சுதந்திரமாகப் பதிவு செய்யமுடிகிறது.

முகநூலின் பயன்மிக்க செயலிகள்

முகநூலைப் பயன்படுத்தும் முன்னர் அதன் முக்கிய செயலிகளை அறிந்து கொள்ள வேண்டும். அவற்றின் பயன்பாட்டை தெளிவாகப் புரிந்து கொள்வதன் வழி, முகநூலை அதிவேகமாகவும் அறிவார்ந்தும் பயன்படுத்திப் பலனடையலாம்.

கீழ்காணும் செயலிகளின் பட்டியல், பயனருக்குப் பயன் அளிப்பது ஆகும்:

முகநூலின் பயன்மிக்க செயலிகள்

Search	- தேடுதளம்
Tag	- மேற்கோள்
Dynamic content	- பல்வகை உள்ளடக்கம்
Link	- தொடுப்பு
Forum	- பொது மன்றம்
hyperlink	- மீத்தொடுப்பு
Authoring	- படைப்பாக்கம்
Emel	- மின்னஞ்சல்
digital gift	- இலக்கமுறை பரிசுகள்

இவற்றையும் தவிர்த்து முகநூல் மீத்தொடுப்பு வழியாகவும் பல செயலிகளை நாம் பயன்படுத்த இயலும். இவையாவும் பயனருக்குப் பயனுள்ளது, எளிமையானதும் கூட.

பள்ளி நிருவாகம் - முகநூல்

முகநூலில் ஒரு பள்ளி, குழுவும் அமைத்து அதன் தனது நிருவாக செயல்முறைக்கு ஏற்ப உருபடுத்திக்கொண்டு செயல்பட்டால், முக்கிய தகவல் பரிமாறிக் கொள்ள, படிவ ஆவணங்களை உடனடி பார்வைக்கு வழங்க; பகிர்ந்து கொள்ள, அரசு கட்டளைகளை விரைந்து நிறைவேற்றிட என பல்வகை வேலைகளை விரைந்து செயலாற்ற முகநூல் பேருதவியே. பள்ளி நேரம் தவிர்த்து இதர நேரங்களிலும் வேலை தடை இல்லாமல் தொடர்ந்திட இது வழிகோலும். அது மட்டுமல்லாது, நிருவாத்தின் செயல்முறை தொடர்பில் உடனடி கருத்து வழங்க ஆசிரியர்க்கும், மாற்று ஏற்பாடுகள் தேவையிருப்பின் செய்து கொள்ள நிருவாகத்திற்கும் சிறந்த வாய்ப்பு. நிருவாகத் திறனும் தரமும் இதனால் மேம்படும். படிவ ஆவணங்கள் மட்டும் அல்ல, பள்ளியில் நடைபெறும் நிகழ்ச்சிகளின் நிழற்படங்களும், காணொளியும் தரவேற்றும் வசதி முகநூலில் உண்டு.

முகநூல் கற்றல் கற்பித்தல் பாடப்பொருளுக்கான கோட்பாடுகள்

1. கற்றலும் அறிவும் பல்வகையானது.
2. ஏற்புடைய தகவல் மூலங்களுக்கு இடையிலான தொடர்புகள் குறித்த தெளிவு.
3. எல்லா தகவல் மூலங்களும் மனிதனை மட்டும் முன்னிருத்துவதில்லை.
4. ஒன்றை அறிந்து கொள்ளும் ஆர்வமானது அறிந்து கொள்ளுதல் என்பதை விட மேலானது.
5. பெறப்படும் தகவலால் தொடர்ச்சியான கற்றல் நடைபெறுகிறது.
6. துறை, சிந்தனை, கருத்துக்கோடல் ஆகியவற்றுக்கிடையிலான தொடர்பை அறியும் ஆற்றல்.
7. பெறப்படும் தகவல் சரியானது; புதியதானது என்பதை உறுதிபடுத்திக்கொள்ளும் திறன்.

மேற்சொன்ன ஏழும் முகநூல் பயனருக்குச் சரியான முடிவெடுக்கும் திறனை உண்டாக்கும்.

ஆசிரியர் - மாணவர்

முகநூல் வழி தமிழ் கற்றல் என்பது வழக்கநிலை கற்றலை விட மேம்பட்டது. முகநூலில் பதியப்படும் தகவல்களை அவ்வப்போது தேவைக்கேற்ப திருத்தம் செய்யவும் மாற்றம் செய்யவும் மேம்படுத்தவும் முடியும். முகநூல் வழி கற்றல் என்பது வகுப்பறைப் பாடத்தின் தொடர்ச்சி என்பதை நாம் புரிந்து கொள்ள வேண்டும்.

முகநூல் வழி தமிழ் கற்பித்தலில் ஆசிரியர் கடப்பாடு

1. வகுப்புகேற்பத் தனி கணக்கு வைத்திருத்தல்
2. ஆசிரியரின் தனிப்பட்ட வாழ்க்கைக்கும் தொழிலுக்கும் வேறுபாட்டை உணர்ந்திருத்தல்.
3. எந்த மாதிரியான படங்களைப் பகிரலாம் என்பதை அறிந்திருத்தல்.
4. எவ்வகை செய்திகளைப் பகிரலாம் (Netiquette) என்பது பற்றி மாணவரிடம் தெளிவாக அறிவுருத்துதல்.
5. மாணவரை முகநூல் கற்றல் கற்பித்தலுக்கும் உட்படுத்தும் முன் பெற்றோரின் அனுமதிப் பெற்றிருத்தல், அவ்வப்போது மாணவரின் வளர்ச்சி குறித்து பெற்றோரிடத்தில் தெரிவித்தல்.
6. எப்போதும் பொறுப்புணர்வோடு பதிவு செய்தல். பள்ளியைப் பற்றித் தூற்றுதல் கூடாது.
7. மாணவர்களை நண்பர்களாக இணைத்துக் கொள்ள வேண்டாம். அளவான இடைவெளி தேவை.

8. கூடுமானவரை ஒரு நாளைக்கு இரு முறையேனும் முகநூல் இணைப்பில் இருங்கள். இஃது உங்கள் வகுப்புக்குக் கூடுதல் மதிப்பளிக்கும்.

வகுப்பில் படித்த தமிழ்ப்பாடத்தின் சுருக்கத்தினை ஆசிரியர் முகநூல் குழுமத்தில் பகிரலாம். இதை அவரது முகநூல் சுவரில் (WALL) எழுதிப் பகிரலாம். அப்பாடம் தொடர்பாக எழும் வினாக்களை மாணவர்கள் அந்த ஆசிரியரின் முகநூல் சுவரிலேயே தொடுக்கலாம். இதற்கான பதில்மொழியை அந்த ஆசிரியரோ அல்லது அந்தக் குழுமத்தில் இருக்கும் வேறு மாணவரோ வழங்கலாம். இதன்வழி, உடனுக்குடன் பாடவிளக்கம் மாணவர்களைச் சென்று அடையும். வகுப்பு கலந்துரையாடலில் பங்கு பெறவியலாத மாணவரும், அச்சம் கூச்சத்தால் பங்கு பெறாத மாணவரும் முகநூலில் தைரியமாக கலந்து கொள்ள முடியும். வேண்டிய விளக்கம் பெறவும் நிறைய வாய்ப்பும் உண்டு. வகுப்பில் நேரப்பற்றாக்குறை காரணமாக எல்லா கேள்விகளுக்கும் விடையளிக்க இயலாத ஆசிரியரும் இந்த முகநூலில் மாணவரின் ஐயங்களுக்குத் தீர்வு காண முடியும். தவிர, ஆசிரியர்-மாணவர் கொண்ட இக்குழுவும், பாடம் சார்ந்த அறிவார்ந்த கலந்துரையாடலையும் இனிதே மேற்கொள்ளலாம். இவ்வாறாக ஒரு பாடத்தின் தொடர்ச்சியும் வளர்ச்சியும் ஆசிரியருக்கும் மாணவருக்கும் நன்மையளிக்கின்றன.

பாடச்சுருக்கம் மட்டுமல்லாது, பாடத் தொடர்பான குறிப்புகள், கேள்விகள், பயிற்சிகள், புதிர்கள், உடனடித் தேர்வுகள் என பல்வகை கூறுகளை இம்முகநூல் வழி ஆசிரியர் பகிரலாம். இதற்கான பல நுட்பமான, பயனருக்கு எளிமையான பல செயலிகளை முகநூல் தன்னகத்தே கொண்டுள்ளது. இணையோர் விரும்பும் வகையில் வண்ணப்படங்களாக, காணொளிகளாக, முப்பரிமாண படைப்பாக வழங்க முடியும். இது மாணவர்களுக்குப் பாடத்தின்பால் புது ஈர்ப்பை உண்டுபண்ணும்.

மேலும், மாணவரே பாடத் தொடர்பான இலக்கண இலக்கியக் குறிப்புகளை மேற்சொன்ன செயலிகளைக் கொண்டு தானே உருவாக்கிக் குழுமத்தில் பகிரலாம். அந்தக் குறிப்புகள் பற்றி பிற மாணவர்கள் கருத்திடலாம். மாற்றம் வேண்டியிருப்பின் குறிப்புகளைத் திருத்தவோ, மேம்படுத்தவோ முகநூலில் வழியிருப்பதால் முழுமையடைந்த பாடக்குறிப்பு மாணவரைச் சென்றடையும். இவ்வாறான பணிகளை மாணவர்கள் தனியாகவோ அல்லது குழுவாகவோ இணைந்து உருவாக்கலாம். இதன் மூலம் அவர்களின் தன்னம்பிக்கை வளரும். எந்நேரமும் ஆசிரியரைச் சார்ந்திராமல், தானே கற்கும் வழிமுறையை மாணவரே கண்டறிந்து சிறக்க முடியும். பாடமும் எந்த அளவு புரிந்து இருக்கின்றது என்பதை ஆசிரியரும் மாணவரும் மதிப்பீடு செய்து கொள்ளவும் முடியும் என்பதால் முகநூல்வழி கற்றல் ஒரு சிறந்த அணுகுமுறையே.

காட்டு 1:-

மாணவருக்கு இணைய வாயிலாக சிறு தேர்வு நடத்தப்படுகிறது. மாணவர் முகநூல் வழியாக பதில் மொழிய வேண்டும். ஒரே மாதிரியான பதில்கள் ஏற்கப்படமாட்டா. பதில்மொழி 100 சொற்களுக்கு மேற்போகாமல் இருத்தல் அவசியம்.

இவ்வாறான சிறு கட்டளைகள் மாணவரின் சிந்தனைத் திறனையும் கருத்துணர் ஆற்றலையும் ஓங்கச் செய்யும். (அப்துல் சமாட், 2011)

இப்படி தேர்வு மட்டுமல்லாது, விடை அளிக்கவியலாத கேள்விகளைக் கூட குழுமத்திடம் வினவலாம். சரியான விடை கிடைக்கப்பெறும் போது அதனை மற்ற மாணவர்களும் அறிந்து கொள்ள முடியும். இப்படியாக மாணவரின் கற்றல் செம்மையாக நடைபெறும்.

காட்டு 2 :-

இடுபணியும் அது தொடர்பான கலந்துரையாடலும் சிக்கலுக்குத் தீர்வாக அமையும்.

தகவல் / செய்தி தொடுத்தல்

முகநூல் வழியாக ஆசிரியர் அவ்வப்போது மாணவருக்குத் தகவல் தொடுத்தல் அவசியமாகிறது. பாடக்குறிப்பு தவிர்த்து, பாட வளர்ச்சி, தேர்வுக் குறிப்புகள், பாடத் தொடுப்புகள் என பல்வேறு செய்திகளை விடுக்கலாம். மாணவரோடு ஆசிரியர் நன்முறையில் அக்கறையோடு இருப்பதை இஃது உறுதி செய்யும்.

கருத்து சேகரித்தல்

பள்ளியில் நடத்தப்பட்ட தேர்வு குறித்தும் மாணவர்களின் கருத்துக் கேட்கலாம். கேள்வி முறைகள், நேரப் பயன்பாடு, தேர்வு தயார்நிலை குறித்தும் மாணவர் கருத்திட வைக்கலாம். பல்வேறான கருத்துக்கள் ஆசிரியரின் பணிக்கு வளமை சேர்க்கும்.

வாக்கெடுப்பு

இது தவிர்த்து, வாக்கெடுப்பு முறையும் முகநூலில் முக்கிய கூறு ஆகும். கேள்விகளையோ அல்லது கருத்துக்களையோ ஆசிரியர் தன் முகநூலில் Ask Question எனும் சுட்டியின் மூலம் எழுதி வாக்கெடுப்பு நடத்தலாம். இதன்வழி சொல்லப்படும் செய்திக்கு மாணவரின் ஆதரவும் எதிர்ப்பும் எவ்வளவு என்று கணித்திட ஆசிரியருக்கு எளிதாக இருக்கும். இக்கேள்வியும் கருத்தும் பாடத் தொடர்புள்ளதாய் மட்டும் அல்லாமல் பொதுவறிவு வடிவமாய்க் கூட இருக்கலாம்.

மாணவர் – மாணவர்

முகநூலில் ஆசிரியர் துணையின்றி மாணவர் கற்கவும் முடியும். வேறு மாணவர்களோடும் குழுவும் அமைத்துப் பயிலலாம். ஆனால் ஒரே வகுப்பு மாணவராக இருத்தல் சிறப்பு. காரணம் ஆசிரியர் வேறாக இருந்துவிட்டால் இக்குழுவும் சரியாக இயங்க முடியாது. பாடத் தொடர்ச்சியும் வளர்ச்சியும் சீராக அமையாது. எனினும் குழுவும் அமைத்த பின் மற்ற குழுமத்தின் உதவியை நாடுவதில் தவறில்லை.

பாடம் தொடர்பான ஒளிப்படங்களை, படத்தொகுப்புகளைக் குழுமத்தில் பகிரலாம். கூடுதல் தகவல்கள் ஒரு பாடத்தை நன்றாய் விளங்கிக் கொள்ள உதவும். சிறப்பாய் பாட விளக்கம் வழங்கும் நண்பருக்கு உடனடியாக பாராட்டு தெரிவிக்க வேண்டும். இது அவரை மேலும் உற்சாகப்படுத்தும். இன்னும் அதிக தகவல்களை அவர் பெற்றுத் தர முனைவார்.

ஆசிரியரிடம் கேட்க தயங்கிய வினாக்களை இந்தக் குழுமத்தில் விடுக்கலாம். நட்பு என்பதால் இங்கே தயக்கத்துக்கு விடுப்பு இருக்கும். ஆசிரியர் சொல்லி விளங்காதது பல நேரங்களில் நட்பு சொல்லி எளிதாய் விளங்குவது இயல்பே.

வகுப்புக்கு வெளியே என்பதாலும் தனக்கு விருப்பமான, நாள்தோறும் பயன்படுத்தும் முகநூல் என்பதாலும் கற்றல் இங்கே விருப்பமான ஒன்றாக மாறுகிறது. நினைத்த நேரத்தில் படிக்கலாம் என்பதும், பிடித்த நேரத்தில் கற்கலாம் என்பதும் மாணவரைக் கவர்ந்ததே.

ஆசிரியர்- ஆசிரியர்

இன்று கல்வி இணையமயமாகிவிட்ட நிலையில், பெருவாரியான தகவல்கள் இணைய தரவேற்றம் காண்கின்றன. எனவே ஆசிரியரும் இணைய அலையில் இணைவது காலத்தின் கட்டாயம்.

பாடத் தொடர்புடைய குறிப்புகளைச் செய்திகளைக் கருத்துகளைப் பரிமாறிக் கொள்ள முகநூல் ஓர் அரிய கருவியே. மொழி, இனம், ஊர், நாடு கடந்து ஆசிரியர்கள் ஒரே குடையின் கீழ் சந்திக்க முகநூல் ஓர் அருமையான தளம். மேலும், கல்வி தொடர்பான செய்திகளையும் அதன் வளர்ச்சிதனையும் தொடர்ந்து கண்காணிக்க முகநூல் ஓர் ஆவணமாகவும் செயல்படுகிறது.

கற்றல் கற்பித்தலின் சிக்கல்களை ஆய்ந்திட முகநூல் வாய்ப்பளிக்கிறது. அவற்றைக் களையும் முறை பற்றி அறிய குழுமப் பயனர்களைக் கேட்டாலே 100 மேற்பட்ட தகவல்கள் கிட்டும். அவற்றில் சிறந்தவற்றை ஆசிரியர் பயன்படுத்தவும் முயலலாம். தானே ஆய்வு செய்து, பயன் அறிந்து பயன்படுத்த நேரமும் காலமும் போதாத நிலையில் இவ்வாறான செயலிகளின் மூலம் ஆசிரியர் எளிதில் பயனடையலாம்.

மலேசியக் கல்வி சூழலில், நிறைய ஆவணங்கள் கற்றல் கற்பித்தலுக்குத் தேவைப்படுகின்றன. அவற்றைப் பெறுவதில் பல சிக்கல்களை பல ஆசிரியர்கள் எதிர்நோக்குகின்றனர். பாடத்திட்டத்தின் ஆவணங்கள், மதிப்பீடுப் படிவங்கள், தமிழ் இலக்கியம் இப்போது அதிகம் முகநூலில் பகிரப்படுகின்றன. எனவே முகநூலைப் பயன்படுத்தும் ஆசிரியர்கள், இதை எளிதாய்ப் பெற்றுக்கொள்ள முடிகிறது. இதுகாறும் முகநூலில் அதிகமான ஆசிரியர் குழுமங்கள் செயல்படுகின்றன. அதில் ஏதாவது ஒன்றிரண்டில் தம்மை இணைத்துக் கொள்ளும் ஆசிரியர் கட்டாயமாக பயன் அடைகிறார் என்பது உறுதி.

இவை தவிர்த்து, ஓர் ஆசிரியருக்குச் சமுதாயக் கடப்பாடும் உண்டு. தான் எதிர்பார்க்கும் சமுதாயத்தை உருவாக்குவதில் ஆசிரியர் பெரிய பொறுப்பு வகிக்கிறார். அவ்வகையில் சமுதாயம் தொடர்பான கல்விச் சிந்தனையோடு கலந்துரையாடவும் குழுவும் அமைத்துச் செயல்படலாம். நல்விளைவு கொண்ட கருத்துக்களால் விழிப்புணர்வு மேலோங்க இது வழிகோலும். அவ்வாறு செயல்படும் குழுமங்களும் இருக்கவே செய்கின்றன. அவற்றிலும் தம்மை இணைத்துக் கொண்டு சமூகப் பணியாற்றலாம்.

நிறைவுக்கு முன்...

தமிழ் கற்றல் கற்பித்தலில் முகநூல் பற்றி இதுவரை கண்டோம். கல்வி சார்ந்து இருக்கும் எவரும் முகநூலைக் கற்றல் கற்பித்தலுக்குப் பயன்படுத்தலாம். எனினும் முறையாக, சமூகப் பொறுப்புணர்வுடன் செயல்படுவது மிகமிக கட்டாயம். அதே வேளை ஒரு புதுமையான அணுகுமுறையைக் கையாளுவதில் சுணக்கம் காட்டுவதும் தீமை விளைவிக்கும். கல்வி இன்று அதிவேகத்தில் பயணித்துக் கொண்டிருக்கிறது. அந்த வேகத்துக்கு ஈடுகொடுத்துப் பயணிப்பது காலத்தின் தேவை என்பதை நாம் உணர வேண்டும். சுருங்கச் சொன்னால்,

செய்தக்க அல்ல செயக்கெடும் செய்தக்க

செய்யாமை யானும் கெடும்.

முகநூலுக்கும் பொருந்தும் இந்தக் குறள்.

துணை நூல்கள்

1. Penggunaan Facebook Sebagai Alat Bantu Dalam Pengajaran dan Pembelajaran di FTMK, Dr Abdul Samad bin Shibghatullah, 2011
2. Facebook Statistics by Country. URL: www.SocialBakers.com
3. Facebook. URL: <http://ms.wikipedia.org/wiki/Facebook>
4. <http://havefunteaching.com/>
5. <http://www.educatorstechnology.com/2011/10/best-facebook-learning-apps.html>
6. <http://www.educatorstechnology.com/2012/06/ultimate-guide-to-use-of-facebook-in.html>
7. <http://www.fractuslearning.com/2013/01/30/facebook-for-education/>
8. <http://www.fretech4teachers.com/2010/09/using-facebook-to-connect-with-students.html#.UUNZVlezyqF>
9. <http://www.onlinedegree.net/50-useful-facebook-tips-for-teachers/>
10. <http://www.slideshare.net/CPappasOnline/facebook-for-elearning-20-educational-applications>
11. <http://www.teachthought.com/technology/100-ways-to-use-facebook-in-education-by-category/>
12. <https://www.facebook.com/thiruvavur.thiru/notes>

பள்ளிகளுக்கான இணையவழிக்கல்வி முறையும் தமிழ் அகராதி உருவாக்கமும்

S.Maniam, New Wave Consulting Pte Ltd, Singapore, smaniam@starhub.net.sg

கட்டுரைச்சுருக்கம் :

இன்றைய சூழ்நிலையில் வெளிநாட்டு சூழ்நிலையில் தமிழ் படிக்கும் மாணவ மாணவியர்களுக்கு தமிழினை எளிதில் புரிந்துகொள்ளக்கூடிய நிலையும் அதை இணையத்தில் ஆங்காங்கே தேடி திரட்டி வரும் ஏற்பட்டுள்ளது.

அந்த சூழ்நிலையினை போக்க தமிழ் அகராதிக்கு என்று தனியாக ஒரு அகராதியை கருவியை உருவாக்கவேண்டிய கட்டாயம் ஏற்பட்டுள்ளது. மேலும் புலம்பெயர்ந்து தமிழ் படிக்கும் மாணவமாணவியர்களுக்கு முழுமையான கல்வி கற்பித்தலை பிரதானமாக கொண்டு இணைய வழியாக படிக்கவேண்டிய அவசியமும் நம் சமுதாயத்திற்கு எழுந்துள்ளது. எல்லா நாட்டு கல்விமுறைகளையும் தமிழ்நாட்டினை அடிப்படையாக வைத்தும் தமிழ் பண்பாட்டினை அடிப்பவையாக வைத்தும் எழுதவேண்டியுள்ளது. ஆனால் அந்த பாடத்திட்டத்தில் முழுமையான தகவல்களை கூற வியலாத நிலை ஏற்பட்டுள்ளது. எனவே குறையினை போக்க ஒரு மேம்படுத்தப்பட்ட இணைய வழிக்கல்வி முறை அவசியமாகிறது. அதை எப்படி செய்யலாம் என்பதே இக்கட்டுரையின் நோக்கம்

கல்வி என்றால் என்ன ?

- அறிவை விரிவாக்குதல்
- நேரடித் தகவல்களை சேகரித்தல்
- தொலைநோக்கு முறையிலான கல்வி
- சமுதாயத்தின் வளர்ச்சி சார்ந்த கல்வி
- தனிப்பட்ட வளர்ச்சி சார்ந்த கல்வி

ஆனால் இன்றைய காலகட்டத்தில் யாரும் இதுபோன்ற கல்விமுறையை சொல்லித்தருவதில்லை.

தொழில்நுட்பம் வழியாக நுட்பக் கல்வி கற்பிக்கும் முறை

தற்போது பள்ளி மாணவர்களுக்கு டிஜிட்டல் முறையில் பிரஜக்டர்களின் துணைகொண்டு செய்யப்பட்டு மாணவர்களுக்கு வகுப்புகள் எடுக்கப்பட்டு வருகின்றன என்றாலும் அது வகுப்பறையில் மட்டுமே உண்டு. மேலும் மாணவர்கள் விருப்பப்பட்ட நேரத்தில் படிக்க தற்போதைய முறைகள் உதவுவதில்லை எனவே மாணவர்களுக்கு கையில் எடுத்துச்செல்லவேண்டிய கையடக்க கணினிகள் அவசியமாகிறது.

தற்போதைய பாடப்புத்தகங்கள் உரை/காணொளி வழியாக இருந்தாலும் தனிப்பட்ட முறையில் மாணவர்களின் திறன் மேம்படுத்துதல் அவசியமாகிறது.

ஐபேடு மற்றும் ஆன்டிராய்டு அடிப்படையிலான டேப்லட் பிசிக்கள் எனப்படும் கையடக்க கணினிகள் பரவலான பயன்பாட்டில் உள்ளன. மேலும் தற்கால மாணவர்கள் நவீன நுட்பங்களான ஐபாடு மற்றும் இதர நுட்பங்களில் தங்கள் நேரத்தினை முழுமையாக செலவழிக்கின்றனர். அப்படியிருக்க அதே ஐபோன் மற்றும் டேப்ளேட்களில் நம் கல்விமுறையை மேம்படுத்தப்பட்ட நுட்பக்கல்வியை வழங்கவேண்டும்

ஒவ்வொரு பாடத்திட்டத்திற்கும் காணொளிகளையும் அதற்கேற்றார்ப்போல் மாணவ/மாணவியர்களின் கல்வித்திறனை அறியும் மென்பொருட்களை உருவாக்கவேண்டும். மேலும் சாணக்கியர் தன் அர்த்தசாஸ்திரத்தில் கூறியதுபோல் நடைமுறை வாழ்க்கையில் ஏற்படும் சிக்கல்களை சமாளிக்க உதவுவதே சிறந்த கல்விமுறை. ஆனால் இன்றைய கல்வி முறை அப்படிப்பட்ட கல்வியைதான் போதிக்கின்றதா என்பது கேள்விக்குறி?

எப்படி கற்பிக்கலாம்

காமிக்ஸ் புத்தகங்கள் மாயாவி, கமுகுவீரன், முக்கிய அறிஞர்களின் வரலாறுகளை காமிக்ஸ் வடிவில் கொடுத்துவருவது போன்று டிஜிட்டல் காமிக்ஸ் புத்தகங்களை உருவாக்கி மாணவர்களுக்கு தரலாம்

கற்பனை பாடங்கள்

கற்பனை பாடங்கள் என்பது மாணவர்களின் கற்பனை திறனை மேம்படுத்தும் ஒரு சாதனமாக பயன்படுத்தலாம்

உதாரணத்திற்கு மரம் தன் வாழ்க்கை மற்றும் அதன் பயனை கூறுவது போன்று, ஆறு தன் வாழ்க்கை மற்றும் அதன் பயனை கூறுவது போன்று முழுமையாக கற்பனையாகவும் ஆனாலும் காணொளி மற்றும் குரல் வடிவில் மாணவர்கள் கற்கும்போது மாணவர்களுக்கு மிகச்சிறப்பாக புரியும்

இதுபோன்ற முறைகள் மூலம் நாம் நிறைய விஷயங்களை மாணவர்கள் முழுமையாக கற்கும்படி செய்யலாம்.



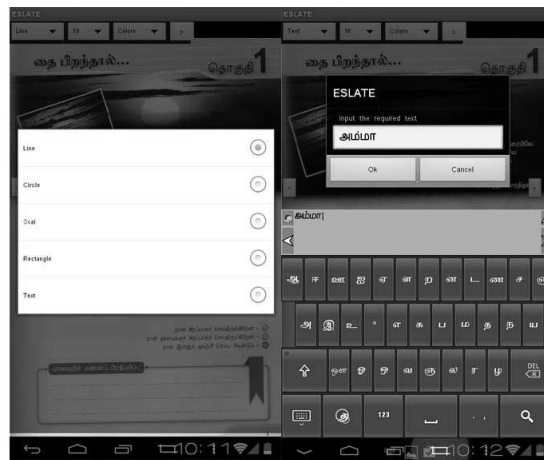
இன்றைய 21ம் நூற்றாண்டு தலைமுறைக்கு ஆங்கில அறிவு என்பது வெகு எளிதாக கிட்டிவிடுகிறது. ஆனால் தமிழ் அறிவு என்பது சற்று கடினமாக உள்ளது. இதற்கான காரணம் எல்லாரும் ஆங்கிலத்தினை நோக்கிவிடுவதால். அப்படியெனில் தமிழை கற்பிக்க என்னதான் வழி ?

ஆம் தமிழை கற்பிக்க இதே தகவல் தொழில்நுட்பத்தினைக்கொண்டே தமிழை போதிக்கலாம். இதற்காக உருவாக்கப்பட்டதுதான் பயில்தமிழ் என்ற இணையத்தளம். இந்த இணையத்தளத்தில் பாடத்திட்டம் முதல் பொது அறிவு வரை எல்லா வகையான தகவல்களையும் தெரிந்துகொள்ளும் அளவிற்கு தகவல் களஞ்சியமாக உருவாக்கப்பட்டு வருகிறது.

மேலும் மாணவர்கள் பயன்படுத்தும் ஐபாடு மற்றும் டேப்ளேட் கணினிகளில் தமிழ் மொழி சார்ந்த காமிக்ஸ் புத்தகம், பாரம்பரிய விளையாட்டுக்கள், தமிழர் பண்பாடுகள், தமிழ் இலக்கணங்களை எளிமையாக புரியவைக்கும் காணொளி, தமிழ் - மற்ற மொழிகளுக்காக அகராதிகள் என மென்பொருட்கள் மாணவ/மாணவியர்களின் கையடக்க கணினிகளில் நிறுவி அதன் வழியாக அவர்களின் கற்கும் திறனை மேம்படுத்தவும் பயில்தமிழ் உதவுகிறது.

சிறப்பம்சம்

டிஜிட்டல் பாடப்புத்தகங்கள்



தற்போதைய நம்முடைய பாடத்திட்டங்கள் முழுமையாக வெறும் எழுத்து வடிவிலான பாடத்திட்டங்களாகவே உள்ளன. இந்தப்பாடத்திட்டன் வழியேயே எதை படிக்கின்றார்களோ அதை அறியும் மேலும் அறிகின்ற வகையில் டிஜிட்டல் முறையில் கற்கும் போது மாணவர்களின் அறிவுத்திறன் மேலும் மேம்படும்.

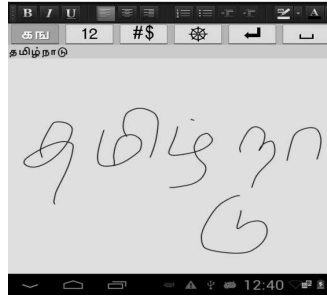
உதாரணம்

பொங்கல் பற்றிய பாடம் அதன் கொண்டாட்டங்கள் பற்றி எழுத்து வடிவில் இருக்கும். அதையே பொங்கல் பற்றிய முழுமையான மூன்று நாட்கள் பண்டிகையின் போது என்னென்ன நடக்கும் என்பதினை காணொளி வழியே காட்டும்போது மாணவர்களுக்கு பொங்கலின் முழுமையான சாரம்சம் தெரிந்துவிடும்.

டிஜிட்டல் முறையில் எழுதுதல்

இன்றைய சூழ்நிலையில் கணினியில் தமிழில் எழுதுவது என்பது எட்டாக்கனியாக உள்ளது. ஆனால் சிபேடு (www.cpadindia.com) என்ற கையடக்க கணினி வழியாக எல்லாமே சாத்தியம். தமிழில் எழுத்துவதற்கு அண்ணா பல்கலைக்கழகத்தின் பேராசிரியர் முனைவர். கிருஷ்ணமூர்த்தி அவர்கள் பொன்மடல் என்ற ஐசிஆர் - இன்டெலிஜெண்ட் கேரக்டர் ரெகனைசேசன் என்ற மென்பொருளை உருவாக்கியுள்ளார். இதன் வழியாக சாதாரண புத்தகங்களில் செய்யப்பட்டும், அடிக் கோடிதல், முக்கியமானவற்றை சேமித்தல் மற்றும் மாணவர்கள் தங்கள் பாடங்களில் வீட்டுப்பாடங்களை திரையில் எழுதலாம். சேமிக்கலாம். இந்த மென்பொருள் பள்ளி மாணவர்களுக்கு மிகவும் பயன்படும் ஒரு மென்பொருள்.

படம் -



இதன் மூலம் மாணவர்களின் கையெழுத்தினை ஒப்பிட்டுப் பார்க்கவும் வசதிகள் ஏற்படுத்தப்பட்டுள்ளது. வருங்காலங்களில் கையடக்க கணினிகள் வழியாகவே தேர்வுகளையும் எழுதலாம்.

மின்னணு தமிழ் அகராதி

Bring the WORLD in your hands with our CPAD

தமிழ் கருத்துப்பதிவுகளை.....

CPAD

Our CPAD Powered By

Android Jellybeans 4.2

10.1 inch

13246600

10.1 inch

Core i3

A9

Quad Core

1GB

RAM

30MB

Wi-Fi

HDMI

Full HD

1080p

8 GB

32 GB EXT

TWO

mini USB

Ports

Connect

Your

Projector &

Monitor

1 Year

Warranty

Audio out / Mic in, maver32 SDC, 100 MB/s LAN Port, 3rd Booting Allowed External MS, Bluetooth, GPS, Mouse, Keyboard and Connect

தமிழ் பழமும் மொழிபெருள்

தமிழ் மின்னணு அகராதி

மொழி அகராதி

குறந்தகளுக்கான கல்வி

மொழிபெருள்

அனுபவக்கல்வி மொழிபெருள்

தமிழ்நாடு

Online Reader for Tamil Children

www.ariyathalvi.com

வாழ்க்கை குழந்தைகளை கல்வித்திறமை

நாங்கள் செய்கிறோம்

அனுபவக்கல்வி

www.anubavakkalvi.com

Redefined Education

மின்னணு தமிழ் அகராதி என்பது தற்போதைய காலத்திற்கு மிக முக்கியமான தேவை.
தற்போது எங்கள் நிறுவனம் தமிழ் – தமிழ் | ஆங்கிலம் தமிழ் | தமிழ் ஆங்கிலம்
ஆகியவற்றுக்கான கையடக்க மின்னணு தமிழ் அகராதியை உருவாக்கிவருகிறது. இதில் இந்தியா,
சிங்கப்பூர். மலேசியா போன்ற நாடுகளில் வார்த்தைகள் அடங்கிய மென்பொருள்.
இதில் 27000க்கும் மேலான சொற்கள் உள்ளது

ல, ந, ர தேடும் வசதி

தமிழில் இதுவரை யாரும் தந்திராத ல, ந, ரா என்ற தேடும் வசதி வார்த்தைகளில் எங்கே ல, ந, ர
போன்றவை வரவேண்டும் என்ற வழிகாட்டும்
நம்மில் பலரும் எங்கே சின்ன ர வரவேண்டும், எங்கே பெரிய ற வரவேண்டும் என்ற சந்தேகம்
எப்போதும் உண்டு அதற்கு வழிகாட்டும் வகையில் இந்த மென்பொருள் உதவும்
பழமொழி + தமிழில் உள்ள பழமொழிகளை அதற்கான விளக்கங்களோடு காட்டும்வசதி
மரபுத்தொடர் மற்றும் உவமைகளுடன் உள்ள தொடர்கள் .
மேற்கண்ட இந்த தமிழ் மென்பொருள் உலகெங்கும் உள்ள தமிழ் ஆசிரியர்களுக்கும்,
மாணவர்களுக்கும் மிக அதிகமாக பயன்படும். மாணவர்களின் கற்கும் திறனை மேம்படுத்துவதே
எங்களின் முக்கியநோக்கமாக இருப்பதால் சாதாரண கற்கும் முறையில் உள்ள குறைகளை தவிர்த்து
பல்வேறு ஆராய்ச்சிகளுடன் இந்த மென்பொருள்கள் மாணவர்களுக்காக உருவாக்கப்பட்டுள்ளன.

மாணவர்களின் பேச்சுத்திறன்

மாணவர்களின் கல்விக்கற்கும் திறனுடன் அவர்களின் பேச்சுத்திறனை மேம்படுத்துவதும் மிக முக்கியமான தேவை.
எனவே அவர்கள் எப்படி தமிழை உச்சரிக்கவேண்டும் என்பது பற்றிய மென்பொருளையும் எங்கள் நிறுவனம்
உருவாக்கிவருகிறது. இதன் மூலம் தமிழை சரியாக உச்சரிப்பதோடு மாணவர்களின் கேட்டு அதையே திரும்ப பேசி
அதை மாணவர்கள் கேட்டு கேட்டு திருத்திகொள்ளும் அடிப்படையில் இம்மென்பொருள்கள்
உருவாக்கப்பட்டுள்ளன.

எனவே நவீன நுட்பத்தின் வழிகாட்டுதல்களுடன் 21ம் நூற்றாண்டில் இளையர்களுக்கு தமிழை கொண்டுசெல்வதே
இக்கட்டுரை மற்றும் நம்மின் நோக்கம்

மூல ஆய்வுக்கட்டுரை:

செல்வ.முரளி - கையடக்க கணினி வழியாக நுட்பக்கல்வி கற்பித்தல் – உலகத்தமிழ் ஆசிரியர்கள் மாநாடு 2013 –
மலேசியா

Computer Assisted Tamil Teaching and Learning

Louis Isack Kumar, Education Officer, Ministry of Education, Singapore

Objective (s)

- The objective of this paper is to share on how to create an interactive forum and to be self-directed and collaborative learners through the use of wiki space - PB works (Bartley Tamil Arangam webpage).
- To observe Tamil Language students' engagement and enthusiasm in the learning and to engage Tamil language students with wiki space-PB works for Self Directed Learning and Collaborative Learning.
- To bring classroom into the 21st Century by encouraging critical thinking and real world communication, with an online workspace, where students are engaged outside the classroom.

Background:

- ✍ Bartley Secondary School Tamil Language students learn to collaborate with their peers using wiki space-PB works to generate their ideas in a group and create an interactive forum.
- ✍ This brings the classroom into the 21st Century by encouraging critical thinking and real world communication, with an online workspace, where students are engaged outside the classroom.

The learning activities are as follows:

To leverage on the use of ICT, Bartley Tamil Arangam webpage encompasses the interactive forum is created to address the problems faced by both self and peers.

The Bartley Tamil Arangam webpage provides a pool of phrases for pupils to select and use in their comments. Pupils are engaged through this interactive platform through collaborative learning with their respective groups.

The collaborative portal (<http://bssmttl.pbworks.com>) is a platform where pupils can share their essays, email letters online and their teacher or classmates can give feedback, comments or edit the essay to help them become better writers.



(Home page – www.bssmttl.pbwork.com)

Key Learning point

We found that collaborative learning provides opportunities for group of students to work together to search meanings or solutions. Interactive tasks can be completed faster because pupils do not have to wait for their peers to respond by asynchronous means like email. Furthermore, it allows pupils to access from any place with an internet connection. Pupils can benefit from the inclusion of perspectives from the language teacher and pupils.

Approach/Method/Design

The PB works work space (<http://www.bssmttl.pbworks.com/>) is packed with relevant resource such as word bank, question banks, compositions.

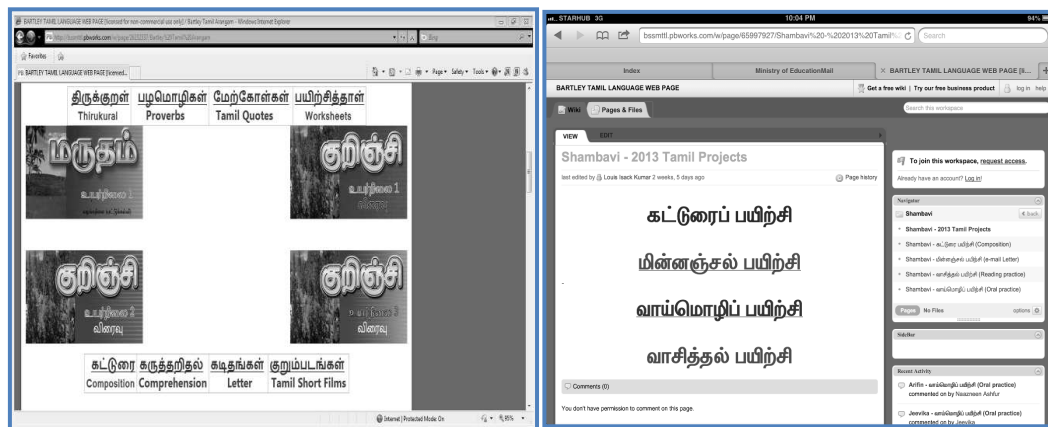
PB works hosts over 300, 000 educational workspaces worldwide, and has helped transform teaching and learning for millions of students, parents and teachers. In Singapore, Educators from MOE schools and individual teachers trust PB works as their collaborative learning environment.

PB works lets students to capture knowledge, share files, and manage projects. It tracks every change, and automatically notifies you and your team to keep everyone in the loop. PB works is secure, reliable, and accessible from any computer or mobile device, so your team can use it anywhere they go, you don't need to download any software or manage any servers. Whatever you're working on, you can customize PB works to make your team more productive and as engaged learners.

- ✍ PB works encourages student-centered learning. Even young students can build web pages, embed images & video, and post documents.
- ✍ Provide access to information sources, book lists, and links to good articles. Have resources stored for future use.
- ✍ Host and share information between students, faculty and staff. Encourage staff development and shared resources across schools.
- ✍ Make distant learning more interactive and collaborative, support research teams, and improve inter-departmental coordination.

What can we do with a Classroom Workspace?

- **Class Resources** - Publish class notes, PowerPoint lectures, schedules and policies; show off examples of great student work.
- **Group Projects** - Build collaborative pages, start discussions and encourage comments.
- **Student Portfolios** - Give students their own page to post content, upload homework, and share their work.



How are we using PB works?

PB works just took off like wildfire, and a big part of that is because of how easy it is to use PB works. It's so similar to what we do in word that the teachers just took to it. They were looking for mechanisms for having kids collaborate online, and PB works was just so much easier than other tools like the blogging services. Bartley Secondary School Tamil Students use PB

works as their homepage, they post their homework assignments...part of the reason we have such broad adoption is that we have teachers, and peers using it.

Bartley Secondary School TL teacher and students have also explored creative ways of using it in PB Works. For example, our Sec 1 Express students are composing stories online, and posting their compositions, and other project works in workspaces to make it easier for children to keep working on their assignments when they are at home. Besides providing a pool of phrases to support the writing of good sentences, these workspaces are also a collaborative tool to encourage collaborative learning in class. In addition, this collaborative writing portal will also support the alternative mode of assessment whereby feedback and comments can be generated online by peers and teachers.

Conclusion and Recommendations

In summary, we recognize the importance of providing our students with the opportunity to learn creatively and engaged to take ownership of their learning using PBworks. We found that the workspace could be used for home-based learning as the teacher is able to monitor the progress of the learning on-line or as a regular feature to promote further learning without sacrificing much curriculum time.

The time and effort put in planning and executing these lessons were fruitful as the students would greatly benefit from working closely with others and to always recognize the importance of being confident and being responsible for their own learning.

மொழி கற்பித்தலுக்கான இணையத் தகவல் மூலங்களை மதிப்பிடுதலும் தெரிவுசெய்தலும் -ஒரு கண்ணோட்டம்

மணிமாறன் கோவிந்தசாமி

துவான்கு பைனான் ஆசிரியர் கல்விக் கழகம், பிளாங்கு

முன்னுரை

தற்காலத்தில் ஆசிரியர்களின் வாழ்வில் இணையத்தின் பயன்பாடு இன்றியமையாத ஒன்றாகிவிட்டது. தொடர்புக் கொள்ளவும் தகவல்களைத் தேடவும் கற்றல் கற்பித்தலுக்குத் தேவைப்படும் மூலங்களை வலைத்தளங்களிலிருந்து திரட்டவும் இணையம் பள்ளிகளிலும் உயர்க் கல்விக் கூடங்களிலும் பயன்படுத்தப்படுகிறது.

உலகில் 377.65 கோடி மக்கள் இரண்டாயிரத்தாம் ஆண்டு செப்டம்பர் மாதம் இணையத்தைப் பயன்படுத்தி உள்ளனர் (Nua Internet Surveys, 2000). மனிதனின் வாழ்க்கையில் இணையம் பெரும் விளைபயனைக் கொண்டு வரும் என்பதனை எவரும் மறுக்கவியலாது. www. @ இணையத்தளம் (web) உலகில் மிக துரிதமான வளர்ச்சி கண்டு வரும் புத்தாக்கத் தொழில்நுட்பமாகும். ஆகவே, கல்விமான்கள் ஒவ்வொருவரும் தங்களது கல்வி அடிப்படைகளையும் கற்றல் கற்பித்தல் கொள்கைகளையும் மீட்டுணர்வு செய்ய வேண்டிய நிலையில் தள்ளப்பட்டுள்ளனர். ஆசிரியர்களின் செயற்பாடுகள் மாற்றம் கண்டுள்ளன. ஆசிரியர்கள் வகுப்பறையில் கற்றலை அமலாக்கம் செய்யும் ஒருவராக மட்டும் இல்லாமல் தொலைநோக்குக் கற்றல் கற்பித்தலை அமலாக்கம் செய்யும் ஒருத்தராகவும் மாற வேண்டிய நிலை இருப்பதை உணர வேண்டும்.

லான்ஸ்பெர்கர் (2000) கூற்றுப்படி இணையத்திற்கு விதிமுறைகளும் இல்லை விதிமுறைகளை விதிக்கக் கூடிய நிலையும் இல்லை. ஆதலால்தான், அங்கே தகவல்கள் மலைபோல் குவிந்தும் மலிந்தும் கிடக்கின்றன. அத்தகவல்கள் அனைத்தும் கற்றல் கற்பித்தலுக்குப் பொருத்தமானவையா எனத் தெரிவுசெய்வதில் கல்வியாளர்கள் பெரும் குழப்பத்தை எதிர்நோக்குகின்றனர். ஆகவே, இணையத்தைக் கட்டுப்பாட்டுக்குள் வைத்துக்கொள்வது அவசியமாகிறது. அத்தகவல்களைக் கற்றல் கற்பித்தலுக்குத் தேவையானவையைத் தெரிவுசெய்வதற்கு முன் அவற்றை மதிப்பிட வேண்டியது மிகமிக அத்தியாவசியமாகிறது. ரசியா, மியன்மார், சீனா முதலிய நாடுகள் (செர்மேன், 2000) இணையத் தகவல்களைத் தணிக்கை செய்தாலும், 60% கணினி நுகர்வோர்களைக் கொண்டிருக்கும் அமெரிக்கா இந்நடவடிக்கையை மேற்கொள்ளவில்லை. மலேசியாவில் மிகத் தீவிரமில்லா இல்லா நிலையில் இத்தணிக்கை நடவடிக்கை மேற்கொள்ளப்பட்டாலும், இணையத்தின்வழி சுயக் கருத்துகளைத் தெரிவிப்போரின் எண்ணிக்கை நாளுக்கு நாள் பெருகிக்கொண்டே வருகின்றது.

இணையமும் உள்ளடக்கமும்

இணையக் கழகம், internet society (2000) என்னும் தன்னார்வ இயக்கம் ஒன்று, இரண்டாயிரத்து இருபதாம் ஆண்டு இறுதிக்குள் இணைய நுகர்வோர்களின் எண்ணிக்கை 400 கோடியைத் தாண்டும் எனக் கணித்துள்ளது. இது உலக மக்கள் தொகையில் 6% ஆகும். உலகில் உள்ள அனைத்துப் பள்ளிகளிலும் இணையத்தைப் பயன்படுத்துவோரின் எண்ணிக்கை நாளுக்கு நாள் அதிகரித்து வருகின்றன. பிரிட்டனில், cyberAtlas (2000), பள்ளிகளில் 43% இணையப் பயன்பாடும் இடைநிலைப்பள்ளி மாணவர்களின் எண்ணிக்கையும் 8.7:1 எனும் அளவில் உள்ளது. 86.55% ஆங்கில மொழியிலேயே இணையத் தகவல்கள் வெளிவருகின்றன. மலேசியாவில் 75% இணையத் தகவல்கள் வெளிநாடு அகப்பக்கங்கள் மூலமாகவே பெறப்படுகின்றன (Utusan Malaysia, 2010).

இணைய உள்ளடக்கங்கள் பல்வகைப்படும். Harris (1997), இணைய உள்ளடக்கங்களை ஐந்து பிரிவுகளாக வகைப்படுத்துகிறார். அவையாவன;

- அ. கருத்து விளக்கம்
- ஆ. சுயக்கருத்துகள்
- இ. கதைகள்
- ஈ. விமர்சனங்கள்
- உ. புள்ளியல் விவரங்கள்

ஹரிஸின் கூற்றுப்படி, மேற்கண்ட தகவல்கள் அனைத்தும் இணையத்தில் எங்கு வேண்டுமென்றாலும் எவ்வளவு வேண்டுமென்றாலும் பெற்றுக்கொள்ளலாம். இணையத்தின் அனைத்து உள்ளடக்கங்களும் கீழ்க்காணும் நோக்கங்களுக்காக வெளியிடப்படுகின்றன என்கிறார்.

- அ. தெரியப்படுத்த (to inform)
- ஆ. நயப்படுத்த (to console)
- இ. விற்பனை செய்ய (to do business)
- ஈ. கருத்துகளைத் தெரிவிக்க (to express idea)
- உ. நம்பிக்கைகளை உருவாக்கவும் மாற்றவும் (to create and change beliefs)

Lansberger (2000), 8 வகையான வலைத்தளங்களைக் கீழ்க்காணும் நோக்கங்களின் அடிப்படையில் பட்டியலிடுகிறார்.

- அ. சுய வலைத்தளம்
- ஆ. விளம்பரத் வலைத்தளம்
- இ. நடப்புச் செய்தி வலைத்தளம்
- ஈ. தலைப்பு வாரியான தகவல்கள் வலைத்தளம்
- உ. பிரச்சார வலைத்தளம்
- ஊ. கற்பித்தல் வலைத்தளம்
- எ. தரவுகள் திரட்டு வலைத்தளம்
- ஏ. மனமகிழ்வு வலைத்தளம்

Ed's Oasis (2002) வலைத்தளம், கல்விக்குத் தொடர்பான 3 வலைத்தளங்களைப் பட்டியலிட்டுள்ளது.

- அ. தகவல் மூலங்கள் (Resources)
- ஆ. செயல்திட்டங்கள் (Projects)
- இ. கூடுதல் தகவல் மூலங்கள் (Extra Resources)

தகவல் மூலமானது மாணவர்களால் ஆய்வுக்குட்படுத்தப்படும் இணைய விவரங்களாகும். நடவடிக்கைகளைப் பற்றிய தகவல்கள் அடங்கிய பகுதியே செயல்திட்டங்கள் எனப்படுகின்றன. கூடுதல் தகவல் மூலங்கள் எனப்படுபவை ஆசிரியர்களின் உதவியுடன் கற்றல் கற்பிதலுக்குப் பயன்படும் தகவல்கள் ஆகும்.

இணையத் தகவல்களை தெரிவுசெய்தலும் மதிப்பிடுதலும் ஏன் அவசியமாகிறது?

மாணவர்களிடையே இணையம் பெரும் விளைபயனைத் தருகின்றது. ஆதலால், கற்றல் கற்பித்தல் தொடர்பான இணையத் தகவல்களைத் தெரிவுசெய்தலிலும் மதிப்பிடுதலிலும் ஆசிரியர் மிகுந்த கவனத்தைச் செலுத்த வேண்டிய நிலை இருக்கின்றது. O'Neil (1999), இணையத் தகவல் மூலங்களைத் தெரிவுசெய்தல் மற்றும் மதிப்பிடுதல் ஏன் அவசியமாகிறது என்பதனைக் கீழே விளக்குகிறார்.

அ. அளவுக்கதிகமான தகவல் மூலங்கள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஆய்வின்படி, ஜனவரி 2000 ஆண்டின் இறுதிக்குள் இணையத்தில் தகவல் மூலங்கள் 1 பில்லியனையும் தாண்டிவிட்டன எனப் பகர்கின்றது. தகவல் மூலங்களின் எண்ணிக்கை அளவுக்கதிகமாக இருப்பதால், ஆசிரியர் தமது பாடத்திற்கேற்பவும் கலைத்திட்டத்திற்கேற்பவும் கற்பித்தல் நோக்கங்களுக்கேற்பவும் அத்தகவல்களைத் தெரிவுசெய்வதிலும் மதிப்பிடுவதிலும் மிகுந்த கவனத்தை செலுத்த வேண்டியது அவசியமாகிறது.

ஆ. இணையத்தில் தகவல்களை வெளியிட வரையறையற்ற சுதந்திரம்

இணையத்திற்கு ஒரு விசித்திரமான தன்மை உண்டு. தனியாள், இயக்கங்கள், கழகங்கள், உயர்க்கல்விக் கூடங்கள் அல்லது எவர் வேண்டுமென்றாலும் எந்நேரத்திலும் எத்தகவல்களைத் தங்கு தடையின்றிச் சுதந்திரமாக வெளியிடலாம். கணினியின் விலை மதிப்பு அதிகமாக இருந்தாலும், <http://www.geocities.com>, <http://www.tripod.com> போன்ற வலைத்தளங்கள், இணைய நுகர்வோர் தத்தம் கருத்துகளை இலவசமாக வெளியிட பற்பல வசதிகளை ஏற்பாடு செய்து தருவதால், கட்டுப்பாடற்ற நிலையில் நன்தகவல்களும் தீந்தகவல்களும் குன்றுபோல் குவிந்து விடுகின்றன. இந்நிலையில் ஓர் ஆசிரியர் பயன்விளைவுமிக்க தகவல்களைத் தெரிவுசெய்வதிலும் மதிப்பீடு செய்வதிலும் மிகுந்த கவனத்தைச் செலுத்துவது அவசியமாகிறது.

இ. எவராலும் உறுதிபடுத்த/ அங்கீகரிக்க வேண்டிய அவசியமில்லை

இணையத்தில் வெளியிடப்படும் தகவல்கள் அனைத்தும் சரிபார்த்தல் மற்றும் அங்கீகரித்தல் செயற்பாங்கினைக் கடந்து வருவதில்லை என்பதால் இணையத்தின்வழி தகவல் தேடுவோர்க்கு இது பெரும் பிரச்சனையாக இருக்கின்றது. இது அச்சிடப்படும் ஊடகங்களுக்கு முரணாகும். காரணம், அச்சிடப்படும் ஊடகங்களான நூல்,

ஆய்வடங்கல், நாளிதழ் போன்றவை வெளியீடு காண்பதற்குமுன் திருத்தப்பட்டுத் தணிக்கைக்குட்படுத்தப்படுகிறது. மாறாக, இணையத் தகவல்களுக்கு அச்செயற்பாங்கு அமலாக்கம் செய்யப்படுவதில்லை.

இத்தகைய பண்டைய செயற்பாங்கின்வழி அங்கீகரிக்கப்படாத தகவல் மூலங்கள் கற்றல் கற்பித்தலுக்குப் பொருத்தமில்லாதவையாக இருக்கலாம்; ஏரணமற்று இருக்கலாம்; நம்பகத்தன்மையில்லாமலும் இருக்கலாம். ஆகவே, ஓர் ஆசிரியர் இணையத் தகவல்களை மதிப்பீடு செய்தல் இங்கு மிக முக்கியமாகக் கருதப்படுகிறது.

இணையத் தகவல்களை மதிப்பீடு செய்வதில் கவனிக்கப்பட வேண்டிய கூறுகள்

இணையத்தில் வெளியிடப்படும் தகவல்களை மதிப்பீடு செய்வதில் கவனிக்கப்பட வேண்டிய கூறுகளைப் பற்றி பல்வேறு சிந்தனைகள் உள். Harris (1997), CARS எனப்படும் சரிபார் பட்டியல் இத்தகைய மதிப்பீட்டிற்கு மிகப் பொருந்தி வரும் என்கிறார்.

1. ஆளுமை (Credibility)
2. துல்லியம் (Accuracy)
3. ஏற்புடைமை (Reasonableness)
4. சார்புடைமை (Support)

வலைத்தளங்களில் வெளியிடப்படும் தகவல் மூலங்களின் ஆளுமையை (Credibility) தகவலின் உரிமையாளர் அல்லது வலைத்தள உரிமையாளர் ஆகியோரது ஆளுமையோடு ஒப்பிட்டு மதிப்பீடு செய்யலாம். சீரான நல்ல தகவல்களைத் தரும் தனியாளர் அல்லது வலைத்தளக் குழுமத்தின் தகவல்கள் நாம் அவர்களிடமிருந்தும் ஆளுமையின் அளவை அதிகரிக்கச் செய்யும். காட்டாக, ஆற்றல்மிகு பதிப்பாசிரியர்களின் (editors) தரக் கட்டுப்பாட்டுடன் கூடிய தகவல் மூலங்கள், தெளிவாகவும் சரியாகவும் குறிப்பிடப்பட்டுள்ள மேற்கோள் நூல்கள், இலகுவாக மேற்கோள்களைத் தேடும் வழிமுறைகள் போன்ற தன்மைகள் கொண்ட ஒரு வலைத்தளம் விழுமிய ஆளுமையை உடையதாகக் கருதப்படும். மாறாக, எழுத்தாளரின் பெயர் குறிப்பிடப்படாத, இலக்கணப் பிழைகள் மலிந்து கிடக்கும், தரக் கட்டுப்பாடில்லாத இணையத் தகவல் மூலங்களை மதிப்பீடு செய்யுங்கால் அவை நிச்சயம் கல்வி நோக்கங்களுக்குக் கிஞ்சிற்றும் பயனைத் தாரா.

தேதி, கருத்து, தகவல் முதலியவை ஒரு தகவல் மூலத்தில் சரியாக எழுதப்பட்டுள்ளதா? இல்லையா? என்பதனை உணர்த்துவதே துல்லியத்தன்மையாகும். வரலாறு தொடர்பான கருத்துகளை இணையப் பயனீட்டாளர்கள் குழம்பாத வகையில் முறையாகவும் வரிசைக்கிரமமாகவும் எழுதப்படுதல் மிகமிக அவசியம். வலைத்தளங்கள் உருவாக்கப்பட்டதன் நோக்கம் தெளிவாக இருக்க வேண்டும். அந்நோக்கங்கள் தெளிவுபடாது இருக்கும் நிலையில் எவ்வகைப் பயனீட்டாளர்களுக்கு அத்தகவல் சென்றடையும் என்பதும் தெளிவில்லாமல்தான் இருக்கும்.

நீதி, சரிவிகித கருத்துகள், மிதமான மற்றும் தகவல்களின் வலிமை ஆகிய கூறுகள் ஏற்புடைமை தன்மையைக் காட்டும். ஒருவேளை, தகவல் மூலங்கள் ஒரு சாராரையோ செயலையோ பெரும்பான்மை நிலையில் சார்ந்து இருந்தால், அத்தகைய தகவல்களைக் கல்வித் துறையில் பயன்படுத்துவதை ஆசிரியர்கள் தவிர்க்க வேண்டும்.

ஓர் எழுத்தாளரின் சுய தகவல் தொடர்பான மேற்கோள் நூல்கள் மிகத் தெளிவாக இருந்தால், சார்புடைமை தன்மையைக் காட்டும். காட்டாக, புள்ளியல் தொடர்பான தரவுகள் கையாளப்பட்டிருந்தால், அத்தகவலின் மூலமும் தெரிவுசெய்யப்பட்ட அணுகுமுறைகளும் தெளிவாகக் குறிக்கப்பட்டிருக்க வேண்டும். பனுவலாக இருப்பின் மேற்கோள் நூல்கள் சரியாகப் பட்டியலிடப்பட்டிருக்க வேண்டும். கொடுக்கப்பட்ட தகவலைச் சார்ந்துவரும் தரவுகள் இல்லையென்றால் அத்தகவல் மேற்கோளுக்கு ஆளுமை நிறைந்ததாக ஏற்றுக்கொள்ளப்பட மாட்டா.

McIntyre நூலகம் (2000), இணையத்தில் வெளியிடப்படும் தகவல் மூலங்களை மதிப்பீடு செய்ய 'பத்து C' (Ten C's) என்னும் ஒரு மதிப்பீட்டுக் கருவியைப் பரிந்துரை செய்துள்ளது. அம்மதிப்பீட்டுக் கருவியின் 10 வகையான மதிப்பீட்டுக் கூறுகளைக் கீழே காண்க.

1. உள்ளடக்கம் (Content)

ஒவ்வொரு தகவல் மூலங்களும் வெளியீடு செய்ததற்கான நோக்கங்கள், வெளியீடு கண்ட திகதி மற்றும் அது புதிய வெளியீடா அல்லது புதுப்பிக்கப்படாத பழைய வெளியீடா போன்ற விவரங்களை ஆராய வேண்டும்.

2. ஆளுமை (Capable)

எழுத்தாளரை அடையாளங் காணுதல் மற்றும் நம்பக்கூடிய தன்மை முதலிய உரிமைகளை உள்ளடக்கிய கூறுகளை மதிப்பீடு செய்யலாம்.

3. புத்தாக்கச் சிந்தனை (Creativity)

ஒரு மதிப்பீட்டாளர் தமது அறிவையும் அனுபவத்தையும் கொண்டு எழுத்தாளர், அவர்தம் வெளியீடு போன்ற கூறுகளை மதிப்பீடு செய்யலாம்.

4. பதிப்புரிமை (Copy Right)

வலைத்தளங்களில் அடங்கியிருக்கும் ஆவணங்களின் பயன்பாடு, வரைகலை, ஒலி அல்லது பிம்பம் போன்ற கூறுகளின் பதிப்புரிமைகளில் கவனம் செலுத்த வேண்டும்.

5. மதிப்பளித்தல் (Commendation)

அசல் ஆவணப் பதிப்புரிமையாளர் அல்லது அவ்வாணத்தில் பயன்படுத்தப்பட்ட மூலங்கள் போன்ற கூறுகளுக்கு உகந்த மதிப்பளிக்கப்பட்டதா என்பதை மதிப்பீட்டாளர் கவனத்தில் கொள்ள வேண்டும்.

6. தொடர்ச்சி (Continuity)

வலைத்தளங்களில் அடங்கியிருக்கும் தகவல்கள் முறையாகப் பராமரிக்கப்பட்டிருக்கின்றதா காலத்திற்கேற்ப கூடுதல் தகவல்கள் சேர்க்கப்பட்டிருக்கின்றதா என்பதை மதிப்பீட்டாளர் ஆராய வேண்டும்.

7. தணிக்கை (Censore)

ஒரு மூலத்தின் உள்ளடக்கத்தைப் பயனீட்டாளர்கள் தணிக்கை செய்ய முடியுமா அல்லது தடைவிதிக்கப் பட்டுள்ளதா போன்ற விவரங்களைக் கருத்தில் கொள்ள வேண்டும்.

8. தொடர்பாற்றல் (Connection)

பயன்படுத்தப்படுத்தப்படும் மூலத்திற்கு வேறு துணைமென்பொருள்கள் தேவைபடுமா போன்ற தொழில்நுட்ப நிலைகளை உணர்ந்து மதிப்பீடு செய்ய வேண்டும்.

9. ஒப்பிட்டாற்றல் (Comparison)

இணையத்தில் உள்ள தகவல் மூலங்களை அச்சு ஊடகங்கள் மற்றும் குறுந்தடுத்த தகவல்களோடு ஒப்பீடு செய்யலாம். புதிய வெளியீட்டைப் பழைய வெளியீட்டோடும் ஒப்பீடு செய்யலாம்.

10. சூழல் (Context)

மூலங்களைத் தேடுதல் மற்றும் வருங்காலங்களில் அம்மூலங்களின் பயன்பாடு போன்ற கூறுகளையும் மதிப்பீட்டாளர் கவனத்தில் கொள்ள வேண்டும். தெரிவுசெய்யப்பட்ட மூலம் கொள்கையா வரலாறா, புள்ளியிலா என்பதனைக் கருத்தில்கொண்டு அவை கற்றல் கற்பித்தலுக்கு உகந்த சூழல் என்ன என்பதனையும் நினைவிற்கொண்டு மதிப்பிடுதல் வேண்டும்.

முடிவுரை

இணையத்தில் தகவல்கள் மலிந்து கிடந்தாலும் எல்லாத் தகவல்களும் மொழிக் கற்பித்தலுக்குப் பொருந்தி வருமா? என்பதனை மதிப்பீட்டாளரான ஆசிரியர் தெளிவாக உணர்தல் அவசியம். வலைத்தளங்களை தெரிவுசெய்வதிலும் மதிப்பிடுதலிலும் முறையான வழிகளைக் கடைப்பிடித்தல் மிகமிக முக்கியம். இல்லையேல் ஆசிரியர்கள் தரமான தகவல்களை மாணவர்களுக்குத் தருவதில் தப்புச் செய்துவிடுவர். நல்லாசிரியர் கையில் இணையம் இருக்கும்போது அது கற்றல் கற்பித்தலுக்கு நல்ல விளைபயனைத் தரும். மாறாக, அவ்விணையம் அலட்சியப் போக்கைக்கொண்ட ஓர் ஆசிரியர் கையில் தவறும்போது அது கற்றல் கற்பித்தலின் உன்னத நோக்கிலிருந்து தடம்புரளும்.

மேற்கோள்கள்

- Alexander, Jan, Tate, Marsha Ann (1999). Checklist for an Informational Web Page. <http://www2.widener.edu/wolfgram-Memorial-Webevaluation/inform.htm>.
- CyberAtlas (1998).43 Percent of UK Schools on Net.<http://cyberatlas.internet.com/big picture>.
- CyberAtlas (1999).Half of Teachers say Web Essential.<http://cyberatlas.internet.com/bigpicture>.
- ED's Oasis (2000).ED's Oasis Online Web Site Evaluation Educators. <http://www.edoasis.org/2guide3.html>

- Harris, Robert (1997).Evaluating Internet Research Sources. <http://www.Vanguard.edu/rharris/evalu8it.htm>
- Inktomi.com (2000). <http://inktomi.com/webmap/>
- Lansberger, Joe (2000).Evaluating Web Site Content <http://www.iss.stthomas.edu/webtruth/evaluate.htm>
- McIntyre Library (2000). Ten C's for Evaluating Internet Source. <http://www.uwec.edu/Admin/Guides/tencs.html>
- NUA (2000). Internet Surveys. How Many Online? [http://nua.ie/survey/how many online/index.html](http://nua.ie/survey/how_many_online/index.html)

இணையம் வாயிலாக தமிழ்க்கல்வி கற்க பிற மொழி மாணவர்கள் எதிர்கொள்ளும் சிக்கல்கள்

கங்காதுரை கணேசன்

விரிவுரையாளர், துங்கு அப்துல் ரஹ்மான் பல்கலைக்கழகம், கம்பார், பேரா – மலேசியா gangadurai@utar.edu.my

முன்னுரை

துங்கு அப்துல் ரஹ்மான் பல்கலைக்கழகத்தில் ஐந்தாண்டுகளாகத் தமிழ்மொழி தொடக்க நிலை பாடமாக கற்பிக்கப்பட்டு வருகிறது. கட்டாயத் தேர்வுப் பாடங்களின் வரிசையில் தமிழ் மொழி ஒரு விருப்பப் பாடமாக இருந்து வருகிறது. இப்பாடம் பிற மொழி மாணவர்களான மலாய், சீன மாணவர்கள் மட்டுமே விருப்பப் பாடமாகத் தேர்ந்தெடுத்துப் படிக்க அனுமதிக்கப்படுகின்றனர். இவ்வகுப்பு ஒவ்வொரு பருவத்திற்குமே வருடத்தில் மூன்று முறை வழங்கப்பட்டு வருகின்றது. ஒரு வாரத்தில் இரண்டு மணி நேர விரிவுரையும் (Lecture) ஒன்றரை மணி நேர கலந்துரையாடலும் (Tutorial) என பருவத்திற்குப் பதினான்கு வாரங்கள் தமிழ் மொழி போதிக்கப்படுகிறது.

“தமிழ் கற்பித்தலை முதல் மொழி நிலை, இரண்டாம் மொழி நிலை, அயல்மொழி நிலை என்று மூன்று வகையாகப் பகுக்கலாம். தமிழைத் தாய் மொழியாய்க் கொண்டவர்களுக்குத் தமிழ் கற்பித்தல் முதல் மொழி நிலை என கருதப்படும். தமிழைத் தாய்மொழியாகக் கொள்ளாத பிற மொழியாளருக்குத் தமிழ் புழங்கும் புறச் சூழலில் தமிழ் கற்பித்தல் இரண்டாம் மொழி நிலை எனப்படும். தமிழ் புழங்காத புறச் சூழலில் பிற மொழியாளருக்குத் தமிழ் கற்பித்தல் அயல்மொழி நிலை என வழங்கப்படும்” (சுப.திண்ணப்பன், 1992: 68). துங்கு அப்துல் ரஹ்மான் பல்கலைக்கழகத்தில் அயல் மொழி என்னும் நிலையிலேயே தமிழ் கற்பிக்கப்படுகின்றது.

தமிழ் மொழிப் பாடத்தைத் தேர்ந்தெடுக்கும் மாணவர்கள் இருபத்தெட்டு மணி நேர விரிவுரையும், இருபத்தொன்று மணி நேர கலந்துரையாடலும், எழுபத்தொன்று மணி நேர சுயக் கற்றலும் என மொத்த கற்றல் நேரம் (Total Learning Time) நூற்று இருபது மணி நேரம் பூர்த்தி செய்தாக வேண்டும். மாணவர்களுக்கு, வகுப்பறைச் சூழலைத் தொடர்ந்து தமிழ் மொழியைக் கற்றுக்கொள்ள இணையத் தளங்கள் மேற்கோள் காட்டப்படுகின்றன. ஆயினும் மாணவர்கள் இணையத் தளங்களில் தமிழ்க்கல்வி கற்க சிக்கல்களை எதிர்கொள்கின்றனர். மாணவர்கள் எதிர்கொள்ளும் சிக்கல்களை ஆராய்ந்து விளக்குவதே இக்கட்டுரையின் நோக்கமாகும்.

மாணவர்களுக்கு மேற்கோள் காட்டப்படும் இணையத் தளங்கள் (பின்னிணைப்பில் இணைக்கப்பட்டுள்ளன) பிற மொழி மாணவர்கள் கல்வி கற்பதற்கான முழுமையான, முறையான தளங்கள் என குறிப்பிடலாகாது. இணைய தமிழ்க்கல்வி, முதல் மொழி நிலை மற்றும் இரண்டாம் மொழி நிலை மாணவர்களுக்கு ஏதுவாக அமைகிறது. பிற மொழி மாணவர்களுக்கு இத்தளங்கள் பல சிக்கல்கள் நிறைந்தவையாக இருக்கின்றன. தொடர்ந்து சிக்கல்கள் குறித்து இங்கு எடுத்துரைக்கப் பெருகின்றன.

1. எழுத்து முறை

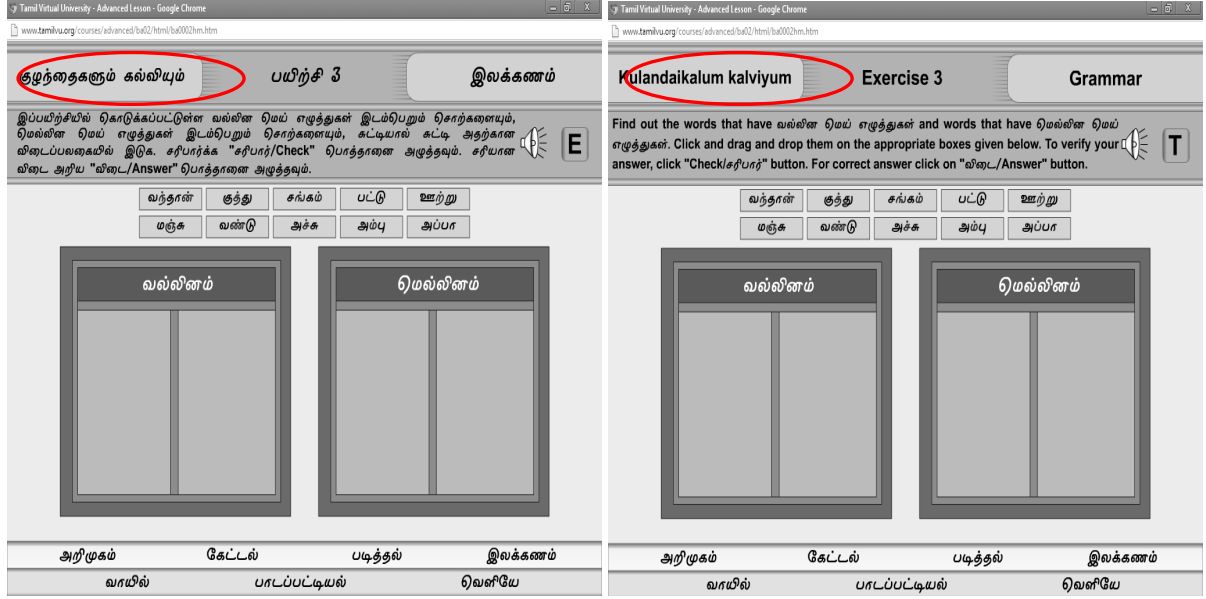
பல்கலைக்கழகத்தில் பிற மொழி மாணவர்களுக்குத் தமிழ் பயிற்றுவிக்கும்போது தமிழ் எழுத்தும், அதை சரியாக உச்சரிக்க ரோமன் எழுத்து முறையும் பயன்படுத்தப் படுகின்றது. தமிழ் மொழியில் குறில் நெடில் வேறுபாடுகள் இருப்பதால், அதை விளக்கப்படுத்த ரோமன் பயன்படுத்தியாக வேண்டும்.

எடுத்துக்காட்டு: என் பெயர் ராமு

En peyar Raamu

My name is Ramu

இப்படியாக பிற மொழி மாணவர்களுக்குத் தமிழ் போதிக்கும்போது அவர்கள் தமிழ் எழுத்துக்களையும், பிழையற்ற உச்சரிப்பும், அதற்கான பொருள் விளக்கத்தையும் முழுமையாக கற்றுக் கொள்கின்றனர். அதோடுமட்டுமின்றி பிற மொழி மாணவர்கள் குறுகிய காலத்தில் இலகுவாக தமிழில் உரையாடவும் வழிவகுக்கும். இந்நிலை இணைய தமிழ்க்கல்வியில் கிடையாது. இணையத்தில் எல்லாமே தமிழில் இருக்கின்றன. கட்டளையும் விளக்கமும் ஆங்கிலத்தில் இருக்கின்றன. விசை அழுத்தினால் தமிழ் உச்சரிப்பு ஒலிக்கிறது. இது பிற மொழி மாணவர்களுக்கு உகந்தது அல்ல. பிற மொழி மாணவர்கள் தொடக்க நிலையில் தமிழ் வாக்கியங்களை வாசிப்பதும் அதை புரிந்து கொள்வதும் எளிதல்ல.



படம் 1-இல், மாணவர்கள் தலைப்பின் அர்த்தம் அறியாது, தமிழில் உச்சரிக்க மட்டுமே இயலும். அது மட்டுமே ரோமன் எழுத்து முறையில் வழங்கப்பட்டுள்ளது. கட்டளையின் சாரம் ஆங்கிலத்தில் விளக்கப் படுத்தப்பட்டுள்ளது. மற்றவை தமிழ்மொழியிலேயே இருக்கின்றன. பிற மொழி மாணவர்களால் உடனடியாக தமிழ் சொற்களைப் படிக்க இயலாது. அவர்களுக்குச் சொற்களின் அர்த்தம் கூட அறிந்துகொள்ள இயலாது. அவர்கள் வெறுமன பயிற்சியைச் செய்ய முடியுமே தவிர அதை படித்து புரிந்து செயல்பட முடியாது. அதனால் மாணவர்களுக்கு எந்த பயனும் இல்லை. ஆனால் www.ilanguages.org இணையத்தளத்தில் தமிழ், தமிழ் ரோமன், ஆங்கிலம் மற்றும் உச்சரிப்பு முறை (ஒலி விசை) அனைத்தும் வழங்கப்பட்டுள்ளன. ஆயினும் அது ஒரு முழுமையான தளமாக குறிப்பிடலாகாது.



படம் 2: www.ilangauages.org இணையப் பாடப்பகுதியிலிருந்து

2. எழுத்து மொழியும் பேச்சு மொழியும்

பிற மொழி மாணவர்கள் எதிர்கொள்ளும் மற்றொரு சிக்கல் யாதெனில் மொழி. தமிழில் இரு வழக்கு மொழி உண்டு. ஒன்று எழுத்து மொழி, மற்றொன்று பேச்சு மொழி. மாணவர்களுக்குக் கற்பிக்கும் போது எழுத்து மொழியே பயன்படுத்தப்படுகிறது. அந்த முறையில் கற்கும் மாணவர்கள், சில இணையத் தமிழ்க்கல்வியில் இணையும்போது பேச்சுமொழி குறுக்கிடுக்கிறது.

படம் 3: www.kidsone.in இணையப் பாடப்பகுதியிலிருந்து

Learn Tamil	Rhymes	Stories	Games	Levels	Worksheet	Wallpaper
Learn Tamil						
Vowels						
Consonants						
Sub Letters						
Vowels & Consonants						
Animals						
Birds						
Colours						
Common Insects						
Clothes						
Directions						
Days & Months						
Edibles						
Education						
Fruits						
Feelings						
Flowers						
Great Teachers						
Household Things						

Tamil Meanings For English Words	
Previous - 30	
English Word	Tamil Meanings
How many apples are there in my hand?	Ethana apple enn kaiella irukku?
How many did you take?	Ethana nee edutha?
How much did he pay you?	Evalavu panam kodutha?
How much distance to go?	Evalavu thooram poakanum?
How was the journey yesterday?	Nethu payanam epadi irunthathu?
Which way did you come?	Entha valiya vanthe?
Which is your favorite color?	Virupamana niram enna?
In which room did you sleep?	Entha araila thoongina?
Which story did you tell?	Entha kathaya sonna?
Which is the sweetest fruit?	Unnaku rumba inipaana palam ethu?
Which is the best newspaper in Hindi?	Entha siethithaai nalla irukkum?
Which Indian state has the largest population?	Entha indhiya naattu maanilathula romba janathogai iruku?
Where did you come from?	Enge irunthu vandha?

பிற மொழி மாணவர்களுக்கு, மேற்கண்ட பாடப்பகுதி மொழியியல் திறன்களின் ஒன்றான பேசுதல் (Speaking) பகுதிக்கு ஏற்புடைய பகுதியாகும். ஆயினும் இப்பகுதியில் காணப்படும் சிக்கலானது பேச்சு மொழி. பிற மொழி மாணவர்கள் எழுத்து மொழியில் பயின்று பேச்சு மொழிக்கு செல்லும் போது, உச்சரிப்பு, பொருள் புரியாமை போன்ற சிக்கல்களை எதிர்கொள்ள வேண்டியுள்ளது.

எடுத்துக்காட்டு: Where did you come from? – Enge irunthu vandha? (பேச்சுமொழி)

Nii enggirundhu vandhaay? (எழுத்து மொழி)

கொடுக்கப்பட்டுள்ள எடுத்துக்காட்டிலிருந்து பேச்சு மொழிக்கும் எழுத்து மொழிக்கும் உள்ள வேறுபாடுகளை நன்கு உணர முடியும்.

3. மரபு சார்ந்த சொற்களஞ்சியங்கள்

பல்கலைக்கழகத்தில் பிற மொழி மாணவர்களுக்குப் போதிக்கப்படும் தமிழ் மொழியானது மலேசியச் சூழலை மையப்படுத்தி அதற்கு ஏற்றவாறு எடுத்துக்காட்டுகள் கொடுக்கப்படுகின்றன. இணையத்தில் இயங்கும் இணையத்தளங்கள் பெரும்பாலும் இந்தியாவை மையமாக கொண்டு செயல்படுகின்றன. இத்தளங்களில் கொடுக்கப்பட்டிருக்கும் சில உதாரணங்கள் இந்திய நாட்டின் மரபுக்கு ஏற்றவாறு அமைந்திருக்கும். சில இணையத்தளங்கள் மேற்கத்திய நாடுகளை மையமாக கொண்டு இயங்குகின்றன. எடுத்துக்காட்டாக www.southasia.sas.upenn.edu இணையத்தளம் அமெரிக்காவை மையமாக கொண்டு இயங்குகிறது. இணையத்தளங்களில் கொடுக்கப்பட்டிருக்கும் சொற்களஞ்சியங்களில் மரபு சார்ந்த சொற்கள் இடம்பெற்றிருப்பது பிற மொழி மாணவர்களுக்கு சிக்கலை ஏற்படுத்துகிறது. அதற்கான விளக்கங்கள் கொடுக்கப்பட்டாலும், அச்சொற்களைப் பயன்படுத்துவதற்கு ஏற்ற சூழல் இங்கு கிடையாது.

படம் 4: www.southasia.sas.upenn.edu இணையப் பாடப்பகுதியிலிருந்து

← → ↻	www.southasia.sas.upenn.edu/tamil/unit_02/glossary.html
	Units: 1 2 3 4 5 6
Select Unit > Unit 2: விருந்து > Glossary	Lessons: 1 2 3 4 5 6 Exercises: 1 2 3 4 5 6 Reading: 1 2 Glossary Conversations Test

Unit 2: Glossary

Tamil written form	Tamil spoken form	English
Common Nouns		
அப்பளம்	அப்பளம்	A thin and round wafer made from flour and fried
அவசரம்	அவசரம்	Haste, quickness
அரை	அரை	Half (1/2) of an amount
அன்பு	அன்பு	Platonic love, affection
ஆட்டோ	ஆட்டோ	Auto-rickshaw

கறி	கறி	Meat, curry
கஷ்டம்	கஷ்டம்	Difficulty
சாவு கிராக்கி	சாவ் கிராக்கி	Term of abuse when one tires someone's patience
சினிமா	சினிமா	Cinema, movies
சும்மா	சும்மா	for free, just, casual; see Unit 2, lesson 2
டஜன்	டஜன்	Dozen
டி.வி	டி.வி.	TV
பைசா	பைசா	Paise, a hundredth of a rupee
பெண்	பெண் / பொண்ணு	Girl
பேஜார்	பேஜார்	Nuisance

மேற்கண்ட படம் 4- இருந்து சில சொற்களை மரபு சார்ந்த சொற்களாக குறிப்பிடலாம். ஆட்டோ, கறி, சாவு கிராக்கி, பைசா, பேஜார் போன்ற சொற்கள் மரபு சார்ந்த சொற்கள். இவ்வகை சொற்கள் இங்கே பயன்படுத்த இயலாது. அதற்கான சூழலும் இங்கே கிடையாது.

பரிந்துரைகள்

உலகளாவிய நிலையில் தமிழ்மொழியின் பயன்பாடும் தேவையும் அதிகரிக்க நாம் வழி வகுக்க வேண்டும். இன்றைய சூழலில் சீன மொழிக்கும், ஜப்பானிய மொழிக்கும் இருக்கக் கூடிய வரவேற்பு தமிழ் மொழிக்கும் ஏற்படுத்திக் கொடுக்க நடவடிக்கைகள் மேற்கொள்ளப்பட வேண்டும். பிற மொழி மாணவர்கள் மட்டுமில்லாது பிற இனத்தைச் சேர்ந்தவர்கள் தமிழ்மொழியை அயல்மொழியாக கற்றுக்கொள்ள இணையத்தில் போதுமான வசதியை ஏற்படுத்தித் தரவேண்டும். முதல் மொழி நிலை மாணவர்கள் தமிழ்மொழி கற்க தமிழ் இணையக் கல்விக்கழகம் (Tamil Virtual University) எத்தகையதொரு வாய்ப்பை வழங்கியிருக்கிறதோ, அத்தகையதொரு வாய்ப்பை பிற மொழியாளர்களுக்கும் வழங்கவேண்டும். அதற்காக உலகளாவிய நிலையிலிருந்து வல்லுநர்கள் குழு, மொழி நிறுவனங்களோடு ஒன்றிணைந்து சிறந்ததொரு இணையத் தமிழ்க்கல்வியை உருவாக்க முனைந்திட வேண்டும்.

பின் இணைப்பு: தமிழ்க்கல்வி தொடர்புடைய இணையத்தளங்கள்

இத்தளங்கள் ஆய்விற்கு பயன்பட்டவை

- www.tamilvu.org
- www.southasia.sas.upenn.edu
- www.noolagam.com
- www.kidsone.in
- www.mylanguages.org
- www.ilanguages.org
- ccat.sas.upenn.edu

துணைபுரிந்த நூல்கள்

1. திண்ணப்பன், சுப., ஐந்தாம் உலகத் தமிழாசிரியர் மாநாட்டு மலர் (1992) சிங்கப்பூர்த் தமிழாசிரியர் சங்கம், சிங்கப்பூர்.
2. பரமசிவம், மு., ஏழாவது உலகத் தமிழாசிரியர் மாநாட்டு ஆய்வடங்கல் (2006) கோலாலம்பூர், மலேசியா.

The Need For Integration Of ICT And Physical Infrastructure Of School Development For The Growth Of The Tamil.

Saravanan Mariappan

Nexus NGN Sdn. Bhd., Malaysia. Email: nexusngn@gmail.com

Abstract: Tamil development in Malaysia is entering into its early stage of golden years. In the recent years with the tireless push from various NGOs the Malaysian government has recognized the need for immediate financial aid for Malaysian Tamil schools for their aging infrastructure and facilities. Millions have been made available for the schools for new buildings, repair works, new furniture, etc. These bold push by the government, will indirectly elevate the development of Tamil Language in Malaysia. With the rapid development of Tamil Schools, comes along the need for well planned physical Infrastructure with integrated ICT within and beyond school boundaries. This paper addresses the fundamental need for Tamil schools to be well planned for immediate and future growth. Immediate planning will be planning out the international standard infrastructure for the school with wheelchair access, large fields, clean environment, paperless management, use of ICT for teaching by ICT integration within and beyond school boundaries. The model school proposed in the paper can be implemented in other parts of the world where the growth of Tamil has been overlooked. The new era of Tamil development in Malaysia is hoped to bridge over oceans and reach out to follow Tamils in other parts of the world via integration of ICT and well planned infrastructure for schools.

Keywords : Integration Of ICT, Physical Infrastructure

1.0 INTRODUCTION

This paper addresses the basic and fundamental requirement of design guidelines for the development of Primary Tamil Schools. This design guidance is intended to assist in the design and proper planning of buildings in response to the educational needs of Tamil school's physical infrastructure. However these infrastructure design guide should be read in conjunction with Ministry of Education's school design guide of the country involved.

Each and every Tamil School's Parents Teachers Association (PIBG) and School Board (LPS) should work out their school physical infrastructure development masterplan, for both short and long term. The masterplan should comprise of

- Overall long term development plan
- Future use of existing rooms and facilities
- Extension plan of existing structure
- Renovation to existing rooms and facilities where necessary for the implementation of the masterplan.
- Perform or provide provisional schedule of essential repair or remedial works where required.
- Replace old Asbestos roof which could lead to mesothel or asbestoses.
- Allow for clean and comfortable access for disabled / handicapped people.
- If possible limit the usage of classrooms within ground and first floor in the event there are not enough rooms / space for classroom lockers. Prevent carrying heavy school bags up to second or third floor.
- Allow for sufficient space for car parks, bus zone, and safe traffic management.
- Also allow for area for hard play area such as basketball, badminton, etc.
- Larger green / grass field as soft play area such as soccer, hockey, etc.
- If space permits provide space for Assembly hall or Performance theater / community hall.
- Provide clear path and access for handicapped individuals.

2.0 TYPICAL GUIDE FOR EDUCATIONAL SPACE REQUIREMENT

2.1 Teaching Space Requirement

The number and type of teaching spaces for Tamil schools varies, tables below are the guide room sizes base on student usage.

Type of Space	Students	Minimum Dimensions	Area
General Classroom	30	7 x 7 m	49 m ²
Lecture Rooms (interconnected)	90	15.6 x 7 m	109 m ²
Group Room	15	7 x 4.2 m	29 m ²
Music/Drama Area	30	11.4 x 7 m	80 m ²
Learning Support Room and Store	30	7 x 7 m + 2.4 x 2.4 m	55 m ²
Guidance Suite (1 classroom + 1 Office)	30	9 x 7 m + 4.8 x 3.2 m	78 m ²
Guidance Suite (1 classroom + 2 Office + waiting)	30	9 x 7 m + 2 x 4.8 x 3.2 m + 3 x 2 m	100 m ²
Religion Room, Meditation Area and Chaplain's Office	30	7 x 7m + 7 x 4.2 m + 4.8 x 3.2 m	94 m ²
Home School Community Liaison	24	7 x 7 m + 4.8 x 3.2 m	64 m ²
Multimedia Learning Laboratory (ICT Lab)	30	11.4 x 7 m	80 m ²
Science Laboratory and Preparation Area	24	11.4 x 7 m + 7 x 2.9 m	100 m ²
Demonstration Room (tiered)	48	9 x 7 m	63 m ²
Art/Craft Room and Store	30	14.2 x 7 m	100 m ²
Home Economics Room and Stores	24	14.2 x 7 m	100 m ²
Architectural Technology/Materials Technology (wood)	24		136 m ²
Technology Room	30		136 m ²
Library and Ancillary Stores (200-499 students)			100 m ²
Library and Ancillary Stores (500 + students)			136 m ²

Total: 1609 m²

2.2 Administration space requirement

Type of Space	Dimensions	Area
Principal's Office	7 x 3.2 m	22 m ²
Deputy Principal's Office	7 x 3.2 m	22 m ²
General Office (schools <500 students)	5 x 4 m	20 m ²
General Office (schools >=500 students)	6 x 5 m	30 m ²
Pastoral Offices (Number as per schedule)	3.2 x 3.2 m	10 m ²
Staff Rooms (area as per schedule)	variable	
Meeting Room	7 x 4.8 m	34 m ²
First Aid Room	4.8 x 3.2 m	15 m ²
Caretaker's Work Area	4.2 x 3.2 m	13 m ²
Photocopy Room	3 x 2 m	6 m ²

Note : (Staff Toilets within total WC allocation)

Total: 182 m²

2.3 Ancillary space

Type of Space	Dimension	Area					
General Purpose /Dining Area (0.3 m ² per pupil. Minimum 100 m ²)	As per schedule						
Kitchenette and Store		25 m ²					
Project Storage	As schedule						
General Storage (<350 students)		20 m ²					
General Storage (for 350-699 students)		40 m ²					
General Storage (for >= 700 students)		60 m ²					
Lockers (9 m ² per 100 pupils)	As schedule	90 m ²					
Toilets (20 m ² per 100 pupils and staffs)	As schedule	200 m ²					
Electrical Switch Room	As schedule	10 m ²					

The area of toilets is based on an allocation of 20m² per 100 persons including staffs and to include universal disabled person's sanitary suite with shower.

2.4 Physical Education Space

Type of Space	Dimension	Area
Smaller PE Hall - Court size 24 m x 13 m(200 to 499 students)	15.6 m x 26.0 m	406 m ²
Larger PE Hall – Court size 28 m x 15 m (For 450 + students)	18.0 m x 33.0 m	594 m ²
Changing Rooms	2 x 32 m ²	64 m ²
Showers (2 x 4 NO. With cubicle dividers)	2 x 6 m ²	12 m ²
Toilet Cubicles (opening off Changing Room)	2 x 2 m ²	4 m ²
Toilets (off main circulation) (1male+1female+Universal including shower 5m ²)	(2x2 m ²) + 5 m ²	9 m ²
Control Centre (includes Teacher's facility)		9 m ²
Equipment Store		20 m ²
General Store		18 m ²
Circulation & Internal Division (including foyer stairs & future lift)		64 m ²
Fitness Suite + Store + Universal access WC	120m ² + 10m ² + 6m ²	136 m ²

Total: 1336 m²

2.5 External Space

Type of Space	Dimension	Area
External Store	10 x 5 m	50 m ²
Yard	Min 50 m ²	

Total: 100 m²

2.6 Car Parking

1 per full time staff member +3 per 100 pupils (For guest/visitors etc.). Total includes two or more spaces for the disabled. 6 additional spaces to be provided for 'stand-alone' PE Hall only.

2.7 Hard Play Area

Pupils	Number	Total Area
Less than 250	2	1020 m ²
2500 – 349	3	1530 m ²
350 – 499	4	2040 m ²
500 – 799	5	2550 m ²
800	6	3060m ²

Total: 10200m²

2.8 Soft Play Area

Playing Pitches	Playing Area
Gaelic Games	145 x 90 m ²
Soccer	90 – 120 x 45 – 90 m ²
Hockey	91.44 x 54.86 m ²
Rugby	144 x 69 m ²
Total: 13050 m ²	

2.9 Community Hall

Based on the above, the school requires minimum land area of 5 to 8 acres.

Note: the minimum finished floor to ceiling height for all teaching spaces are 3.15m.

Grouping of spaces are also required to be carried out, such as frequently used by visitors/ community should have access without entering general teaching area used by students. Class rooms used for special education such as ICT labs, science room, Arts room, etc. are proposed to be place on upper-floors.

3.0 CONCLUSION

With the above guide the developed Tamil schools will have clean, dynamic and spaces environment for healthy learning. Also by minimizing the building height teachers and students will not undergo health risk such as fatigue, muscle strain, back pain, distortion of the spine's natural curves and rounding of the shoulders. These health risks are mainly due to carrying heavy school bags up the stairs cases few floor high. It was reported that Malaysian Tamil School students carrying their school bag which is about 22.13% (saravanan, 2012) of their body weight much higher than recommended by Vall and Elimt (1997) which is 5% of body weight.

Further to health related issues are the used of Asbestors roof which could cause mesothelioma or asbestosis which are deadly. Hence it should be one of the priority to replace the old Asbestos roof with clay tile roofs. With healthy environment in school, the children will be able to focus in their education.

Standardization of the Physical Infrastructure and the implementation of ICT labs in the schools, we can hope for effective teaching of Tamil and greater growth of Tamil language.

Typical development layout plan for Tamil schools are as shown below.

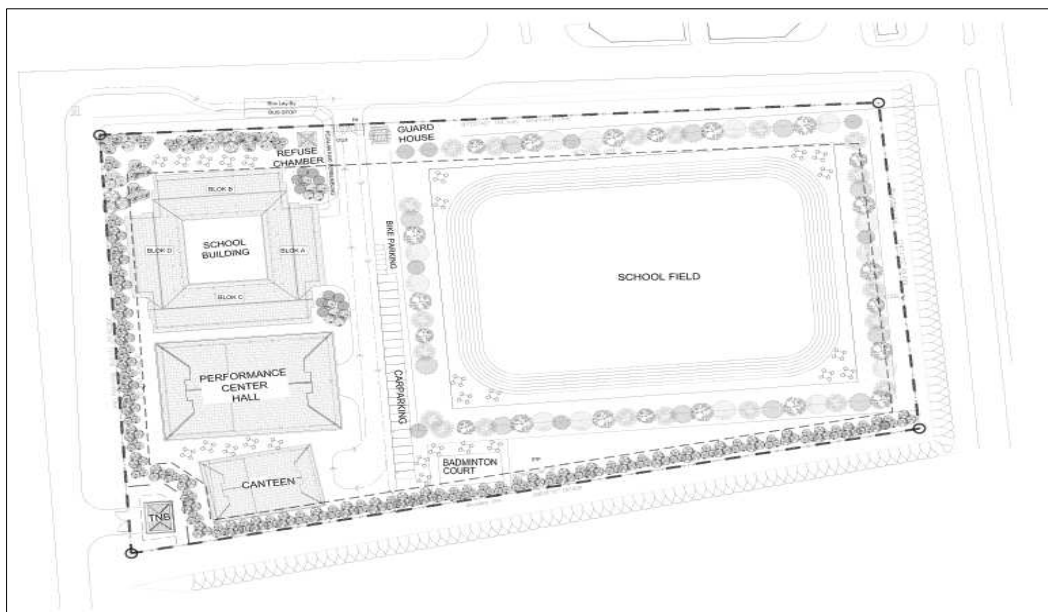


Figure 1 : Typical school development layout plan.

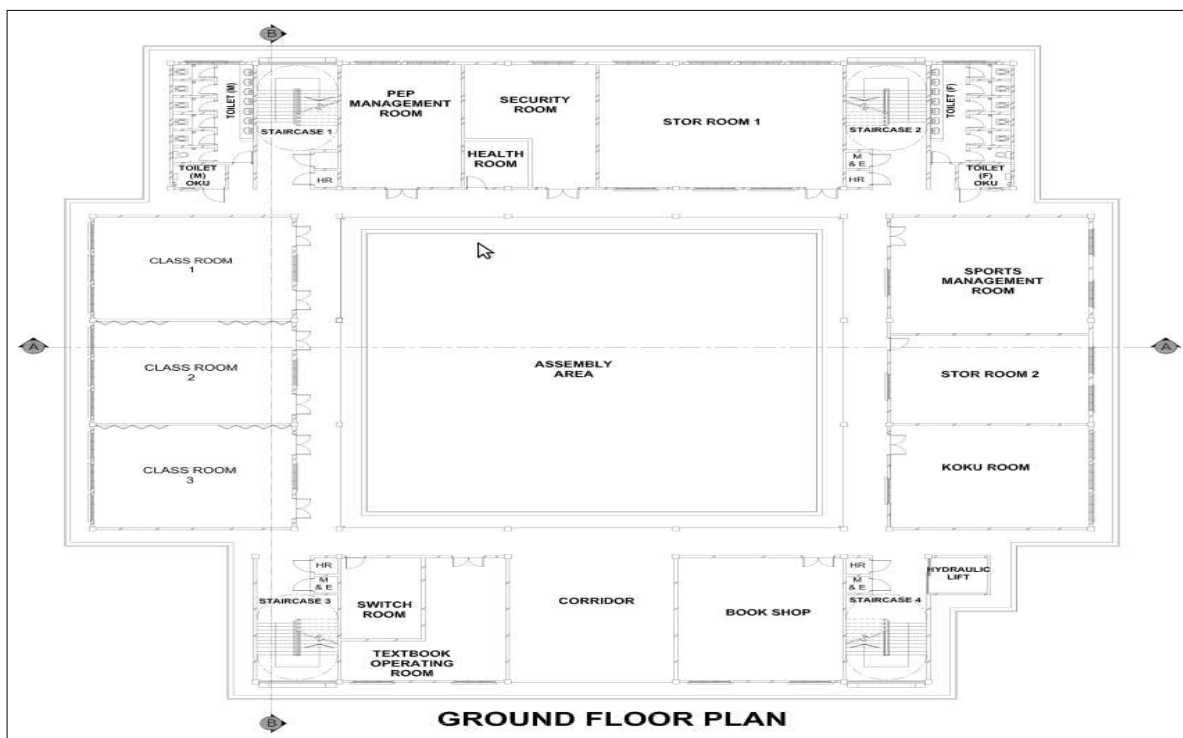


Figure 2 : Typical ground floor plan

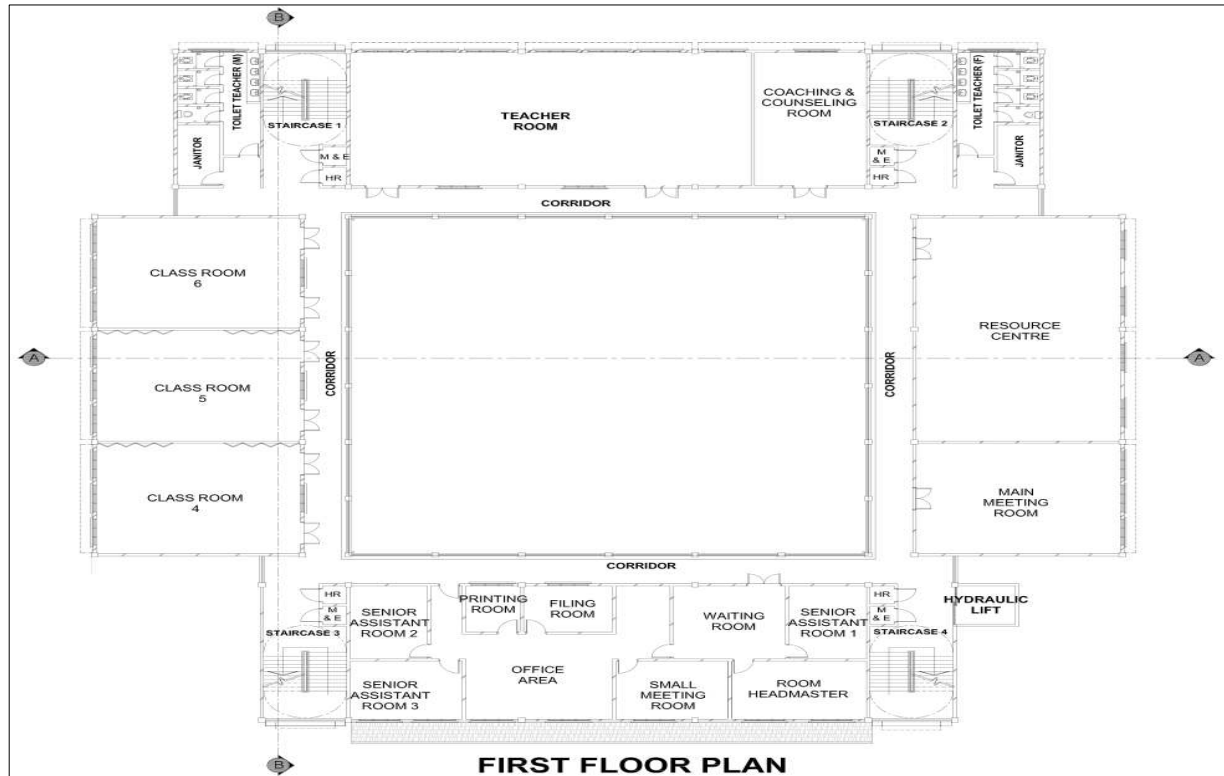


Figure 3 : Typical first floor plan

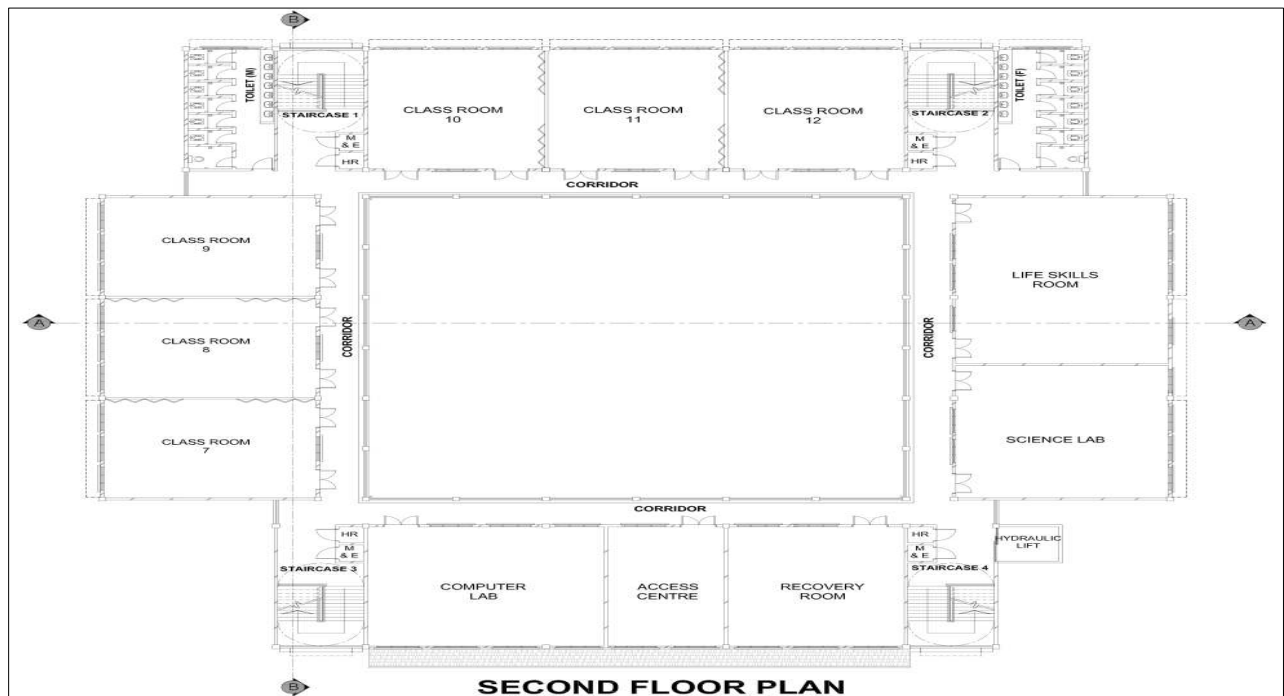


Figure 4 : Typical second floor plan

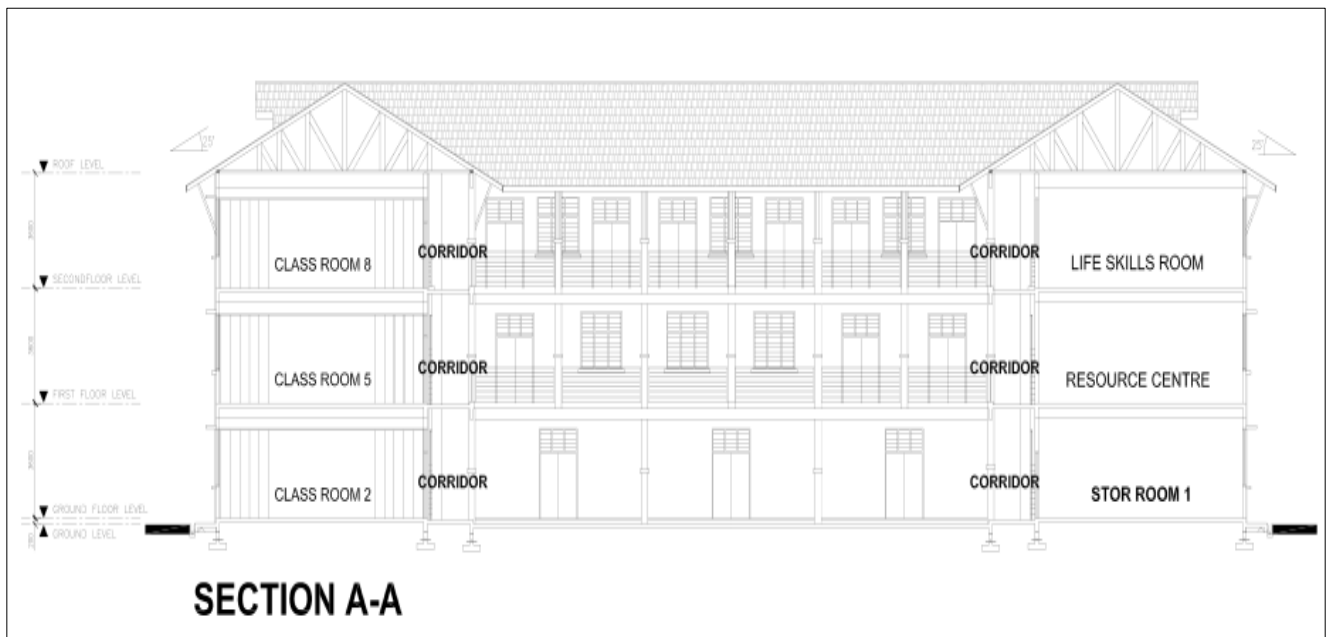


Figure 5 : Typical cross section of the school Block.

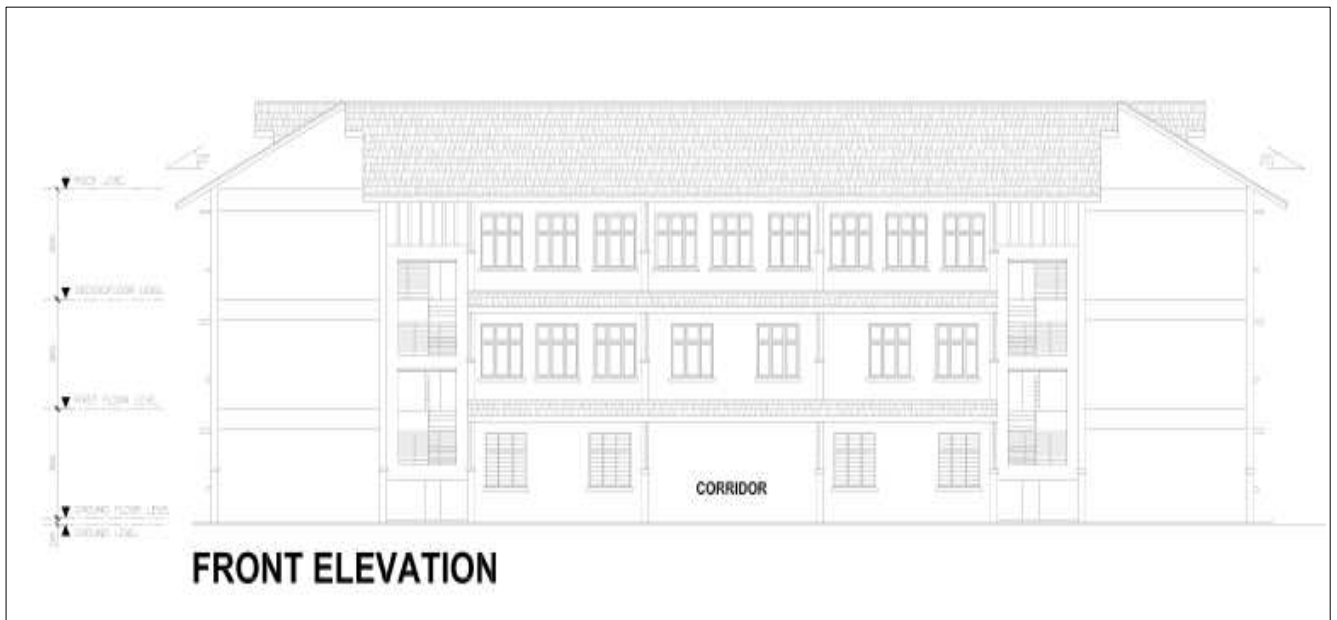


Figure 6 : Typical Front View of the school.

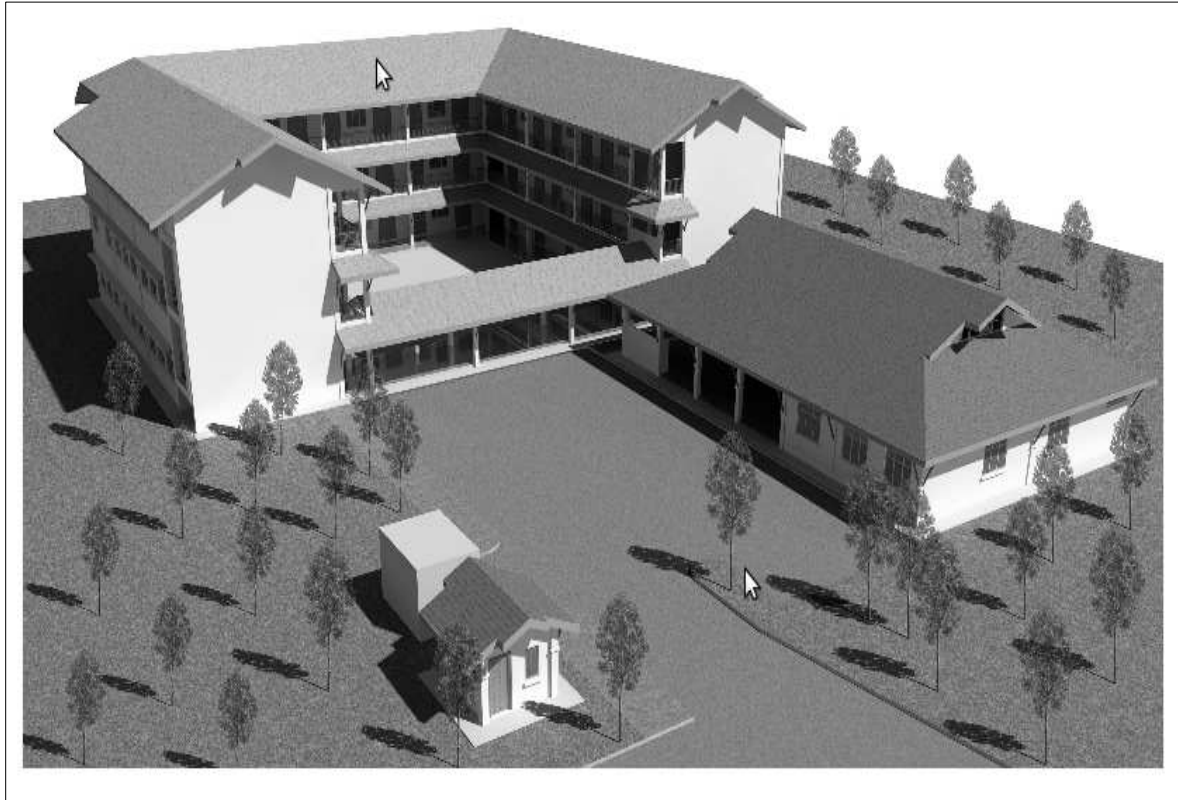


Figure 7 : 3D overall view of Typical Tamil school to accommodate 600 – 800 students.

Sustainability of Teaching Information and Communication Technology (ICT) Skills in Malaysia Tamil Schools using Tamil Open Source Computing And It's Social Impact

Gunasegaran Kandaswamy

*Program Titian Digital, Hindu Youth Organisation Port Klang, Malaysia – Malaysian Community Education Foundation
kkvguna@yahoo.co.uk*

Program Titian Digital – Digital Bridge Programme

Introduction

Information and Communication Technologies (ICT) provide a great development opportunity by contributing to information dissemination, providing an array of communication capabilities, increasing access to technology and knowledge and making teaching and learning more interactive and effective.

However, access to and the cost of ICT continues to be a major obstacle in schools, particularly in the developing countries. In Malaysia, most of the Tamil schools are partially government aided schools and they are located at rural and peri-urban areas, where these areas are home to a great majority of underserved communities. Here, the problems are even bleaker, where the infrastructure is either old or non-existent.

As ICT becoming the crucial means for improved economic conditions, a holistic ICT programme is most needed particularly in tackling those young children who are neglected by macro policy who reside mostly at rural and peri-urban areas. Understanding the need, the **Hindu Youth Organisation Port Klang, Malaysia (HYO Port Klang) & Malaysian Community Education Foundation** launched the **Program Titian Digital (PTD)** in April 2009 by setting up sustainable ICT Labs in Tamil Schools, using open source computing in Tamil and thin client technology which are reliable and cost effective.

Implementation

(i) Identify potential participating schools or receive applications from potential schools.

- after identifying or receiving applications, meetings are held with the School Administration, Parent Teacher Association (PTA), School Board of Governors (SBG) to explain the guidelines and conditions to be followed by the schools.

(ii) Identify sponsors/funders

- Sponsors/funders can be identified from various party (eg: Federal Government, State Government, Foundations, Cooperates, School Board of Governors, Parent Teacher Association and philanthropies)

(iii) Setup ICT Labs

- Labs are set according to the highest number of students in a class. The lab with tiles flooring includes server, terminal, networking, LCD projector, screen, internet service, tables, chairs & air-condition.

(iv) Formation of ICT Development Committee

- This Committee consists of Headmaster, Senior assistance of Curriculum, ICT Coordinator, PTA Representative, SBG Representative and PTD Representative. They will ensure the successful implementation of the programme.

(v) Conduct ICT Classes

- This classes are taught by a tutor employed by the PTA. Students will be attending the ICT Literacy classes minimum once a week (1 hour). This lab is also used to conduct ICT based teaching (eg: teaching of Science)

(vi) Training of Tutors

- Tutors attend compulsory training once a month. The training includes Module Content, Pedagogy and Technical.

(vii) Mentoring of Tutors

- Tutors will be observed and guided regularly by the school administration and periodically by pedagogy experts.

(viii) Development of Curriculum, Tutor's Guide & Students Module

- A team of experts consist of IT experts, Senior ICT tutors, Educationist and Textbook Officers are involved in the above production.

Besides the above tasks, in order to ensure sustainability of the programme, the PTD team also carries out the following:

- (i) Conduct ICT Coordinators meeting twice a year to discuss on the performance of each ICT Lab & issues related to the implementation of the programme.
- (ii) Annual centralise ICT Development Committee meeting. This meeting is to share experience, issues and give feedback and suggestions for the development of PTD.
- (iii) Conduct hardware auditing twice a year to ensure the labs are functioning well.
- (iv) Conduct annual accounts auditing to ensure the financial account are kept well.
- (v) Ensure that the PTA is financially sound. This is to make sure that the PTA is able to pay the salary of the tutor, maintenance are done regularly and upgradements are done when necessary.
- (vi) Empower the local community to support the school to sustain the ICT Labs.

Program Titian Digital is proud to announce that it will unleash the ICT potential of 35, 000 students that will undergo a minimum of an hour ICT training every week besides teaching and learning via ICT at 100 Program Titian Digital ICT Labs throughout West coast of Peninsular Malaysia.

Note: There are 523 Tamil Schools (of which 371 are partially government aided schools) with 103, 000 students in Malaysia

The Project's Social Impact

The success that PTD has achieved in ICT-based education can mainly be attributed to the fact that we have brought the technology and tools at cheapest rate as possible by applying open source and thin client technology. We have setup ICT labs which is comparatively 30% cheaper than conventional lab set up in Malaysia. The cost effective solution become one of the main attractive point of PTD for our funders and even school management to participate in the programme. The project which started in 2009, will be managing **100 schools** nationwide by the end of 2013 compare to our 5 years target to reach out **65 schools**. The achievements will be 35% higher than initial target.

Program Titian Digital holds The National Record in Malaysian book of Record as Largest Open Source User. It became the eye opening among Malaysian especially students, teachers and the general public on the availability of open source software. Even the Ministry of Education of Malaysia just announce last year to adopt the open source software to their ICT programme. We believe PTD remain as a leading open source user especially in ICT teaching and learning in Malaysia which is a breakthrough project initiated by a non profit organisation to reach the community, especially pupils from economically disadvantage society.

PTD has taken a four-prong approach to implement a successful and effective ICT education programme in schools. The approach is consists of the following ;

1. Develop and distribute free materials (teachers & students modules and open source digital contents in Tamil which is the pupils first language).

Program Titian Digital works with subject experts to develop ICT teaching guide and student guide in Tamil. We have successfully published level 1, 2 & 3 Student module (guide book).

2. Prepare teachers on effective integration of ICT in classroom teaching.
3. Build local capacity to monitor, support and sustain the programme.

4. Carry out research and development of appropriate technology and network infrastructure

The above approach enables PTD to create a greater impact on teaching and learning among teachers and pupils in Malaysian Tamil schools.

The PTD ICT programme's key benefits are as follows;

- a. Promotes learning by doing approach.
- b. Enriches learning through a combination of audio, video, images, text and animation
- c. Provides a platform that engages students

The above can be proven with the attendance to PTD-ICT classes, where in average 90% attendance is recorded during ICT classes compare to 75% during conventional classes. Schools with the lab have shown significant growth in student enrolment which school administration declares due to the ICT facilities.

PTD strongly believe that teachers/tutors hold the key to the success of any programme at the schools. The initiative able to create impact in teaching and learning process with new approach of teaching is because of customised ICT curriculum and training for the teachers by pedagogy and teachers trainer experts.

PTD has worked in partnership with various stakeholders to ensure that they build necessary institutional capacity to implement ICT education. Such partnership have also underscored the important roles that various experts within the system in successfully expanding new medium of teaching and learning and integrate teaching of other subject using the ICT.

PTD constantly engage all the stakeholders and opted to work with existing resources within our education system. We believe that such an approach will help scale the programme in the future.

PTD also contributed to small level economic activity in the economically backward community by creating job opportunity to about 100 peoples to become computer tutors.

The minimal fees collection by Parent Teacher Association (PTA) from lab users is contributing to PTA fund which primarily use for lab maintenance and ICT activities.

PTD encourages other community members especially the school teachers, alumni and parents to utilize the ICT labs to enhance their knowledge and skills & also to promote the PTD – ICT labs as a community tele-centre.

Students undergoing PTD-ICT literacy programme will be evaluated at end of each level and a certificate of achievement awarded jointly by University of Science Malaysia - National Advanced IPv6 (USM NaV6) Centre, a premier centre in the area of Next Generation Internet.

Proactive Location Based Services In Tamil

¹Karthikeyan K ²Vinayak U S M P ³Vijayalakshmi M

Department of Information Science and Technology, College of Engineering,

Anna University, Guindy, 600025, Chennai, India

¹ karthiking777@gmail.com ² awesomevinayak@gmail.com ³ vijim@annauniv.edu

ABSTRACT: In present day, Location-Dependent Information Services in native language is a basic necessity among mobile users. Mobile phone users have started using their phone not only to communicate but also to access the information they need from every place at anytime. Now the prime challenge is to provide these location based services proactively in the user's own native language. In this paper, we develop an android based mobile application to provide location-dependent and context aware services proactively, which are always relevant to the user's context to provide a complete satisfaction of Location-Based Services in Tamil. The services are based on the location of the user, the activity that he is involved in at a particular instance (based on user's daily planner) and the time at which the data is retrieved.

Keywords: Location based services in Tamil, Location Dependent Queries, Mobile computing, and Mobile communication.

I INTRODUCTION

Mobile phones have become a common and an unavoidable device to the people all over the world. Over six million people in the world are using mobile phones and the count is increasing day by day. Among the mobile phones, Android operating system which was developed by Android Inc. is fast developing and more appealing to the users than any other mobile OS. Android Smart phones had the features of Personal Digital Assistant, compact digital camera, GPS facility etc. More than just communicating, smart phones have another side- which provides the Location Based Services to the user. In this fast moving world, using the applications in android, people can find the directions to any place, just by sitting in any location through mobile phone. For example, we can find the route from your college to your home. The corresponding application neatly list outs the direction to walk. Also, several day planner android applications are available in English. They help the user to plan their day to day activities and get a remainder at that time. But the most of the application available presently are available currently are in English. This is a biggest disadvantage as a common man who could not understand English cannot use the application. So the biggest challenge for us is to develop an android application for an android phone to provide the Proactive Location Based Services in Tamil language to the users. It is discussed detail in this paper.

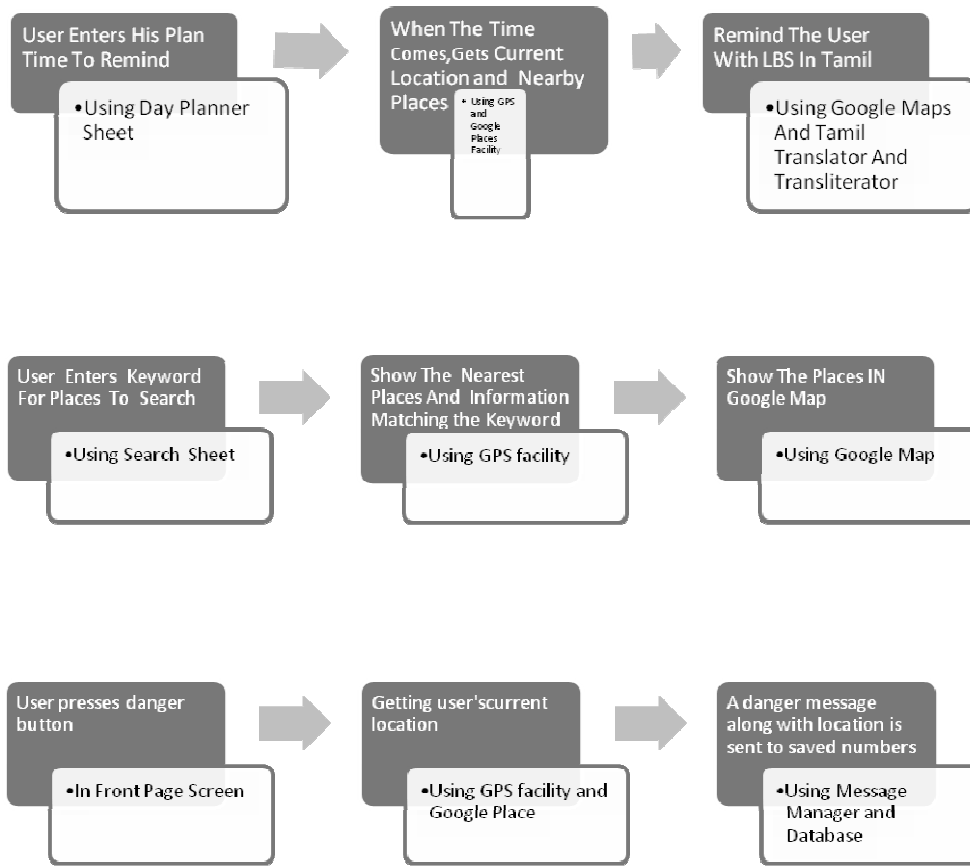
II RELATED WORKS

Several features are available in smart phones such as using Global Positioning System (GPS), several mobile applications, compact digital cameras etc. The use of GPS to locate the current position of the user using satellites in mobile phones has been described briefly in the paper [5]. Since the GPS facility has to be turn on, people worry about the battery life of the phone. The energy efficient position tracking is discussed detail in paper [3]. The Location Based Services based on GPS and Google maps are provided to the user via applications helping the user to know their exact locations. It is discussed in detail in paper [2], [4] and [6]. Also there are certain applications which use the Location Based Services to help the users during emergency. It is discussed in paper [1].

The shortcoming of all the above models is that none of them were implemented in the user's native language. They were all provided in the global language English. We wished to overcome this shortcoming. So the android application which we developed provides "Proactive Location Based Services in Tamil" to the user. Translation from English to Tamil is discussed in paper [7]. This application is the real need for the people of Tamil Nadu as smart phones usage is at its peak among the people of Tamil Nadu.

III SYSTEM DESIGN

The process that happens within our system is as follows: The user enters his/her activity he wishes to perform in Tamil. We search for the key term in the activity mentioned and we start the service. Once the specified time, mentioned by the user comes, the current location of the user is obtained by the use of Global Positioning System (GPS) facility available in the mobile and the key term is passed as a query to the Google Maps and the nearby locations that match the domain specified by the user is retrieved. The retrieved queries are transliterated into Tamil language and delivered to the users. Thus Proactive Location Based Services are provided to the user in Tamil Language.



When the users press the danger button, a text message is sent to the people whom the user has stored in our application that he/she is in danger.

IV IMPLEMENTATION AND RESULT

“A successful person is one who manages his activities”

In the android application which we develop, there is a day planner which the user can use to manage his day-to-day activities. For example, the user can enter that he wish to watch a movie and set the time as 4 pm. The application gives a remainder to the user at 4 pm reminding him to watch movies and also provides the nearby cinema theatres where he can watch the movies. It is achieved by the help of using Global Positioning System (GPS) facility available in the mobile phone and using Google maps. All these services are provided to the user in his native language, here taken as, Tamil.

In the android application which we develop, we provide three options to the user:

1. The day planner (With Remainder)
2. Search
3. Help during danger

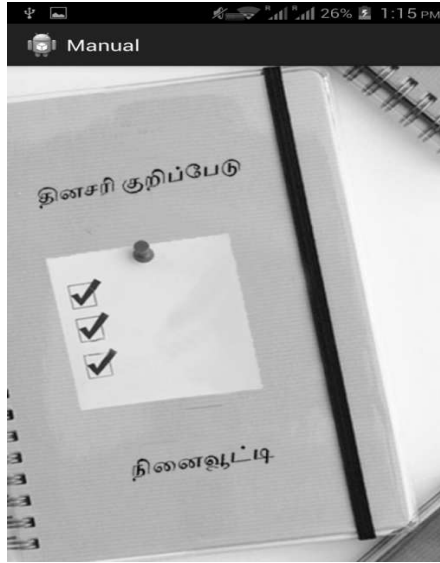


Figure 1



Figure 2

A. THE DAY PLANNER

In the busy world, we may forget any small job; say like buying a pen, due to other pressure or tight job schedule. It would be helpful for any one of us to remind us on our daily activities. It helps us to keep track in time and manage our day-to-day activities. Planning the activities make a one perfect. The 'Day Planner' facility in our android application helps the user to serve this cause. As can be seen in figure 3, when he/she enters the screen, a database is maintained that knows whether the user has already set the plan or he enters freshly. Based on that if he enters freshly he is redirected to activity mentioned in Figure 4. Thus he can set his plan and time here. Also the user can enter the distance range within which he needs the result. It commences the background service of the application.



Figure 3

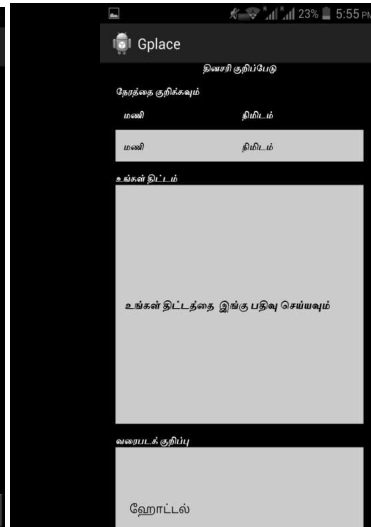


Figure 4

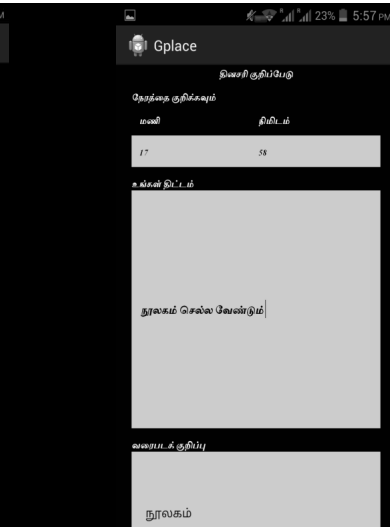


Figure 5

When the time set by the user has come, the android application provides a notification message to the user along with alarm notifying the activity which he/she has planned. Additional to these using LBS, the application lists out the nearby location based on the users requirements. If the user has set "To buy a pen" in Tamil at 10 am, then our application provides a remainder to the user and listing out the nearby stationery shops nearer to the user at his current geographical position. It also provides the additional functionalities such as providing the direction from the user's current location to the spot they wish to

go. It is done using Google Maps. We also provide options for the user by which he can call or text the spot which he/she wishes to go. It improves the communication facility. It is mentioned in Figure 6 to Figure 14.

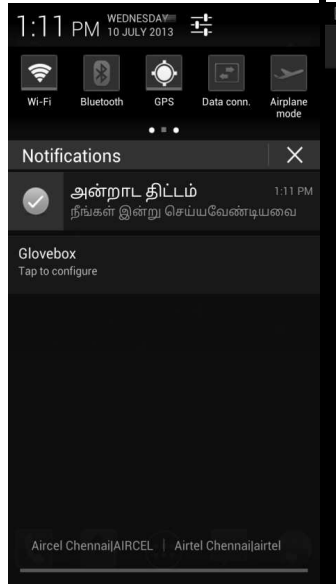


Figure 6

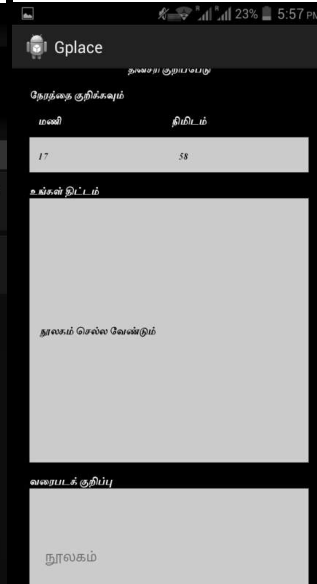


Figure 7

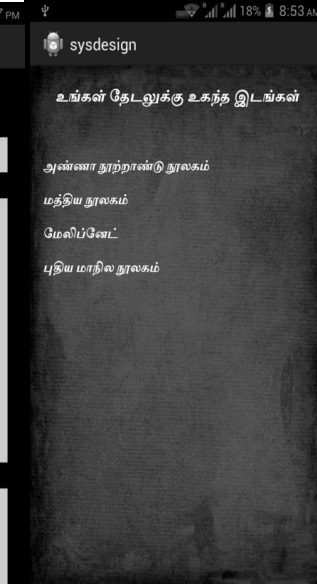


Figure 8

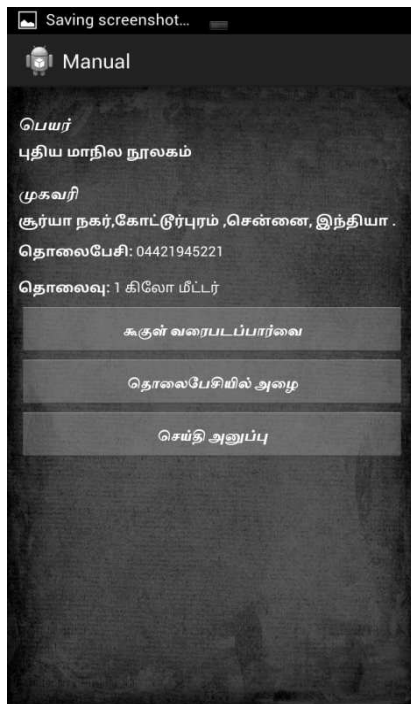


Figure 9



Figure 10



Figure 11

Four Different Views Of Google Maps (Satellite, Hybrid, Normal, Terrain Views)



Figure 12



Figure 13



Figure 14

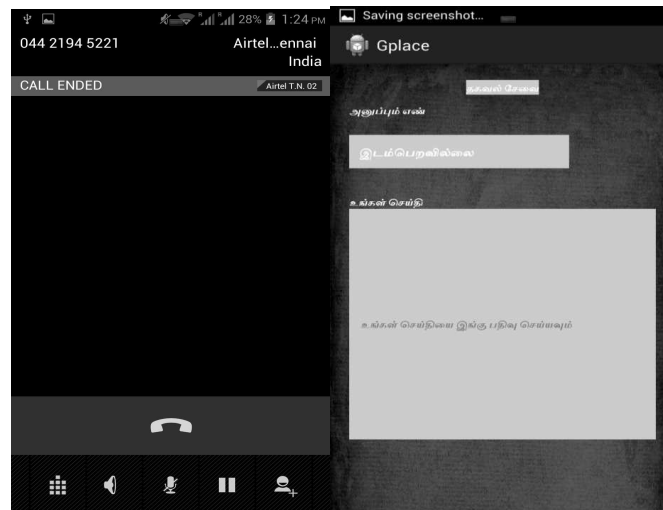


Figure 15

Figure 16

In case if the user had already set the activity, and then, based on the value from the database, he is redirected to the activity shown in Figure 17, where he can Edit his plan, View his Plan and even Cancel his plan

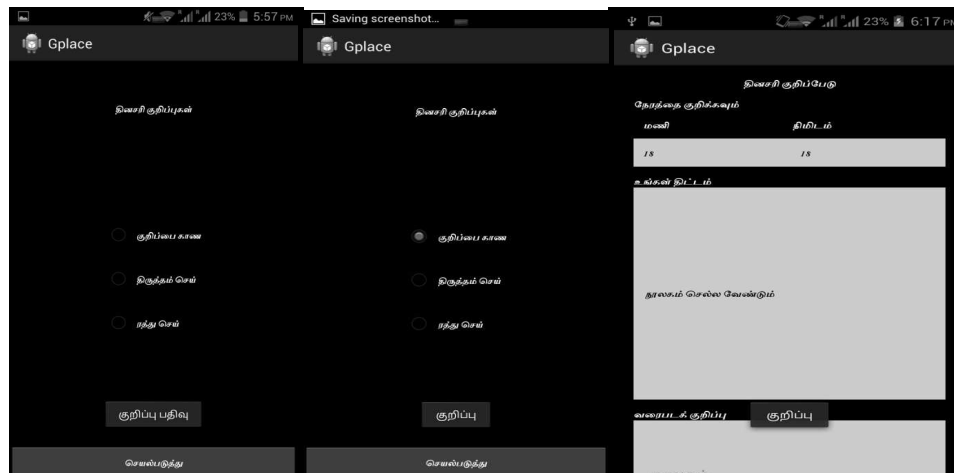


Figure 17

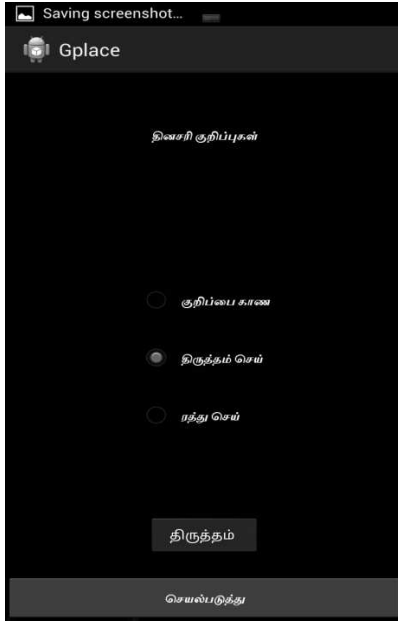


Figure 18



Figure 19

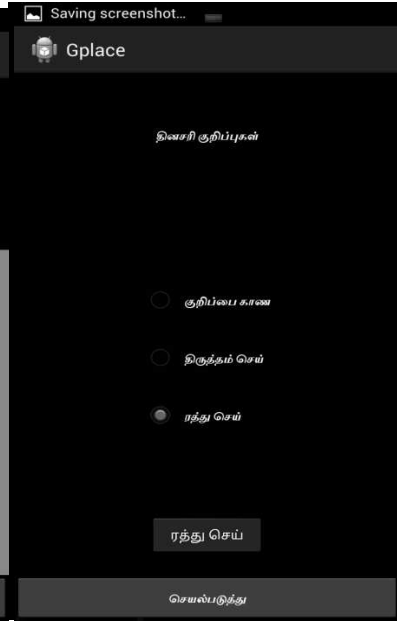


Figure 20

Figure 21

Figure 22

B. SEARCH

Instead of setting remainder and getting the locations by setting the day planner, the user may also come across a situation where he needs to search for some domain from his current location. It can be done easily using the search facility available in our application. The user can enter the domain he wishes to search and get the locations nearby which matches to the domain using the Global Positioning System (GPS) facility in his mobile. The domain options we provide to the user are restaurant, educational institutions, libraries, parks, stationeries etc..., Using this way the user can easily get direction to the spot he/she wish to go using Google Maps, contact the spot directly via his mobile by making call or sending a text. It is described in Figures 23 to 28 and

Figures 8 to 16 (Shown above).

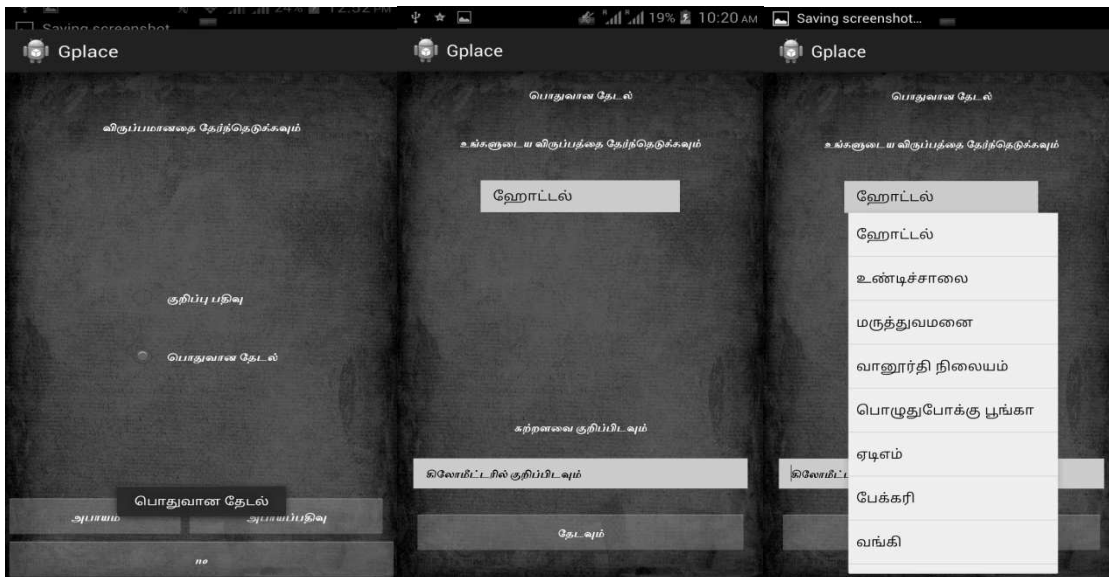


Figure 23

Figure 24

Figure 25

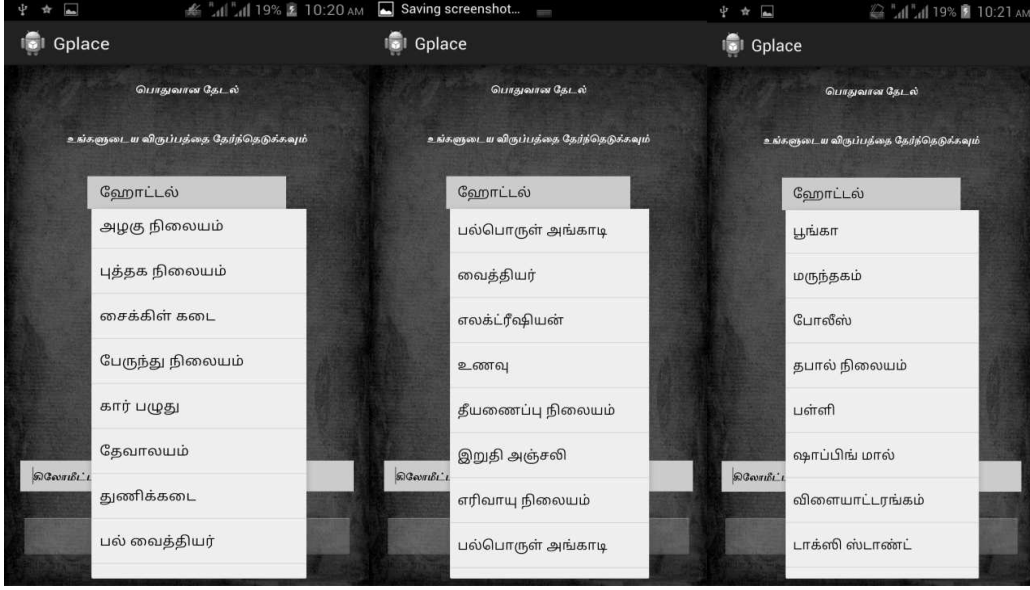


Figure 26

Figure 27

Figure 28

C. HELP DURING DANGER

Today's world is filled with danger and risk everywhere. People need to safeguard themselves in every situation. If there is an application available for them to ensure their safety, it would be a boon for him/her. Here the primary objective of this option is to provide safety at any time to the user by getting help from his close friends/relatives. The user can feel relaxed and safer at any place. Here, using the danger help option in our Tamil application, the user can save five mobile numbers of people which he wishes to contact during danger. When he/she presses the danger button (Available in the First Activity), a message is sent to all the five people and also to the nearby police station that he/she is in danger and also with the current location where he/she is. By receiving the message the message receiver can take any immediate action to help our application user, thus rescuing the user from danger. The user can also modify the mobile numbers he/she has set at any time. It is shown in figures 29 to 33.

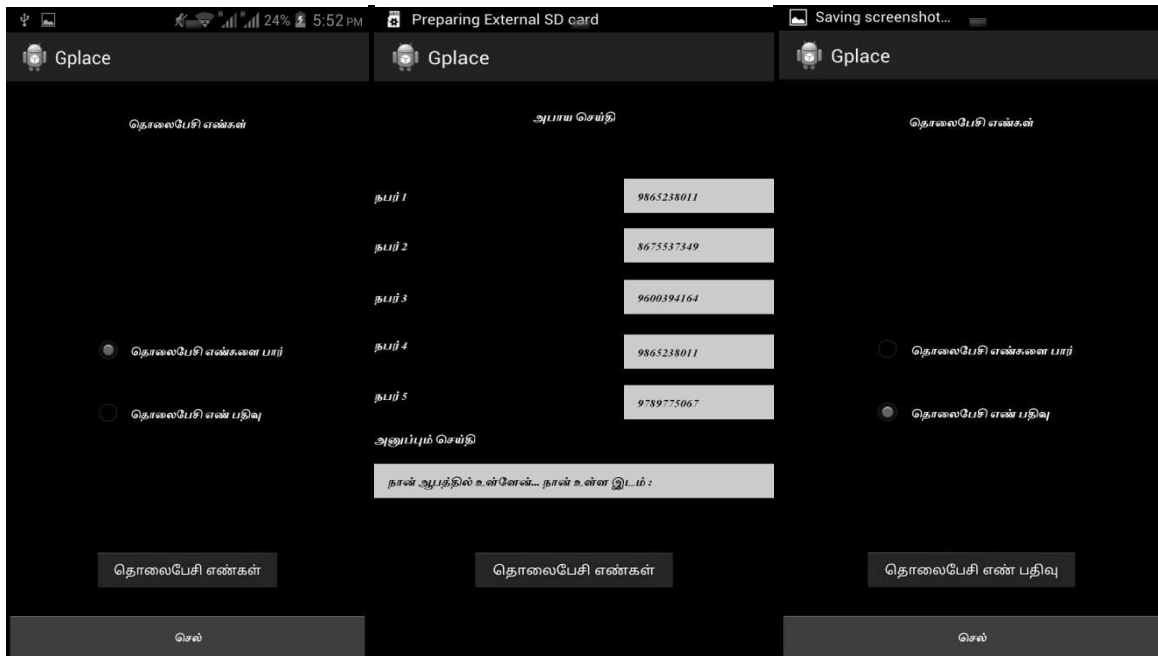


Figure 29

Figure 30

Figure 31

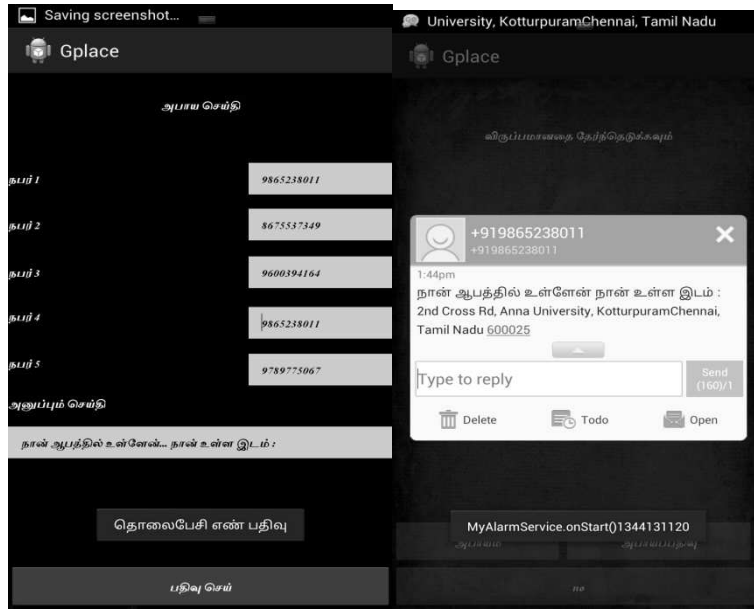


Figure 32

Figure 33

V CONCLUSION

Thus the project implements location based services proactively in the user's own native language-Tamil to help the people who are not much familiar with English. In future, this application can be developed to support multiple languages. By developing this application, it also increases the people's interest towards Tamil computing.

REFERENCES

- [1] Kamarudin, N.; Fac. of Inf. & Commun. Technol., Univ. Teknikal Malaysia, Durian Tunggal, Malaysia ; Salam, S. "Enabling mobile Location Based Services for emergency cases" published in Research and Innovation in Information Systems (ICRIIS), 2011 International Conference
- [2] Krishnamurthy, N. "Using SMS to deliver location-based services" published in Personal Wireless Communications, 2002 IEEE International Conference
- [3] Bareth, U.; Deutsche Telekom Labs., Tech. Univ. Berlin, Berlin, Germany; Kupper, A." Energy-Efficient Position Tracking in Proactive Location-Based Services for Smartphone Environments" published in Computer Software and Applications Conference (COMPSAC), 2011 IEEE 35th Annual
- [4] Kumar, S.; Dept. of Computer. Eng., Aligarh Muslim Univ., Aligarh, India; Qadeer, M.A.; Gupta, A." Location based services using android (LBSOID)" published in Internet Multimedia Services Architecture and Applications (IMSAA), 2009 IEEE International Conference
- [5] Kelly, Damian ; Department of Electronic Engineering, National University of Ireland, Maynooth, Ireland ; Behan, Ross ; Villing, R. ; McLoone, Sean published in "Computationally tractable location estimation on Wi-Fi enabled mobile phones" published in Signals and Systems Conference (ISSC 2009), IET Irish
- [6] Woo-Jin Choi ; New Jersey Inst. of Technol., Newark, NJ, USA ; Tekinay, S. "Location based services for next generation wireless mobile networks" published in Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual
- [7] Suganthi, S.; Comput. Sci. & Eng., Nat. Eng. Coll., Kovilpatti, India ; Bamarukmani, P. ; Srinivasagan, K.G. ; Saravanan "Semantic based orthographic with prepositional phrase for English-Tamil translation". published in Intelligent Human Computer Interaction (IHCI), 2012 4th International Conference

Social Interaction and Content Sharing in Mobile P2P Environment using Tamil Language

L.Libin Lougine and M.Vijayalakshmi

Department of Information Science and Technology, College of Engineering Guindy, Anna University, Chennai
slibinlougine@gmail.com, vijim@annauniv.edu

Abstract- Adaptation of mobile phones from simple phones and SMS capable devices into smartphones made Mobile Peer-to-Peer communication an easily achievable one. In recent year's smartphones users have increased in large number and because of this impact, social interaction and content sharing is possible among those users. Using smart phones, users can share their opinion rating, information or comments about the nearby places, restaurants etc which may be unknown to other users. This paper proposes social interaction in Tamil language that allows the user to message, rate, share information and multimedia data to nearby smartphones when they are in range. Thus smart phone Tamil users can get useful information which makes the users to take important decisions and plan their activities. A protocol called Bluetooth Device-Discovery protocol is implemented for the discovery of smartphones in the environment that comes in range and to communicate between users. An efficient content sharing scheme in Tamil language is proposed in Mobile P2P smart phone social network by using Discover-Predict-Delivery protocol (DPD). Collaborative Filtering is applied to sort the attraction based on the rating.

Keywords- Mobile Peer-to-Peer communication, Bluetooth device-discovery, Discover-predict delivery, Social network

I. INTRODUCTION

Mobile Peer-to-Peer communication is an advanced technique in Mobile Computing that can be used in real time applications. The major limitations in this environment are user mobility, limited battery power of the device, scarce bandwidth and frequent disconnections. Mobile Peer-to-Peer communication is a one in which every node act as server and client. The main advantage of this Mobile Peer-to-Peer environment is that there is no centralized or distributed server to control the mobile devices. Hence mobile nodes can connect to each other whenever they think they need to connect and there will be no server shutdown, busy server, server unavailable kind of things. This Mobile Peer-to-Peer communications are easily achievable one in today's environment, because of the development and usage of new technologies such as Bluetooth and Wi-Fi.

Nowadays number of smartphone users has rapidly increased. Because of this increase in smartphone usage, content sharing becomes popular among smartphone users and Mobile Peer to Peer environment takes social networking into a new form. The main advantage of social networking in Mobile Peer-to-Peer environment is that there will be no server and so it is not required to upload or download the contents to or from centralized servers. Ad-hoc networks can be easily constructed using smartphones as they are equipped with various network interfaces, such as Bluetooth and Wi-Fi. But the connectivity between smartphones is expected to be disconnected frequently, due to the movement patterns of carriers and the signal propagation phenomena.

Torrents are the real-time best examples for the successful Mobile Peer-to-Peer environment. Facebook is the best example for social networking, but with servers. Introducing social interaction along with content sharing among smartphone users without using any centralized server is successful, if smartphone users co-operate with one another and communicate in their native language. Detection of mobile devices that comes in range is done by using Bluetooth device discovery protocol. Content sharing between smartphone users in Tamil language is done efficiently by implementing Discover-Predict-Delivery protocol. The proposed Mobile Peer-to-Peer system, employ mobile nodes which can send contents to another node, when those nodes are in range. Every node in range can receive the content from another node, only if both nodes are interested in communication. Communication between nodes for content sharing and social interaction is done by using Tamil language.

The rest of the paper is organized as follows. Literature survey done for this work is reviewed in Section II. Section III, describes the concepts and overall architecture of the designed system. The evaluations of the designed system are reported in Section IV. Section V concludes with a summary and a discussion about future work.

II. RELATED WORK.

In recent years, the existing idea of proximity sensors is used to share information with neighbouring users. The best example for above idea is the Hummingbird project, [10] which used RF-enabled GameBoy. It is a device to alert users when others are near so as to support social awareness and collaboration, thus social interaction made possible. Another one technique is SocialNet [14], in which mobile devices are equipped with RF-communications for the detection of neighbouring users. In MobiTip [4], user will share their history and other information in peer-to-peer environment, when they are met. Social Serendipity [15] is the advanced technique, which use Bluetooth technology. In this method users will carry a Bluetooth-enabled mobile phone that detects neighbouring users and then user will be triggered spontaneous face-to-face interaction with those who have similar user profiles. The two systems mainly used for listening music in mobile are tunA [1] and SoundPryer [13], but interestingly that system transformed into something called social experience. by synchronizing the listening to music between neighbouring devices. Another such system PushMusic [8] and PushPhoto [16], they are mobile music and also a photo sharing systems, where users automatically receive songs/photos that are already recommended by neighbouring players.

The last among these systems is the system that enables user to express their ratings/credits about the unknown places or attractions or point of interest in the environment. These rating/credits are exchanged among the smartphone users in a social networking manner. These networking are done through mobile peer-to-peer environment. In this social network [17] sharing is proactive, there is no need for the user to aware about the device that is crediting/rating out. The above described system is closely related to our work.

For content sharing in Delay Tolerant Networks [6], existing works use epidemic routing as the solution for the problems. Epidemic routing can achieve highest delivery rate and lowest latency rate. Mobile prediction algorithm [6] is to predict the contact information for smartphone users. Contact information tells about the mobile nodes that have higher possibility of connection with the host smartphones.

III. SYSTEM DESCRIPTION.

Mobile Peer-to-Peer environment use short-range wireless technologies for mobile communication. The range duration for these short-range technologies is from 10 to 100 meters.

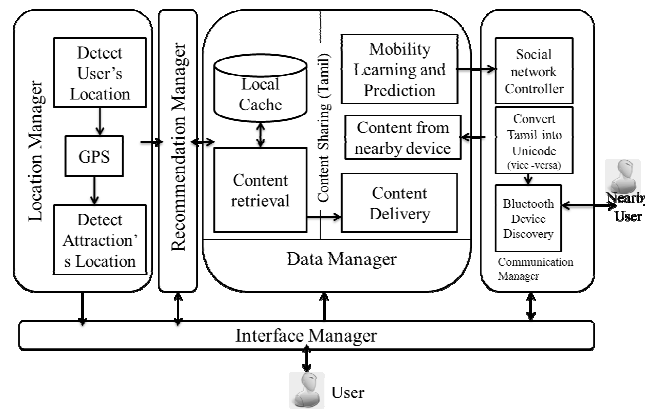


Figure 1. System Architecture

One peer will detect another peer, if their distance is smaller than their transmission range. As long as the two peers stay in the range, they can exchange information. Also peer will get information directly from its neighbours or indirectly from remote peer by multi-hop transmission technique. By using these techniques social interaction and content sharing is made possible in this type of environment. The overall system architecture for the proposed system is shown in Figure 1. This proposed system includes five managers such as Interface Manager, Location Manager, Recommendation Manager, Data Manager and Communication Manager that are described as follows.

E. A. Interface Manager

This manager allows the user to interact with the mobile device in Tamil and through this manager user can control all the other managers. User will control the managers by giving set of instructions, which may be differing from text [Tamil], graphics, and images to any other multimedia format. First step is user can enter the input as text format as well as other formats likes images video, graphics and so on. These inputs are sent to other managers like an instruction. Smartphone users can also use this manager to enter their credits/ratings information of an unknown places or attractions, which will be exchanged among the nearby users later. The architecture for interface manager is shown in Figure 2.

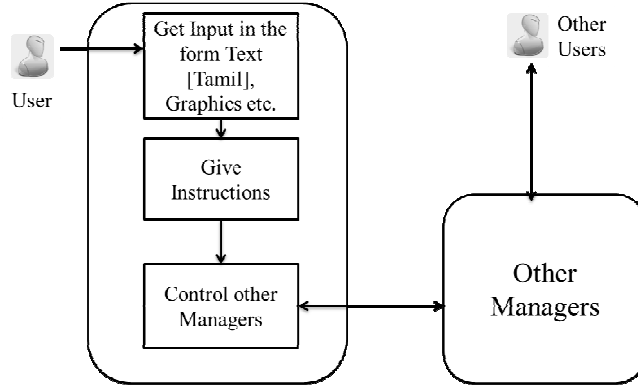


Figure 2. Architecture for Interface Manager

F. B. Location Manager

The Location manager will take care of location detection and tracking of other mobile nodes as well as for the fixed landmarks. For detecting user's location GPS is employed. This technology is available with all smartphones to identify the user's location. This manager communicates with Recommendation Manager and Interface Manager. User will not have direct contact with Location Manager. Detected location information of the host, nearby nodes and other attractions will be sent to the Recommendation Manager as and Interface Manager. Recommendation Manager uses this information for the rating purpose. Interface Manager uses the same information for visualisation purpose. The architecture for Location Manager is shown in Figure 3.

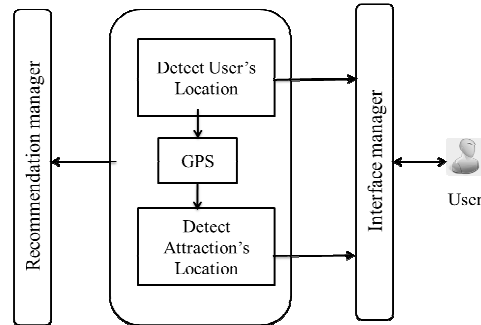


Figure 3. Architecture for Location Manager

C. Communication Manager

The Communication Manager uses short-range wireless technologies for Peer-to-Peer communication rather than broadband wireless technology. Notable Mobile Peer-to-Peer communication technologies include Bluetooth and Wi-Fi. Newly developed mobile phones are equipped with these Mobile Peer-to-Peer communication technologies. The simulation of this system is done using the Bluetooth technology; because of the cost and the availability of Bluetooth package in Java 2 Micro Edition.

The Communication Manager is responsible for the social network formation. This manager uses Bluetooth connections to

communicate with each other. This manager will communicate with Data Manager, Interface Manager and with other smartphone users. The Communication Manager finds out other nodes which have similar interest and group these nodes to form a social network. Then this manager converts the information received in Tamil into Unicode which will be a reverse for neighbor's content and stores the content. This network makes it possible for the connected smartphones to interact socially and also share the contents, opinions and ratings from Data Manager with other nearby users. For exchange of contents to other nodes, Communication Manager will search for nearby devices using Bluetooth device-discovery protocol. The architecture for Communication Manager is shown in Figure 4.

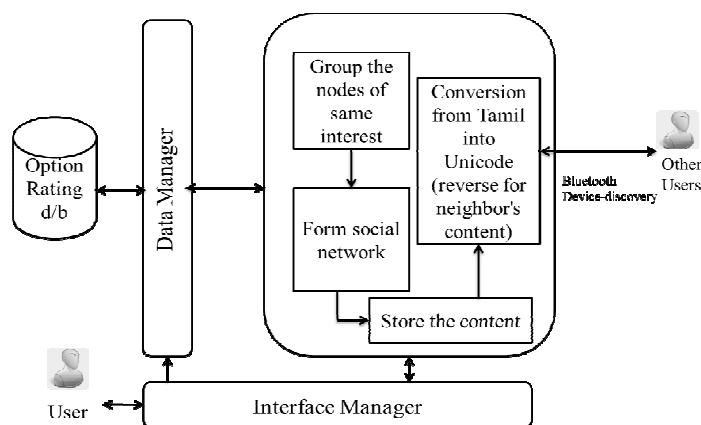


Figure 4. Architecture for Communication Manager

G. D. Data Manager

The Data manager communicates with Recommendation Manager, Interface Manager and Communication Manager. User will control this manager indirectly by controlling the Interface Manager. The Data manager handles user's own credits/ratings for attractions and also the credits/ratings given by other users. Data manager will store the rating data in the opinion rating database directly, in which the ratings about the attraction are saved during his/her visit. When two users are within the communication range and they are staying long time, then they may exchange some opinion ratings. The user's own rating list and the rating lists of other users obtained through data exchange are stored in the opinion rating database.

Data Manager shares the content in Tamil to other nodes through communication manager. This manager will convert each and every Tamil symbol into Unicode and vice-versa for the purpose of exchange. Content sharing is categorised into two phases, they are content retrieval phase and content delivery phase. In content retrieval phase, mobile nodes first search for the contents in its local cache. If content is discovered then content delivery phase will be initiated. In content delivery phase, it will deliver the content to the query sender. Content sharing is made possible by using Discovery-Predict-delivery protocol [6].

Mobile learning and prediction is the next technique used in this data manager. This method also consists of two phase, mobility learning and mobility prediction. Mobility learning is based on the mobile user, which stay long in certain areas and exchange information with nearby users. Mobile user's movement/path information is tracked for the mobility learning purpose. And in mobility prediction phase, it predicts the contact information of the user based on certain location. When a mobile user reaches any location they will check for some previous records about connectivity. Based on the previous record they can find the contact information of other users in their proximity and suggest this information to the user. Then the host users can choose the appropriate devices from the contact information for the connectivity. Discovery-Predict-delivery protocol is used for the prediction of the mobile nodes for the contact information. The architecture for data manager is shown in Figure 5.

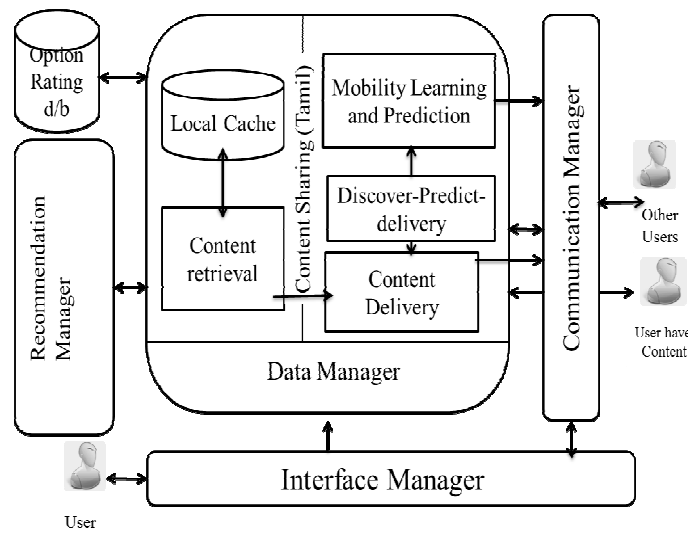


Figure 5. Architecture for Data Manager

E. Recommendation Manager

Recommendation manager has connections with Location Manager, Interface Manager and Data Manager. This manager will look after the opinion rating database for sufficient number of ratings. If the rating is sufficient in number then score for the unvisited attraction is computed and put in the list of rating table. It then collects all the credits/rating and sort the attraction based on the rating. This sorting is done using Collaborative Filtering algorithm.

Collaborative Filtering (CF) algorithm technique is used in this manager. This algorithm will use all collected rated value to form a sorted list of attraction based on rating. This sorted list of attraction will be send to the data manger later for the further process. Collaborative filtering has two way of character. They are narrow filtering mode and a general filtering mode. In general filtering mode, collaborative filtering will filter based on information or patterns. The architecture for recommendation manager is shown in Figure 6.

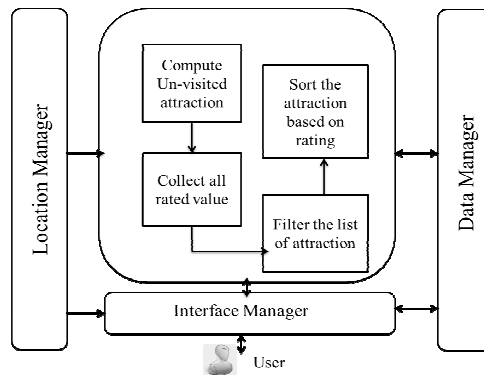


Figure 6..Architecture for Recommendation Manager

These five managers are combined together to form social interaction along with content sharing and mobility prediction in a Mobile Peer-to-Peer environment, where no centralised or distributed server is present.

IV. Evaluation.

Extensive experiments are done to evaluate the performance of this approach. These experiments show that this approach is good for smartphone users in real-time Mobile Peer-to-Peer environment. Evaluation is mainly concentrated on two important factors. They are discovery-delivery period ratio and accuracy analysis. The first parameter checks whether the content discovery and content delivery is proper with user's query. The second parameter checks the accuracy of requests satisfied in social networking. The following section discusses about the simulation setup and the evaluation parameters.

A. Simulation Setup

The experiments are modelled based on a typical scenario of smartphone users, which are moving and communicating between one another inside a city or in any other tourist spots. Smartphones are simulated using Sun's Java 2 Mobile Edition [J2ME] package. And the city is simulated in this J2ME emulator. Movement of every smartphones are simulated by varying the longitude, longitude, altitude values of each device emulator. J2ME has the option of varying the latitude, longitude, altitude values in external event option, which is in MIDLET menu bar. Random movement of the smartphones is simulated by adding a script in XML that can change the location and speed of the devices continuously.

Nearby devices to the host device are found by using Bluetooth package in J2ME emulator. GPS can be simulated using the longitude, latitude and altitude values that are mentioned above. Using these values we can fix the position of smartphones and also the location of the attractions. From this simulation setup social networking with content sharing is established.

B. Evaluation Parameter

Content Delivery Ratio and Accuracy analysis are the two performance metrics, used to evaluate the performance of this framework. They are describes as follows.

H. Content Delivery Ratio (α): The content delivery ratio is the ratio of the number of successfully received contents to the number of generated contents.

$$I. \alpha = ((T-Q)*q) \div T \quad [6]$$

J. Where 'T' refers to Time taken, 'Q' refers to Content delay and 'q' refers to Probability of nodes joined for content sharing.

K. Accuracy analysis: Total number of request satisfied through social networking within given period of time. The Accuracy analysis is computed as follows.

$$L. Aa = 1/n \sum Aa (Qi, t) / ((Ei-Si)) \quad [17]$$

M. Where 'Aa' refers to Accuracy analysis, 'Qi' refers to Number of active Queries, 'Ei' refers to Starting time of query and 'Si' refers to End time of query.

C. Result

Extensive experiments are done to evaluate the performance for this approach. These experiments show that this approach is good for smartphone users in real-time mobile peer-to-peer environment. Evaluation is mainly concentrated on two important factors. They are content delivery ratio and accuracy analysis. First one will check whether the content discovery and content delivery is proper with user's query. Second one will check about the accuracy of requested satisfied in social networking. The experiments are evaluated based on the two main ideas. They are as Content Delivery Ratio v/s Number of nodes and Accuracy analysis v/s No. of Participants for Social Networking.

Content Delivery Ratio V/S Number of Nodes

The content-delivery ratio is used in the forwarding decision process of content delivery. When a user requires receiving content in a given period (which is a query lifetime), a fraction of the lifetime is used for content discovery, and the remaining fraction is used for delivering the content. When the forwarding decision is terminated earlier (i.e., is higher), the sharing overhead is reduced due to a decrease in unnecessary query spread.

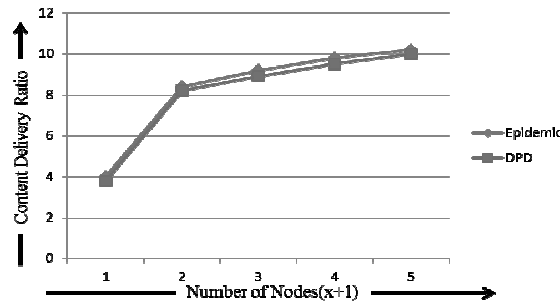


Figure 7. Content Delivery Ratio v/s Number of nodes

In this experimental result, accuracy analysis is evaluated based on the number of participants in social network. Figure 8 shows the values of accuracy analysis for epidemic algorithm and Delivery-Predict Discovery algorithm based on the number of participants in social network. It shows the improvement in the accuracy analysis value for DPD algorithm, when compared with epidemic algorithm.

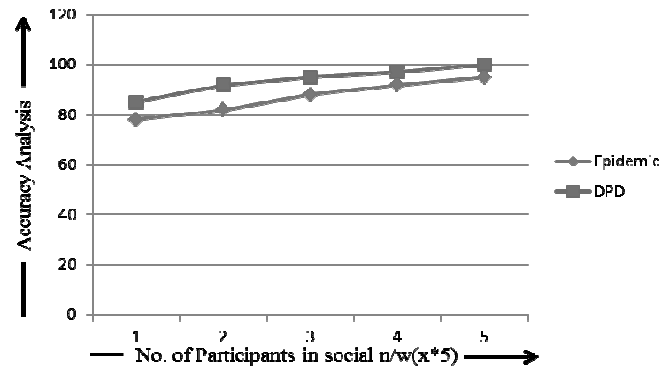


Figure 8. Accuracy Analysis V/S Participants in Social Networking

V. CONCLUSIONS

The proposed system performs social interaction along with content sharing using Tamil language in Mobile Peer-to-Peer environment. This system is made possible, because of the advantages of recent smartphones (i.e., availability of advanced wireless technologies like Bluetooth, Wi-Fi and GPS). The system allows the user to express their own opinions about the unknown places. This information will help the nearby user to know about their surroundings. Smartphones in one area can form a group and create a social network without any server's support. Smartphones can share contents as text or as image formats in their native language and as a result of social interaction, tourists will receive right information and contents at right time.

Future enhancement can be done in several directions. First, the present system allows only text format and image formats as contents to share among the smartphones during social interaction. These formats can be extended with more multimedia formats. Second, grouping of smartphones is only possible with nearby users. It can be extended to support more number of users; hence we can spread social networking in wider range. Finally, it would be perfect if some security policies are developed for content sharing.

REFERENCES

- A. Bassoli, J. Moore, and S. Agamanolis, "tunA: Socialising music sharing on the move," K. O'Hara and B. Brown (eds), *Consuming Music Together: Social and Collaborative Aspects of Music Consumption Technologies*, Springer: New York, 2006.
- A. Felfernig, S. Gordea, D. Jannach, E. Tepan, and M. Zanker, "A short survey of recommendation technologies in travel and tourism," *OGAI Journal*, vol. 25, no. 2, 2007, pp. 17-22.
- A. Garcia-Crespo, J. Chamizo, I. Rivera, M. Mencke, R. Colomo-Palacios, and J. Gomez-Berbis, "SPETA: Social Pervasive e-Tourism Advisor," *Telematics and Informatics*, vol. 26, no. 3, 2009, pp. 306-315.
- A. Rudstrom, M. Svensson, R. Coster, and K. Hook, "MobiTip: Using bluetooth as a mediator of social context," *Proceedings of the International Conference on Ubiquitous Computing*, 2004.
- D.J. Klein, J. Hespanha, and U. Madhow, "A Reaction-Diffusion Model for Epidemic Routing in Sparsely Connected MANETs," *Proc. IEEE INFOCOM*, Mar. 2010.
- Elmurod Talipov, Yohan Chon, and Hojung Cha, "Content Sharing over Smartphone-Based Delay-Tolerant Networks," *IEEE transactions on mobile computing*, vol. 12, no. 3, March 2013, pp. 581-595.
- European Travel Commission (ETC) and World Tourism Organization (UNWTO), *Handbook on E-marketing for Tourism Destinations*, World Tourism Organization, 2008.

- Hakasson, M., Rost, M., Jacobsson, M. and Holmquist, L.E. (2007) Facilitating Mobile Music Sharing and Social Interaction with PushMusic. Proceedings of the Hawaii International Conference on System Sciences, Hawaii, US.
- H.Cao, O. Wolfson, B. Xu, and H. Yin, "MOBI-DIC: Mobile discovery of local resources in peer-to-peer wireless network. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, vol. 28, no. 3, 2005, pp. 11-18.
- J. Paradiso, J. Gips, M. Laibowitz, S. Sadi, D. Merrill, R. Aylward, P. Maes, and A. Pentland, "Identifying and facilitating social interaction with a wearable wireless sensor network, " Personal and Ubiquitous Computing, vol. 14, no. 2, 2010, pp. 137-152.
- K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou, "Developing a context-aware electronic tourist guide: Some issues and experiences, " Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2000, pp. 17-24
- Ardissono, A. Goy, G. Petrone, M. Seqnan, and P. Torasso, "Intrigue: Personalized recommendation of tourism attractions for desktop and handset devices, " Applied Artificial Intelligence, vol. 17, no. 8-9, 2003, pp. 687-714.
- M.Ostergren, and O.Juhlin "Car drivers using Sound Pryer- Field trials on shared music listening in traffic encounters, " K. O'Hara and B. Brown (eds), Consuming Music Together: Social and Collaborative Aspects of Music Consumption Technologies, Springer: New York, 2006.
- M.Terry, and E.Mynatt, "SocialNet: Using patterns of physical proximity over time to infer shared interests, " Proceedings of the Conference on Human Factors in Computing Systems, 2002, pp. 816- 817.
- N.Eagle, and A.Pentland, "Social Serendipity Mobilizing social software, " IEEE Pervasive Computing, vol. 4, no. 2, 2005, pp. 28-34.
- R. Rost, M. Jacobsson, and L. Holmquist, "Push!Photo: Informal photo sharing in ad-hoc networks, " Proceedings of the International Conference on Ubiquitous Computing, 2006.
- Wan-Shiou Yang, San-Yih Hwang, Ying-Wei Shih "Facilitating Information Sharing and Social Interaction in Mobile Peer-to-Peer Environment, " 8th IEEE International Workshop on Pervasive Learning, Life and Leisure, March 2012, pp.673-678.

Android Based Mobile Application in Tamil for Danger Assistance

*Gowtham.N.R, Dhanraj.R, Lokesh Kumar.K,
Tamizhselvi S P, Vijayalakshmi M*

*Department of Information Science and Technology, , College of Engineering Guindy, Anna University, Chennai
nrg.gowtham92@gmail.com, dhanrajdhacks@gmail.com, lokeshshinealways@gmail.com,
tamizh8306@gmail.com, vijim@annauniv.edu*

Abstract - With the help of smartphones and their inbuilt sensors, Danger Assistance System (DAS) can be efficiently built for the users. The number of smartphone users is increasing drastically day by day. Therefore it can be considered that a significant amount of population are using smartphone and providing danger assistance to the users through smartphones will be really helpful and highly efficient. Google's mobile operating system, Android, has a large community of applications that extend the functionality of devices, written primarily in a customized version of the Java programming language. Since Android powered smartphone users are rapidly increasing, this work proposes a location based Android application in Tamil that provides danger assistance to the mobile users. This Android application will guide the user during his danger situations by providing safety related information. Moreover it also gives appropriate notifications to the caretaker when the user is in trouble and is not able to save on his own. This danger assistance and safety monitor system makes use of Global Positioning System to provide safety related location based services. This system is collaboration between the user, the caretaker and safety-providing agencies (hospitals, police etc.,) enabled by the power of smartphones.

Keywords: Danger Assistance, Android, Global Positioning System, Location Based Application

1. INTRODUCTION

Conventional mobiles are designed to make voice call between two users. But due to the revolution in electronics, the functionality of mobile phones increases. Services are extended like GPRS, and GPS in the mobile phones. Meanwhile hardware argument like camera and sensors like Accelerometer, Gyroscope, Digital Compass and Barometer increases the other mobile phones functionality. Also Touch screen brings a new era in the mobile world. Hence Developers look for a flexible, user-friendly operating system for mobile phones.

Android operating system is a one which is initially developed by Android, Inc., and later maintained by Google, Inc ., that became the backbone of smartphone development. Open source OS and the free availability of SDK leads to vast development of application.. This paper focuses on developing a mobile application in Tamil for providing danger Assistance to the smart phone users when they are in trouble.The user can develop this services are implemented in android smart phones to find the safer places, nearby places during danger situations..

There are two methods available to provide LBS. The first method deals how to process the location data received from client and get response from server. Another deals with position of mobile using LBS in real time. The proposed work provides Danger Assistance System for tamil language users. In case of emergency, the system must able to identify the location of the user and send photo to the caretaker. The caretaker processes the image and calls the numbers that are provided in the database. It also sent message about nearest hospital details and location.

II RELATED WORK

Google service known as Google Places API [6] which provides services like spatial location and preferred point of interest. The request is processed through HTTP request and coordinates is got through response as latitude and longitude.

The place details [7] gives more detailed description about the location address, street name, contact number etc. If the reference number and place is provided to search, it initiates rescue operation by alerting contact number stored in database. Location based Services [1] is implemented on Android based smart phones and provides the value-added services like advising clients of current traffic conditions, providing routing information, helping them to find nearby hotels.

Chris Thompson et al. has proposed a solution for using the on board sensors [2] to detect car accident using android mobile. It uses physical information to identify accidents. Kai-florian et al. [3] describes modular architecture [5] which focuses on

the outcomes of LBS for users. The user-centric model highlights the effects of differences in interaction for users. The process of informing the network about the current location of the mobile user is known as location update.

In paper [7] the author integrates the camera with the smart phone for face detection and for the safety of the people in home and ease control of home entrance. In the proposed work, we have designed the danger assistance system to provide location based service and sensor is activated during emergency situation.

Android application fetches the location and address of the user using GPS. The application deduces the mobility of the user with the help of accelerometer. The location and mobility details of the user are updated periodically to the database. During danger situation, the application provides user, the nearby safe places and their details according to their preferences. If the user is in extreme danger and he is unable to use the mobile, the application senses the danger and takes a picture automatically and sends it to the server for image processing. The caretaker will call the emergency number automatically and forwards the image to the number stored in the database. Thus the rescue operation is initiated.

III SYSTEM DESCRIPTION

Mobile application is developed which can be handled by the client. Since the application is developed to be used in danger situations, the user interface is designed in a simple and efficient manner. The application keeps track of the user's location by the use of GPS. Reverse geocoding mechanism is employed which is performed through Google Maps API. With the use of GPS and Google Maps API, a request is sent to the Google Maps server with those parameters and the response is received in XML format. User query is processed by the query processor and if it is in Tamil, the translator is used for conversion. The response is then parsed and then displayed to the user. Ambulances are also provided with smartphones and their locations are updated to the database. Sensors are used by the mobile to sense the dangerous situation automatically, captures the image and send to the server so that the caretaker will make emergency call from the database. The service provider is connected to internet and Google translator is used for converting English to Tamil and send the response to the application. The system architecture is depicted in Fig 1.

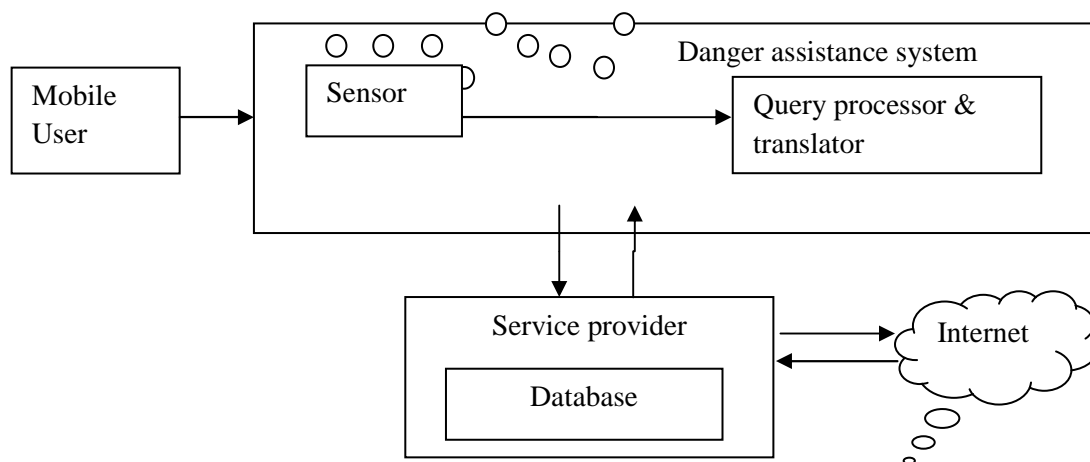


Figure 1. System Architecture

IV IMPLEMENTATION

Location based service is another key functionality that gets used in smart phone applications. It is often combined with maps to give a good experience to the user about their location. Location service allows finding out the current location of the device. The application requests for periodic update of the device location information. The application can also register an intent receiver for proximity alerts like, when the device is entering and existing from an area of given longitude, latitude and radius.

The proposed system has been implemented using Android SDK with the Eclipse IDE and AVD (Android Virtual Device) emulator. AVD is an emulator which provides the android hardware and software environment to test application on computer. In Eclipse, the application is run under Windows -> Android SDK and AVD Manager. The following functionalities are enabled such as Android Location API

These are the different classes present under Location API package to retrieve the Location information of the user. *LocationManager*- The class provides access to the location service. It also provides facility to get the best Location Provider as per the criteria. *LocationProvider*- It's an abstract super class for location providers. A location provider provides periodic reports on the geographical location of the device. *LocationListener*- This class provides callback methods which are called when location gets changed. The listener object has to be registered with the location manager.

Criteria- The class provides the application to choose suitable Location Provider by providing access to set of required properties of the LocationProvider. Android also provide an API to access the google maps. So with the help of the google maps and the location APIs the application can show required places to the user on the map.

V RESULTS

The location of the user is fetched through Global Positioning System (GPS) as in Figure 2. Since the application needs constant monitoring of the user's location, based on the distance of the user from his safe location, time intervals are fixed and the user's location is constantly monitored.



Figure 2. User Interface

The nearby safe places are displayed such as police station, hospitals and restaurants and he is allowed to choose the distance radius as in Figure 3.

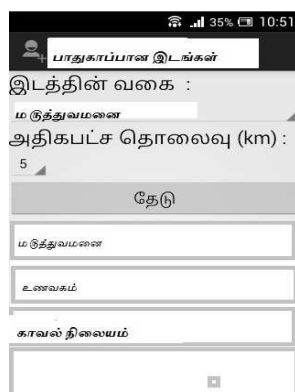


Figure 3. Nearest Safe Place Details

In addition to the safe place details provided to the user, the user is also provided with details of the nearby ambulances. Each ambulance is provided with an android device which updates its location in regular intervals to the server. The user can find the nearest ambulance by accessing the server in emergency situations as shown in Figure 4.



Figure 4. safe place details

VI CONCLUSION AND FUTURE WORK

This work implements an application that enables GPS and other sensors to sense the user's state and provide him appropriate details to come out of the danger situations. Also, it provides the details of the user's state to the caretaker which will enable the caretaker to take appropriate measures. This framework can be the fundamental for danger detection system for business use. The Android application is also developed in Tamil user interface for specific people who does not have English knowledge. The future work would involve implementing the same technique to improve effective modelling of multi-person interaction and with multiple cameras will be employed for capturing the same scene of different angles.

REFERENCES

- [1] Amit Kushwaha, Vineet Kushwaha, "Location-Based Services Design and Implementation Using Android Platforms", Proc. Of International Journal of Advances in Engineering & Technology, Vol. 1, Issue 1, pp.14-20, Mar 2011.
- [2] Chris Thompson, Jules White, Brian Dougherty, Adam Albright, and Douglas C. Schmidt, "Using Smartphones to Detect Car Accidents and Provide Situational Awareness to Emergency Responders", Journal on Mobile Wireless Middleware, Operating Systems, and Applications, Vol. 48, pp. 29-42, 2010.
- [3] Kai-Florian Richter, Drew Dara-Abrams and Martin Raubal, "Navigating and Learning with Location Based Services: A User-Centric Design", Proceedings of the 7th International Symposium on LBS and Telecartography, pp. 261-276, 2010.
- [4] Manav Singhal, Anupam Shukla, "Implementation of Location based Services in Android using GPS and Web Services", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 2, January 2012, pp 237-242.
- [5] Weilun Lao, Jungong Han, and Peter H.N. de, "Automatic Video-Based Human Motion Analyzer for Consumer Surveillance System", IEEE Transactions on Consumer Electronics, Vol. 55, No. 2, pp. 591-598, 2009.
- [6] GoogleMaps API <http://code.google.com/apis/maps/documentation/imageapis/index.html>
- [7] GooglePlaces API <http://code.google.com/apis/maps/documentation/places>.

Mobile Based Rural Road Monitoring System

*Jeyabalan.K+, Iniya Nehru .E+, Rex Aantonny. P**

+Principle system analyst-National Informatics Centre, Chennai, India

+Senior technical Director-National Informatics Centre, Chennai, India

**Software Executive, Cadgraf Digital Pvt Ltd, Chennai, India*

¹j.balan@nic.in ²nehru@nic.in ³joyemperor@gmail.com

Abstract: In India more than 70% population are living in the rural area. Development of infrastructure like road, school in rural area has many challenges due to lack of man power, money and poor infrastructure. It is inevitable to use technology in monitoring the rural roads. We propose a new technology using a smartphone to monitor the rural road projects under execution.

Keywords: Location Tracking, Data compression, Image processing, Mobile technology.

EXISTING SYSTEM

An online system has been implemented for Rural development and Panchayat department, Government of Tamil Nadu, India. This system is used by the Rural development department for monitoring the execution of various road projects executed in rural area of Tamil Nadu.

This system captures the various parameters such as project cost, sanction number, executing agency, location of the project, physical progress of the project at various stages with photographs using digital camera, payment of bills to an executing agency at various stages, etc.

Various MIS (Management Information System) reports are being generated block level, district level and state level for effective implementation of the progress. Though various MIS reports are available in the existing system the quality of the project executed cannot be ensured unless photos of the project under various stages with geo code and time stamps are available to the higher level administrator.

Because of these lacunas a mobile based application which captures the photo of the status of the projects at the project location with Geo Tag (Latitude and Longitude) and time stamps has been proposed.

PROBLEMS WITH EXISTING SYSTEM

- Web based system available for monitoring physical status and financial status of road projects.
- Even in this system photos are being uploaded to see the physical status of the work since the photos are not Geo-coded the photo uploaded to the web based system may not be original because the photos taken from some other road with good physical progress can be uploaded and higher level authorities can't know the real progress of the road project.
- In the existing system the photo is uploaded to the central system doesn't have the time stamp; the higher level authorities can't know what time the photo was taken.
- The quality of the ongoing road project at a particular location can't be visualised in the existing system since the photos are not Geo-coded.
- The engineer can visit more than one road project in a day and capture many photos on different road and brought them to the data entry operators and there is a high chance mixing photo of a road with another road since the photos are not uploaded from the location where it was captured.

PROPOSED SYSTEM

A country like India spends enormous amount of money in rural development projects such as construction of roads, schools, houses, water tanks, etc. The geographical spread of the country is one of the reasons for the motivation of the entire project. The proposed technology helps to solve the monitoring of such projects sitting at their offices.

To avoid the problem faced by administrators sitting at various levels such as State, District and Block level, we are proposing a Mobile Based Rural Road Works Monitoring System. Here, we use a smart phone for capturing the photo of current work status from project location and upload to the server database. In the server, the uploaded information can be

stored in Tamil as well as in English. This method will solve the problems in monitoring the ongoing rural area projects. The architecture diagram of the system is shown in figure (1).

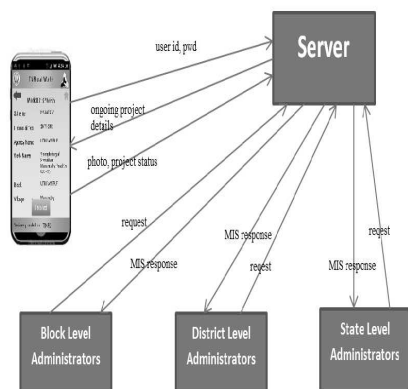


Figure (1): Architecture diagram of mobile based rural road monitoring system.

The architecture diagram shows our proposed system; the user can enter his user id and password to download the ongoing project from the Central Server. The application captures the IMEI number of the mobile device along with user id and password and sends to the Server. If the user and IMEI number are valid then Server pushes the ongoing rural works pertaining to the users' area to the mobile device. While capturing the photos on the ground the system also captures the Geo coordinate of the location, Time stamp and other information selected by the user. The photo captures can be uploading to Central Server. If there is no internet connectivity at the location then the photos along with other information can be stored in local database and then upload to the Server when the internet connectivity is available.

The administrators at various level can download various MIS reports based on their requirements from the Online Scheme Monitoring System for effectively monitoring these ongoing Rural Road works.

PROCESS

This application is integrated with Online Scheme Monitoring System already implemented for RD Dept Tamil Nadu [1]. By using that every Engineer (AE, AEEs) at Block level (sub- district level) has been provided a mobile with GPS / AGPS facility with 3G Connectivity. Every Engineer has provided with user id and password. The IMEI Number of Mobile is registered in the central Server during User Registration. Mobile Application will be distributed through e-mail; Using install option the application package is install in the mobile after that using Login Screen, the Engineer logs in to Central Server after the successful login the pending Road Works in his Block is get downloaded to Mobile from Online Scheme Monitoring System Database, that is showed in figure (2). Mobile Application has a facility to check GPRS and GPS availability and can be turned on by the user. The Mobile application facilitates the Engineer to update status of particular work along with GPS and Photograph of work site when he makes field visit.

TN Rural Works	
268050	Pulipakkam BT Road at Kattuputhur to pullipakkam
326865	Manamathykandigai Strengthening of Manamathy Kandigai - Ilanagar Road Km 0/0 - 2/3
326886	Manamathy Strengthening of Serpakkam - Manamathy Road Km 0/0 - 2/3
326911	Malayankulam Strengthening of Marutham Malayankulam - Sithamalli Road Km 0/0 - 2/4
520400	Maruthavambadi Strengthening of Malliyankaranai Maruthuvampadi - Sithamalli Road Km 0/0 - 3/2
Services provided by: TNRD	

Figure (2): pending works of ongoing process.

The chainage (road location in distance from starting point) details where the photos to be captured is entered in the mobile; while capturing the photos. The captured photo of size range between 2MB to 3MB is compressed using base64 encoding scheme to 70K - 100K with reasonable quality. The compressed photo is then uploaded with GPS Coordinates (LAT, LONG) and other parameters to central server if Internet connectivity is available at the site if connectivity is not available the captured photos along with other parameters are stored in the local database (SQLite).



Figure (3): Work Stage Entry Screen

Once connectivity is available, all the pending photos and its parameters are uploaded to central server. The uploaded photos can be seen from the Central Server and the photos can be viewed on work site on global map using Google maps.

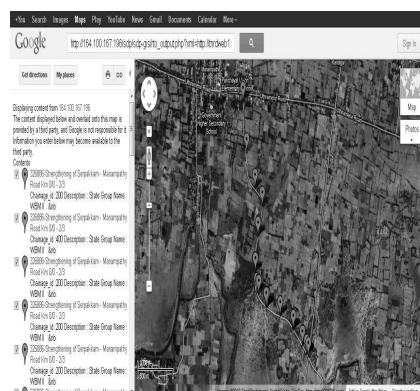


Figure (4): Viewing the actual locations of the photos taken for a Road Work on the Map



Figure (5): Viewing the Photos of a Road Work on the Google Map.

The proposed system has security features. Users who have authorised permission alone can upload photo. A user id and password is used to access the server. Then the application in turn checks the IMEI numbers that has been stored in the

server, verifies and allow the user to enter his user name and password. Each user is assigned a specific portion of the road work to be monitored a pending road works. The pending works alone get downloaded to the mobile of the particular user based on the user profile. Then the user goes to the specific location and the application get the GPS coordinates of that location and it is attached to the photo sent.

In the uploading process we have used **base64 encoding scheme** to reduce the size of the image for fastest processing.

BENEFITS OF THIS SYSTEM

- Administrators at Block, District and State level can see the progress of Rural Road works with real photographs access the web server through internet.
- System brings transparency in the Project Status Report by engineers
- System ensures that the engineer visits the ongoing project site on regular intervals based on the guidelines given by the Government because the time stamp of the visits are stored in the data base.
- Over the period anyone can see the status of the Road on a particular time period
- Helps in administrative decision when people demand for relaying of the same road after some time.
- Various MIS reports can be generated for effective monitoring.
- The system also helps to view the quality of the construction of the road works to some extent.

SOFTWARE

The OS that has been is used UbuntuLinux. In the Server side the database used is PostgreSQL9.1. The script has been done using PHP 5.0 and webserver Apache.

Mobile (clients) side requirements are Android OS, SQLite. The Server and client used XML based web service for data exchange.

STATUS OF IMPLEMENTATION

This system was pilot tested in one of the blocks in Kanchipuram district of Tamil Nadu. Now it's being implemented in Tiruppur district in future the application will be implemented in 385 blocks of Tamil Nadu.

FUTURE EXPANSION

- This system can be used in future for the infrastructure monitoring like construction of schools, buildings, etc.
- It can also use for Dust Bin clearance monitoring project.
- Using HTML5, CSS3, Java Script the same system can be implementing forwork in all types of mobile operating systems including Android.

CONCLUSION

In vast country like India using mobile technology which is prevalent in every nook and corner of the country can be effectively used for monitoring the project executed in remote areas by the higher level monitoring authorities. This system brings more transparency in the project implementation.

When people demand for re-laying of the same road in future the administrator can use this technology to capture photos of present status of the road before taking decision.

REFERENCE

[1] Online scheme monitoring system implemented by for Rural Development & Panchayat Raj Dept. Govt of Tamil Nadu.

ஒரு முழுமையான தட்டல் மற்றும் கையெழுத்து உள்ளீட்டு மென்பொருளை நோக்கி

முனைவர் வெ. கிருஷ்ணமூர்த்தி

(முன்னாள் பேராசிரியர், அண்ணா பல்கலைக்கழகம்)

லேர்ன் ஃபன் சிஸ்டம்ஸ், சென்னை, இந்தியா prof.vkrish@gmail.com

பலவகைத் தட்டச்சு வகைகளையும், கையெழுத்தையும் கொண்ட தமிழுக்கான உள்ளீட்டு மென்பொருள் ஒன்றை முழுமையானதாக உருவாக்கும் முயற்சியில் வெற்றி பெற்றதில், எதிர்கொண்ட பல பிரச்சனைகளையும், அவற்றைத் தீர்க்க எடுத்த முடிவுகளையும் இந்தக் கட்டுரை ஆராய்கிறது. பிரச்சனைகளை, மொழியியல் சார்ந்தது, கணிப்பொறித் தொழில்நுட்பம் சார்ந்தது, பயனர் இடைமுகம் சார்ந்தது என மூன்று வகைகளாகப் பிரிக்கலாம். இந்த மூன்றிலும் சரியான முடிவுகள் எடுக்கப்படும்போதே ஒரு சிறந்த மென்பொருளை உருவாக்க முடியும். இந்தப் பிரச்சனைகளில் சில -

1. மொழியியல் சார்ந்தது - உள்ளிடும்போதே உடனுக்குடன் சொல்திருத்தி செயல்பட வேண்டும். பலருக்குச் சந்தியில் சந்தேகம் ஏற்படுவதால், அவற்றை நீக்க உதவி வேண்டும். பலருக்கு ல, ர, ன என்ற மூன்று எழுத்துக்களில் வரும் சந்தேகங்களைப் போக்க சொற்களின் பொருள்களையும் கொடுக்க வேண்டும். ஒரு சொல் பல மாறுபட்ட வடிவங்களில் வந்தாலும், வேர்ச்சொல்லின் பொருள் கொடுக்கப்பட வேண்டும். 'மதியழந்த' போன்று இரு வேர்ச்சொற்கள் சேர்த்து எழுதப்படும்போது அவற்றையும் கண்டறிய வேண்டும்.

2. கணிப்பொறித்துறை சார்ந்தது - மேஜைக் கணிப்பொறிகளில் மட்டுமில்லாமல், பலகைக் கணிப்பொறியிலும், கைபேசிகளிலும் உடனுக்குடன் செயல்படும் வேகம் வேண்டும். மேஜைக் கணிப்பொறிகளில் சில நூறு பக்கங்களையும் ஒன்றிரண்டு நிமிடங்களில் சொல்திருத்தம் செய்யும் வல்லமை வேண்டும். அளவில் சிறியதாக இருக்க வேண்டும்.

3. பயனர் இடைமுகம் சார்ந்தது - கைபேசிகளில், மற்ற மொழிகள் சார்ந்த மென்பொருள்களுக்கான விசைப்பலகை அமைப்பை ஒத்து இருக்கும்படி அமைய வேண்டும். பயனருக்கு எளிதில் புரியும் வகையில் இருக்க வேண்டும்.

ஒரு உரையை உள்ளிடுவதில் ஒலி, கையெழுத்து, விசை மேல் தட்டல் என மூன்று முறைகள் உள்ளன. இதில், ஒலி வழி உள்ளிடும் முறை இங்கு எடுத்துக்கொள்ளப்படவில்லை. ஏற்கனவே எழுதப்பட்டு அல்லது அச்சிடப்பட்ட தாள்களைப் பார்த்து எழுத்துக்களை உணர்வது உடனடிச் செயல்பாட்டில் வராது. அதுவும் இங்கு சேர்க்கப்படவில்லை. தட்டுதல் மற்றும் எழுதுதல் என்பனவற்றை உடனடியாக மின்உரையாக மாற்றும் செயல்பாடுகள் மட்டும் உள்ள உள்ளீட்டு மென்பொருள்கள் இங்கு பேசப்படுகின்றன.

விசைப்பலகையில் தட்டி எழுத்துக்களை உருவாக்குதல் என்பது வளர்ந்து, இப்போது, வேகமாக உள்ளிட வகை செய்யும் பல உத்திகளை உள்ளடக்கி, 'assisted input method' என்று ஆங்கிலம் போன்ற மொழிகளுக்கு வரத்தொடங்கியுள்ளது. இவை சொற்களை விரைவாகத் தட்ட வகை செய்கின்றன. இது சரியான சொற்களைத் தேர்வு செய்வதால், எழுத்துப் பிழைகளும் தவிர்க்கப்படுகின்றன.

தமிழ் என்று பார்க்கும்போது, ஆங்கிலத்தில் இல்லாத பல அம்சங்கள் சிக்கல்களுக்கு வழி வகுக்கின்றன. முதலாவதாக, எழுத்துக்களின் எண்ணிக்கை. இதில் உயிர்மெய் எழுத்துக்களை, மெய் + உயிர் எனத் தட்டுவது சிக்கலைக் குறைக்கிறது. கைபேசிகளில் சில 12 விசைகளைக் கொண்டுள்ளது. இவற்றில் தமிழ் எழுத்துக்களை வைக்கும்போது, ஒரு விசையில் நூலைந்து எழுத்துக்கள் இருக்கவேண்டி வருகிறது. அப்போது ஆங்கிலத்தில் உள்ள T9 போன்ற 'ஒருதட்டல்' முறை இருந்தால் அது உள்ளீட்டை எளிமையாக்கும்.

ஆனால் இதை உருவாக்குவதில் தென்னிந்திய மொழிகளுக்கு ஒரு சிக்கல் உள்ளது. தமிழில் 'வந்துகொண்டிருந்தவர்களிடம்தானே' என்பது ஒரே சொல். இதில் பத்து பகுதிகள் உள்ளன. தமிழில் ஒரு சொல்லில் இம்மாதிரி பல பகுதிகள் இருக்கலாம். இம்மாதிரிச் சொற்கள் எல்லாவற்றையும் ஒரு அகராதியில் வைத்து, அதில் இருப்பதை எடுத்துக்கொடுக்கும் வழி சரியானதாகத் தெரியவில்லை. இந்த வழி சில இடங்களில் சொல்திருத்தத்திற்கும் பயன்படுத்தப்படுகிறது. இதுவும் சரியான வழியாகத் தெரியவில்லை. வேர்ச் சொற்களை மட்டும் அகராதியில் வைத்துக்கொண்டு, அவற்றின் மாற்றங்கள் எல்லாவற்றையும் கணிப்பின் வழியாக உருவாக்குதல் சிறந்த வழி. இதனால் அகராதியின் அளவு குறையும். எந்தச் சொல்லும் விட்டுப்போகாது.

இந்த வகையில் ஒரு சொற்களின் மாற்றங்களை சரியாகவும் விரைவாகவும் கணக்கிடுவது மிக அவசியம். இதற்கு அகராதியின் வடிவமைப்பு மற்றும் கணிப்பு வழியின் திறன் இரண்டும் முக்கியம். மேலும், 1999 போன்ற

மென்பொருளானது, ஒரு தட்டல் சரத்தில் வரக்கூடிய பல எழுத்துச் சரங்களையும் பார்த்து, அவற்றில் எவையெல்லாம் சரியான சொற்கள் என்று கண்டறிய வேண்டும். தட்டல் சரத்திற்குப் பல சரியான சொற்கள் வர வாய்ப்புள்ளது. எடுத்துக்காட்டாக, எங்களால் வடிவமைக்கப்பட்டிருக்கும் கே12 என்னும் ஒருதட்டல் அமைப்பில், 5326275 என்று தட்டினால், அது - படியில், மணியில், மடியில், பணியில் என்ற 4 சொற்களைக் கொடுக்கும். இந்த ஒருதட்டல் மென்பொருள் வேகம் குறைந்த கைபேசிகளில் பயன்பட வேண்டும் என்பது அடிப்படைத் தேவை. அதனால் எவ்வளவு வேலை செய்தாலும் உடனுக்குடன் விடை கிடைத்தாக வேண்டும். கே12 உள்ளீட்டு முறை இந்த வகையில் வெற்றி பெற்றுள்ளது என்பது குறிப்பிடத்தக்கது.

ஒரு சொல் எப்போது முடியும் என்று கூற முடியாது. எத்தனை தட்டுகளிலும் முடியலாம். அதனால் ஒவ்வொரு தட்டு தட்டப்பட்டதும், அத்தனை தட்டுகளில் முடியக்கூடிய சொற்களை மட்டும் தேர்ந்தெடுக்க வேண்டும். அப்படிப்பட்ட சொல் எதுவும் கிடைக்காத போது மட்டும், சொல்லின் பகுதியைக் காட்டவேண்டும். அதனால் ஒவ்வொரு முறையும் முழுச்சொற்களைத் தேட வேண்டும்.

நண்பர் ஒருவர் எழுதிய உரையைப் படிக்கும் போது அவர் பல இடங்களில் இரு சொற்களை இணைத்து எழுதுவது தெரிந்தது. எடுத்துக்காட்டாக, அறிவிழி, மனங்கொத்தி, மதியிழந்த, மனமிரங்கி போன்றவை. இந்த இரு வேர்கள் கொண்ட சொற்களை இதுவரை எந்த சொல்திருத்தியும் கணக்கில் எடுத்துக்கொண்டதாகத் தெரியவில்லை. இம்மாதிரிச் சொற்கள் எல்லாம் தவறென்றுதான் காட்டப்படும். ஆனால் இந்த இருவேர்ச் சொற்களையும் சரியென்று சொல்லும்படியான ஒரு சொல்திருத்தியை உருவாக்குவது சொல்திருத்தியின் திறனை அதிகரிக்கும். பயனருக்கு நேரம் மிச்சமாகும்.

இருவேர்ச் சொற்கள் எவ்வாறெல்லாம் எழுதப்படுகின்றன என்ற ஒரு ஆராய்ச்சி மேற்கொள்ளப்பட்டு, அதன் அடிப்படையில் பொன்மொழியின் சொல்திருத்தி வடிவமைக்கப்பட்டுள்ளது. இந்த ஆராய்ச்சி ஒரு முதல் முயற்சி. அதனால் இன்னும் பல இடங்களில் இருவேர்ச் சொற்கள் தவறென்று வருகிறது. ஒரு சொல்லுக்கான பரிந்துரைகள் தேடும்போது பல இருவேர்ச் சொற்கள் வருவது ஒரு விரும்பத்தகாத விளைவு. தமிழ் ஆராய்ச்சியாளர் எவரும் இது பற்றி ஆராய்ந்துள்ளார்களா என்பது தெரியவில்லை. அத்தகைய ஆராய்ச்சியின் முடிவுகள் கிடைத்தால் சொல்திருத்தி இன்னும் மேம்படும்.

கே12 உள்ளீட்டு முறையில் ஒருவர் இருவேர்ச் சொற்களைத் தட்டினால் என்ன செய்வது? ஏற்கனவே உள்ள வேலைப் பளுவினால், உடனடியாக இருவேர்ச் சொற்கள் தேடப்பட மாட்டாது. அதனால், ஒருவேர்ச் சொல் இல்லாதபோது எதுவும் காண்பிக்கப்படாது. இந்த சமயங்களில், சொல் காண்பிக்கப்படும் உரைப்பெட்டியின் வலது ஓரத்தில் உள்ள அம்புக்குறியைத் தட்டினால், வரும் பட்டியலில் இருவேர்ச் சொற்கள் காண்பிக்கப்படும். எடுத்துக்காட்டாக, 'மனங்கொத்திப்' என்று தட்டினால், முதலில் எதுவும் வராது. ஆனால், அம்புக்குறியைத் தட்டியதும், இருவேர்கள் உள்ள 8 சொற்கள் காண்பிக்கப்படும்.

'மனங்கொத்த' எனத் தட்டும்போது, மறக்கவந்த, பறக்கவந்த போன்ற சொற்கள் காண்பிக்கப்படும். அம்புக்குறியைத் தட்டியதும், வரும் பட்டியலில் இந்த இரு சொற்கள் மட்டுமே இருக்கும். இந்தப் பட்டியலின் மேற்புறம் உள்ள 'இருவேர்' என்னும் பொத்தானைத் தட்டினால், 'மனங்கொத்த' உள்பட பல இருவேர்ச் சொற்கள் காண்பிக்கப்படும்.

சிலருக்கு ல, ர, ன என்ற மூன்று எழுத்துக்களில் அவ்வப்போது சந்தேகம் எழும். கல், கள் என்ற இரண்டு மட்டும் சரி. கழ் சரியல்ல. வலி, வளி, வழி என்ற மூன்றுமே சரி. இவற்றில் எது நமக்குத் தேவையான சொல் என்ற சந்தேகம் வரும். கல்லிலிருந்து என்று எழுதும்போது, அதன் வேர்ச்சொல்லுக்கான பொருள் பற்றிய சிறு விளக்கம் இருந்தால் நல்லது. இதுவும் கே12 முறையில் செயல்படுத்தப்பட்டுள்ளது. ல, ர, ன ஐயங்களை நீக்கி தவறில்லாமல் எழுத, கே12 விசை அமைப்பு சிறிது மாற்றத்துடன் 'கே12லரன' என்னும் முறையாக செயல்படுத்தப்பட்டுள்ளது. இதில், ல, ள, ழ என்ற மூன்று எழுத்துக்களுமே ஒரே விசையில் வைக்கப்பட்டுள்ளன. இது போல், ன, ண, ந என்ற மூன்றும் ஒரே விசையில் உள்ளன. ர, ற என்ற இரு எழுத்துக்களும் ஒரே விசையில். இதனால் லரன எழுத்துக்களில் தவறு செய்ய வாய்ப்பே இல்லை. ஒரு சொல் மட்டும் இருக்கும்போது அது மட்டும் காண்பிக்கப்படும். ஒன்றுக்கு மேற்பட்டது இருந்தால், அவற்றுக்கான குறிப்பைப் பார்த்து சரியான சொல்லைத் தேர்ந்தெடுக்கலாம். இருவேர்ச்சொல்லாக இருக்கும்போது, இரு வேர்களுக்கும் உள்ள குறிப்புகள் காட்டப்படும். குறிப்புகள் இல்லாத இடங்கள் சிறு கோட்டால் காட்டப்படும். இந்த கே12, கே12லரன அமைப்புகள் படத்தில் காட்டப்பட்டுள்ளன.

லலச				பறபப			
ஒலி		வலி		பணம்		பணப்	
கங அஆ1	சௌ இஈ2	டண உஊ3	ல	கங அஆ1	சௌ இஈ2	ட உஊ3	ல
தந ஏஏ4	பம '5	யரஐ ஃ6	ஃ	த ஏஏ4	பம '5	யரஐ ஃ6	ஃ
லவஒ ஒஒ௭7	ழளஸ ஷஐ8	றளஹ கூபூ9	ஃ	லழளவ ஒஒஒ௭7	ஷஐ 8	ளணநஹ கூபூ9	ஃ
ப.	ப.	12#	ப.	ப.	ப.	12#	ப.

படம் 1 கே12 மற்றும் கே12லரன விசைப்பலகை அமைப்புகள்

இந்த கே12 முறைகளில் சொல்திருத்தம் உள்ளடங்கியுள்ளதால் தவறுகள் குறையும். இந்த விசைப்பலகைகளின் இட அமைப்பு மிகவும் சீரானதாக இருப்பதால், எவரும் ஒரு சில நிமிடங்களிலேயே கற்றுக்கொள்ளலாம். ஐந்தாம் வகுப்பு மாணவர்கள் எளிதாகக் கற்றுக்கொண்டுள்ளனர்.

எந்த அகராதியிலும் மனிதர்களின் மற்றும் ஊர்களின் பெயர்கள் அத்தனையும் இருக்க முடியாது. அதனால் அகராதியில் புதிய சொற்களைச் சேர்க்கும் வசதி இருப்பது அவசியம். ஆங்கிலம் போன்ற மொழிகளில் சொற்களை அப்படியே சேர்ப்பது போதுமானது. ஆனால், தமிழில் அப்படியல்ல. கம்பர் என்ற பெயரைச் சேர்த்தால் போதாது. கம்பருக்குத்தான், கம்பரிடம் என்பது போன்ற பல மாறுபட்ட வடிவங்களில் பின்னால் பயன்படுத்தப்படலாம். அதனால் தமிழில், ஒரு சொல்லை அப்படியே சேர்க்காமல், அதை ஒரு பெயர்ச்சொல்லாகச் சேர்க்க வேண்டும். அத்துடன், அதன் பல வடிவங்களையும் கணிக்கும் வசதியும் கொடுக்கப்பட வேண்டும். இது அவ்வளவு எளிதல்ல. ஏனென்றால், ஒரு பெயர்ச்சொல்லானது பல விதங்களில் மாறுபடலாம். வசதி என்பது வசதியாக என்று மாறும்போது ய என்னும் உடம்படுமெய் கூட வருகிறது. பணம் என்பது பணமாக என்று வரும்போது ம என் என மாறுகிறது. இப்படியான மாற்றங்களில் நாம் கொடுக்கும் சொல் எப்படி மாறும் என்பதைச் சரியாகக் கணித்துச் செயல்பட வேண்டும். இதுதான் சிக்கல். இதைத் தீர்க்கவும் இங்கு வழி காணப்பட்டுள்ளது. அதனால், இங்கு ஒரு சொல்லை அப்படியே சேர்க்கலாம், அல்லது, பெயர்ச்சொல்லாகச் சேர்க்கலாம். ஒரு பெயர்ச்சொல்லைச் சேர்ப்பது அதன் ஆயிரக்கணக்கான மாற்றங்களையும் சேர்ப்பதற்குச் சமம். இந்த வசதி வேறு எங்கும் இருப்பதாகத் தெரியவில்லை.

கே12 விசைப்பலகையில் ஒரு தட்டல் சரத்திற்குப் பல சொற்கள் வரலாம். அப்போது நாம் தேர்ந்தெடுப்பதை நினைவில் வைத்துக்கொண்டு, அடுத்த முறை அதையே முதல் பரிந்துரையாகக் கொடுப்பது உள்ளிட்ட எளிமையாக்கும். இதுவும் கே12இல் செயல்படுத்தப்பட்டுள்ளது. இன்னும் கே12லரனவில் செயல்படுத்தப்படவில்லை. இதற்கு ஆகும் இடம், காலத் தேவைகளைக் கருத்தில் கொண்டு சரியான செயல்முறை வடிவமைக்கப்பட்டு வருகிறது.

தொழில் நுட்பம் சார்ந்த சிக்கல்களை இப்போது பார்க்கலாம். மேஜைக் கணிப்பொறி, பலகைக் கணிப்பொறி, கைபேசி என இப்போது மூன்று வகை கணிப்பான்கள் அதிகம் பயன்படுகின்றன. இப்போது கடிகாரக் கணிப்பொறிகளும் வரத்தொடங்கியுள்ளன. இவற்றில் ஒவ்வொன்றிலும், இருக்கும் நினைவக அளவும், விசைப்பலகை அளவும், தொடுதிரை அளவும், செயலாக்கியின் வேகமும் பெரிதும் மாறுபடுகின்றன. தற்போதைக்கு, முதல் மூன்றிலும் திறம்படச் செயல்பட வேண்டும். இந்தக் கணிப்பொறிகளில் தற்போது விண்டோஸ் (32 பிட் மற்றும் 64 பிட்), ஆண்டிராய்ட், ஐஓஎஸ், லினக்ஸ், ப்ளாக்பெரி போன்ற ஓஎஸ்கள் அதிகம் பயன்படுகின்றன. இவற்றில் ஒவ்வொன்றிற்கும் தனித்தனியாக மென்பொருள் தயாரிக்க வேண்டும். ஒவ்வொன்றிலும் சிறிது மாற்றங்கள் இருக்கும். தற்போதைக்கு, விண்டோஸ், ஆண்டிராய்ட் மற்றும் ஐஓஎஸ்ஸுக்கு மட்டும் மென்பொருள்கள் உருவாக்கப்பட்டுள்ளன.

விசைப்பலகையைக் கருத்தில் கொண்டால், தமிழ்99, தட்டச்சு, பாமினி, ஒலிமாற்றம் (ஆங்கில எழுத்துக்களில் இருந்து) என்பவை பரவலாகப் பயன்பாட்டில் உள்ளவை. இவை எல்லாமே கொடுக்கப்பட்டிருக்கின்றன. இவை ஒவ்வொன்றிலும் எழுத்துக்களை உருவாக்கும் முறையில் உள்ள வித்தியாசங்கள் கணக்கில் எடுத்துக்கொள்ளப்பட வேண்டும். சாதாரண உள்ளீடு என்றால் சிக்கல் இல்லை. ஆனால், தட்டலின்போதே சொல் சரிபார்க்கப்பட வேண்டும் என்றாலும், லரன எழுத்துக்களில் உள்ள பிழைகள் சரி செய்யப்படவேண்டும் என்றாலும், லரன எழுத்துக்களில் மாற்றங்களுடன் வேறு சொற்களை எடுத்துக் கொடுக்க வேண்டும் என்றாலும் சிக்கல் அதிகரிக்கிறது. ஒரு எடுத்துக்காட்டு பார்ப்போம்.

தமிழ்99 விசைப்பலகையில் கடல் என்று கிடைக்க, கடல் புள்ளி என்று தட்டுவோம். அப்போது, க, கட, கடல், கடல் என்று வரிசையாகக் கிடைக்கும். இதில் அகரமேறிய உயிர் முதலிலும், மெய் அடுத்ததாகவும் கிடைக்கும். ஆனால், ஒலிமாற்ற வகையில், katal என்று தட்டும்போது, க், க, கட், கட, கடல் என்று கிடைக்கும். இங்கு முதலில் மெய்யும், பிறகு அகரமேறிய உயிர்மெய்யும் கிடைக்கிறது. இங்கு, எந்த விதத்தில் அடிப்படை தரவினை வைக்க வேண்டும் என்பது சிக்கலை ஏற்படுத்துகிறது. மெய் + உயிர் என்று வைப்பது எழுத்துக்களின் இலக்கணத்தை ஒட்டி இருக்கிறது. இது ஒலிமாற்றத்தில் வேகமாகச் செயல்படும். ஆனால், மற்ற விசைப்பலகைகளில் சற்று அதிக நேரம் எடுக்கும். எல்லா விசைப்பலகைகளிலும் வேகமாகச் செயல்பட சரியான உத்திகள் அவசியம். இது, ஒரே தரவே பல விதங்களில் வைப்பதைத் தவிர்க்கும். கே12 மற்றும் கே12லரன் என்ற இரு விசைப்பலகைகளிலும் செய்யவேண்டிய வேலை அதிகம். அதனால், விசைப்பலகைக்கு ஏற்றபடி தரவு இருப்பது தவிர்க்க முடியாததாக ஆகிறது.

பொன்மொழி வேர்ட்பேட் போன்ற ஒரு சொல் செயலி. தமிழ் சொல்திருத்தமும், பல அகராதிகளும் இதன் தனிச் சிறப்பு. ஆனால், வேர்ட், இன்டிசைன், பவர்பாயின்ட் போன்ற எல்லா மென்பொருள்களிலும் நேரடியாகத் தமிழை உள்ளிடும் தேவையும் உள்ளது. அப்போது, எழுதிய உரையின் மீது உள்ளீட்டு மென்பொருளுக்கான ஆளுமை மிகவும் குறைவு. இத்தகைய உள்ளீட்டிலும், உடனடியாகச் சொல்திருத்தம் செய்யவும், லரன் மற்றும் சந்தி சந்தேகங்களைப் போக்கிக்கொள்ளவும் முடியும் வகையில் தொழில் நுட்பத்தை அடுத்த நிலைக்குக் கொண்டு சென்றுள்ளோம். பொன்மடல் என்பது பல மென்பொருள்களிலும் தமிழை உள்ளிட உதவுகிறது.

இதன் சிறப்பு என்னவென்றால், ஒரு சொல்லைத் தட்டியவுடன், இடைவெளியைத் தட்டுவதற்கு முன்னால், + குறியீட்டைத் தட்டினால், ஒரு பட்டியல் வரும். அதில் தட்டிய சொல்லுக்கான லரன் பரிந்துரைகளுடன், சந்திப்பிழை திருத்தப்பட்ட பரிந்துரைகளும், இருவேர்ச்சொல் பரிந்துரைகளும் காட்டப்படும். இந்தப் பட்டியலில் சரியான சொற்கள் மட்டுமே காண்பிக்கப்படும். + குறியீட்டைத் தட்டியவுடன் பட்டியல் காண்பிக்கப்படாவிட்டால், கொடுத்த சொல் தவறென்று பொருள். + குறியீட்டைத் தட்டுவது விண்டோஸில் மட்டுமே. ஆண்டிராய்டில் அதுவும் தேவையில்லை. பரிந்துரைகள் தாமாகவே ஒரு வரியில் வரிசையாகக் காட்டப்படும்.

விண்டோஸில் கே12 விசைப்பலகையில், + அல்லது கீழ் அம்புக்குறியைத் தட்டி பரிந்துரைகள் பட்டியலைப் பார்க்கலாம். இதில் முதலில் ஒருவேர்ச் சொற்கள் காண்பிக்கப்படும். கீழே உள்ள 'லரன்', 'இருவேர்' பொத்தான்களைத் தட்டி அவற்றையும் பார்க்கலாம்.

தேர்வுகள்	
மனத்தால்	
மறந்தால்	நினைவிலிருந்து நீங்கு
பறந்தால்	ஆகாயத்தில் செல்
மறந்தால்	வீரம்

தேர்வுகள்	
பணத்தால்	காசு
மனத்தால்	உள்ளம்
மனத்தால்	வாசனை, திருமணம்

தேர்வுகள்	
பணத்தால்	காசு
மனத்தால்	உள்ளம்
மனத்தால்	வாசனை, திருமணம்

தேர்வுகள்	
பணத்தால்	[காசு] + [மரம்]
பணத்தால்	[காசு] + [விளக்கு]
மனத்தால்	[உள்ளம்] + [மரம்]
மனத்தால்	[உள்ளம்] + [விளக்கு]
மனத்தால்	[வாசனை கமழ், திருமணம்] + [பாறைத் துண்டு]
மனத்தால்	[வாசனை கமழ், திருமணம்] + [-]
மனத்தால்	[வாசனை, திருமணம்] + [மரம்]
மனத்தால்	[வாசனை, திருமணம்] + [விளக்கு]

படம் 2 பொன்மடலில் பரிந்துரைகள்

தற்போது ஐஓஎஸ்ஸில் நேரடி உள்ளீடு அனுமதிக்கப்படுவதில்லை. அதனால் பொன்மொழியில் உரையை எழுதி அனுப்ப வேண்டும். ஃபஸ்புக், டுவிட்டர், வாட்ஸ்ஆப், கூகுள் ஸர்ச், எஸ்எம்எஸ், ஈமெயில் போன்றவற்றிற்கு உரையை, கிளிக்போர்டிற்கு அனுப்பாமலேயே அனுப்பலாம்.

விண்டோஸில் வேர்ட், இன்டிசைன் போன்றவை பதிப்பக உலகில் உரையெழுத பெரிதும் பயன்படுகின்றன. அவற்றில் எழுதப்பட்ட உரையை சொல்திருத்தும் தேவை உள்ளது. இதற்காக தனியான மென்பொருள்கள் உருவாக்கப்படவேண்டும். தற்போது பொன்மொழி என்ற பெயரில் வேர்டுக்கு உருவாக்கப்பட்டுள்ளது. விரைவில்

இன்டிசைனுக்கும் வரும். இந்த இடத்தில் நேரம் மிக முக்கியம். நூற்றுக்கணக்கான பக்கங்களை ஒன்றிரண்டு நிமிடங்களில் செய்யும் திறன் இருத்தல் மிக அவசியம். அதாவது ஒரு நொடியில் நூற்றுக்கணக்கான சொற்களைத் திருத்தும் வல்லமை வேண்டும். இதற்கு ஏற்றாற்போல் மென்பொருள் வடிவமைக்கப்பட வேண்டும். பொன்மொழி மற்றும் பொன்சொல் இவ்வாறு வடிவமைக்கப்பட்டுள்ளது. தவறுகள் அவற்றின் தன்மைக்கேற்ப பல நிறங்களில் காட்டப்படுகின்றன. இவற்றைப் பார்த்துத் தேவையான இடங்களில் மட்டும் திருத்தம் செய்துகொள்ளலாம். இதனால் நேரம் மிச்சமாகும். லரன் சொற்களில் மாற்றுச் சொற்கள் இருந்தால் அவற்றை ஆரஞ்சு நிறத்தில் காண்பி என்னும் விருப்பம் தேர்த்தெடுக்கப்பட்டிருந்தால், சொல்திருத்த நேரம் சுமார் 4 பங்காகிறது. இது தேவையில்லையென்றால் ஒரு நொடியில் சுமார் 500 சொற்களுக்கும் மேலாகச் சரி பார்க்கும் திறன் பொன்மொழிக்கு உள்ளது.

தமிழில் சொல்திருத்தம் அவ்வளவு எளிதல்ல. ஆயிரக்கணக்கான மாற்றங்களை ஒரு சொல் ஏற்பதால், அவற்றில் சில விட்டுப்போக வாய்ப்பு உள்ளது. இவை காணப்படும்போது இந்த மாற்றங்களையும் எளிதில் சேர்க்கும் வண்ணம் இந்த மென்பொருள்கள் வடிவமைக்கப்பட்டுள்ளன.

தட்டச்சு பழகாதவரும், எப்போதாவது தமிழ் எழுதுபவரும் விசைப்பலகையைப் பயன்படுத்த மாட்டார்கள். அவர்களும் கணிப்பு உலகில் தமிழைப் பயன்படுத்தி, பயன்பெற வேண்டுமானால் உடனடிக் கையெழுத்து அறியும் மென்பொருள் அவசியம். பொன்மடலிலும், பொன்மொழி ஐஓஎஸ்ஸிலும் ஒரு தமிழ் எழுத்து உணரி சேர்க்கப்பட்டுள்ளது. இது மிகச்சில மில்லி செகண்டுகளிலேயே எழுதிய எழுத்தை அறிந்து சொல்லும். இது பலரும் பல விதமாக எழுதுவதை அறியும். குண்டாக, ஒல்லியாக, சாய்த்து என்று பல வகைகளிலும் எழுதலாம். இதற்கு எந்தவிதப் பயிற்சியும் தேவையில்லை என்பது ஒரு சிறப்பம்சம்.

கணிப்பொறித்துறை சார்ந்த சிக்கல்கள் என்று பார்த்தோமானால், பல வேகங்கள் உள்ள செயலிகளில் வேலை செய்யவேண்டும் என்பதும், எல்லாவற்றிலும் விரைவாகச் செயல்பட வேண்டும் என்பதும் முக்கியமானவை. இவற்றை கருத்தில் கொண்டு கீழ்க்கண்ட முடிவுகள் எடுக்கப்பட்டன.

டேடாபேஸ் எதுவும் பயன்படுத்துவதில்லை. இதனால் அந்த டேடாபேஸின் மிகை நேரம் (overheards) குறையும். முடிந்த அளவு இடத் தேவையைக் குறைக்க வேண்டும். இதற்கு, ஒரு சில பிட்டுகள் தேவைப்படும் இடங்களில் பிட்டுகளாகப் பயன்படுத்துவது, சேமிக்கும்போது சுருக்கிச் சேமிப்பது, தேவையானபோது தேவையானதை மட்டும் விரித்துக் கொள்வது, மிக விரைவாகச் செயல்பட வேண்டிய இடங்களில் ஒரே தகவலை பல விதங்களில் சேமித்து வைப்பது - இத்தகைய உத்திகள் பல செயல்படுத்தப்பட்டுள்ளதால்தான் ஒருதட்டல் முறை, சொல்திருத்தி, எழுத்துஉணரி என்று எல்லாமும் மிக விரைவாகச் செயல்படுகின்றன.

பயனர் இடைமுகம் பற்றிப் பார்க்கும்போது, விண்டோஸ் அமைப்பில் சிக்கல் இல்லை. எல்லா கணிப்பொறிகளும் ஒரே மாதிரியான விசைப்பலகையைப் பயன்படுத்துகின்றன. கைபேசிகளில் விசைப்பலகைகளை வடிவமைப்பதில் சிக்கல்கள் உள்ளன. இன்னும் செந்தரம் என்று எதுவும் முடிவு செய்யப்படாததால், ஒவ்வொரு தயாரிப்பாளரும் ஒவ்வொரு விதமாகச் செயல்பட்டு வருகின்றனர். ஒருவர் ஒரு வரிசையில் 10 விசைகள் இருக்கும்படி வடிவமைக்கிறார். ஒருவர் 11 விசைகள் பயன்படுத்துகிறார். மொழிகளுக்கு ஏற்றபடி வரிசைகளின் எண்ணிக்கையும் மாறுகிறது. மேலும் கடைசி வரிசையில் இடைவெளிச் சாவியுடன் வேறு எத்தனை சாவிகள் இருக்கலாம் என்பதும் வேறுபடுகிறது. இவற்றில் பலருக்கும் ஒத்துப்போகும் வண்ணம் விசைப்பலகைகளை வடிவமைக்க வேண்டியுள்ளது. பொத்தான்களின் நிறங்களும் மற்றவர்கள் பயன்படுத்துவதைப் போன்று இருப்பது அவசியமாகிறது. இவற்றைச் சரியாக அமைக்கவே அதிக நேரம் ஒதுக்க வேண்டியுள்ளது.

ஒவ்வொருவருக்கும் ஏற்ற வகையில் ஒரு விசைப்பலகை. விசைப்பலகையைக் கையாள விரும்பாதவருக்கு கையெழுத்து உணரி என அனைவருக்கும் பயன்படும் வகையில் மென்பொருள்கள் உருவாக்கப்பட்டுள்ளன. இவை விண்டோஸ், ஆண்டிராய்ட், ஐஓஎஸ் என்ற மூன்று இயங்குதளங்களில் வேலை செய்யும். மேஜைக் கணிப்பொறி, டேப்லட் கணிப்பொறி, கைபேசி என்று பல கணிப்பாண்களிலும் செயல்படும். உரையை உள்ளிடுவதை விரைவாக்க தரமான, வேகமான சொல்திருத்தி உள்ளது. எந்த மென்பொருளில் தமிழை உள்ளிட்டாலும் சொல்திருத்த வசதிகள் - சந்தி, லரன், இருவேர் திருத்த/பரிந்துரை வசதி உள்ளன. மற்ற சில மென்பொருளில் உள்ள உரையைச் சொல்திருத்தவும் முடியும்.

தமிழில் உள்ளிட பல உதவிகளுடன் கூடிய தரமான மென்பொருள்கள் உருவாக்கத்தில் உள்ள சிக்கல்களையும், அவற்றைத் தீர்க்க மேற்கொண்ட சில தொழில்நுட்ப முடிவுகளையும் இங்கு கண்டோம்.

Translating Tamil Speech (SL) as English Text Message (TL) in Android Mobile Phones

Vikram Srinivasan

Department Of Computer Science, Sree Sastha College of Engineering, Chennai 600 123, India
svikram984@gmail.com

Abstract— Mobiles phones are used every nook and corner and every man, hence innovative technological applications are needed. Moreover, in the scenario of android mobiles not only professionals but even common users expect ample innovations. The paper focuses on translating Tamil speech (SL) as English text message (TL). Even though there are some applications used for translating SL to TL, its one step ahead to convert user's speech in Tamil to English text by applying Hidden Markov Model Tool Kit (HTK). The performance of the system is manipulated in terms of various measures like Word Error Rate, Accuracy Percentage and Correctness Percentage. Building a speech recognition system for machine translation expect target language output to be grammatically correct. The outcome from speech recognition, target language output (English) can often be ungrammatical. To solve the problem, we propose an ongoing chunk-based grammar checker by using trigram language model and rule based model. Finally the HTK translates user's Tamil speech to English text and recipient receives the message in English as well as audio or sentences in Tamil.

Keywords—Hidden Markov Model (HMM), HMM Tool Kit (HTK), Speech recognition, Trigram Language Model, Rule based Model, Segmentation and tagging, chunking and Morphological analyzer.

1. INTRODUCTION

The best mode of sharing knowledge among human beings all this ages has been only speech. Speech production and perception are one of the means of communication. Researches on replication of natural human speech by machines are on line for many years. The issue of recognition of speech can be now termed as "speech to text conversion problem". The predominant research topic that prevails for more than four decades is ASR (Automatic Speech Recognition). For countries with low literary rates and with rural background, implying the speech recognition system for their regional language is on the developing side. Two basic elements of a language checker are spell checker and grammar checker. Inspection and correction of each text word is carried out by the spell-checker. In places where there are larger than the word the grammar checker identifies it and thereby corrects it.

One of the important features of the word processor is the implementation of grammar checker as a feature of a larger program. In the case of machine translation, the grammar checker is not available as a individual free program. Thus, grammar checker is carried out as an excellent translation feature by using trigram language model and rule base model. The process is initiated with the translated English sentence as the input. To begin with the input sentence is tokenized and tagged POS to each word.

A chunk based sentence structure is created by grouping the tagged words into chunks by parsing them. Following the formation into chunks, the connectivity of chunks is identified by using trained sentence patterns. In case of incorrect sentence pattern, the chunk errors are analysed and it's corrected using English grammar rules.

The challenges in translating the sentences from English to Tamil language are

1. The structure and morphology of Tamil language is quite rich. It has in it the SOV order which is highly flexible according to the sentences.
2. In tamil language framing of meaningful and grammatically correct sentences even in a word is feasible. Eg. "odinaan". It explains the action performed, tells the tense and the gender of the person executing the action. The tedious task here is the inculcation of the gender ending. Ambiguity problem such as words that house various meaning when they occur in parts of speech. Translation in such laid words accurately is also a great task.

3. Rule based technique and statistical technique gives more accuracy in translation, but there are no such parallel schemes available for its implementation.

2. RELATED WORK

A number of researchers are working on Speech recognition systems (SRS). They use algorithms like Hidden Markov Model (HMM), Dynamic Time Warping (DTW), and Artificial Neural Networks (ANN) for the implementation of SRS. Among them the best choice would be DTK, since, DTK based SRS has a good accuracy for isolated Tamil digits (Dharun and Karnan, 2012) (Karpagavalli et.al, 2012). The performance of DTW for small number of templates is good and its dependent of the speaker (Anusuya and Katti, 2011). In (Murugan et. al, 2012) SRS, medium size vocabulary was built. The system was trained with 21 speakers and the recognition accuracy was 84 %. The Syllable based Continuous speech recognizer done by Lakshmi.A and Hema A Murthy (Lakshmi and Murthy, 2006) uses group delay based two level segmentation to extract the syllable segment from the speech data. R.Thangarajan and A.M.Natarajan (Thangarajan and Natarajan, 2008), in their work on syllable based Continuous Speech Recognition for tamil have framed an algorithm which details the linguistic rules of tamil to recognize prosodic syllables in a tamil word.

The tamil syllable comes with itself combination of short vowel (Kuril), long word (Nedil) and consonant (Mei eluthu). The syllabic rules of tamil language which consists of 8 patterns plays a vital role of segregating the given utterance into syllables. The performance of the algorithm is 80 %, Saraswathi Selvarajan & T.V.Geetha (Selvarajan and Geetha, 2007) designed a morpheme based language model for Tamil speech Recognition system. This system is used to minimize the size of vocabulary. It carries out this by decomposing the words that are formed from a base word into stem and ending and the subword morpheme is stored for training the language model. The Swedish grammar was checked by Stymne and Ahrenberg (2010) for evaluation tool and post process tool of statistical machine translation. The English-swedish translation was carried out by a factored phase-based statistical machine translation (PBSMT) system which is base on Moses (Koehn et al., 2007) and chiefly on rule-based Swedish grammar checker Granska (Dome ij et al., 2000; Knutsson, 2001). A detailed study about the frequently encountered style and text structure errors were carried out by Buscail and Dizier (2009). These errors were let in by a number of authors while producing texts. In case of any arguments for or against certain correction an argumentation system can be made use of.

Sensing the need of the hour Google has devised a highly efficient and appreciated translation service named “Google translation”. This serves as an excellent translation for all languages, specifically for the languages that followed SVO (Subject-Verb-Object) sentence agreement. For languages such as Tamil, Telugu etc. which does not obey the SVO rule Google still finds it difficult to provide a translational service. These languages are begin termed as ALPHA languages because it's still on the developing phase in forms of translation.

3. METHODOLOGY

3.1 Hidden Markov Model Toolkit (HTK)

The HTK is used as a tool kit for building Hidden Markov models. It is predominantly used in Speech Recognition research. It also has its space in other applications like research into speech synthesis, character recognition and DNA sequencing. HTK is a package of library modules and tools available in C source form (Steve Young et al, 2006). These tools help in analysis of speech, HMM training, testing and results analysis. It also has various other tools, each designed with a proposed function. The toolkits consisting of single mixture Gaussian distribution and multiple mixtures Gaussian supports HMMs. The processing stages are divided into two, the training phase and recognition phase involved in HTK. To estimate the parameters of the set of HMMs to tools of HTK from the training phase are used which involves the training utterances and their associated transcriptions. Unknown utterances are transcribed using HTK recognition tools in the recognition phase.

3.1.1 Hidden Markov Model (HMM)

HMM is a Statistical Model in which the system being modeled follows a Markov process. Markov process consists of unknown parameters, by which the hidden parameters are determined from the observable parameters (Chandra and Akila, 2012). It also unfolds a sequence of possible events in which the probability of each event depends only on the state of the previous event (Rabiner and Jung, 1993). All these qualities are confined as the Markov property. There are different states in

HMM that which satisfies the Markov property. The output distribution for each state is defines as a collection of Gaussian densities. The various applications of HMM includes many areas like Signature verification, Speech and Speaker Recognition, Bio Informatics and Bar code reading (Anusuya and Katti, 2011).

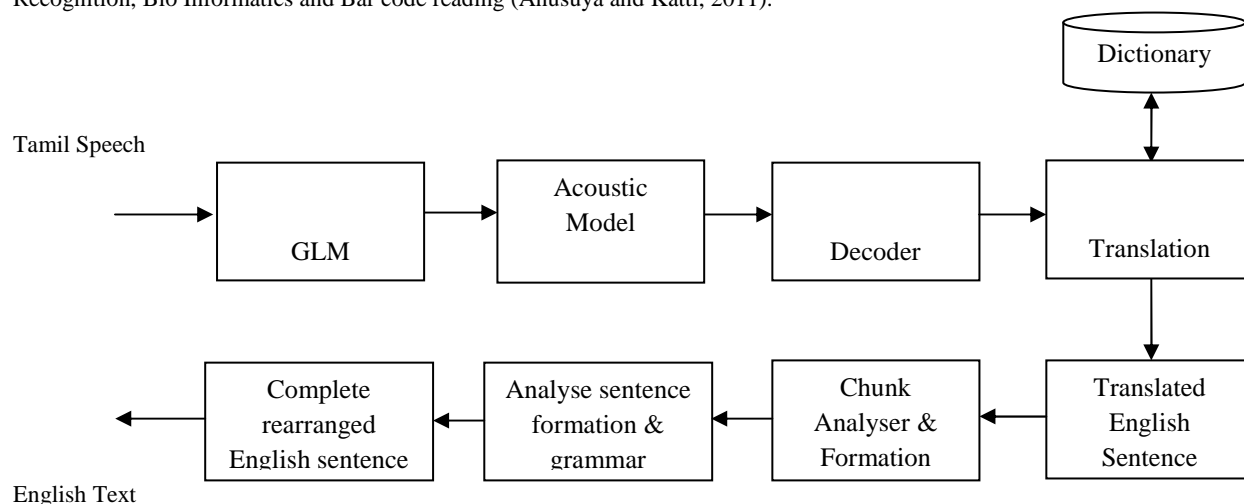


Figure 1: Overview of Proposed System – Phase I

The figure 1 represents the overview of proposed system for phase I i.e conversion of Tamil speech to English text.

3.2 Speech Recognition for Tamil Languages

Tamil language includes 12 vowels and 18 consonants. These combines to form 216 compound characters. The vowels are placed inside or both sides of the consonant to form the compound character. One special character in classical Tamil is ayutha ezhuthu (ஃ). In total there are 247 letters in standard Tamil alphabet.

Components of speech recognition engines

1. Language Model or Grammar- Language Models consists of a very large list of words. The frequency of occurrence of these words is in a given training sequence. Another smaller file of predefined combinations of words is called the grammar. Every single word in it is associated with a list of words or subword units.
2. Acoustic Model - Certain distinct words or the sounds which make up each word in the language model or grammar is represented statistically in the acoustic model.
3. Decoder - The words or subword units from the user is screened by the software program. The software program finds the acoustic model for the equivalent word or its unit.

The following steps are carried out to build simple isolated word recognition with whole word models using HTK.

1. Constructing the grammar
2. Constructing a dictionary for the models
3. Creating transcription files for training data.
4. Extracting the features
5. Training the Acoustic Model.
6. Evaluating the recognizer result from the test data.
7. Reporting recognition result.

3.3 Grammar and Dictionary Construction

The speech recognition engine (SRE) expects every input to be a constraint one. The presences of constraints are exclusively unfolded by the recognition grammar. The SRE listens for a list of words. The words trained in the acoustic model are only included in the grammar. A set of variable definitions and a regular expression that helps to recognize the words are included in the grammar. It is for used convenience word network defined using low level notations are needed for the speech recognizer. These low level notations are called Standard Lattice Format (SLF) (Young et al, 2006). Each word to word

transition is listed explicitly in SLF. The dictionary determines the way in which each word in the training data is expanded. It comprises of a collection of words in the training set. When the word is spotted the applicable symbol to output is also found in the dictionary, stepping towards construction of dictionary is sorting a list of required words. A simple list created by hand is enough for a small ASR system, but for a larger ASR system, the list of words is prepared from the training text corpus.

3.4 English Statistical Machine Translation System

Myanmar-English statistical machine translation system has framed source language model, alignment model, translation model and target language model to complete translation. In English statistical machine translation system (SMT) the input for source language model is an English sentence.

- Part-of-Speech (POS) tags and function tags for Myanmar word in the source language model and then it searches for the grammatical relations of Myanmar sentence.
- The translation model comprises of phrase extraction, translation from Myanmar sentences to English sentences by using Myanmar-English bilingual corpus. The Word Sense Disambiguation (WSD) system and the translation model combines to solve discrepancies when a phrase of a Myanmar sentence has more than one meaning.
- Another model that works along with all the other models is the alignment model. The primary function of which is to build the word and phrase, thereby aligning Myanmar-English bilingual corpus.
- The target language model includes two parts such as reordering the translated English sentences and smoothing it by using English grammar checker to reduce grammar errors.

The target language model is the proposed system used to monitor the grammatical errors for translated English sentences.

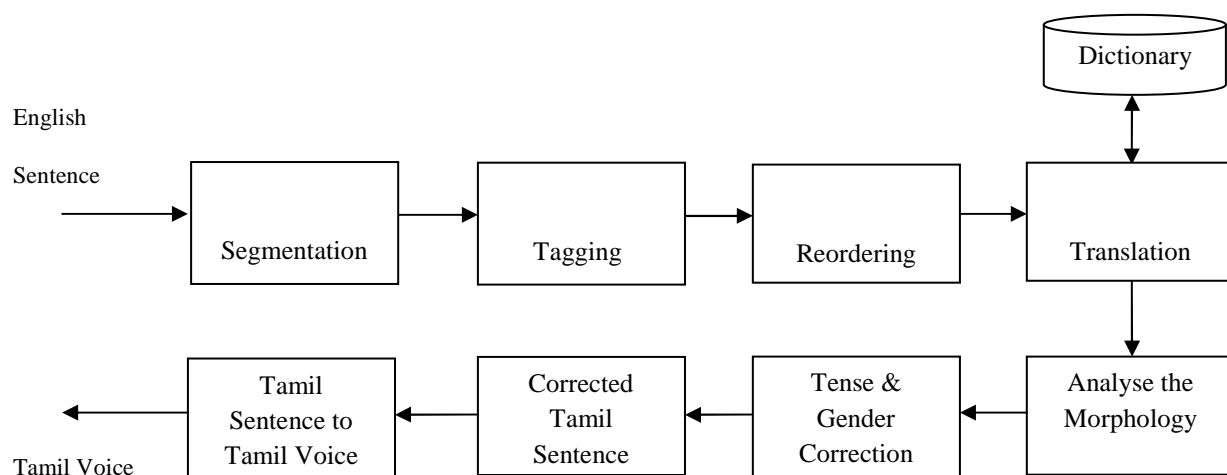


Figure 2: Overview of Proposed System – Phase II

The figure 2 represents the overview of proposed system for phase II (i.e) conversion of English text to Tamil voice.

3.5 Translation from English to Tamil:

In this phase the main objective of is to build a translation system that translates English sentences to Tamil Sentences. Due to the less availability of parallel corpora for English to Tamil the system is implemented using a Hybrid Technique. The Hybrid Technique is the combination of both Rule Based Technique and Statistical Technique. The system is first implemented in a Rule Based approach which involves segmentation and tagging, Rule Based Reordering, Morphological Analysing, and dictionary based translation to the Target language. Then the errors in the translated sentences are corrected by applying Statistical technique.

The segmentation and tagging of the source sentence is done using the Parts Of speech tagger (here Stanford POS tagger is used). The English sentences are taken as an input to the parts of speech tagger. The tagger tokenizes each word in a sentence

and identifies the parts of speech information such as verb, noun, adjective etc. of that word. Then the words and their tagged information are stored in a separate file which is used for reordering of sentences.

The tagged words are arranged according to the order which is mentioned below.

UH/ PP/ WP/WRB/WDT/ NNP/ PRP/ RB/ DT /CC/ JJ/ PP\$/ WP/JJR/ JJS/ IN/ NN/ NNS/TO/ VB/ VBD/ VBG/ VBN/ VBP/ VBZ/ MD

The Translation is done with the help of word dictionary file. The word dictionary file contains around 5000 words. If a word in an input sentence does not exist in the dictionary file then it will be considered as a proper noun. So those words are transliterated to the target language. Transliteration is implemented using the rule based technique. After the sentences are translated into the target language the morphological analysing has to be done. Because in Tamil language a single word shows the tenses, gender and the action performed which appears like a complete sentence. So the gender ending with the proper tenses has to be added with the verbs. The error in the target sentence is corrected statistically using the file which contains collection of Tamil verbs with proper tense and gender endings. Finally the Tamil sentence for the corresponding English sentence is generated, then the Tamil Sentences is converted into Tamil voice.

4. EXPERIMENT RESULTS

The design details are explained by using the following parameters

1. Extent of Measure
2. Chunk-based Grammar Checker
3. Use of Hybrid Techniques for translation from English to Tamil

1. Extent of Measure

There are several criteria's that measure the performance of an ASR. To list a few are Word Error Rate, Word Recognition rate or Accuracy and Correctness rate. Eq. 1 is made use to calculate the Correctness percentage of the test data, where N is the total number of test utterance, H is the number of word hits (utterance recognized), I, S and D are the number of Insertion errors, Substitution Error and Deletion Error.

1.1 Insertion Error (I)

It is denoted by I, which occurs when a data is inserted in the transcription (Morris A.C et. al, 2004) and it is not in the reference (training) set. An example for insertion error is shown in Figure 3.



Figure 3: Example for Insertion Error

1.2 Substitution Error (S)

The transcription data at times are substituted (Morris A.C et. al, 2004) by another Data as shown in the Figure 4 example. These kinds of errors are named to be the substitution error (S)

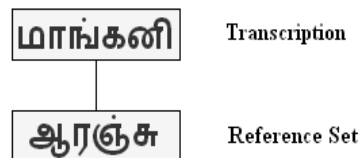


Figure 4: Example for Substitution Error

1.3 Deletion Error (S)

The deletion error (D) occurs when a data is not in transcription but present in the reference set (Morris et. al, 2004). The deletion error is shown with an example in Figure 5.

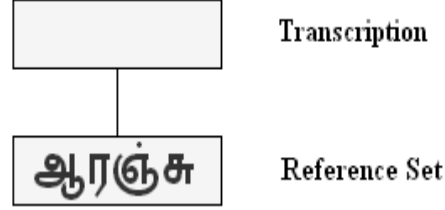


Figure 5: Example for Deletion Error

The correctness percentage is calculated using Eq. 1, with the number of Deletion Error and Substitution Error.

$$\text{Correctness Percentage} = \frac{N-D-S}{N} \times 100 \quad (1)$$

The accuracy percentage (Steve Young et al, 2006) is calculated using Eq. 2 where all the three errors are used for manipulation.

$$\text{Accuracy Percentage} = \frac{N-D-S-I}{N} \times 100 \quad (2)$$

The word error rate is calculated using Eq. 3.

$$W ER = \frac{S+D+I}{N} \times 100 \quad (3)$$

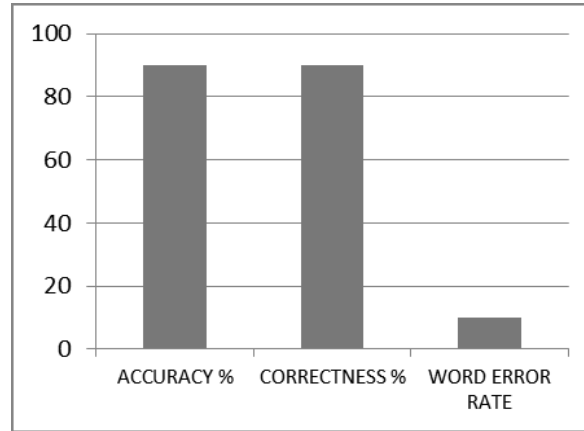


Figure 6: Graphical Representation of the Performance of ASR for the given test data

Chunk-based Grammar Checker

All the words used in SMT system comes from the corpus, hence the possibility of spelling errors are minimal in the translation output. Thus a target-dominant grammar checker formulated for Myanmar-English statistical machine translation system.

The SMT system is verified by testing on about 1800 sentences. The input sentence is analysed by the system and classified as simple, compound and complex sentences. Further information's about the sentence type such as interrogative or declarative also obtained. The possible errors prone to occur are subject verb agreement, missing chunks and incorrect verb form.

Use of Hybrid Techniques for translation from English to Tamil

The given source sentence is parsed and tagged using POS tagger and the resulting information is stored in a separate file. The reordering of the sentence is executed by following the above formulated order by using the tagged information. The bi-gram model is used as a tool for chunking of the source sentence. With the reference of the word dictionary file all these bi-grams are translated into Tamil language. The word by word translations are done with the bilingual dictionary. The absence of a word in the dictionary denotes that it is a proper noun which is to be transliterated to Tamil language. Following this the gender ending rules are verified with each sentence and target output is obtained. A file containing a collection of Tamil verbs with proper tense and gender endings will correct an error that still exists in the target sentence. To the end Tamil sentence for the corresponding English sentence is generated.

CONCLUSIONS

With the help of HTK tool an Isolated Tamil word Recognizer was built. Figure 6 pictures out the result obtained from the recognizer and it was found to have good accuracy rate for isolated Tamil words. The training set consist about 80 data and a test set of 10 data. These data's confined the accuracy and correctness percentage to be high. A word based model will not be suitable as the vocabulary size grows in the thousands range. As the vocabulary size grows it needs large storage space to store the acoustic models of all training data words thereby increasing the complexity of the system. A word out of the trained data is not identified by the isolated word ASR. At this instant a chunk-based grammar checker for translating English sentences making use of trigram language model and rule based model is implied. The various sentence patterns are spotted out with the help of Context Free Grammar rules and further to divide a text into chunk types relating to certain syntactic units. The training sentence patterns applied were our own formulations. Stanford POS tagger was used for identifying the Parts Of Speech information of the input sentences. So far we have created the word dictionary file which contains around 4500 words from English to Tamil. Gender ending rules for simple sentences such as simple present, simple past, simple future were created. Gender ending verbs for all possible tenses were created for the purpose of statistically correcting the errors. Finally Tamil speech is converted to English text and its sent to recipient. In recipient side the English sentence is converted into Tamil voice.

References

- [1] V.S.Dharun, M.Karnan, 2012, "*Voice and Speech Recognition for Tamil Words and Numerals*", IJMER, Vol. 2, 5, pp 3406-3414.
- [2] Dr.E.Chandra and A.Akila, 2012, "*An Overview of Speech Recognition and Speech Synthesis Algorithms*", Int.J.Computer Technology & Applications, Vol 3, 4, pp 1426-1430
- [3] M.A.Anusuya, S.K.Katti, 2011, "*Classification Techniques used in Speech Recognition Applications: A Review*", Int. J. Comp.Tech. Appl., Vol 2, 4, pp 910-954
- [4] Dhanalakshmi V and Rajendran S. (2010). "*Natural Language processing Tools for Tamil grammar Learning and Teaching*", International journal of Computer Applications (0975-8887), Volume 8, No.14.
- [5] Poornima C , Dhanalakshmi V, Anand Kumar M andv Soman K P, "*Rule based Sentence Simplification for English to Tamil Machine Translation System*", International Journal of Computer Applications (0975 – 8887), Volume 25– No.8, July 2011.
- [6] R. Harshawardhan, Mridula Sara Augustine and K.P. Soman, "*Phrase based English – Tamil Translation System by Concept Labeling using Translation Memory*", International Journal of Computer Applications (0975 – 8887), Volume 20– No.3, April 2011.
- [7] S. Sripirakas, A. R. Weerasinghe, D. L. Herath, University of Colombo School of Computing, "*Statistical Machine Translation of Systems for Sinhala - Tamil*", In the proceedings of the International Conference on Advances in ICT for Emerging Regions (ICTer), pp.62-68, 2010.

- [8] Zhemín Zhu, Delphine Bernhard and Iryna Gurevych 2010, “*A Monolingual Tree-based Translation Model for Sentence Simplification*”, Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010).
- [9] Sharma, A. and N. Jaiswal. 2010, “*Reducing Errors in Translation using Pre-editor for Indian English Sentences*”, Proceedings of Annual Seminar of CDAC-Noida Technologies (ASCNT), Noida, pp.70-76, India.
- [10] Crysmann, B., N. Bertomeu, P. Adolphs, D. Flickinger and T. Kluwer. August 2008, “*Hybrid processing for grammar and style checking*”, Proceedings of the 22nd International Conference on Computational Linguistics, pp. 153-160, Manchester.
- [11] Saraswathi Selvarajan and T.V.Geetha, 2007, “*Morpheme Based Language Model for Tamil Speech Recognition System*”, The International Arab Journal of Information Technology, Vol 4.No.3, pp 214-219.
- [12] R.Thagarajan et. al, 2008, “*Word and Triphone Based Approaches in Continuous speech Recognition for Tamil Language*”, WSEAS TRANSACTIONS on SIGNAL PROCESSING, Vol 4, 3, pp 76-85.

தன்னார்வ முனைந்து கற்றலுக்கான தமிழக ஆறாம் வகுப்பு தமிழ் பாடக் குறுஞ்செயலி

Abraham Meshach Ponraj, T.Priyanka, R.R.Jeyshree & B.Padma

Madurai Kamaraj University, Palkalai Nagar, Madurai 625 021

Email: ponrajcomputer@gmail.com

T.V.G. DURGASREE, Sri Vidya College Of Education, Virudhunagar, Tamil Nadu INDIA Sridurga_2008@rediffmail.com

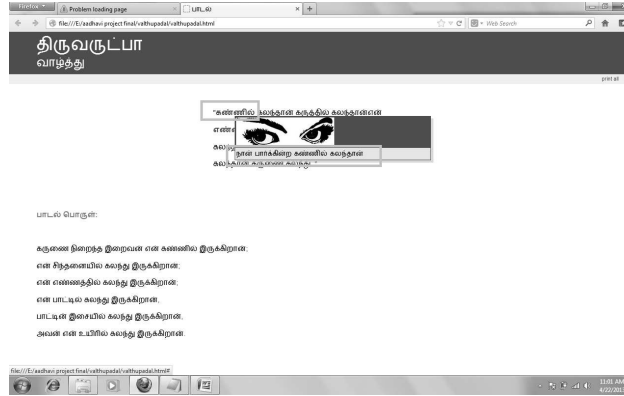
Abstract

As per the samacheer kalvi introduced by tamil nadu government, the CBSE students whose mother tongue is not tamil, should also learn tamil. It is difficult enough to learn a language that one hardly speaks at home. It is even more challenging when the spoken and written forms of the language are not the same, as is the case with the Tamil language. How can we get our students to use a language they do not own? In this paper, We address this question using activity based learning.

Introduction

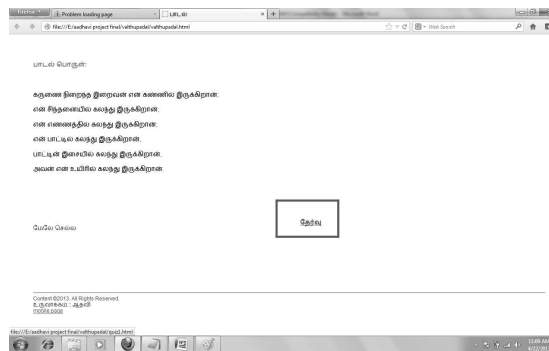
Activity-based learning or ABL describes a range of pedagogical approaches to teaching. Its core premises include the requirement that learning should be based on doing some hands-on experiments and activities. The idea of activity-based learning is rooted in the common notion that children are active learners rather than passive recipients of information. If child is provided the opportunity to explore by their own and provided an optimum learning environment then the learning becomes joyful and long-lasting. Here we explain various activities as follows:

ONLINE DOUBT CLEARING



Here valthupadal have 4 lines of poem. This session is called online doubt clearing this page will be redirect within a 30 min. Because student first read all the lines in this page then he/she understand the meaning of the poem then if they have any doubt clear it using text poppers when they move cursor over an appropriate word and popup text appears.

Skill based Evaluation



After 30mins these pages automatically redirect to next quiz page or students click that **தேர்வு** link. Then they attend that quiz session.

Next that Novice stage automatically direct to the quiz when they fix the puzzle words. That quiz contains missing word and மயங்கொலி பிழை. It is used to correct their spelling mistakes. This quiz related to the previous section so student easily identify the correct letter that word. Here also have time limitation it is based on that school teaching schedule. After pressing score button it will be again evaluates that student is qualified to next level or not. If else it will be come back to the novice stage. Otherwise it will be move on to next level that is AVERAGE stage.

Here given same puzzle game but time is less then novice stage. And here puzzle blocks contains just two types of colors. Those are dark blue and light blue. First row contain dark blue blocks. And second row contain light blue same color combinations helps to find next row of words. It is for average knowledge students. Then it move to next page that page contains that poem authors details then click quiz. That page contains questions about that author so student must answer correctly then only they move to next level else they read again that authors details. It is best way to know the particular authors detail correctly.



If student pass this level student move on to next level that level is EXPERT stage.

This level also contains a same puzzle but here no color combination that word with in a white box expert student can fix the pieces into a correct place. Then they attend the final quiz program that contain same questions about author here student must answer all questions then only he/she complete the lesson.



MOTIVATION VIDEO and AUDIO



Here motivation part includes the video content for பறவைகள் பலவிதம் lesson it is used to students know about that bird sanctuary. It is motivates them to read that lessons more interested.

Conclusion

Studying is no longer confined inside the four walls of the classroom. As iPhones, BlackBerrys, iPads and other devices grow in number; learning is becoming more mobile and easier for the average student.

Online learning environment is a ubiquitous in which students can become totally immersed in the learning process. So, online learning environment is a situation or setting of pervasive *or omnipresent* education *or learning*. Education is happening all around the student but the student may not even be conscious of the learning process.

BIBLIOGRAPHY

WEBSITE REFERENCES:

- ⇒ <https://en.wikipedia.org/wiki/Pedagogy>
- ⇒ <http://www.preservearticles.com/2012010920333/use-of-play-way-in-language-teaching.html>
- ⇒ <http://www.merlot.org/merlot/index.html>
- ⇒ <http://www.w3schools.com/>
- ⇒ <http://www.echoecho.com/flash.htm>
- ⇒ <http://www.tamiluv.com/>

BOOK REFERENCES:

- ⇒ **Learning Web Design: A Beginner's Guide to HTML, CSS, JavaScript, and Web Graphics** BY *Jennifer Niederst Robbins*
- ⇒ **JavaScript in Easy Steps** by Mike McGrath

செல்பேசியில் தமிழ்க் கலைச்சொல்லாக்கம்,

தரப்படுத்தம், செந்தரப்படுத்தம்

முனைவர் நா.ஜானகிராமன்

தமிழ்த்துறைத்தலைவர், பாரதிதாசன் பல்கலைக்கழகக் கல்லூரி, பெரம்பலூர்
janakirambdu@gmail.com

தமிழ்மொழியின் பயன்பாடு இன்று உலகளாவிய நிலையில் பல்வேறு தளங்களில் பரந்து விரிந்துள்ளது. அறிவியல் துறையில் நுழைந்த தமிழ் இன்று கணினித்துறையில் வெகுவாக முன்னேற்றம் கண்டு வருகின்றது. இதனைப்போல் செல்பேசியைப் பயன்படுத்தாதவர் யாரும் இல்லை என்று கூறும் அளவிற்கு செல்பேசியின் பயன்பாடு தவிர்க்க முடியாத ஒன்றாக உள்ளது. வலைத்தளங்களைக் கூட தமிழில் காண மென்பொருட்கள் வந்துவிட்டன. ஆனால் இன்று நாம் கடைகளில் வாங்கும் அனைத்து செல்பேசிகளில் நிறுவன செல்பேசிகளிலும் தமிழ்பயன்பாட்டுக்கு வந்துவிடவில்லை. நோக்கியா போன்ற நிறுவனமும், சில சீனத்தயாரிப்புகளிலும் மட்டுமே தமிழில் தமிழில் இயங்குதளம் உள்ளது. ஏனையவற்றுக்கு இல்லை. தமிழர்கள் வாங்கும் அனைத்து செல்பேசிகளிலும் தமிழ்மொழி இடம்பெற்றால் நாம் தமிழின் பெருமையை உயர்த்தியவர்களாகிவிடுவோம். அவ்வகையில் செல்பேசிகளில் அனைத்து கூறுகளிலும் தமிழ் இடம்பெற்றால் தமிழ் வளரும்.

கலைச்சொற்கள் இதற்கு அடிப்படையாகும். சில செல்பேசிகளில் இன்பாக்ஸ் என்று ஒலிபெயர்த்து எழுதப்பெறுகின்றன. கருவியின் செயல்பாடு, அமைப்பு, கருத்து எனப் பல்வேறு கோணங்களில் நின்று கலைச்சொல்லினை உருவாக்கலாம். ஏற்கனவே, பல்வேறு கலைச்சொற்கள் கணினி சார்ந்து உருவாக்கப்பெற்றுள்ளன.

இன்று பயன்பாட்டிலிருக்கும் சாம்சங் என்ற நிறுவனம் வெளியிட்ட செல்பேசியில் இடம்பெற்றுள்ள அறிவியல் கலைச்சொற்களுக்கு ஏற்ற தமிழ்ச்சொற்கள் படைக்கும் முயற்சியில் கலைச்சொல்லாக்கமும், தரப்படுத்தமும், செந்தரப்படுத்தம் இக்கட்டுரையின் ஆய்வாக அமைகின்றது. சென்ற ஆண்டு (2012) நடைபெற்ற தகவல் தொழில்நுட்ப மாநாட்டில் வழங்கப்பெற்று அதன் அடிப்படையில் அறிஞர்கள் வழங்கிய கலைச்சொற்கள் தொடர்பான கருத்துக்களைக் கொண்டு மேலும் சில கலைச்சொற்கள் மாற்றப்பெற்று செழுமைப்படுத்தப்பட்டுள்ளன.

கலைச்சொல் உருவாக்கம் –தரப்படுத்தம்- செந்தரப்படுத்தம்

கலைச்சொல் உருவாக்கம் வளர்ந்து வரும் அறிவியல் உலகில் மிக முக்கியமாகின்றது. உருவாக்கப்பட்ட கலைச்சொல்லாக்கம் தரமுடையதாகவும், எல்லோராலும் ஏற்றுக்கொள்ளக் கூடியதாகவும் இருக்கவேண்டும். வழக்குச்சொற்களைப் பயன்படுத்தியும், ஏற்கனவே வழக்கில் உள்ள கலைச்சொற்களை ஏற்றுக்கொண்டும் தமிழ் மொழியில் அறிவியல் கலைச்சொற்கள் வளர்தல் சிறப்பாகும். எங்கும் தமிழ் எதிலும் தமிழ் என்பது செல்பேசி வழி இன்று விரிவுபடுத்தலாம். காரணம், இதைவிட வேறு ஊடகம் தமிழுக்கு இன்று இல்லை. உருவாக்கம் என்பது முதல் நிலை. அதனைத் தரப்படுத்தி செம்மைப்படுத்துவது அடுத்தநிலை. செந்தரப்படுத்தல் என்பது இறுதிநிலை. இதற்குப்பிறகு உருவாக்கும் ஒரு சொல் எவ்வித விவாதங்களும் இன்றி ஏற்றுக்கொள்வதாக இருத்தல் வேண்டும். இந்த அளவிலான கொள்கையிலும் கோட்பாட்டிலும் கலைச்சொல்லியலார் கவனம் செலுத்த வேண்டும்.

கலைச்சொல் உருவாக்கலின் வரலாறு-நோக்கம்

தமிழில் கலைச்சொற்களைப் படைத்துக்கொள்வதின் மூலம் தமிழ்மொழியைப் பல்வேறு தளங்களில் வளர்த்தெடுக்க முடியும் என்ற நோக்கில் கலைச்சொற்கள் படைக்கப்படுகின்றன. வினை, பண்பு, பெயர், தொழில், அடை என்ற கூறுகளில் கலைச்சொற்கள் தோற்றம் பெறுகின்றன. அறிவியலையும் தொழில்நுட்பத்தையும் தமிழில் கொண்டு வர தமிழகத்தாரும் புலம்பெயர்ந்தாரும் எண்ணற்ற பணிகளில் ஈடுபட்டுவருகின்றனர். இவற்றில் தமிழ்நாட்டினரைவிடவும் புலம்பெயர்ந்தோர் பங்களிப்பு மகத்தானது. தமிழில் அறிவியலை வளர்க்கவும், அது தமிழில் நிலைபெறவும் தரமான கலைச்சொற்கள் மொழிக்குத் தேவையாகின்றன. எல்லோரும் கலைச்சொல்லாக்கத்துறையில் ஈடுபட்டு உழைத்தாலும் அவர்கள் உருவாக்குகின்ற சொற்கள் தரமுடன், செந்தரத்துடன் இயங்குகின்றனவா? என்பதிலும் அச்சொற்களை கையாள்கின்ற நபர்கள் எங்ஙனம் அதனைப் பயன்படுத்துகின்றனர் என்ற நோக்கிலும் கலைச்சொல்லியலாளர்கள் ஆராயவேண்டும் (இராதா செல்லப்பன் (2006:34). அந்தவகையில் இன்றைய பயன்பாட்டில் கலைச்சொற்களின் பயன்பாடு பல்வேறு இடங்களில் இடம்பெறிலும் செல்பேசியில் இடம்பெறல் என்பது மொழியினைப் பயன்படுத்துகின்றவர்கள் மேலும் மொழியினை

அறிவதற்கும் அறிவியலைத் தமிழில் தெரிந்துகொள்வதற்கும் வாய்ப்பாக இருக்கும் என்பதில் கருத்து வேற்றுமை இருக்க முடியாது.

அறிவியல் தமிழ் பயில் என்று சொன்ன பாரதிதாசனும், அறிவியலை தகவல் தொடர்பிற்கு எவ்விதம் பயன்படுத்தவேண்டும் என்று அன்றே சொன்ன பாரதியாரும் கலைச்சொல்லாக்கப் பாதைக்கு வழிகாட்டியவர்கள் எனலாம்.

கலைச்சொல்லாக்க முறைகள்

கலைச்சொல்லாக்கத்திற்குப் பல்வேறு நெறிகள் பின்பற்றப்படுகின்றன. எனினும் யாவரும் எளிதில் பயன்படுத்தும் வகையில், ஏற்றுக்கொள்ளும் வகையில் அமையப்பெறல் சிறந்தது. ஒலிபெயர்ப்பு, மொழிபெயர்ப்பு, கருத்துப்பெயர்ப்பு என்று மூன்றும் அடிப்படையானதாகும். இவற்றில் முன்னொட்டு, பின்னொட்டு என்ற வகையால் மேலும் பல பகுப்புக்களில் உருவாக்கலாம். எவ்வாறாயினும் கருத்துச்சிதைவு கூடாது என்பதில் தெளிவு வேண்டும். (இராதா செல்லப்பன் :2011) சொல்லுருவாக்கம் பல நிலைகளில் நிகழும் ஆயினும் அறிவியல் மற்றும் தொழில்நுட்பத்துறைக்கு கலைச்சொற்கள் தவிர்க்கமுடியாத ஒன்றாகும் (சுந்தரம். இராம (1995:34) கலைச்சொல்லாக்கத் துறையில் பல்வேறு அறிஞர்களின் கருத்துக்களை ஒருங்கிணைத்து பார்க்க வேண்டும்.

தமிழ்ச்சொற்கள் - கலைச்சொற்கள்

செல்பேசியைத் திறந்தவுடன் நம் கண்ணில் படுவன கீழ்க்கண்ட சொற்களாகும். கணினியில் ஏராளமான இடங்களில் நுழைந்த தமிழ் இந்த செல்பேசிக்குள்ளும் நுழைந்து தமிழ்மணம் வீசவேண்டும் என்ற நோக்கில் அவற்றுக்குரிய கலைச்சொற்கள் இங்கே உருவாக்கப்பெறுகின்றன. செல்பேசி இயங்கும் தன்மை, மற்றும் பொருண்மை, கருத்து, பயன்பாடு ஆகியவற்றின் அடிப்படையில் கலைச்சொல்லாக்கம் நிகழ்த்தப்பெறவேண்டும். சூழல் பயன்பாடு இங்கு முக்கியமாகின்றது.

உருவாக்கிய கலைச்சொற்கள்

sweep screen to unlock-திரை விலக்கி நுழைக.	Service numbers- சேவை இலக்கம்
On-தொடங்கு	Settings-அமைப்புகள்
Off-நிறுத்து	Message-செய்தி
logs -அழைப்புகள்	Send- அனுப்பு
all calls-அனைத்து அழைப்புகள்	Add media-ஊடக இணைவு
missed calls-தவறிய அழைப்புகள்	Picture-படம்
received calls-ஏற்ற அழைப்புகள்	Video-காட்சிப்படம்
dialled calls- முயன்ற அழைப்புகள்	Sound-ஒலி
copy to dialling screen-திரைவழி பேச படியெடு.	Photo-நிழற்படம் எடு
add to reject list - தவிர்ப்புப் பட்டியல்	Video record-காட்சிப்பதிவன்
Log manager- அழைப்பு ஆளுகை	Sound record-ஒலிப்பதிவன்
block by address-முகவரி தடு	Add text- செய்தி கூட்டு
Message-செய்தி	Add subject-பல்லாடகச் செய்திவழி
Delete-அழி	Attach item-இணைப்பு வகை
More-பிற	File-கோப்பு
Contact- தொடர்புகள்	Name card-முகவரி அட்டை
Groups-குழுக்கள்	Calendar-நாட்காட்டி
Favourite-விருப்பமானவை	Task-திட்டமிடல்
Assign-தேர்ந்தெடு	Memo-குறிப்பாணை
Create contact-தொடர்பு உருவாக்கு	Sending options-அனுப்பு வசதி
Copy-ஒட்டு	Reply path-பதில்வழி
Send name card via- முகவரி பட்டை அனுப்பு வழி	Request delivery-சேர்ந்த நிலை
Keep a copy-படியெடு	
Create contact-தொடர்பு உருவாக்கம்	Sending options-அனுப்பும் வசதி
Copy-படியெடு	Reply path-பதிலனுப்பு வழி
Send name card via- முகவரி பட்டை அனுப்பு வழி	Request delivery-சேர்ந்த நிலை
Keep a copy-படியெடு	
Music- இசை	Email- மின்னஞ்சல்
Push message-அனுப்பிய செய்திகள்	

All tracks-அனைத்துப் பாடல்	Voicemail number- குரல்ஞ்சல் எண்
Settings- அமைப்பு	Sending options- அனுப்பு வசதிகள்
Effects- கேட்பு இனிமை	Receiving options- ஏற்றல் வசதிகள்
Internet- இணையம்	Help- உதவி
Go to home page- முதல் பக்கம் செல்.	Memory status- நினைவக நிலை
History- விளக்கமுகவரி	Available- காலியிடம்
Messages-செய்திகள்	Busy count- பயன்படுத்திய பகுதி
Create message-செய்தி உருவாக்கு	Phone- செல்பேசியில்
Create email -மின்னஞ்சல் உருவாக்கு	Sim- சிம் அட்டை, எண்குறி அட்டை
Conversations-உரையாடல்	General settings- பொது அமைப்புகள்
Inbox-செய்திப்பெட்டி	Keep a copy- படியெடுக்க
Email inbox-மின்னஞ்சல் பெட்டி	Delete by folder- கோப்பினை அழிக்க
Sentbox-அனுப்பிய செய்திகள்	Block messages- செய்திகளைத் தடுக்க
Outbox-வெளிப்புறச் செய்திகள்	Alert repetitions- மறுமுறை நினைவூட்டு
Drafts-செய்திக்கோர்வை	Once- ஒருமுறை
Myfolder-செய்திக்கோப்பு	Every2minutues- இரண்டுநிமிடத்திற்கு ஒருமுறை.
Voicemail-குரல்வழி அஞ்சல்	My files- கோப்புகள்
Broadcast message-ஒளிபரப்பு வகை	Images- படங்கள்
Spam box-அழித்த செய்திகள்	Video-காட்சிப்படம்
Settings-அமைப்பு	Sound- ஒலிகள்
General settings- பொது அமைப்பு	Other- பிற
Text message- எழுத்துச்செய்தி	Create folder- கோப்பு உருவாக்கு
Multimedia message- பல்லாடகச் செய்தி	Wallpaper- முகப்புத்தாள்
Details- விபரங்கள்	Theme- பின்னொளிர் படம்
Time- நேரம்	Widgets- அமைப்பு வசதி
Type-வகை	Shortcuts- எளிதில் செல்ல
Name-பெயர்	Font type- எழுத்துரு
Size-அளவு	Brightness- வண்ண அளவு
Calendar-நாள்காட்டி	Backlight- பின்னொளிச்செறிவு
Create- உருவாக்கு	Call- அழைப்பு
Appointment- நாளுக்குறிப்பு	Voice call- குரல்ழைப்பு
Anniversary- ஆண்டுவிழா	Fake call- முகவரியற்ற அழைப்பு
Event-நிகழ்வு	Backup manager- மீள் பாதுகாப்பு
List by- வரிசைமுறை	View by- வகைகாணல்
Week- வாராந்திர	Backup- மீளெடுப்பு
Day-நாள்	Restore-மறு உள்ளீடு, மறுபதிவு
More-பிற	Help- உதவி
Go to day- நாளுக்குள் செல்	Security- பாதுகாக்க
Go to date- தேதிக்குள் செல்	Reset setting-மறுசீரமைப்பு
Delete- அழி	Master reset- மைய மறுசீர்
This month- இந்த மாதம்	Alarms-நினைவு ஒலி
Before date- தேதிக்கு முன்	Communities- சமுதாய ஒருங்கிணைவு
All events- அனைத்து நிகழ்வும்	Facebook- முகநூல்
Camera- படக்கருவி	Bluetooth- அருகமை பரிமாற்றம்
Sim card management- சிம்கார்டு நிர்வாகம்	Memo- குறிப்பாணை
Sound profile- ஒலிக்கோப்பு	Task- திட்டமிடு
Display-காண்திரை	FM Radio- பண்பலை
Time and date- நேரமும் நாளும்	Clock-கடிகாரம்
Network- இணைப்பு	Games- விளையாட்டு
Sim auto network- தானியங்கு வலையம்	Applications-பயன்பாடு
Registration- பதிவு	Voice recorder- குரல் பதிவன்
Sound-ஒலி	Calculator-கணக்கி
Normal- இயல்பு	Convertor-மாற்றி

Silent-ஒலியிலா	Dictionary- அகராதி
Driving-வாகன நேரம்	Display- காண்திரை, காட்சித்திரை
Outdoor- வெளிவேலை	---
Offline-இயங்காநிலை	--

ஏறத்தாழ நூற்று எழுபது கலைச்சொற்கள் இங்கே உருவாக்கிக் கொடுக்கப்பட்டுள்ளன. இவையனைத்தையும் செல்பேசியில் கணினி நிரல் வல்லுனர்கள் ஏற்றினால் தமிழ் உலகமெலாம் பரவும் வகையில் எளிதில் சென்றடையும்.

முடிவுரை

அறிவியலும் தொழில் நுட்பமும் வளர்ந்த இந்த சூழலில் செல்பேசிப் பயன்பாட்டுக்குத் தனித்தமிழ்ச்சொற்கள் உருவாக்கப்பெறல் வேண்டும். அதற்குரிய முயற்சியில் தமிழார்வலர்களும் மொழிப்பற்றாளர்களும் செயல்படவேண்டும். முன்பே உருவாக்கப்பெற்ற சில கலைச்சொற்களைப் பயன்படுத்தியும் வழக்குச்சொற்களுக்கு முக்கியத்துவம் கொடுத்தும் கலைச்சொற்களை ஆக்கவேண்டும். அதற்குரிய பணியினை தகவல் தொழில்நுட்பவியலாளர்களும் தமிழ் வல்லுனர்களும் ஏற்றால் தமிழ் தகவல் தொடர்புத்துறையில் மேலும் வளம்பெறும். உருவாக்கப்பெற்ற கலைச்சொற்கள் பல முதல்நிலையில் இருந்து தரப்படுத்தி, செந்தரப்படுத்தவேண்டும். அதனை மக்கள் வழக்கில் பயன்படுத்தவேண்டும்.

ஆதாரநூல்கள்

1. இராதா செல்லப்பன், (1985) கலைச்சொல்லாக்கம், மங்கை, வாத்தியார் தோட்டம், சென்னை
2. இராதா செல்லப்பன் (1992), அறிவியல் தமிழ் இன்றையநிலை, உலகத்தமிழராய்ச்சி நிறுவனம், சென்னை.
3. இராதா செல்லப்பன் (2008), மொழியியல், கவிதை அமுதம், திருச்சி
4. இராதா செல்லப்பன் (2011), தமிழும் கணினியும், கவிதை அமுதம், திருச்சி
5. சண்முகம்.செ.வை. (1994), அறிவியல் தமிழாக்கம், என்.சி.பி.எச். சென்னை.
6. சுந்தரம்.இராம (1999), பொருள் புதிது, வளம் புதிது (1999), வசந்தம் பதிப்பகம், தஞ்சாவூர்
7. ஜானகிராமன் (2003), அறிவியல் கலைச்சொல்லாக்கம், இராகவன் பதிப்பகம். விருத்தாசலம்
8. ஜானகிராமன் (2004) அறிவியல் தமிழ் வரலாறு (2004), இராகவன் பதிப்பகம், விருத்தாசலம்.

அலைபேசிகளிலும் கைக்கணினிகளிலும் குறுஞ்செயலிகளுக்கான இணையத் தொழில்நுட்ப நியமனங்கள் சுகந்திவெங்கடேஷ்

உலகத்தை உள்ளங்கையிலேயே வலம்வரும் திறனை ஒவ்வொரு மனிதனுக்கும் இன்றைய சாமர்த்திய அலைபேசிகள் கொடுத்துள்ளன. இணையத்தை உலவுவதும் மின்னஞ்சல் பார்க்க, பாடல் கேட்க, புகைப்படம் பிடிக்க, விளையாட என்று சில காரணங்களுக்காக முதலில் பயன்படுத்தப்பட்டு வந்த நிலையில் சாமர்த்திய அலைபேசிகளும், கைக்கணினிகளும், மின்வாசிப்பான்களும் குறுஞ்செயலிகளை தற்போது அறிமுகப்படுத்தி ஒரு மேசைக்கணினியின் வேலையைச் செய்ய இன்று முனைகின்றன. இந்தத் தொழில்நுட்பம் எதிர்காலத்தில் இன்னும் வேகமாகவளரும்.

வணிக நோக்கத்துடன் பார்த்தாலும் தமிழ்மொழியின் வளர்ச்சி என்ற பரிணாமத்தின் இன்னொரு முகமாக இதை எடுத்துக்கொண்டாலும் தொழில்நுட்பத்தின் தொடக்கநிலையிலேயே இதைக் கணினித்தமிழும் பயன்படுத்திக்கொள்வது மிக அவசியம்.

பல்வேறு கையடக்கச் சாதனங்களும் சாமர்த்திய அலைபேசிகளும் ஒவ்வொரு தனிமனிதனின் தேவையைப் பூர்த்திசெய்யும் வகையில் இன்று சந்தையில் கிடைக்கின்றன. அவை ஒவ்வொன்றுக்கும் தனித்தனி நடமாடும் இயங்குதளங்கள் உள்ளன. Android, iOS, Symbian, RIM's BlackBerry OS, Samsung's Bada, Microsoft's Windows Phone, webOS, MaemoMeeGo ஆகிய இயங்குதளங்கள்தவிர Mozilla's Firefox OS, Canonical Ltd.'s Ubuntu Phone ஆகிய புதிய நடமாடும் இயங்குதளங்களும் சந்தையில் விரைவில் வர உள்ளன. இப்போது சந்தையில் உள்ள மின்வாசிப்பான்களும் கையடக்கக் கணினிகளாகச் செயல்படத் தொடங்கியுள்ளன. இந்தக் கருவிகள் சாதாரணமாகச் செயல்பட இந்தக் குறுஞ்செயலிகள் உதவுகின்றன.

சாமர்த்திய அலைபேசிகள் உருவான தொடக்கத்தில் ஒவ்வொரு இயங்குதளத்திற்கும் தனித்தனியாகக் குறுஞ்செயலிகள் உருவாக்கப்பட்டன. அந்தக் காலக்கட்டத்தில் எல்லா அலைபேசிகளிலும் கைக்கணினிகளிலும் குறுஞ்செயலிகள் செய்ய ஒருவருக்குத் தெரிந்து இருக்கவேண்டிய தொழில்நுட்பங்களைக் கீழே உள்ள அட்டவணை காட்டுகிறது.

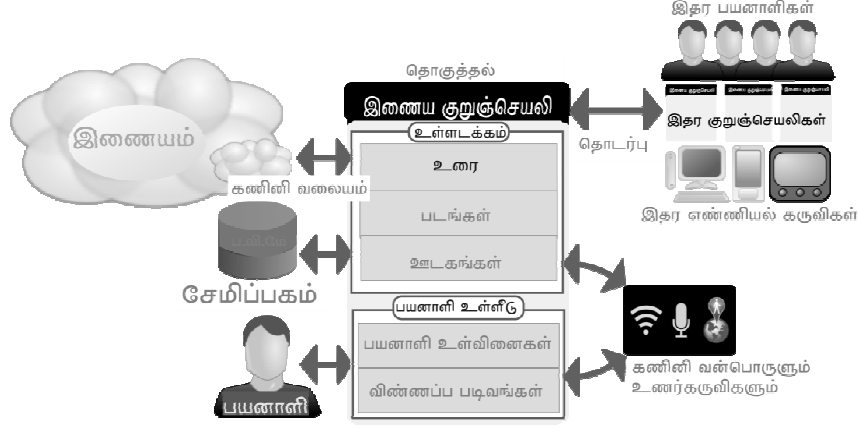
இயங்குதளம் (platform)	மென்பொருள் உருவாக்கச் சூழ்நிலை (Integrated development environment)	கணினிமொழி (language)
IOS	XCode	Objective-C
Android	Eclipse	Java
Windows Phone	Visual Studio	C# or VB. NET

இதனால் மூலதனம், நேரம், மனித ஆற்றல் ஆகியவை வீணடிக்கப்பட்டன. ஒரே உட்பொருளை வேவ்வேறு அலைபேசிகளிலும் கணினிகளிலும் கொண்டுவருவதில் பல பிரச்சனைகள் எழுந்தன. அதே நேரத்தில் எல்லா இயங்குதளங்களிலும் ஒரேமாதிரியாக இணையப்பக்கங்கள் தெரிவதில் ஒரு பிரச்சனையும் இருக்கவில்லை. அலைபேசிகளிலும் கைக்கணினிகளும் உட்பொருளை நல்லவிதமாகக் காட்ட இணைய உலாவிகள் முன்வந்தன.

இதனால் இணையத் தொழில்நுட்பங்களான ஹெடிஎம்எல் 5, சி.எஸ்.எஸ் 3 ஜாவாஸ்கிரிப்ட் ஆகிய தொழில்நுட்பங்கள் கையடக்கக் குறுஞ்செயலிகள் அமைப்பதில் மிக முக்கியப் பங்கு வகிக்க ஆரம்பித்தன. அவைகளைப் பயன்படுத்தி இணையத்தளங்களே குறுஞ்செயலிகள் போலச் செயல்படத் தொடங்கின.

குறுஞ்செயலிகள் என்றால் என்ன என்று சிந்தித்தால் இவை மேஜைக்கணினிகளில் செயல்படும் மென்பொருள்களை ஒத்தவை. குறுஞ்செயலிகள் என்று சொல்லும்போது தகவல் தருபவையாகவும், வேலை செய்ய உதவுவதாகவும் இரு பிரிவாகப் பிரிக்கலாம். இந்த இரு செயல்களையுமே இணைய உலாவிகள் ஏற்கனவே சரியாகச் செய்து கொண்டிருந்த காரணத்தால் இணையத்தளத் தொழில்நுட்பங்களே குறுஞ்செயலிகள் அமைப்பதற்குச் சிறந்த தொழில்நுட்பங்களாக இப்போது கருதப்படுகின்றன.

ஒரு குறுஞ்செயலி இணையத் தொழில்நுட்பங்களைக் கொண்டு சிறப்பாகச் செயல்பட வேண்டும் என்றால் அதற்கு என்னென்ன தேவை என்பதைக் கீழே உள்ள படம் காட்டுகிறது.



பயனாளிகள் பார்க்கவேண்டிய படங்கள், உரைகள் மற்றும் பல ஊடகங்கள் தெரியும் விதம், பயனாளிகள் உள்ளிடும் விவரங்களும் விண்ணப்பப் படிவங்களும் அமையும் விதம், கையடக்கக் கணினியின் புகைப்படக்கருவி, ஒலி உள்வாங்கி, புவி இருப்பிடக் குறிப்புக்குறி (geolocation) ஆகியவற்றிலிருந்து விவரங்களைப் பெறும் வகை, பயனாளிகளின் சொந்த விவரங்களை அறிந்து அவற்றைச் சேமித்துவைக்கும் முறை, கணினிக் கட்டமைப்பு, தொலைத்தொடர்பு, கையடக்கக் கருவியின் செயல்திறன், அதை மேம்படுத்தும் வகை என்று பல வளராக்கங்களை இந்த இணையத் தொழில்நுட்பங்கள் செய்ய உதவுகின்றது என்று உலக இணையத்தளங்களின் கூட்டமைப்பு கூறுகிறது. மேகக்கணினிப் பயன்பாட்டிலும் குறுஞ்செயலிகளைப் பயன்படுத்த இணையத் தொழில்நுட்பங்கள் உதவுகின்றன.

ஹெடிஎம்எல் 5

இன்று உலக இணையத்தளங்களின் கூட்டமைப்பினால் பரிந்துரைக்கப்படும் ஒரு முக்கியமான இணையப்பக்கத் தொழில்நுட்பமாகும். இது பலவகைகளில் குறுஞ்செயலிகள் அமைப்பதில் உதவுகிறது. இணைய நியமனங்களுக்கு ஏற்றவகையில் குறுஞ்செயலிகளை அமைக்கும்போது அது மக்களை அடைய எளிதாகிறது. இதனால் எல்லாக் கையடக்கக் கருவிகளிலும் பயனாளிகள் ஒரே மாதிரியான அனுபவத்தைப் பெற முடிகிறது. ஒவ்வொரு கைக்கணினிகள், சாமர்த்திய அலைபேசிகளுக்கு இடையில் உள்ள வேறுபாட்டை விரைவாகவும் வேகமாகவும் தகர்த்தெறிய ஹெடிஎம்எல் 5 உதவுகிறது.

<!doctype html> என்ற எளிதான அறிவிப்பு எல்லாக் கையடக்க உலாவிகளிலும் குறுஞ்செயலிகள் பிரச்சனையின்றி வேலை செய்ய உதவுகிறது. எந்த எழுத்துருவின் உதவியில்லாமல் ஒருங்குறி எழுத்து என்ற குறிப்பிடும்படும் கொண்டு தமிழில் எளிதாக எழுதலாம். <audio> <video> என்ற இரு இழைகளும் ஒலி, காணொளி ஊடகங்களை எல்லாக் கருவிகளிலும் எந்தத் தடையின்றிக் காட்ட உதவுகிறது. கையடக்கக் கருவிகளில் உள்ள ஊடகங்களைப் பதிவு செய்யும் வசதிகளான புகைப்படம் பிடித்தல், ஒலி, ஒளியைப் பதிவு செய்தல் ஆகியவற்றைக் கையடக்க உலாவிகளில் சேர்க்கும் ஆராய்ச்சிகள் உலக இணையத்தளங்களின் கூட்டமைப்பினால் நடத்தப்பட்டுவருகின்றன. அது மட்டுமல்ல, கையடக்கக் கருவிகளின் நுண்ணிய விவரங்களின் களஞ்சியத்தைப் பற்றிய ஆராய்ச்சியும் நடந்துவருகிறது.

<canvas> என்ற இழை கையடக்கக் கருவிகளைப் பயன்படுத்திப் படங்களை வரையவும் புகைப்படங்களை மாற்றிஅமைக்கவும் செய்யலாம். இது குழந்தைகளுக்கான விளையாட்டுகள் தயாரிக்கவும் பயன்படுத்தப்படுகிறது. <svg> என்ற இழைகொண்டு பயன்படுத்தப்படும் ஆயக்கணிதமுறையில் வரைந்த படங்கள் கையடக்கக் கருவிகளின் திரை அளவிற்கு ஏற்றபடி சரியாக மாறிக்கொள்கிறது.

கையடக்கக் கருவிகளின் மிக முக்கியமான தேவை பயனாளர்களிடமிருந்து விவரங்களைத் திரட்டும் விண்ணப்பப் படிவங்கள் ஆகும். மேசைக்கணினிகளில் விண்ணப்பப் படிவம் செம்மையாகச் செயல்படுவதற்கு இணையப் பரிமாற்றகத்திற்கு விவரங்கள் செல்லவேண்டிய தேவை இருந்தது. தற்போது ஹெடிஎம்எல் 5 விண்ணப்ப வடிவ இழைகளில், கையடக்கக் கருவிகளே எளிதாக முதல் கட்டச் சோதனையைச் செய்து முடிக்க உதவுகின்றன. <input required> என்ற பண்பு எந்த விதமான அதிகப்படியான விதிகள் இல்லாமல் பயனாளிகள் தங்கள் விவரங்களைக் கொடுத்துள்ளனரா என்று ஆய்வுசெய்து தெரிவித்துவிடுகிறது. <input type="email" pattern="^[^ @]*@[^ @]*" value=""> என்ற சிறு குறிப்பு, பயனாளிகள் மின்னஞ்சலைச் சரியாக இட்டு இருக்கின்றனரா என்று சரிபார்க்கிறது. Place holder என்ற பண்பு என்ன மாதிரி விவரம் எதிர்பார்க்கப்படுகிறது என்ற விவரத்தைச் சுருக்கமாகத் தெரிவிக்க

உதவுகிறது. ஹெடிஎம்எல் 5 தொழில்நுட்பத்தைப் பயன்படுத்தி எழுதும்போது குறுஞ்செயலிகள் பயனாளிகளின் வேகத்திற்கும் அவசரத்திற்கும் ஏற்றபடிச் செயல்படும்.

சி.எஸ்.எஸ் 3

இன்னுமொரு முக்கியமான இணையத் தொழில்நுட்பம் கையடக்கக் கணினிகளில் குறுஞ்செயலிகள் உருவாக்க உதவுகிறது. சி.எஸ்.எஸ் 3 கொண்டு குறுஞ்செயலிகளை கையடக்கக் கருவிகளுக்கென்று பிரத்தியேகமாக உருவாக்கப்பட்டவைகளாக அலங்கரித்துக்கொள்ளலாம். மிக முக்கியமாக சி.எஸ்.எஸ் 3ல் உள்ள ஊடகத்தைப் பற்றிய விவரங்கள் (CSS Media Queries) கையடக்கக் கணினியின் வடிவமைப்பின் அளவுக்கு ஏற்றபடி குறுஞ்செயலியின் அளவை மாற்றிப் பயனாளிகளுக்குக் காட்டும். சி.எஸ்.எஸ் 3ல் அளவுமுறைகளை விழுக்காடாகவோ em என்று அழைக்கப்படும் எழுத்துருக்களின் அளவையொத்தோ அமைக்கும்போது எல்லாக் கையடக்கக் கருவிகளிலும் சாமர்த்திய அலைபேசிகளிலும் பக்கங்களும் படங்களும் எழுத்துருக்களும் தானாகவே மாறிக்கொள்ளும்.

சி.எஸ்.எஸ் 3ன் நிலைமாற்றும்தன்மை (Transitions), உருமாற்றும்தன்மை (transform) ஆகியவை மூலமாகச் சிறு சிறு அசைவுகளை (animation) குறுஞ்செயலிகளுக்குள் உருவாக்கலாம். இருபரிமாண, முப்பரிமாண (2D and 3D) அசைவுகளும் இதில் அடங்கும். தற்போது நாம் பயன்படுத்திவரும் True Type, Open Type எழுத்துருக்கள்என்னும் WOFF என்ற எழுத்துரு கோப்பு வடிவமாக எளிதில் மாற்றி நாம் தமிழ்மொழியில் தற்போது உள்ள அச்சப் பிரதி எழுத்துருக்களை இணைய எழுத்துருக்களாக மாற்றி இணையத்தில் பயன்படுத்தலாம். @font-face என்ற அறிவிப்பைப் பயன்படுத்தி நம்முடைய எழுத்துருக்களை எளிதாகக் கையடக்கக் கருவிகளில் இணைக்க இயலும்.

ஜாவாஸ்கிரிப்ட் என்னும் நிரலர் மொழி மூலம் பயனாளிகளிடம் கையடக்கக் கணினியையும் எளிதாகப் பேசவைக்க முடியும். இவ்வாறு பேசவைக்கும் மென் ஒருங்கு இயக்கத் திட்டச் செயலாக்க இடைமுகப்புக்கள் (application programming interface) அனைத்தும் இந்தத் தொழில்நுட்பத்தைச் சார்ந்து உள்ளன.

முக்கியமாகப் பலவிதமான விவரங்களைச் சேமித்துவைப்பதில் ஜாவாஸ்கிரிப்ட் முக்கியப் பங்கு வகிக்கிறது. கையடக்கக் கருவிகள் தங்களுக்குள்ளே சேமிக்கும் விவரங்களாகட்டும், கணினிப் பரிமாற்றகத்தில் சேமிக்கவேண்டிய விவரங்களாகட்டும், அனைத்தும் எளிமையான ஜாவாஸ்கிரிப்ட் கொண்டு அமைக்க முடியும். அடுத்ததாக இணையத்தொடர்பு இல்லாத நிலையிலும் இணைய உலாவிகள் சரியாக வேலை செய்யவேண்டிய விவரங்களைச் சேமிப்பதும் ஜாவாஸ்கிரிப்ட்டின் வேலைதான்.

குறுஞ்செயலிகளில் முக்கியமான இடத்தைப் பிடித்திருப்பது விளையாட்டுகள். அவற்றை ஹெ.டி.எம்.எல் 5 இன் <canvas> இழையைப் பயன்படுத்திப் பல வரையும் மென் ஒருங்கு இயக்கத் திட்டச் செயலாக்க இடைமுகப்புக்களை ஜாவாஸ்கிரிப்ட் மூலம் எளிதாகவும் விரைவாகவும் அமைக்க முடியும். ஜாவாஸ்கிரிப்ட் அடிப்படையில் எழுதப்படும் மென் ஒருங்கு இயக்கத் திட்டச் செயலாக்க இடைமுகப்புகளைப் பயன்படுத்துவதால் பயனாளிகள் செறுகிகளைத் (plugin) தரவிருக்கச் செய்யவேண்டிய தேவை இல்லை. உடன்நிகழ் திட்டச் செயலாக்க முறைகளையும் (Multi-threaded or concurrent programming) ஜாவாஸ்கிரிப்ட் மூலம் அடைய முடியும். அதனால் கைக்கணினிகள் வேகமாக வேலை செய்ய முடியும். கோப்புகளைப் படிப்பதற்கும் கோப்புகளை எழுதுவதற்கும் அவற்றைப் பிரித்து வரைமுறைபடுத்துவதற்கும் ஜாவாஸ்கிரிப்ட் கொண்டு தயாரிக்கப்படும் மென் ஒருங்கு இயக்கத் திட்டச்செயலாக்க இடைமுகப்புகளை உருவாக்க முடியும்.

CreateJS Suite என்பது ஹெ.டி.எம்.எல் 5 அடிப்படையில் அமைந்த ஓர் ஊடாடும் திறவூற்று ஜாவாஸ்கிரிப்ட் கட்டமைப்பு. இதனுள்ளிருக்கும் EaselJS மூலம் பலவிதமான படங்கள் வரையும் வகை எளிதாக்கப்பட்டுள்ளது. <canvas> இழையின் அடிப்படையில் அமைந்து இருந்தாலும் flash தொழில்நுட்பத்தைப் போல அமைக்கப்பட்டு, நிரலர்களுக்கு

உதவியாக இருக்கிறது. இதில் இருக்கும் TweenJS, EaselJS-ஓடு இணைந்து அசைவூட்டங்களை அமைக்க உதவுகிறது. ஹெ.டி.எம்.எல் 5ல் உள்ள ஒலிக்கான <audio> இழை விளையாட்டுகளுக்கோ பயனாளிகளுடன் ஊடாடுவதற்கோ சரியாக உதவுவதில்லை. இந்தக் குறையைப் போக்க நிரலர்கள் SoundJS என்ற கட்டமைப்பைப் பயன்படுத்துகின்றனர். PreloadJS படங்கள், ஒலிகள் ஆகியவற்றைக் கையடக்கக் கணினிகளில் ஏற்ற உதவுகிறது. அடுத்து உள்ள படம் ஒரு அசைவூட்டம் எப்படி preloadjs மூலம் தயாரிக்கப்படுகிறது என்பதைக் காட்டுகிறது.



பயனாளிகள் பயன்படுத்தும் நாள்காட்டி, விலாசப்பத்தகம் ஆகியவற்றில் உள்ள விவரங்களைக் குறுஞ்செயலிகள் பெறுவதற்கும் ஜாவாஸ்கிரிப்ட் வழிசெய்கிறது. புவி இருப்பிடக் குறிப்புக்குறிகளின் மென் ஒருங்கு இயக்கத்திட்டச் செயலாக்க இடைமுகப்புகளைக் கையடக்கக் கணினிகளின் இருப்பிடத்தைத் துல்லியமாக அறிந்து பயனாளிகளின் தேவையை இடத்திற்கு ஏற்றபடி பூர்த்திசெய்ய உதவுகிறது. கையடக்கக் கணினிகள் முக்கியமாகப் பயன்படுத்தப்படுவது தொடர்புகொள்வதற்காகத்தான். பயனாளிகள் ஒருவரை ஒருவர் குறுஞ்செய்திகள் அல்லது இணையச்செய்திகள் மூலம் தொடர்புகொள்கின்றனர். அதற்கு உதவுவது ஹெச்.டி.எம்.எல்லின் இணையச் செய்திமுறைகள்.

ஜாவாஸ்கிரிப்ட்விர அஜெக்ஸ், பைதான், எக்ஸ்.எம்.எல், ஆகிய தொழில்நுட்பங்களும் குறுஞ்செயலிகள் அமைக்க உதவியாய் அமைகின்றன.

குறுஞ்செயலிகளின் பகிர்ந்தளிப்பு என்பது குறுஞ்செயலிகள் எழுதுவோரின் ஒவ்வோருவரின் கவனமும் இருக்கவேண்டும். ஒவ்வொரு இயங்குதளத்திற்கும் ஏற்றபடி அந்தந்த நிறுவனங்களில் நிரலர்கள் பதிவு செய்துகொள்ளவேண்டும். அந்தந்தக் கையடக்கக் கணினிகளுக்கு ஏற்றார்போல் குறுஞ்செயலிகளை அமைக்க உதவுவது JQuery Mobile என்ற ஜாவாஸ்கிரிப்ட் கட்டமைப்பு. phone gap என்ற ஜாவாஸ்கிரிப்ட் கட்டமைப்பு இணையத் தொழில்நுட்பங்களைக் கொண்டு அமைந்த ஒரு குறுஞ்செயலியை ஆப்பிள், ஆண்ட்ராய்டு, விண்டோஸ், பிளாக்பெர்ரி ஆகிய இயங்குதளங்களுக்காக விரைவாக மாற்றிவிடுகிறது. மேலே சொன்ன இரு கட்டமைப்புகளும் ஒரு திறவூற்றுக் கட்டமைப்பு என்பது முக்கியம்.

சுருக்கமாகச் சொன்னால் ஒரு குறுஞ்செயலியின் உள்ளடக்கங்களான உரைகள், படங்கள், ஊடகங்கள் ஆகியவற்றைப் பிழையின்றிக் காட்ட உதவுவது ஹெ.டி.எம்.எல்5. காட்டப்படும் உட்பொருள்களுக்கு வண்ணம் கொடுப்பது, அவற்றை மேம்படுத்திக் காட்டும் எழுத்துருக்களை இணைத்து அலங்காரப்படுத்துவது போன்ற வேலைகளை சி.எச்.எச்.3 செய்கிறது. பயனாளிகளிடமிருந்து விவரங்கள் சேகரிக்கவும், சேகரித்த விவரங்களைச் சேமித்துப் பகிர்ந்துகொள்ளவும், ஒரு பயனாளியின் கையடக்கச் சாதனம் மற்றக் குறுஞ்செயலிகளுடனும், மற்ற உணர்கருவிகளுடனும் வன்பொருட்களுடனும் உறவாடச் செய்யவும் அடிப்படையாக இருப்பது ஜாவாஸ்கிரிப்ட்.

முடிவுரையாக சொல்ல வேண்டியது கையடக்கக் கருவிகள் இன்று மனிதனுக்கு இன்றியமையாதவையாக உள்ளன. இணையத்தின் வளர்ச்சியால் இணைய உலாவிகளே மென்பொருள்களாகப் பயன்படுத்தப்படுகின்றன. இந்தநிலையில் இணையத் தொழில்நுட்பங்களே இன்றைய சூழ்நிலையில் குறுஞ்செயலிகள் தயாரிக்க ஏதுவானவை. பல்வேறு கையடக்கக் கணினிகளில் உள்ள வேறுபாட்டை இவை எளிதில் தாண்ட உதவுகின்றன என்பதால் இவையே ஒரு பொது நியமனமாக எதிர்காலத்தில் அமையும். இணையத் தொழில்நுட்பங்களைப் பற்றிய அறிவு மட்டும் இருந்தால் போதும், அவற்றிற்கு என்று தனியாக விலையுயர்ந்த மென்பொருள்களை வாங்கவேண்டியதில்லை. இதனால் இந்த வளரும் தொழில்நுட்பங்கள்மூலம் தமிழில் குறுஞ்செயலிகளைச் சிக்கனமாக அமைத்து தொழில் அளவில் தமிழால் வெற்றி அடையலாம்.

தமிழ்க் குறுஞ்செயலிகள் உருவாக்கம், உள்ளடக்கம், பயன்பாடு - ஒரு மதிப்பீடு

முனைவர் சி. சிதம்பரம் M.A., M. Phil., Ph. D.,

துணைப்பேராசிரியர், தமிழ்ப்புலம், காந்திகிராம கிராமியப் பல்கலைக்கழகம்,
மின்னஞ்சல் : mudalvaa@gmail.com

வளர்ந்துவரும் அறிவியல் தொழில்நுட்ப உலகில் சூட்டிகைப் பேசிகளின் (Smart Phones) வருகை மனிதனின் செயல்பாடுகளை விரைவுபடுத்தியுள்ளன. இணையத்தைப் பயன்படுத்தி கணினிவழி செயல்பட்ட மனிதன் அதே பணிகளை நவீன கையடக்கக் கருவிகளிலும் பயன்படுத்தத் தொடங்கிவிட்டான். ஆரம்பக் காலகட்டங்களில், செல்பேசி சந்தையில் நோக்கியா (Nokia), மோட்டரோலா (Motorola), சோனி எரிக்சன் (Sony Ericsson), சாம்சங் (Samsung), எல்.ஜி (L.G) ஆகிய நிறுவனங்கள் மட்டுமே ஆதிக்கம் செலுத்தி வந்தன. இந்நிறுவனங்கள் தயாரித்த ஆரம்பகாலச் செல்பேசிகள் ஒலிவழித் தகவல் தொழில்நுட்ப வசதி கொண்டதாக மட்டுமே இருந்தன. இவை செல்பேசிகள் கண்டுபிடித்த காலகட்டங்களில் மக்களால் பரவலாகப் பயன்படுத்தப்பட்டு வந்தன. ஆரம்ப கால கட்டத்தில் செல்பேசி தொலைவில் உள்ளவர்களைக் குரல்வழித் தொடர்புகொள்ளும் நிலையில் மட்டும் (தூரத்தின் அளவைக் குறைக்கும் வகையில்) அமைந்திருந்தது.

செல்பேசி முதலில் ஒரு தகவல் தொடர்பு சாதனமாக மட்டுமே இருந்தது. ஆனால், இன்று செல்பேசியில் உள்ள வசதிகளால், பல அறிவியல் சாதனங்கள் பயன்பாட்டில் இல்லாமல்போய்விட்டன என்று சொல்லும் அளவிற்குச் செல்பேசியில் கணிப்பான் (Calculator), கடிகாரம் (Watch), குறிப்பேடுகள் (Notes), செய்தித்தாள்கள் (News Papers), இணைய வானொலிகள் (Internet Radios), வரைபடங்கள் (Maps), கடக்க வேண்டிய வழிகளை, திசைகளைக் காட்டும் கருவி (GPS), புகைப்படக் கருவி (Camera), வானிலையை அறிந்து கொள்ளும் வசதி (Whether), மின்னூல்கள் (E-Books), தேடுபொறிகள் (Search Engine), பேருந்து, புகைவண்டிச் சீட்டுகள் முன்பதிவு செய்யும் வசதி (Online Ticket Booking), தமிழ், ஆங்கில மொழிகளில் தட்டச்சு செய்யும் வசதி (Typewriting), குறுஞ்செய்தி, நிழற்படம், குறிப்புகள், குரல் அனுப்பும் வசதி (SMS, Multimedia Massage), நேரடி வங்கி சேவைகள் (Online Banking), இணைய அகராதிகள் (Internet Dictionary), நேரடித் தொலைக்காட்சிகள் (Online Television), சமுதாய வலைத்தளங்கள் (Social Networks), நேரடி வர்த்தகம் (Online Trading), தகவல் சேமிக்கும் பெட்டி (Drop Box), நேரடி மொழிபெயர்க்கும் வசதி (Online Translation), இதர இணைய சேவைகள் (Other Internet Service) என அனைத்தையும் ஒருங்கே அமையப்பெற்ற அதிநவீன செல்பேசிகள் சந்தையில் வலம்வருகின்றன.

‘தேவைகள்தான் கண்டுபிடிப்புகளுக்கான தாய்’ என்பார்கள். இந்த நிலையில் மனிதனின் பல்வேறு தேவைகளை உணர்ந்து பல்வேறு வகையான குறுஞ்செயலிகள் சந்தையில் வெளிவந்த வண்ணம் உள்ளன. சந்தையில் பல முன்னணி நிறுவனங்கள் கைபேசி மற்றும் கையடக்கக் கணினிகளை அறிமுகப்படுத்தி வருகின்றன. இவ்வகை சாதனங்கள் இயங்குவதற்குப் பல இயங்குதளங்கள் (OS) பயன்பாட்டில் உள்ளன. அவற்றில் மிகவும் குறிப்பிடத்தக்க இயங்குதளங்களான, ஆப்பிள் IOS, ஆண்ட்ராய்டு, பிளாக்பெரி, விண்டோஸ், படா ஆகியவை அதிகமாகப் புழக்கத்தில் உள்ளன. இவ்வகை இயங்குதளங்களில் தமிழ்மொழியில் அமைந்துள்ள செயலிகளை உருவாக்கும் முறைகள், அதில் ஏற்படும் சிக்கல்கள், தமிழ்வழிச் செயலிகளின் உள்ளடக்கம், அதன் பயன்பாடுகள் குறித்த ஒட்டுமொத்த மதிப்பீடாக இக்கட்டுரை அமைகிறது. முதல்வன் என்ற ஆப்பிள் IOS குறுஞ்செயலி உருவாக்கும்போது ஏற்பட்ட அனுபவத்தின் அடிப்படையில் இக்கட்டுரை வடிவமைக்கப்பட்டுள்ளது.

செயலி – விளக்கம்

ஒரு செயலைச் செய்யக்கூடிய செய்கையை விவரிக்கும் நிரல் துண்டைச் செயலி என்ற சொல் குறிக்கிறது. அதாவது ஒரு செயலை அல்லது வினையை செய்யவல்ல ஆணைத்தொடர்களைக் குறிக்க, செயலி என்ற சொல் தற்போது அதிக அளவில் பயன்படுத்தப்பட்டுவருகிறது. நவீனத் தொலைபேசிகளில் ஒரு குறிப்பிட்ட செயலைச் (Action) செய்துமுடிக்க மனிதர்களால் பல கட்டளைகளைப் பிறப்பித்து உருவாக்கப்படும் நிரல் தொகுப்பை குறுஞ்செயலி (Apps) என்று அழைக்கப்படுகிறது. கணினியில் பல செயல்களைச் செய்யும் செயலிகளை, மென்பொருள் (Applications- Software’s) என்றும், குறுகிய நிலையில் செய்யக்கூடியதைக் குறுஞ்செயலி என்றும் வழங்கப்படுகிறது.

தமிழ்ச் செயலிகள் உருவாக்க முறைகளும், சிக்கல்களும்

ஆப்பிள் நிறுவனம் அமெரிக்காவில் ஐ ஃபோன்களை உற்பத்தி செய்தாலும் உலகில் உள்ள அனைவரும் அதனை விரும்பி வாங்கக் காரணம் அந்தந்த மொழிகளில் பயன்படுத்தக்கூடியதாக அதன் செயலிகள் (Apps) அமைந்திருப்பதே. இந்தியாவில் ஆப்பிள் நிறுவனம் தனது பொருட்களை விற்பதற்குக்கூட நேரடி முகவர்கள் (Dealer)

இல்லாத சூழலில் இந்தியாவில், குறிப்பாகத் தமிழகத்தில் அதிக அளவில் ஆப்பிள் ஐ ஃபோன்கள் பயன்படுத்தப்பட்டு வருவதற்கு தமிழ் வழிச் செயலிகளும் மிக முக்கிய காரணமாக அமைகின்றன. ஆப்பிள் IOS இயங்குதளத்தில் செயல்படும் தமிழ்ச்செயலிகள் X-code என்ற அமைப்பில் Objective c என்ற மொழியில் எழுதப்பட்டு mac மூலம் வடிவமைக்கப்படுகிறது. இவ்வாறு வடிவமைக்கப்பெறும்போது அதில் இயல்பாகவே **இணைமதி** என்ற தமிழ் ஒழுங்குகுறி எழுத்துரு அமைந்துள்ளதால் தமிழ் எழுத்துகள் மிகத்தெளிவாகவும், அழகாகவும் எந்தவிதக் குழப்பமும் (Junk) இல்லாமல் தெரிகிறது என்பது இதன் தனிச்சிறப்பு ஆகும். ஐ ஃ போனில் செயலிகள் கூடம் (Apps Store) பிரிவிற்குச் சென்றால் அங்கு ஏறக்குறைய 600-க்கும் மேற்பட்ட செயலிகள் இடம்பெற்றுள்ளன. இது ஆண்ட்ராய்டு செல்பேசிகளைக் காட்டிலும் எண்ணிக்கையில் குறைவானதாக இருந்தாலும், அனைத்துத் தமிழ்ச்செயலிகளும் தரமானதாகவும், முழுப் பயன்பாட்டிற்கு உட்படுத்தக்கூடியதாகவும் அமைந்துள்ளது குறிப்பிடத்தக்கது.

ஆண்ட்ராய்டு இயங்குதளத்தில் செயல்படும் தமிழ்ச்செயலிகள் Andstk அமைப்பில் Java ping என்ற மொழியில் eclipse என்ற கருவியைப் பயன்படுத்தி வடிவமைக்கப்படுகிறது. இந்த முறையில் இயங்கும் தமிழ்ச்செயலிகளில் தமிழ் எழுத்துருக்கள் தெளிவாகத் தெரிவதில்லை. இதற்குக் காரணம் ஆரம்பகாலப் பதிப்புகளில் ஒழுங்குகுறி எழுத்துருக்கள் இல்லாத நிலையில் சில எழுத்துரு மாற்றிகளைப் பயன்படுத்தித் தமிழ் எழுத்துகள் கொண்டுவரப்பட்டன. தற்போதைய பதிப்புகளில் (And4.0 Vesion) தமிழ் எழுத்துக்கள் இடம்பெறுகின்றன. எனினும் இவ்வகை சாதனங்களில் தமிழ் எழுத்துருக்கள் அந்த அளவிற்கு (வடிவம், அழகு) சிறப்பானதாக இல்லை. தமிழ்க் குறுஞ்செயலிகள் ஆண்ட்ராய்டு சாதனங்களில் ஆயிரக்கணக்கில் இடம்பெற்றிருந்தாலும் அவை போதுமான அளவு தரமானதாக இல்லை.

விண்டோஸ் இயங்குதளங்களில் Winstk அமைப்பில் .NET மொழியில் Visuval Studio, Xpress Edition 2012 என்ற கருவிகளைப் பயன்படுத்தி வடிவமைக்கப்படுகிறது. விண்டோஸ், பிளாக்பெரி, படா போன்ற இயங்குதளங்களில் தமிழ்க் குறுஞ்செயலிகள் அதிக அளவில் வடிவமைக்கப்படுவதில்லை. ஏனெனில், இவ்வகை இயங்குதளங்கள் மக்கள் மத்தியில் பெரும் வரவேற்பைப் பெறவில்லை என்பது மிகமுக்கிய காரணமாகும்.

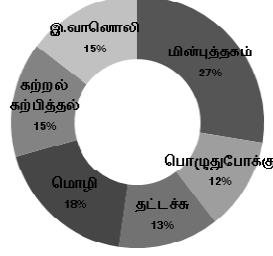
ஐ ஃபோன் மற்றும் ஆண்ட்ராய்டு உள்ளிட்ட ஸ்மார்ட் போன்களின் பயன்பாடு அதிகரித்துவரும் நிலையில், இருந்த இடத்தில் இருந்தே உள்ளங்கையில் எல்லாவற்றையும் அணுக / பெற உதவும் செயலிகளும் பிரபலமாகி வருகின்றன. இருப்பினும் இவ்வகைச் செயலிகளில் அதிக அளவில் (Size) உள்ளடக்கங்களைத் தர இயலாது. ஏனெனில், இந்தச் செயலிகளைத் தரவிறக்கம் செய்யும்போது அதிக கால அவகாசம் தேவைப்படுகிறது. இந்தியாவில் இணைய வேகம் குறைவாக உள்ளதே இதற்குக் காரணமாகும். இதனால் பயனாளிகளுக்குச் சலிப்பு ஏற்படுகிறது. மேலும், பல வகையான கோப்பு வடிவங்களை இணைப்பதில் சில தொழில்நுட்பக் கோளாறுகள் உள்ளன. (எ.கா. pdf, flash files) மேலும், தமிழ்க்குறுஞ்செயலிகள் உருவாக்கத்தின்போது போதிய கால அவகாசம், பொருளாதார வசதி தேவைப்படுகிறது. தமிழ் மொழியில் குறுஞ்செயலிகளை நிறுவும்போது அந்தந்த நிறுவனங்களில் பதிவுசெய்யும் நடைமுறைகள், பதிவுத்தொகை மற்றும் புதுப்பித்தல் கட்டணம் (Windows, IOS - 99 \$/year, Android-40\$/year), சரிபார்த்தல், ஒப்புதல் பெற காத்திருக்கும் காலம் (15 நாட்கள்) முதலிய நடைமுறைச் சிக்கல்கள் ஏற்படுகின்றன. தமிழில் ஒரு குறுஞ்செயலியை உருவாக்க மூன்று மாத கால எல்லையும், 30, 000 ரூபாய் முதல் 2 இலட்சம் ரூபாய் வரை பொருட்செலவும் தேவைப்படுகிறது.

தமிழ்ச் செயலிகள் - உள்ளடக்கம் (Tamil Applications - Content)

அமைப்பியல் முறையில் தமிழ்ச் செயலிகள் நிலையானவை (Static Content), புதுப்பித்தல் நிலையில் மாறக் கூடியவை (Rss Feeds) என இருவகைப்படும். மின்னூல்கள் போன்றவை நிலையானதாகவும், செய்திகள் போன்றவை அடிக்கடி மாறக்கூடிய உள்ளடக்கமாகவும் உள்ளன. தமிழ்மொழியில் இடம்பெற்றுள்ள செயலிகள் பல்வேறு பிரிவுகளில் அமைந்துள்ளன. அவையாவன : மின் இதழ்கள், இணைய வானொலிகள், தமிழ்மொழி கற்றல் கற்பித்தல் கருவிகள், தமிழ்த் தட்டச்சுக் கருவிகள், தமிழ் நேரடித் தொலைக்காட்சிகள், தமிழ் நேரடி மொழிபெயர்ப்புக் கருவிகள், நாளிதழ்கள் - செய்திகள், சமுதாய வலைத்தொடர்புகள், தேடுபொறிகள் போன்றன. இவற்றின்மூலம் தரவுகளைப் பெறுதல், தமிழ் போட்காஸ்ட் (Podcast), அகராதிகள், தமிழ் மரபு சார்ந்த செயலிகள், தனி நபர், பிரபலங்கள் சாந்தவை, பொழுதுபோக்கு முதலிய பிரிவுகளில் அடங்கும்.

தமிழ் இணைய இதழ்கள், குறுஞ்செயலிகளாக, பொன்னியின் செல்வன், ஜெயகாந்தன் கதைகள், திருக்குறள், புத்தகப் பூங்கா, தமிழ் குர் ஆன், தமிழ் பைபிள், பகவத் கீதை, இந்து மந்திரங்கள், தேவாரம், சுவாமி விவேகானந்தர், புகாரீ ஹதீஸ் முதலியவை அமைந்துள்ளன. இவை இலக்கியம், ஆன்மீகம், கலை, பண்பாடு போன்ற பிரிவுகளில் இடம்பெற்றுள்ளன. I-Kural என்ற செயலி திருக்குறள் அமைப்புடன், மணக்குடவர், பரிமேலழகர், மு.வ, கருணாநிதி, சாலமன் பாப்பையா, ஆகியோரது உரைகளுடன் ஜி.யு போப் அவர்களின் மொழிபெயர்ப்புடன் ஒருசேரக் கிடைப்பதால் திருக்குறளை ஒப்புநோக்கி ஆய்வு செய்யப் பேருதவியாக அமைந்துள்ளது.

TAMIL APPS CONTENT



மின் இதழ்கள்

இன்று அனைத்து நாளிதழ்களும், வார, மாத இதழ்களும் மின்னிதழ் வடிவிலும் வெளிவருகின்றன. விகடன் குழும இதழ்கள், பிரிமியம் தமிழ், இந்தியா டுடே, நக்கீரன், பெமினா தமிழ், தி சண்டே இந்தியன் தமிழ், சக்தி, பெரியார் - பிஞ்சு, சத்தியம், மயன். செய்திகள் வரிசையில் தினமலர், தினகரன், மாலைமலர், ஒன் இந்தியா தமிழ், காலக்கதிர், தமிழ் தினசரி செய்திகள், தமிழ் செய்திகள், நீயூஸ் ஹண்ட், புதிய தலைமுறை முதலியவை குறிப்பிடத்தக்கன. நாளிதழ்கள் பெரும்பாலும் விலையில்லா வசதியுடன் கிடைக்கின்றன. ஆனால், பெரும்பாலான வார, மாத இதழ்கள் கட்டணம் செலுத்திப் படிக்கும் நிலையில் அமைந்துள்ளன.

தமிழ் தட்டச்சுச் செயலிகள்

செல்பேசியில் தமிழில் தட்டச்சுசெய்து அவற்றைக் குறுஞ்செய்தி, மின்னஞ்சல், கட்டுரை எனப் பல்வேறு வடிவங்களில் பயன்படுத்துவதற்குப் பின்வரும் குறுஞ்செயலிகள் வழிவகை செய்கின்றன. முரசுவின் செல்லினம், தமிழ் எடிட்டர், தமிழ் டைப் SMS, தமிழ் மின்னஞ்சல் திருத்தி, தமிழ் விசைப்பலகை-IOS, தமிழ் பானினி விசைப்பலகை, யுனிகோட் தமிழ் SMS, தமிழ் குறுஞ்செய்தி, தமிழ் விசைப்பலகை என்பன குறிப்பிடத்தக்கவாகும். இவை அனைத்தும் ஒருங்குறி குறியேற்றத்தில் அமைந்த எழுத்துருக்கள் என்பது குறிப்பிடத்தக்கது. எனினும், இந்த எழுத்துருக்களைத் தட்டச்சு செய்யும் முறையில் தரப்படுத்தப்பட்ட விசைப்பலகைகளாக இல்லை. தமிழ் ஒலிபெயர்ப்பு (Phonetic) முறையில் தட்டச்சு செய்யும் வகையிலேயே பெரும்பாலும் உள்ளன. இவ்வாறு தமிழில் தட்டச்சு செய்த சொல்லைக் கொண்டு விக்கிப்பீடியா தமிழ் செயலி, தமிழ் கூகுள் தேடுதல் பகுதியில் இட்டு நமக்குத் தேவையான தரவுகளைத் தேடுவதற்கு மேலே குறிப்பிட்டுள்ள விசைப்பலகைகள் அனைத்தும் பயன்படுகின்றன.

தமிழ்மொழி கற்றல்- கற்பித்தல் செயலிகள்

இன்றைய காலகட்டத்திலும் பழைய கல்வித்திட்ட முறையே அமைந்திருப்பது மாணவர்களிடையே சலிப்பைத் தருகின்றது. புதிய சூழல், புதிய கற்றல் கருவிகள் எனப் புதுமையான வகைகளில் பல பள்ளிகளில் வகுப்பறைகள் (Smart Class Rooms) அமையத் தொடங்கிவிட்டன. அந்த வகையில் சூட்டிகைப் செல்பேசியில் குழந்தைகளுக்கு தமிழ்மொழியை கற்றுத்தருவதில் பல்வேறு செயலிகள் உதவி செய்கின்றன. அவை, I write Tamil, Trace Tamil Alphapets-1, 2, Tamil Rhymes, Fun 2 Learn Tamil, Tamil Lessons-1, 2, Avyayiaar Books, Tamil Numbers, Tamil Slate, I Tamil Kids, Flash cards Tamil Lessons, ஆத்திசூடி, மழைப்பாடம், செல்லமே செல்லம், தமிழ்ப் பழமொழிகள், சங்கம், தமிழ் எழுத்து, பாட்டி சொன்ன கதைகள் என இவை அனைத்தும் குழந்தைகள் ஆர்வமுடன் தமிழ் எழுத்துகளையும் எண்களையும் எளிதில் கற்றுக்கொள்ள உதவுகின்றன.

தமிழ்மொழி வளர்ச்சிக்கு உதவும் குறுஞ்செயலிகள்

இணையத்தில் தமிழ் அகராதிகள் இருப்பதுபோல தமிழ் அகராதிகளான, Tamil Dictionary, English- Tamil Dictionary, Lifco தமிழ் அகராதி ஆகியவை குறுஞ்செயலிகளாக இடம்பெற்றுள்ளன. இவ்வகை செயலிகள் தமிழ் மொழி சார்ந்த அறிவினையும், தெளிவினையும் தருகின்றன.

தமிழ் இணைய வானொலிகள்

வானொலியின் மிக முக்கியமான பணி அறிவித்தல் பணியாகும். இப்பணியைத் தமிழ்மொழியில் அமைந்த ஆயிரக்கணக்கான இணைய வானொலிகள் செய்துவருகின்றன. இவற்றைக் கேட்க, பல்வேறு செயலிகள் ஐ ஃபோன், ஆண்ட்ராய்டு செல்பேசிகளில் உள்ளன. இந்தியாவில் உள்ள பண்பலைகள் (FM) எனப்படும் வானொலிகள் மட்டுமல்லாது உலகில் எந்தப் பகுதியில் இருந்தும் ஒலிபரப்பாகும் தமிழ் இணைய வானொலிகளையும் கேட்கும்

வசதி இந்தச் செல்பேசிகளில் இடம்பெற்றுள்ளன. இவற்றில் தமிழ் உரைகள், உரையாடல், திரையிசைப் பாடல்கள், செவ்வி, பட்டிமன்றம், கலந்துரையாடல், பேட்டிகள், தொழில்நுட்பம், கலை, பண்பாடு, தமிழ் மரபுகள், மாண்புகள் முதலியவற்றையும் ஒலிபரப்பி, தமிழ்மொழி வளர்ச்சிக்கு உறுதுணையாக அமைந்துள்ளன. Hallo FM, Indian Radio, Tambura, Tamil Hits, Tamil Beat, Tamil Top, Thuthi FM, New Tamil Mp3, Tamil Radio, Kalasam, Sivan Koil, Tamil Christian முதலியவை ஆகும். SBS வானொலி ஆஸ்திரேலியாவிலிருந்து ஒலிபரப்பு செய்கின்ற போதிலும் தமிழகத்தில் நடக்கும் நிகழ்வுகளைக் கூர்ந்து ஆராய்ந்து செய்திகளை வழங்குகிறது.

பிற தமிழ்ச்செயலிகள்

தமிழ் நாட்காட்டி, மாலைமலர் நாட்காட்டி முதலிய செயலிகள் தமிழ் வழியிலேயே அமைந்துள்ளன. தமிழ் நாட்குறிப்புகள், Tamil Astrology, I Translate, I Publisher, Tamil Game, Ads Tamil, Tamil E- Greetings, Tamil Notepad, Tamil Pages, Tamil Time, Tamil Facebook, SBS Podcast, Tamil Online TV முதலிய பொழுதுபோக்கு நிறைந்த பல செயலிகள் தமிழ்மொழி வளர்ச்சிக்கு உறுதுணையாக அமைந்துள்ளன.

தமிழ்க் குடையின்கீழ்...

உலகத் தமிழர்களை ஒன்றிணைக்கும் வகையில் பல்வேறு செயலிகள் உள்ளன. யாழ்-யுகே, யாழ்-உலகம், தமிழ் சங்கம்-ஹாங்காங், சமூகம் ஆகிய செயலிகள் உலகத் தமிழர்களை ஒரு குடையின்கீழ் கொண்டுவரும் முயற்சியில் ஈடுபட்டுவருகின்றன. இவை தவிர இலங்கைத் தமிழர்களைப் பற்றி, அடெடெரனா தமிழ், தமிழ் Mirror ஆகிய இரு செயலிகள் செய்திகளை வெளியிட்டுவருகின்றன.

தமிழ்ச் செயலிகள் வளர், வளம் பெற...

தமிழ்வழிச் செயல்படும் குறுஞ்செயலிகள் கற்றல், கற்பித்தல் நிலைகளில் ஆசிரியர்களுக்கும், மாணவர்களுக்கும் தமிழ்த் தரவுகளை பெறும் நிலையில் ஆய்வாளர்களுக்கும், செய்திகள், நாட்டுநடப்புகள், இசை, இணைய வானொலிகள், திரைப்படம், ஒளிக்கோர்வைகள், பொழுதுபோக்கு முதலியவை பொதுமக்களுக்கும் பயன்படும் வகையில் அமைந்துள்ளன. மேலும் இத்துறைகள் சார்ந்த பல குறுஞ்செயலிகள் புதிதாக உருவாக்கவேண்டிய தேவை ஏற்பட்டுள்ளது. மலேசியாவில் நடைபெற இருக்கும் 12-ஆவது உலகத்தமிழ் இணைய மாநாட்டை ஒட்டி சிறந்த தமிழ்க் குறுஞ்செயலிகளுக்குக்கான போட்டிகள் அறிவித்ததைத் தொடர்ந்து, தமிழக அரசும் (தமிழ்வளர்ச்சித்துறை), இந்திய அரசும் (மனிதவளமேம்பாட்டுத்துறை), பிற தமிழ் அமைப்புகளும் (தமிழ்ப்பேராயம்) தமிழில் சிறந்த குறுஞ்செயலிகளை உருவாக்குபவர்களுக்கும், வடிவமைப்பவர்களுக்கும் பரிசுகள் வழங்கி, அவர்களை ஊக்கப்படுத்தும் வகையில் பல்வேறு சலுகைகளை அறிவிக்க வேண்டும். இவ்வாறு அறிவிப்பதன் மூலம் பயனுள்ள பல குறுஞ்செயலிகள் தமிழில் உருவாக்கம் பெற வாய்ப்புகள் உள்ளன.

கைக்கணினியும் தமிழும்

பேரா. முனைவர் வெ.இராமன்

e-mail: raman600@gmail.com, Prof.Dr.V.Raman@gmail.com

தமிழானது உலகெங்கும் பல்வேறு நாடுகளில் 80 மில்லியன் மக்களுக்கு மேல் பேசப்பட்டு வரும் மொழியாகும். இம்மொழியின் தாக்கமானது பல்வேறு மொழிகளிலும் உள்ளது. இது செம்மொழி எனும் சிறப்பைப் பெற்றுள்ளது. இம்மொழியின் பயன்பாடு கணினி யுகத்தில் பல்வேறு நிலைகளிலும் சிறப்புடன் விளங்குகிறது. அப்படியிருக்க இன்று மக்களின் கைகளில் தஞ்சம் பெற்று சிறப்புடன் விளங்கும் கைக்கணினிகளில் தமிழ்மொழியின் பயன்பாடு எவ்வாறு உள்ளது என்பதைப் பற்றி ஆராய்வதுடன் மேலும் தமிழ் மொழியின் பயன்பாட்டை எவ்வகையில் கைக்கணிகளில் மேம்படுத்தலாம் என்று ஆராய்வதும் இக்கட்டுரையின் நோக்கமாகும்.

கணினியின் வரலாறு

சுமார் இரண்டாயிரம் வருடங்களுக்கு முன் சீனர்கள், ஜப்பானியர்கள், பெரிய எண்களை கூட்ட, பெருக்க "அபாக்கஸ்" என்ற மணிச்சட்டத்தை பயன்படுத்த தொடங்கினர். இதுதான் கணித செயல்முறைகளை இயக்க உதவும் முதல் சாதனமாக கருதப்படுகிறது. கி.பி.1619-ல் ஸ்காட்லாந்து நாட்டை சேர்ந்த "ஜான் நேப்பியர்" என்ற அறிஞர் சில எலும்பு துண்டுகளை குடைந்து ஒரு கருவியை உருவாக்கினார். அதை "நேப்பியர் போன்ஸ்" என்று அழைத்தனர். இதன் மூலம் பெருக்கல், வகுத்தல் போன்ற செயல்முறைகளை எளிதாக செய்ய முடிந்தது. கி.பி.1622-ல் வில்லியம் ஆல்ட்ரெட் என்பவர் "ஸ்லைட் ரூல்" என்ற கருவியை கண்டுபிடித்தார். கால்குலேட்டர் கண்டுபிடிப்பதற்கு முன் இந்த கருவிதான் கணிதப்பயன்பாட்டிற்கு பலர் பயன்படுத்தினர். பெருக்கல் குறியீட்டுக்கு முதன் முதலில் "X" -ஐ பயன்படுத்தியது இவர்தான்.

கணினியின் தந்தை என கருதப்படும் சார்ல்ஸ் பாபேஜ் (Charles Babbage, 1791 - 1871) பிரித்தானிய கணிதவியலாளர், கண்டுபிடிப்பாளர். இன்றைய கணினிகள் பயன்படுத்தும் எந்திர கணக்கியல் இயந்திரங்களைக் கண்டுபிடித்தவர். 1837யில் முதல் முழுமையான செய்நிரல் கணினியை (Programmable Mechanical Computer) சார்ல்ஸ் பாபேஜ் வடிவமைத்தார். ஆனால் அக்கால தொழில்நுட்ப எல்லை, நிதி பற்றாக்குறை, மற்றும் தன்னுடைய வடிவமைப்புடன் தனகுதலை நிறுத்தமுடியாமை (ஆயிரக்கணக்கான கணினி சம்பந்தப்பட்ட பொறியியல் செயற்திட்டங்களின் முடிப்புக்கு காரணமாக பண்பு) போன்ற காரணங்களின் கலப்பால் இந்த சாதனத்தை அவரால் முழுமையாக உருவாக்க முடியவில்லை. இவரது இயந்திரத்தில் துளையிடப்பட்ட அட்டைகள் (Punch Card) பயன்படுத்தப்பட்டன. லேடி அடா லோவேலசே (Lady Ada Lovelace) என்ற பெண்தான் உலகின் முதலாவது கணினி செயல்நிரலாளர் (Computer Programmer) ஆவார். இவர் சார்ல்ஸ் பாபேஜின் நண்பர் என்பது குறிப்பிடத்தக்க விடயமாகும்.

கணினியின் தலைமுறைகள்

ஆரம்பத்தில் உருவாகியிருந்த கணினிகள் மிகக் குறைந்த பயனைத் தருவதாகவே இருந்தது. மேலும் இதன் அளவு ஒரு விளையாட்டு மைதானத்தின் அளவை ஒத்திருந்தது. இதற்கு பிறகு அடுத்தடுத்த கட்ட வளர்ச்சிகளின் போது கணினிகளின் செயல்திறன் மற்றும் அதன் வேகமும் அதிகரித்தது. மேலும் அதன் உருவ அளவும் படிப்படியாக குறைவடைந்து சென்றது. இம்முறையில் நிகழ்ந்த மாற்றங்களை கணினியின் தலைமுறைகள் (Computer Generation) என அழைத்தனர். மேலும் கணினியின் தலைமுறைகள் ஐந்தாகப் பிரிக்கப்பட்டன.

முதலாவது தலைமுறை

கி.பி 1946 - 1959 வரையிலான காலப்பகுதிகளில் அறிமுகப்படுத்தப்பட்ட கணினிகள் முதல் தலைமுறையைச் (First Generation) சேர்ந்ததாகும். இக்காலப்பகுதிகளில் உருவாக்கப்பட்ட கணினிகளில் வெற்றிடக் குழாய்கள் (Vacuum Tube) எனும் தொழில் நுட்பம் பயன்படுத்தப்பட்டன. இது அமைப்பில் மிகப் பெரியதாகவும், அதிக திறனுடைய மின்சாரமும் இதற்கு தேவைப்பட்டது. மேலும் இது அதிக இரைச்சலையும் உருவாக்கியது. இவ்வகையான கணினிகளை இயக்குவதற்கு மிகப் பெரிய குளிரூட்டிகள் (Air Conditioners) தேவைப்பட்டது.

எனியாக் (ENIAC - Electronic Numerical Indicator Analyzer Computer) என்பதே உலகின் முதலாவது முழுமையான மின்னுவியல் கணினியாக (Electronic Computer) கருதப்பட்டது. இது 30 டன் எடையும் 100 அடி நீளமும் 8 அடி உயரமுமாகக் காணப்பட்டது. இதில் 17468 வெற்றிடக் குழாய்கள் (Vacuum Tube) பயன்படுத்தப்பட்டன. இவற்றை புரிந்து கொள்ள மிகக் கடினமான இயந்திர மொழியிலேயே (Machine Level Language) இயக்க முடியும்.

இரண்டாவது தலைமுறை

கி. பி. 1959 -1964 வரையிலான காலப்பகுதியில் அறிமுகப்படுத்திய கணினிகள் இரண்டாம் தலைமுறையைச் (Second Generation) சேர்ந்ததாகும். இக்காலப்பகுதியில் அமைக்கப்பட்ட கணினிகளுக்கு திரிதடையம் (Transistor) பயன்படுத்தப்பட்டது. திரிதடையம் (Transistor) ஆனது வெற்றிடக் குலைகளைகளிலிருந்து (Vacuum Tube) முற்றிலும் மாறுபட்ட தொழில்நுட்பத்தைக் கொண்டிருந்தது. இது அளவில் சிறியதாகவும், வேகமாகவும் (Faster), நம்பகமானதாகவும் (Reliable), குறைந்த விலைகளிலும் தயாரிக்கக் கூடியதாகவும் இருந்தது. இதனை ஜோன் பார்டீன் (Jhon Bardeen), வில்லியம் சொக்கலே (William Shockley), வால்ட்டர் பிரட்டின் (Walter Brattain) என்பவர்கள் கண்டுபிடித்திருந்தனர். இந்தக் கணினிகளில் இயந்திர மொழி (Machine Level Language) மட்டுமன்றி assembly language -ம் பயன்படுத்தப்பட்டன.

மூன்றாவது தலைமுறை

கி. பி. 1965 - 1971 வரையிலான காலப்பகுதிகளில் அறிமுகப்படுத்திய கணினிகள் மூன்றாம் தலைமுறையைச் (Third Generation) சேர்ந்ததாகும். இக்காலத்தில் அறிமுகப்படுத்திய கணினிகளில் ஒருங்கிணைச் சுற்று (Integrated Circuit) எனும் தொழில் நுட்பம் பயன்படுத்தப்பட்டது. இக்கணினிகளில் மிக அதிகமான திரிதடையத்திற்கு (Transistor) பதிலாக ஒரு ஒருங்கிணைச் சுற்று (Integrated Circuit) பயன்படுத்தப்பட்டது. இதனால் கணினியின் அளவு சிறியதாகப்பட்டதுடன் மின்சார பாவனையளவும் குறைக்கப்பட்டது. ஒருங்கிணைச் சுற்று (Integrated Circuit) ஆனது செயல் திறன் அதிகமானதாகும். இன்றைய இலத்திரனியல் சாதனங்களில் ஒருங்கிணைச் சுற்று (Integrated Circuit) பயன்படுத்தப்படுகின்றமை குறிப்பிடத்தக்க ஒரு விஷயமாகும். இதனை ராபர்ட் நய்ஸ் (Robert Noyce) உம் ஜாக் கில்பி (Jack Kilby) உம் இணைந்து கண்டுபிடித்தனர். மேலும் இந்த தலை முறைகளில்தான் High Level Language என்ற நவீன பயன்பாட்டு மொழி அறிமுகம் செய்யப்பட்டது.

நான்காம் தலைமுறை

கி. பி. 1971 முதல் தற்போது வரை உள்ள கணினிகள் அனைத்தும் நான்காம் தலைமுறையைச் (Fourth Generation) சேர்ந்ததாகும். இதில் நுண்செயலி (Microprocessor) எனப்படும் தொழில்நுட்பம் பயன்படுத்தப்படுகின்றது. இன்டெல் (Intel) நிறுவனத்தில் பணி புரிந்த டேட் ஹோப் (Ted Hoff) என்பவர்தான் இதனைக் கண்டுபிடித்தார். இது ஒருங்கிணைச் சுற்றை (Integrated Circuit) விட மிக அதீத திறன் கொண்டதாக நுண்செயலி (Microprocessor) விளங்குகின்றது.

ஐந்தாம் தலைமுறை

ஐந்தாம் தலைமுறை (Fifth Generation) என்பது புதிய தொழில்நுட்ப முறையிலான கணினிகளைக் குறிக்கின்றது. இது இன்றைய அதி நவீன கணினிகளைக் காட்டிலும் இவை சிறந்தவையாக காணப்படும். எப்படி என்றால் இவற்றால் சுயமாக சிந்திக்க முடியும். இவற்றிற்கு தட்டச்சுப் பலகை, சுட்டி (Keyboard, Mouse) போன்ற பாகங்கள் அவசியமில்லை. இவற்றின் திரை தொடு தன்மை (Touch Screen) உடையன. இவ்வகையான கணினிகளை Surface Computer என அழைக்கப்படும்.

கைக்கணினிகளின் (Tablets) வகைகள்

மைக்ரோசாப்ட் நிறுவனம் டேப்லெட் பிசி எனும் பதத்தை உபயோகிக்க ஆரம்பித்ததே கைக்கணினியின் முதல்படி எனலாம். தொடுத்திரையைக் கொண்ட, சற்றே மாற்றியமைக்கப்பட்ட இயங்குதளத்தைக் கொண்ட கையடக்க கணினியே கைக்கணினியின் ஆரம்பம்.

1987ம் ஆண்டு ஆப்பிள் நிறுவனம் கைக்கணினி உற்பத்திக்கான முயற்சியைத் துவங்கியது. பின்னர் 1993ம் ஆண்டு தனது கைக்கணினியை வெளியிட்டது. ஸ்ட்ராங் ஏஆர்எம் நுண்செயலியைக் கொண்ட தொடு திரைக் கணினியை வெப்பேட் (WebPAD) எனும் பெயரில் இன்டெல் நிறுவனம் 1999ம் ஆண்டு அறிமுகப்படுத்தியது. இது பின்னர் இன்டெல் வெப் டேப்லெட் (Intel Web Tablet) என்று அழைக்கப்பட்டது. 2002ம் ஆண்டு மைக்ரோசாப்ட் நிறுவனம் மைக்ரோசாப்ட் டேப்லெட் பிசி (Microsoft Tablet PC) -யை அறிமுகப்படுத்தியது. அதீத விலையாலும் உபயோகப் பிரச்சனைகளாலும் இது சந்தைப்படுத்துதலில் வெற்றியடையவில்லை. 2005ம் ஆண்டு நோக்கியா நிறுவனம் நோக்கியா770 எனும் லினக்ஸ் இயங்குதளத்தை அடிப்படையாகக் கொண்ட முதல் இணைய உபயோகம் கொண்ட கைக்கணினியை வெளியிட்டது.

ஆண்ட்ராய்ட் இயங்குதளத்தின் வரவு கைக்கணினிகளின் வரலாற்றையே மாற்றி சந்தையை ஆக்கிரமித்தது எனலாம். 2008-ல் ஆண்ட்ராய்ட் அடிப்படையிலான கைக்கணினி தோன்றலாயிற்று எனினும் 2009ம் வருடம் ஆண்ட்ராய்ட் 1.4

இயங்குதளம் கொண்ட கைக்கணினி வெளியாயிற்று. 2011ம் வருடம் வெளியான ஆண்ட்ராய்ட் ஹனிகோம்ப் மிகப் பிரபலமடைந்து எங்கெங்கும் ஆண்ட்ராய்ட் உபயோகத்தைப் பரப்பிற்று எனலாம்.

2010ம் வருடம் சாம்சங் கேலக்ஸி வெளியாயிற்று. இதன் உடனடி துவக்கம் (instant on), பலமுனைத் தொடுதல் (multi touch), மற்றும் பல வசதிகள் பயனாளர்களைக் கவர்ந்தது எனலாம். இதே வருடம் ஆப்பிள் ஐஓஎஸ் இயங்குதளம் கொண்ட ஆப்பிள் நிறுவனத்தின் கைக்கணினி வெளியாகி மக்களைப் பெரிதும் கவர்ந்தது. இது புதிய வகை மக்கள் உபயோகக்கருவி எனும் பாராட்டைப் பெற்றது.

2013ம் வருடம் ஏசஸ் மற்றும் சாம்சங் நிறுவனங்கள் ஆண்ட்ராய்ட் மற்றும் விண்டோஸ் 8 இயங்குதளங்களைக் கொண்ட இரட்டை இயங்குதளத்துவக்க கைக்கணினிகளை வெளியிட்டது. பயன்பாட்டை நிறுத்தாமலே இயங்குதளங்களை மாற்றிக் கொள்ளும் வசதி இக்கைக்கணிகளில் உள்ளது குறிப்பிடத்தக்கது.

கைக்கணிகளின் இயங்குதளங்கள்

கைக்கணினிகள், மற்ற கணினிகள் (மேசைக் கணினிகள், மடிக்கணினிகள் போன்றவை) போல் அல்லாது உடனடித் துவக்கம், தொடுதிரை உட்புகுத்தல் என பல பயனர் வசதி கொண்டவையாக இருப்பதால் இதன் இயங்குதளங்களும் மாறுபட்டவையாகவே உள்ளன.

ஆண்ட்ராய்ட்

கூகிளின் ஆண்ட்ராய்ட் இயங்குதளமே உலகில் அதிகமாக உபயோகிக்கப்படும் கைக்கணினிகளுக்கான இயங்குதளம் எனலாம். இது திறவூற்று மென்பொருளின் அடிப்படையில் லினக்ஸ் அடிப்படையிலான தொடுதிரைக்கான இயங்குதளமாகும்.

பிளாக்பெர்ரி

இது ஃயுளனெக்ஸ் (QNX) அடிப்படையிலான பிளாக்பெர்ரிக்கான பிரத்யேகமான இயங்குதளமாகும்.

ஐஓஎஸ்

இது ஆப்பிள் நிறுவனத்தால் தயாரிக்கப்படும் ஐபோன் மற்றும் ஐபோட் டச் மற்றும் ஐபேட் ஆகியவற்றிற்கான யுனிக்ஸ் செயல்பாட்டை உட்கொண்ட மேக் இயங்குதள அடிப்படையிலான இயங்குதளமாகும். தொடுதிரைக் கைக்கணினிகளுக்காக இது சிறப்பாக வடிவமைக்கப்பட்டுள்ளது. மிகக்குறைந்த மின்சக்தியை செலவிடும் வகையில் இது அமைக்கப்பட்டுள்ளது.

மோட்புக்

ஆப்பிள் நிறுவனத்தால் ஐஓஎஸ் வெளியீட்டிற்கு முன் வாகாம் நிறுவனத்தின் மென்பொருளோடு இணைந்து ஆப்பிள் மேக்புக் கணினிக்காக தயாரிக்கப்பட்ட இயங்குதளமே மோட் புக் இயங்குதளமாகும். எனினும் ஏறத்தாழ இது வழக்கொழிந்து விட்டது எனலாம்.

லினக்ஸ்

லினக்ஸ் அடிப்படையிலான இயங்குதளங்கள் பலவகைக் கைக்கணினிகளில் செயல்பட்டுக் கொண்டிருக்கின்றன. மாயிமோ என்ற டெபியான் லினக்ஸ் அடிப்படையிலான நோக்கியாவிற்கான இயங்குதளம், உபுண்டு அடிப்படையிலான இயங்குதளம், ஓப்பன் சுசி லினக்ஸ் அடிப்படையிலான இயங்குதளம் ஆகியவற்றை இங்கு குறிப்பிடலாம்.

இண்டெல்

இண்டெல் நிறுவனமானது மொப்லின் (மொபைல் லினக்ஸ்) எனும் இயங்குதளத்தை தனது ஆட்டம் சிபியு நெட்பிக்கிற்கான இயங்குதளமாக வடிவமைத்துள்ளது.

மீகோ

மீகோ என்பது லினக்ஸ் அடிப்படையிலான இண்டெல் மற்றும் நோக்கியோ கைக்கணிகளுக்கான இயங்குதளமாகும். இதுவும் ஒரு திறவூற்று அடிப்படையிலான இயங்குதள மென்பொருளாகும்.

விண்டோஸ்

தற்போது விண்டோஸ் 8 மற்றும் விண்டோஸ் ஆர்டி எனும் இரு இயங்குதளங்களை வடிவமைத்து பயன்பாட்டிற்கு வெளியிட்டுள்ளது. இதில் விண்டோஸ் 8 இயங்குதளம் கைக்கணிகள் பயன்பாட்டிற்கும் ஏற்றவாறும் விண்டோஸ் ஆர்டி கைக்கணிகள் பயன்பாட்டிற்காக மட்டும் உள்ளபடி வடிவமைக்கப்பட்டுள்ளன. பல்வகையான கைக்கணிகள் தயாரிப்பாளர்களும் இவ்வகை இயங்குதளங்களை தங்களது கைக்கணிகளில் பயன்படுத்தி வருகின்றனர்.

வெப் ஓஎஸ்

ஹேவ்லெட் பாக்யார்ட் நிறுவனம் தனது கைக்கணினிக்கான இயங்குதளமாக வெப் ஓஎஸ் திறவூற்று மென்பொருள் அடிப்படையில் தயாரித்து வெளியிட்டது எனினும் தற்போது இவ்வகையான இயங்குதளத்தை எல்ஜி எலக்ட்ரானிக்ஸ் நிறுவனத்திற்கு விற்றுவிட்டதாக அறிவித்துள்ளது.

கைக்கணியில் தமிழின் பயன்பாடு

கைக்கணியில் தமிழின் பயன்பாடு என்பது தற்போது பல்கிப்பெருகி வருகிறது எனலாம். மேசைக்கணினி, மடிக்கணினி ஆகியவற்றில் உள்ள அளவு தமிழின் பயன்பாடு மடிக்கணியில் உள்ளதா என்று பார்க்கும்காலை பெரும் ஏமாற்றத்தைச் சந்திக்கவேண்டியுள்ளது. இதற்குக் காரணம் தமிழின் பயன்பாடு என்பது கைக்கணினியைப் பொறுத்தவரை மிகவும் சிரமம் தரும் செயலாகவே உள்ளது. இருப்பினும் கைக்கணியின் வளர்ச்சி என்பது தற்போது பெருகி வரும் நிலையில் தமிழின் பயன்பாடு உடன் வளர்ந்து வருகிறது என்பது சற்று ஆறுதல் தரும் விஷயமாகும்.

தமிழின் பயன்பாடானது கைக்கணினிகளில் பல்வகைகளில் வெளிப்படுகிறது. செய்திகள் படிக்க, புத்தகங்கள் படிக்க, குறுஞ்செய்திகள் அனுப்ப, மின்னஞ்சல் அனுப்ப, சமூக வலைத்தளங்களில் பயன்படுத்த என்று பல்வகைகளிலும் தமிழ்ப்பயன்பாட்டுத் தாக்கம் விரிவடைந்துள்ளது. எனினும் அலுவலகப்பயன்பாட்டிற்கான தமிழ் உபயோகம் மிகவும் சிரமம் தரும் செயலாகவே உள்ளது.

புத்தகங்கள் படிப்பதற்கு என்ஹெச்எம் ரீடர், விகடன் போன்ற மென்பொருட்களை அவர்களது தளத்திலிருந்து இறக்கி பயன்படுத்தும்போது, எவ்வகையான சிரமம் இன்றி அவர்களது புத்தகங்களைப் படிக்க இயலுகிறது.

செய்திகள் படிக்க தமிழகத்தின் பல்வேறு செய்தி நிறுவனங்கள் அவர்களது செய்திகளைப் படிக்க வழிவகை செய்துள்ளது. செல்லினம் மென்பொருளை இறக்கிக் கொள்வதன்மூலம் மின்னஞ்சல், குறுஞ்செய்தி ஆகியவற்றை தமிழில் அனுப்ப இயலும். இதுதவிர பல கட்டணம் செலுத்திப் பெறப்படும் மென்பொருட்களின் வழியாகவும் இவ்வகையான தமிழ்ப் பயன்பாட்டினைப் பெற இயலும்.

கைக்கணியில் அலுவலகப்பயன்பாடும் தமிழும்

ஆண்ட்ராய்ட், ஆப்பிள் ஐபேட் போன்றவற்றில் தமிழை அலுவலக பயன்பாட்டிற்கான மென்பொருட்களில் பயன்படுத்தினால் கைக்கணினியினால் தமிழை சரிவர காண்பிக்க முடிவதில்லை. எழுத்துருக்கள் பிரச்சனை இங்கு வெகுவாக தலைதாக்குகிறது. மேசைக்கணினிகள், மடிக்கணினிகள் ஆகியவற்றில் ஏறத்தாழ மறைந்துவிட்ட எழுத்துருப்பிரச்சனை இங்கு இன்னும் தீர்க்க இயலாத பிரச்சனையாக உள்ளது.

பெரும்பாலும் அலுவலகப்பயன்பாட்டிற்காக பயன்படுத்தப்படும் அலுவலகத் தொகுப்புகளில் தமிழ் எழுத்துருக்கள் பிரச்சனை இன்னும் தீர்க்கப்படவில்லை. இதனால் மேசைக்கணினி, மடிக்கணினி ஆகியவற்றில் உருவாக்கப்பட்ட கோப்புகளை கைக்கணியில் படிக்க இயலாத சூழ்நிலையும், புதிய கோப்புகளை கைக்கணினியினால் உருவாக்க முடியாத சூழ்நிலையும் உள்ளது. இதனால் மீண்டும் மடிக்கணினியையே தமிழ் உபயோகத்திற்காக எடுத்துச் செல்ல வேண்டியுள்ளது.

எனவே எழுத்துரு, தட்டச்சு முறைகள் ஆகியவற்றை உடனடியாக கைக்கணினிக்கென்று வடிவமைத்து எந்த அலுவலகப் பயன்பாட்டிலும் சிக்கல் வராதவாறு பயன்படுத்தும் நிலையை உருவாக்க வேண்டிய கட்டாயம் எழுந்துள்ளது. இதனை உடனே சீரமைக்கும் பொறுப்பு தமிழ் கணினி வல்லுநர்களுடையதாகும்.

முடிவுரை

இவ்வாறு தமிழின் பயன்பாடு கைக்கணினிகளில் பல்கிப் பெருகி இருப்பினும் அலுவலகப்பயன்பாட்டில் சிக்கல் இருப்பது வருந்தத்தக்க செய்தியாகும். கைக்கணிகளின் வளர்ச்சி என்பது உபயோகிப்பாளர்களிடையே பெறும் வரவேற்பைப் பெற்றுள்ளது. எளிமையான உபயோகம், கொண்டு செல்லும் தன்மை, உடனடித் துவக்கம், நீண்ட மின்சக்தி, மிகக்குறைவான எடை, பல்வகை இணையத் தொடர்புத் தன்மை போன்றவையே கைக்கணினி உபயோகிப்பாளர்களிடையே வரவேற்புப் பெற்றுள்ள வசதிகளாகும். இதனால் ஏறக்குறைய மடிக்கணினிகளைத்

துறந்து கைக்கணினிகளுக்கு பலரும் மாறிக் கொண்டுள்ளனர். இந்நிலையில் தமிழ் அலுவலகப் பயன்பாட்டில் ஏற்படும் சிக்கல் தமிழ் உபயோகத்தை குறைக்கும் என்பதில் ஐயமில்லை. எனவே தமிழின் பயன்பாடு அதிகரிக்க தமிழ்ப் பயன்பாட்டில் உள்ள சிக்கல்களைக் களையவேண்டிய பொறுப்பு தமிழ்க் கணினி வல்லுநர்களுடையதாகும். இவ்வாறு சிக்கல்கள் களையப்பட்டால் தமிழின் பயன்பாடு கைக்கணினியின் வழியாகவும் பல்கிப் பெருகும் என்பதில் ஐயமில்லை.

மேற்கோள்கள்:

1. தமிழ் விக்கிபீடியா இணைய தளம் (www.ta.wikipedia.com)
2. விக்கிபீடியா இணைய தளம் (<http://en.wikipedia.org>)
3. ஆண்ட்ராய்ட் இணைய தளம் (www.android.com)
4. ஆப்பிள் இணையதளம் (www.apple.com)
5. இண்டெல் இணையதளம் (www.intel.com)
6. V.Raman, “Standardising Tamil language encoding method; Analysis of technical issues” APAN International conference (2001), Malaysia.

மென்பொருள்கள், தளங்கள் மற்றும் தொழில்நுட்பங்கள் வழி கையடக்க கணினி சாதனங்களில் தமிழ் மின் புத்தகங்கள் மற்றும் மின் இதழ்கள்

சி.குணசேகரன், புதிய ஊடகப் பயிற்றுவிப்பாளர், மென்பொருள் உருவாக்குநர், சிங்கப்பூர்

powerguna@gmail.com

பலகோடி ஆண்டுகளுக்கு முன் தோன்றிய நம் புவியுலகை சிலநூறு ஆண்டுகளில் தம் விரல்நுனியில் கொண்டு வர மனிதகுலத்திற்கு அறிவியலும், தொழில் நுட்பமும் பேருதவி புரிந்து வருகிறது.

காலம் காலமாகப் பல்வேறு மாற்றங்களை அடைந்து இன்றுதொழில்நுட்பத்தின் வழியாகத் தமிழை இனிமையாகவும் எளிமையாகவும் மின் ஊடகம் மூலம் கற்றுக்கொடுப்பதற்குரிய சூழல் உருவாகி உள்ளது.

கையடக்கத் கணினித் தொழில் நுட்பங்கள் வினாடிக்கு வினாடி பல்கிப்பெருகிக் கொண்டிருந்தாலும் 21 நூற்றாண்டில், மின்புத்தகம் மற்றும் மின் இதழ் தமிழ் வளர்ச்சிக்கு எவ்வகையில் உறுதுணையாக இருக்கும் என்பதை விளக்குகிறது இந்த ஆய்வுக்கட்டுரை.

கையடக்க கணினி தொழில் நுட்பங்கள்- அறிமுகம்

விலையாலும் வடிவத்தாலும் ஆரம்பத்தில் பலருக்கும் எட்டக்கனியாக இருந்த கையடக்கக் கணினி இன்று, அனைவரது வாழ்விலும் நீங்காத இடத்தைப் பிடித்துள்ளது. அண்மைய கணக்கெடுப்பின்படி இக்காலY தலைமுறை (Generation Y) அதிகம் பயன்படுத்துவது கையடக்க கணினி சார்ந்த சாதனங்களாகவே உள்ளன.*1, இதன் வழி மின் நூல் வாசிப்பவர்களின் எண்ணிக்கையும் பெரும் அளவில் கூடி வருகிறது.

மின்னூல் தரும் நன்மைகள்

காலம் காலமாகக் கல்வெட்டுக்களிலும் ஓலைச்சுவடிகளிலும் காகிதங்களிலும் நூல்களிலும் தமிழ்மொழியைக் கற்று வந்த காலம் மாறி இன்று அறிவியலின் அதீத வளர்ச்சியில் மின்னூல் உருவாக்கப் பட்டிருக்கிறது.

கணினி, மடிக் கணினி, தொட்டுணர் கருவி (ipad) (Android), (Window8) போன்றவற்றில் தமிழ் நூல்களைப் பதிவேற்றம் செய்து, இருந்த இடத்தில் இருந்துகொண்டு வாசிப்பதற்கு ஏற்புடையதாக அமைவதே மின்னூல் எனப்படும்.

இன்றைய சூழலில் கணினிநம் அன்றாட வாழ்வில் இன்றியமையாத ஒன்று. அதன் வழி பதிவேற்றம் காணும் மின்னூல்கள், நூலகம் சென்று புத்தகம் தேடும் நேரத்தைக் குறைத்து கற்கும் வாய்ப்புகளை அதிகரிக்கின்றன. தமிழில் உள்ள பல்வேறு வகையான நூல்களை நினைத்த நேரத்தில் கற்பதற்கும், கற்கும் போது ஏற்படும் ஐயப்பாடுகளை உடனுக்குடன் தீர்ப்பதற்கும், நூலில் உள்ள கருத்துக்களைப் பற்றிக் கலந்துரையாடுவதற்கும் உரிய வழி முறைகள் மின்னூலில் அதிகமாக உள்ளன. அனைவரின் விருப்பத்துக்கு ஏற்ப நூல்களை தேர்வு செய்து கற்பதற்குரிய வாய்ப்பு இருப்பதால் அவர்களுக்குத் தமிழ் மொழியைக் கற்க வேண்டும் என்ற ஆர்வம் அதிகரிக்கிறது. மேலும், தனக்குள் மறைந்து இருக்கும் எழுத்துப் படைப்புக்களை உலகோருக்கு எடுத்து இயம்பும் வாய்ப்பும் நமக்கு ஏற்பட்டுள்ளது.

மேலும் பொருளாதார ரீதியாக தரமானத் தமிழ்ப் படைப்புகள் படிக்கும் வாசகர்களை உடனுக்குடன் சென்றடைவதுடன், அதனைச் சார்ந்த நல்ல கருத்துகளும், இணைய கருத்துப் பரிமாற்றங்களும் தரமான தமிழ் மின்னூல்கள் படைக்க அதிகமான வாய்ப்புக்களை ஏற்படுத்தியுள்ளன.

மின்னூல்வழி கற்பதால் பல புத்தாக்கச்சிந்தனைகளும் மாற்றங்களும் உருவாகும் வாய்ப்புகள் அதிகரித்துள்ளன. பசுமை நிறைந்த உலகை உருவாக்குவதற்குக் கணினி சார்ந்த பயன்பாடு மூலம் காகிதங்களின் தேவை பெரிதும் குறைவதால், மரங்களின் ஆயுள்காலம் அதிகரிக்கப்பட்டு வருங்காலச் சமுதாயம் செழுமையுடன் வாழ வாய்ப்பு ஏற்பட்டுள்ளது. தெளிவான எழுத்துகள், உடனுக்குடன் மின்னூல்கில் உலா வருவதன் மூலம் அதிகமான செய்திகளைக் கற்று அறியும் வாய்ப்பு, நிறைய நூல்களைத் திரட்டிப் படிக்கும் எளிய முறை போன்றவை, மின்னூல் வசம் நம்மைச் சுண்டி இழுக்கும் சக்திகளாக மாறி உள்ளன.

காணொளி, இசை, வண்ணமயமான விளக்கப்படங்கள், இணையப்பக்கம்வழி இருவழி தகவல் பறிமாற்றங்கள் மூலம் தமிழில் எழுதவும், படிக்கவும் எண்ணற்ற வாய்ப்புகளை மின்புத்தகங்கள் வழங்குகின்றன.

மின் இதழ்களும் நீங்களும்

மின்இதழ்செய்திஒருநவீனமற்றும்வேகமாகவளர்ந்துவரும்வடிவம்மூலம்உங்கள்வாசகர்களை அடையவடிவமைக்கப்பட்டஒருபயனுள்ளவிற்பனைதளமாகும்.

பேஸ்புக், சமூக ஊடகங்கள் அல்லது மின்னஞ்சல் மூலம் எளிதாகப் பகிர்ந்து கொள்வதற்கு இது ஒரு சிறந்த மாற்று தளமாக உருவெடுத்துள்ளது.அச்ச மற்றும் விநியோகம் செலவுகளைக்சேமிப்பதுடன் கணினிகள், ஐபாட் மற்றும் ஐபோன்கள் உட்பட அநேக ஊடகங்களில் வாசிக்க முடியும்.

உலகளாவியவலைமூலம்தமிழில் பல இணைய இதழ்கள் வழங்கப்படுகின்றன. தமிழ்நாட்டில் உள்ள பல பெரிய அச்ச வெளியீட்டாளர்கள் இப்போது ஒரு கட்டணம் பல்வேறு ஆன்லைன் சேவைகள் மூலம் அவர்களின் அச்ச இதழ்களை விற்பனைக்கு வழங்கத் தொடங்கி உள்ளனர்.*2

கையடக்க கணினி வழி கற்பதன் பயன்

இன்றைய சூழலில் நம்மிடையே கைத்தொலைபேசி என்பது உயிரும் உடலும் போல இரண்டறக் கலந்த ஒன்றாக அமைந்திருக்கிறது. அதன் வழியாகத் தமிழ்மொழியை மிகுந்த ஆர்வத்துடன் பயன்படுத்த முயல்வதைக் காண முடிகிறது.அனைவரிடமும் கைத்தொலைபேசி பயன்பாட்டில் இருப்பதாலும் அடிப்படையில் தமிழ் எழுத்துகள் அதனுள் இணைக்கப் பட்டுள்ளதாலும் மொழிப் பரிமாற்றம் எளிமையாக நடைபெற வழி வகுக்கிறது.

உலக நாடுகளில் அறிமுகப் படுத்தப்பட்டுள்ள தடையில்லா இணையச்சேவை, கைத்தொலைபேசி பயன்படுத்தும் முறை, தமிழ் மொழி வளர்ச்சிக்கு புதுப்பாதை அமைத்துத் தந்துள்ளது.

எளிமையான கட்டமைப்பு மூலம் தமிழ்ப் பயனீட்டாளர்களை ஒருமுகப்படுத்துவதற்கு குறுந்தகவல் அனுப்புவதன் மூலம் தமிழில் கலந்துரையாட முடிகிறது. மின் புத்தகம் வழி பாடத்திட்டங்கள், வாசிக்கும் வழி முறைகள், கணினி, இணையம், கைத்தொலைபேசி இம்மூன்றையும் ஒருங்கிணைப்பதன் மூலம் தமிழ்மொழியை எளிமையாகவும் இனிமையாகவும் பயன்படுத்த முற்படுகின்றனர்.

தமிழ் மின்புத்தகம், மின் இதழ், கையடக்கக் கணினிகள் வழி பயன்கள்

கையடக்கக் கணினிகள் தமிழ் மொழிப் பயன்பாட்டை ஊக்குவிப்பதோடு, மொழித் திறன்களில் ஏற்படுகின்ற படிப்படியான வளர்ச்சி நிலையைக் கண்காணிப்பதற்கும் ஏதுவாக அமைகிறது. மின் புத்தகம் வழி தரமான செய்திகளைப் பதிவிறக்கம் செய்து முழுமையான அறிந்து கொள்ள இக்கருவிகள் பெருந்துணை புரிகிறது. எந்நேரமும் கைகளில் தவழ்ந்து கொண்டிருப்பதால் நினைத்த நேரத்தில் அதனை காலவிரயம் ஏற்படாமல் பயன்படுத்த முடிகிறது.

இருவழிக்கலந்துரையாடலுக்கு இக்கருவி பயன்படுவதோடு, தயக்கமின்றித் தங்களுடைய பதில்களை வழங்குவதற்கு ஏதுவாக அமைகிறது. விளையாட்டு உத்தி முறைகளின் வழி, மின் புத்தகங்கள் வாசிக்கப்படும் போது தமிழைஎளிமையாகச்சுபரிந்துகொள்ள முடிகிறது. பல்வேறு வகையான விளையாட்டு உத்திகளைக்கைத் தொலைபேசிகளில் பதிவிறக்கம் செய்து அதன்வழி புத்தாக்கமும் ஊக்கமும் ஏற்பட வழி வகை செய்கிறது.

கைத்தொலைபேசியில் ஜிபிஎஸ் (GPS), வீடியோ, புகைப்படக்கருவி, 4ஜி கட்டமைப்பு போன்ற வன்பொருள்கள் மின் புத்தகம், மின் இதழ்கள் உருவாக்கத்திற்கும் பகிர்வுக்கும் மிகவும் பயனுள்ளதாக அமைந்திருப்பது அனைவரும் அறிந்த ஒன்றாகும்.

தமிழ் மின்புத்தகம், மின் இதழ்வளர்ச்சியில் புத்தாக்க சிந்தனை

கணினி உலகில் ஒருவழித்தகவல் பரிமாற்றமே அதிகமாக நடைபெற்ற நிலை கடந்த பத்தாண்டுகளில் இருவழித்தகவல் பரிமாற்றமாக மாறும் முறைமேம்பட்ட மின்புத்தகங்களை, பல்லாடக தொழில்நுட்பங்கள் அடங்கிய மின் இதழ்களை உருவாக்கும் உத்திகளை அறிமுகப்படுத்தி உள்ளது.

இம்முறை தமிழ்மொழியின் உயர்வுக்கும் அதிகப் பயன்பாட்டிற்கும் வழி வகுத்துள்ளது. தாமாகவே அறிந்து உள்வாங்கிக் கற்கும் வகையில் கையடக்க கணினிகள் மறு உருவாக்கம் கண்டுள்ளன.

இம்முறையின் வழி ஒரே நேரத்தில் உலகின் பல்வேறு பகுதிகளில் வாழ்பவர்கள் தமிழ்மீன் நூல்களையும் வியாபார ரீதியில் வெற்றி பெரும் மீன் இதழ்களையும் தத்தம் கையடக்க கணினிகளில் பதிவிறக்கம் செய்து படித்து இன்புறும் அளப்பரிய வாய்ப்புக்களை ஏற்படுத்தி உள்ளது.

விரும்பிய நேரமெல்லாம்படித்துப் பயனடைவதற்கு ஏதுவாகத் தொழில்நுட்பம் அமைகிறது.செய்யும் தவறுகள் மற்றவர்கள் முன்னிலையில் சரி செய்யப்படும் அவல நிலை இம்முறையில் இல்லை. பகிர்ந்து கொள்ளப்படும் தகவல்களில் தவறுகள் ஏற்படும் தருணத்தில் அப்பொழுதே திருத்தப் படுவதற்கும் சரியான உறுதியான விடைகளை இதன்வழி பெறமுடிகிறது. சுயமாகச் சிந்திப்பதற்கும் கருத்துக்களைப் பரிமாறிக்கொள்வதற்கும் கையடக்க கணினிகள் பேருதவியாக அமைகின்றன. படைப்பாளர், எழுத்தாளர்களும் தகவல்களைப் பரிமாறிக்கொள்வதற்கும் கையடக்க மிகுந்த ஆர்வம் ஏற்படுவதற்கும் இம்முறை உதவியாக அமைகிறது. நூல்களையோ ஏடுகளையோ சுமந்து செல்ல வேண்டும் என்ற நிலையையின்புத்தகங்களும், மின் இதழ்களும் மாற்றி உள்ளன.

கிளவுட்கம்ப்யூட்டிங் (cloud computing)எனும் கணினி கட்டமைப்பின் மூலம், பல்லுடகம், ஊடாடல், உருவாக்கம், நிர்வாகம் மற்றும் விநியோகம், கருவிகளின் தொகுப்பு மின்கற்றல் படிப்புகள் மற்றும் தகவல்தொடர்புமற்றும் கருத்துக்களை, இணையம் மூலம் தமிழ்மொழியை உலகளாவிய முறையில் அறிமுகப்படுத்த அடித்தளம்இட்டுள்ளது.

பரிந்துரைகள்

அன்டிராய்ட் (Android), விண்டோஸ்8 (Windows8) முறையில் தமிழ் பயன்பாடு இப்போது முன்னேற்றம் கண்டுள்ளது.ஐ-பட் போன்ற மடிக் கணினிகளில் IOS 7செயலியில், தமிழ் தட்டச்சு முறை வழங்கப்படும் செய்தி, தமிழை தரமாக மின்னூல்கில் கொண்டு சேர்க்க புதிய வழி அமைத்துள்ளது.

தமிழில், இதுவரை பொருளாதார அடிப்படையில் பயன் உண்டாகும் வகையில் முயற்சிகள் மேற்கொள்ளப்படாததால் பல அரிய செயல் பாடுகள் முனையிலேயே கருகிப்போய் விடுகின்றன.அந்த நிலை தற்போது ஆக்ககரமான மாற்றங்களை பரவலாகக் காண முடிகிறது.

தொன்மைமிக்க செம்மொழியில்உள்ள உன்னதமான புதினங்கள், காப்பியங்களை மின்னூல்களாக இன்றைய தலைமுறையினர் விரும்பும் வகையில் வெளிக்கொணர முன்னணி படைப்பாளர்கள் முயலலாம்.

21 நூற்றாண்டில் தமிழ் சார்த்த தொழில் தொடங்கவும் தமிழை வாழ்க்கை மொழியாகவும் மாற்றுவதற்கு புத்தாக்க வாழ்நாள் பயிற்சிகள், உதவித்திட்டங்கள், மின்னூல் தயாரிப்பு வேலை வாய்ப்புகள் மூலம் புத்தாக்கத் தொழில் முனைவர்களை உருவாக்குதற்குப் பேருதவியாக அமைவதற்கு உலகளாவிய முறையில் முயற்சிகள் மேற்கொள்ளப்பட வேண்டும்.

முடிவுரை

தொன்மை மிக்க தமிழ், தனக்கே உரிய அழகை, செழுமையை நம்மிடம் அடையாளம்காட்ட இதுவரை இணையம், ஒரு வழித்தடமாக அமைந்துள்ளது.இனி வரும் காலம், இளைய சமுதாயம் தமிழை சுவாசிக்க, நேசிக்க புதுவகையான சிந்தனைகளை உள்ளடக்கிப் புத்தாக்க எழுச்சியோடு செம்மொழியை உலகோர் அனைவரும் மின்னூல், மின் இதழ் வழி, 21 நூற்றாண்டையும் கடந்து அதன்படி நிற்க, புதிய ஊடகமும் இணையமும் புத்தாக்கச்சிந்தனையும் பேருதவி புரியும் என மனமார நம்புவோம்.

*1 flurryfindings, <http://www.newmediatrendwatch.com/world-overview/98-mobile-devices>

*2<http://www.onlinemagazineshub.com/tamil-magazines.html>

References Websites

<http://www.adobe.com/sea/>

<http://www.apple.com/education/>

www.nearpod.com

www.google.com

<http://tamilmurasu.com.sg/>

<http://tamilelibrary.org/teli/sintalk1.html>

[https://wiki.smu.edu.sg/digitalmediaasia/Digital Media in Singapore](https://wiki.smu.edu.sg/digitalmediaasia/Digital%20Media%20in%20Singapore)

www.ida.gov.sg

www.mda.gov.sg

<http://www.nac.gov.sg/>

Reference Books

68 Tips for eLearning Engagement and Interactivity

Karen Forni 5.20.13 | Publications Library | Guild eBooks

61 Tips on mLearning: Making Learning Mobile

Karen Forni 10.24.12 | Publications Library | Guild eBooks

53 Tips on Using the Cloud for eLearning

Karen Forni 5.7.12 | Publications Library | Guild eBooks

, Sandie Gay, Using e-Books and e-Readers for Adult Learning

Shubhangi Harsha, Sumit Kataria, Teaching with iPad How-To

Salmon, Gilly (2002) E-tivities: The Key to Active Online Learning, London, Kogan Page

Paul Chin, Using C&IT to Support Teaching

Slater, Paul and Varney-Burch, Srah (2001) Multimedia in Language

Susan Ko, Steve Rossen : Teaching Online - A Practical Guide, 2nd Edition

Reynol Junco & Jeanna Mastrodicasa : Connecting to the net.generation

Les Lloyd, Editor : Technology and Teaching

Kristen J. Nelson; 2nd Edition : Teaching in the Digital Age

Shirley Bennett, Debra Marsh and Calre Killen : Handbook of Online Education

பயன்பாட்டு நோக்கில் தமிழ்க் குறுஞ்செயலிகள்

முனைவர் துரை. மணிகண்டன், தமிழ்த்துறைத் தலைவர், பாரதிதாசன் பல்கலைக்கழக உறுப்புக்
கலை மற்றும் அறிவியல் கல்லூரி, இனாம்குளத்தூர், திருச்சிராப்பள்ளி.
E-mail: mkduraimani@gmail.com

முன்னுரை

”யாதும் ஊரே யாவரும் கேளிர்” என்ற பொன்மொழிக்கு நிகராக இன்று உலகில் இருக்கும் அனைவரும் உறவினர்களாக உருவாக்கம் பெற்று வருகின்றனர். காரணம் இணையம். இணையம் இன்று பல்புறப்பணிகளையும் செய்யும் அன்பு கடவுள்.

இணையம் முதலில் பாதுகாப்பிற்காகப் பயன்படுத்தப்பட்டு வந்தாலும் பிறகு அது பல்வேறு பணிகளையும் ஒரே நேரத்தில் செய்யும் ஆற்றலுடன் வலம் வந்தது.

இணையத்தில் நாம் ஒரு செய்தியைத் தேடி எடுக்க வேண்டுமென்றால் அதற்குப் பல இணைய முகவரிகளோயோ, வலைப்பதிவு முகவரிகளையோ, பொது தளங்களையோ சென்று நாம் காண வேண்டிய சூழலில் உள்ளோம். இதனால் அதிகமான நேரத்தையும், பொருள்செலவையும் இழக்க நேரிடுகிறது. ஆனால் செல்பேசி மற்றும் கையடக்கக் கணினிகளில் என்ன செயலுக்கு எந்த வலைமுகவரியோ அதைவிட அதிகமான, மிகத் துல்லியமானச் செய்திகளை நமக்குத் தரும் தமிழ்க் குறுஞ்செயலிகள் இன்று ஆயிரத்திற்கும் மேல் உள்ளன.

குறுஞ்செயலி – விளக்கம்

கையடக்கக் கணினிகளிலும் செல்பேசிகளிலும் பயன்படுத்தப்படும் சிறிய மென்பொருளையே குறுஞ்செயலிகள் என்று அழைக்கலாம். மேலும் நவீனத் தொலைபேசிகளில் ஒரு குறிப்பிட்ட செயலைச் (Action) செய்துமுடிக்க மனிதர்களால் பல கட்டளைகளைப் பிறப்பித்து உருவாக்கப்படும் நிரல் தொகுப்பை குறுஞ்செயலி (Apps) என்று அழைக்கப்படுகிறது. கணினியில் பல செயல்களைச் செய்யும் செயலிகளை, மென்பொருள் (Applications - Software) என்றும், குறுகிய நிலையில் செய்யக்கூடியதை குறுஞ்செயலி என்றும் வழங்கப்படுகிறது. மென்பொருள் என்பது திரைப்படம் போன்றது குறுஞ்செயலிகள் என்பது குறும்படம் போன்றது.

பயன்பாடுகள்

இந்த குறுஞ்செயலிகள் அதிகமாக ஆப்பிள் ஐ.ஓ.எஸ், ஆண்ட்ராய்டு செல்போன்களிலும், கையடக்க கணினிகளிலும் பயன்படுத்தப்பட்டு வருகின்றன. இதில் பயன்படுத்தப்படும் தமிழ்க் குறுஞ்செயலிகள் ஆப்பிள் செல்பேசி மற்றும் கையடக்க கணினிகளில் அறுநூறும் (600), ஆண்ட்ராய்டு செல்பேசி மற்றும் கையடக்கக் கணினிகளில் ஆயிரத்திற்கும் (1000) மேலும் உள்ளன.

செல்பேசிகளிலும் கையடக்கக் கணினிகளில் தமிழ்க் குறுஞ்செயலிகள் ஆயிரத்திற்கும் மேல் இருந்தாலும் பலருக்குத் இதன் பயன்பாடுகள் தெரியவருவதில்லை. எனவே அதன் பயன்பாட்டை எடுத்து விளக்கும் முறையில்

I-Kural என்ற திருக்குறள் செயலி, தமிழ்ஸ் கோட்ஸ் (Tamil Quotes), தமிழ்க் காப்பியம் (Tamil epic), tamil Indian-news papper, தேவாரம், மின் இதழ்கள், சுவாமி விவேகநந்தர், கவிஞர்கள் (Tamil poet bio), ஜெயகாந்தன் (tamil story), செல்லினம், பூங்கா, தமிழ் குர் ஆன், தமிழ் பைபிள், பகவத் கீதை, இந்து மந்திரங்கள், என்ற இதுபோன்ற எண்ணற்ற தமிழ்க் குறுஞ்செயலிகளின் பயன்பாடுகளை இக்கட்டுரை எடுத்து விளக்கவுள்ளது.

I-குறள் (I-kural)

இந்த குறுஞ்செயலியை வடிவமைத்தவர் திரு.அரவிந்த என்பவர். கல்வி என்ற பாகுபாட்டில் இதில் அறத்துப்பால், பொருட்பால், இன்பத்துப்பால் என்ற திருக்குறள் இயல் அமைப்பில் பிரித்துக் கொடுத்துள்ளார்.

ஒவ்வொரு குறளுக்கும் உரை மிகத்தெளிவாக அமைந்திருப்பது சிறப்பு. மணக்குடவர், பரிமேழலகர், G.U.போப், போன்றோரின் உரை விளக்கம் மற்றும் மொழிபெயர்ப்புடன் இடம்பெற்றிருக்கிறது. மு.வரதராசனார், கலைஞர் கருணாநிதி, சாலமன் பாப்பயா போன்றோரின் உரைகளும் இதில் உள்ளன. இந்த குறுஞ்செயலி உலகப் பொதுமறைக்குக் கிடைத்த புதுமையான வரவேற்பு என்று கூறலாம். இந்த உரை விளக்கம் பழைய உரையாசிரியர்களின் உரையோடு புதிய உரையாசிரியர்களின் உரையைப் பொருத்திப் பார்க்கவும் இந்த குறுஞ்செயலி பயன்படுகின்றது.

தமிழ்ஸ் கோட்ஸ் (Tamil Quotes)

இந்த குறுஞ்செயலியில் தமிழ்மொழியில் உள்ள முதுமொழிகள் என்று கூறப்படுகின்ற புகழ்பெற்ற ஒருசில பழமொழிகள் இடம்பெற்றுள்ளன.

ஐந்தில் விளையாதாது

ஐம்பதில் விளையுமா

போன்ற பழமொழிகள் இடம்பெற்றிருக்கிறது. இந்த செயலியில் பல இடங்களில் ஒற்றுப்பிழை உள்ளன. அவற்றை நீக்கி இந்த செயலியை வெளியிட்டால் தமிழ் வளத்திற்குப் பெருமை சேர்க்கும்.

தமிழ் நூலகம் (Tamil Library)

அமுத ரஞ்சனி இரகசியங்கள் தொடங்கி சப்தகிரி தமிழ் முடிய 1000 த்திற்கும் மேற்பட்ட பழமையான நூல்கள் மின்புத்தகமாக இந்த குறுஞ்செயலியில் உள்ளன. இவற்றில் இடம்பெற்றுள்ள மின்புத்தகங்கள் 100 முதல் 400 பக்கத்திற்கு மிகாமல் உள்ளன. இந்த குறுஞ்செயலியில் நூல்களை தரவிறக்கம் செய்து படிப்பது மிகப்பயனுள்ளதாக இருக்கிறது. இந்தச் செயலியில் நூல்களைத் தரவிறக்கம் செய்வதில் கால அதிகமாகப் பிடிக்கிறது. இதனைப்போக்கும் விதமாக பக்கத்தை இயலாகப் பிரித்து 30 பக்கமாக கொடுக்க முன்வந்தால் இந்த செயலி பயன்பாட்டாளர்களுக்கு மிகப் பயனுள்ளதாக அமையும்.

தமிழ்க்காப்பியம் (Tamil Ebic)

அவ்வையாரின் நல்வழித் தொடங்கி சிலப்பதிகாரம், பாரதிதாசன் படைப்புகள், கல்கியின் பொன்னியின் செல்வன் (1145 பக்கம்), திருமந்திரம், திருக்குறள் வாசகம் போன்ற காலத்தால் அழியாதா தமிழ்ப்படைப்புகள் இந்த செயலியில் இடம்பெற்றுள்ளன. சுஜாதாவின் விஞ்ஞானப்பார்வையிலிருந்து மிகப்பழமையான நூல்களும் உள்ளன. இந்த செயலியில் மதுரைத் திட்டத்திலிருக்கும் ஒருசில மின் நூல்களை இணைத்துள்ளார்கள். இந்தச் செயலியில் உள்ள நூல்கள் அனைத்தையும் மற்ற நண்பர்களுக்கு மின்னஞல் மூலமாக அனுப்பலாம். கணினியிலும் இதனைப் பார்க்கலாம். திரையில் வாசிக்கும்போது இடையில் நிறுத்தி மீண்டும் படிக்கலாம். நாம் எந்த பக்கத்தில் இடையில் நிறுத்தியுள்ளோம் என்பதைக் குறித்துக்கொள்ளும் வசதியும் இந்த செயலியில் இடம்பெற்றிருப்பது மற்ற தமிழ்க் குறுஞ்செயலியிலிருந்து முற்றிலும் மாறுபட்ட வடிவில் வடிவமைக்கப்பட்டுள்ளது தனிச்சிறப்பாகும்.

தேவாரம்

திருஞானசம்பந்தர், திருநாவுக்கரசர், சுந்தர் பாடிய (மூவர்தேவாரம்) தேவாரப் பதிகங்கள் பொழிப்புரை மற்றும் ஆங்கில மொழிபெயர்ப்பில் வெளியிட்டுள்ளனர். இந்த செயலி சைவை அடியார்களின் விருப்பத்தேர்வாக அமைந்துள்ளது. மேலும் பலரும் இதனை பயன்படுத்தும் எளிமையான முறையில் இந்த செயலி வடிவமைப்பட்டுள்ளது.

ஜெயகாந்தன் (Tamil Story)

Tamil Story என்ற குறுஞ்செயலியில் ஜெயகாந்தனின் தேர்ந்தெடுக்கப்பட்ட பத்து சிறுகதைகள் இடம்பெற்றுள்ளன. இதுபோன்று பிற தமிழ் எழுத்தாளர்களின் படைப்புகள் பல தமிழ்க் குறுஞ்செயலிகளாக உருவாக்கப்பட வேண்டும்.

கவிஞர்கள் (Tamil poet Bio)

Tamil poet Bio- என்ற தமிழ்க்குறுசெயலியில் அகத்தியர், நக்கீரர், ஒட்டக்கூத்தர், கம்பர், சேக்கிழார், சிவப்பிரகாச சாமிகள், வள்ளலார், பாரதியார், பாரதிதாசன், சுப்பிரமணிய சிவா, திரு.வி.கல்யாணசுந்தரம், கவிமணி தேசிய விநாயகம்பிள்ளை, மறைமலை அடிகள், புதுமைபித்தன், செய்குதம்பி பாவலர், கண்ணதாசன், வாலி, வைரமுத்து, போன்றோர்களின் தன்விபரக் குறிப்புகள் தமிழ், ஆங்கில இருமொழிகளிலும் இடம்பெற்றுள்ளன.

மின் இதழ்கள்

விகடன் குழுமம் தனித்தனியாக ஒவ்வொரு பிரிவிற்கும் ஒவ்வொரு செயலியை வடிவமைத்து வெளியிட்டுள்ளனர். மேலும் இந்தியாடுடே தமிழ், நக்கீரன், பெரியார் இதழ் போன்ற தனி இதழ்களும் தங்களுக்கான தமிழ்க்குறுஞ்செயலிகளை உருவாக்கி வெளியிடுகின்றன.

Tamil, Indian News paper

Tamil, Indian News paper+ என்ற பெயரில் தமிழ் ஆங்கிலம், இந்தி, உள்ளிட்ட மொழிகளில் அமைந்துள்ள 31 நாளிதழ்கள் இந்த செயலியில் இடம்பெற்றிருக்கின்றன. தினகரன், தினமலர், தினத்தந்தி நாளிதழ்கள் தமிழ்மொழியில் இடம்பெற்றவைகள். இவை மட்டுமின்றி முன்னனி செய்தித்தாள்கள் தனித்தனியாகத் தமிழ்க்குறுஞ்செயலிகளை வடிவமைத்து வெளியிட்டுள்ளனர். எ.கா தினமணி, மாலைமலர், காலைக்கதிர் போன்றவையாகும்.

செல்லினம்

முரசு முத்துநெடுமாறன் அவர்களால் செல்பேசிகளில் தமிழ் எழுத்துருக்களைத் தட்டச்சு செய்வதற்குப் பயன்படும் செல்லினம் என்ற தமிழ் குறியீட்டு முறையை அறிமுகம் செய்து அதனை உலகப்புகழ்பெற்ற செல்பேசி நிறுவனங்களில் கொடுத்துள்ளார். அவரின் செல்லினம் குறுஞ்செயலியில் மலேசிய, சிங்கப்பூர் தமிழ்

எழுத்தாளர்களின் படைப்புகளை உலகத்தமிழர்களுக்கு அறிமுகம் செய்துவைத்துள்ளார். இதுபோன்று இன்னும் பல வெளிநாடுகளில் வாழும் தமிழ் ஆர்வலர்கள் அந்நாட்டுத் தமிழ் எழுத்தாளர்களைக்கொண்ட தமிழ்க் குறுஞ்செயலிகளை உருவாக்கி வெளியிட்டால் தமிழ் உலகத்திற்குப் பெருமை சேர்க்கும்.

சுவாமி விவேகநந்தர்

சுவாமி விவேகநந்தர் செயலியில் சிக்காக்கோ சொற்பொழிவுகள், வரலாறு, ஆன்மீக வரலாறு, பொன்மொழிகள், ஐயாவின் அனுபவங்கள் அடங்கிய கருத்துப்பேழை தமிழ்க்குறுஞ்செயலியாக செல்பேசியில் கிடைக்கின்றன. இதுபோன்ற இன்னும் பல அறிஞர்களின் படைப்புகள், வாழ்க்கைக் குறிப்புகள் கொண்ட தமிழ்க்குறுஞ்செயலிகள் தமிழில் உருவாக்கப்படவேண்டும்.

தமிழ்க்குறுஞ்செயலியால் பயன்பெறுவோர்

தமிழ்க் குறுஞ்செயலியால் பயன்பெறுவோர் செல்பேசி வைத்திருப்பவர்கள் மற்றும் தமிழ் ஆர்வலர்கள்தான். இது மேலும் வளர்ந்து அனைத்து மக்களையும் சென்றடைவ வேண்டும். அதற்கு இந்த உத்தமம் நடத்தும் மாநாடு வழிவகை செய்யும் என்பதில் எமக்கு நம்பிக்கை உள்ளது.

கட்டுரையின் நோக்கம்

1. தமிழ்க் குறுஞ்செயலிகளில் எவை எவை எந்தெந்த பயன்பாட்டிற்குப் பயன்படுத்தப்படுகின்றன.
2. தமிழ்க்குறுஞ்செயலிகளின் பயன்பாடுகள் இன்று எந்த அளவிற்கு முன்னேற்றம் பெற்றுள்ளன.
3. இதனால் பயன்பெரும் பயனாளர்கள் யார்? யார்?
4. இந்த குறுஞ்செயலிகளைப் பயன்படுத்துவதால் காலவிரயம் ஏற்படாமல் தவிர்க்கப்படுகிறது.
5. தற்பொழுது இருக்கும் தமிழ்க் குறுஞ்செயலிகள் மட்டுமின்றி புதிய தமிழ்க் குறுஞ்செயலிகள் உருவாக்கப்படவேண்டும்
6. உலகில் உள்ள பெரிய வீரர்களின் கதைகளை உள்ளடக்கியும், இராமயணம், மகாபாரத கதைத்தலைவர்களின் வீரதீரச் செயல்களை குறுஞ்செயலிகளாக உருவாக்கியுள்ளனர். அதுபோல நமது மன்னர்களான சேர, சோழ, பாண்டியர்களின் வீரதீர விளையாட்டுகளை உள்ளடக்கிய தமிழ்க் குறுஞ்செயலிகள் உருவாக்கப்பட வேண்டும்.
7. இணையப்பயன்பாட்டிலும் மென்பொருள் வளர்ச்சியிலும் ஆங்கிலத்திற்கு அடுத்தப்படியாக சீனா இருக்கிறது. அதற்குக் காரணம் அவர்களின் இணையதளமென்பொருள்கள் வடிவமைப்பு அனைத்தும் ஒரே குறியீட்டுமுறையில் அமைந்திருப்பதால்தான். அதுபோல இனி உருவாக்கும் தமிழ்க்குறுஞ்செயலிகள் அனைத்தும் ஒரே குறியீட்டு முறையில் அமைந்தவையாக இருக்கவேண்டும்.
8. பயன்பாட்டில் உள்ள ஒருசில தமிழ்க்குறுஞ்செயலிகளில் எழுத்துப் பிழைகள் வாக்கியப்பிழைகள் உள்ளன (தமிழ்ஸ் கோட்ஸ் (Tamil Quotes) அவற்றை நாம் போக்க வேண்டும்.

முடிவுரை

இதுவரை தமிழ்க் குறுஞ்செயலிகள் பற்றிய சிறிய அறிமுக நிலையில் இக்கட்டுரை அமைந்திருக்கும். இது வருங்காலத்தில் வளர்ந்த நிலையை அடைய வேண்டும். அதற்கு நாம் புதிய புதிய தமிழ்க் குறுஞ்செயலிகளை உருவாக்கம் செய்ய வேண்டும். அவ்வாறு உருவாக்கும் குறுஞ்செயலிகளை நாம் பயன்படுத்த வேண்டும். எதையும் பயன்படுத்தினால்தான் அதன் பயன்பாடு தெரியவரும்.

கூகிள் இணையக் கடைகளில் (Google Play) தமிழ்மொழி சார்ந்த அண்ரோய்ட் குறுஞ்செயலிகள் (Android Apps): ஒரு பார்வை

மு. மணியரசன், விரிவுரையாளர், துவான்கு பைனான் ஆசிரியர் கல்விக்கழகம், பினாங்கு.மலேசியா
maniyarasan.muniandy@gmail.com.

முன்னுரை:



தற்போது இணையத்தில் கூகிள் தேடுதளம் முன்னணியில் இருக்கிறது என்றால் அதனை மறுப்பார் இல்லை. கூகுளின் எண்ணற்ற வசதிகளான கூகுள்+, கூகுள் பஸ், கூகுள் குரோம், கூகுள் ஓஎஸ், கூகுள் மேப்ஸ், கூகுள் சாட், கூகுள் அப்ஸ், கூகுள் ப்ளாக்கர், கூகுள் பிக்நிக், கூகுள் ஜிமெயில், கூகுள் எர்த், கூகுள் யூடியூப், கூகுள் பிகாசோ, கூகுள் டிரைவ், கூகுள் ப்ரபைல், கூகுள் ஆண்ட்ராய்டு, கூகுள் எபிசோடிக், கூகுள் லேப்ஸ், கூகுள் ட்ரேன்சிலேஷன், கூகுள் ஆர்குட் போன்ற அனைத்து அம்சங்களும் இணையதளத்தில் இருந்தபடியே நமக்கு உதவுகின்றன.

நம் வாழ்வில் ஒரு அங்கமாகத் திகழும் கூகிள் நிறுவனம் தினமும் பற்பல மாற்றங்களைச் செய்து, பல மில்லியன் மக்களை தன் பக்கம் இழுத்து வைத்து இருக்கின்றது. எளிமை, வேகம், புதுமை, பாதுகாப்பு ஆகியவற்றால் கூகிள் தனித் தன்மை பெற்று விளங்குகின்றது. இணையப் பயனாளர்களுக்கும் கூகிளின்பால் ஆர்வம் ஏற்பட்டுத் தொடர்ந்து நாட்டம் செலுத்தி வருகின்றனர்.

கூகிள் நிறுவனம் வழங்கும் வசதிகளுள் 'கூகிள் பிளே' (Google Play) என்பது மிகவும் குறிப்பிடத்தக்க சேவையாகும். இது இலக்கமுறைத் தகவல்களை வழங்கும் இணையக் கடையாகும். இச்சேவையின் வழியாக அன்டிராப்டு பயன்பாடுகள், இசைக்கோப்புகள், புத்தகங்கள், திரைப்படங்கள், விளையாட்டுகள் போன்றவை பயனாளர்களுக்குக் கிடைக்கின்றது. கடந்த 2012, மார்ச் முதல் கூகிள் நிறுவனம் தனது ஆன்டிராப்டு அங்காடியையும், இசைச் சேவையையும் இணைத்து கூகிள் பிளே என்ற இணையக் கடையைத் தொடங்கியது. இவற்றில் பெரும்பாலானவை பயனாளர்களுக்கு இலவசமாக அளிக்கப்படுகின்றது. ஒரு சில குறுஞ்செயலிகளை மட்டும் பயனாளர்கள் அவசியம் கருதி பணம் கொடுத்து வாங்க வேண்டிய நிலை இருக்கும். இக்கடைகளில் தமிழ் மற்றும் தமிழ் இலக்கியம் தொடர்பான குறுஞ்செயலிகள் மற்றும் அவற்றின் பயன்பாடு ஆகியனவற்றை ஆராய்வதே இக்கட்டுரையின் நோக்கமாகும்.

கூகிள் இணையக் கடைகளில் (GOOGLE PLAY) தமிழ்மொழி சார்ந்த அண்ரோய்ட் குறுஞ்செயலிகள் (ANDROID APPS):

தமிழ்மொழி சார்ந்த இவ்வகைக் குறுஞ்செயலிகள் <http://play.google.com/store/search> என்ற இணைய முகவரியில் கிடைக்கின்றன. இக்கட்டுரை வரையப்பட்ட நாள் வரையில் (10.7.2013) கூகிள் இணையக் கடைகளில் கிடைக்கப்பெற்ற தமிழ் சார்ந்த குறுஞ்செயலிகள் சில அடையாளங் காணப்பட்டுள்ளன. அவற்றை இங்கே காணலாம்.

1. இயல்

அ) தமிழ் கற்பிக்கும் குறுஞ்செயலிகள்

தமிழ் ஆசான்



Tamil Aasaan Learn on the Go

தமிழ் கற்போம்



Tamil Karpom

தமிழ் கற்றல் துணைப்பொருள்-1



Tamil Learning kit - 1

தமிழ் கற்றல், தமிழில் எழுதுதல்



Learn and Write Tamil Letters

தமிழ் எழுத்துக்கள்



Tamil Alphabets

தமிழ் கற்போம்



Lets Learn Tamil

தமிழ் எழுத்துக்கள்



Tamil Alphabets Lite

எளிதில் தமிழ் பேச

lets Learn Tamil



Tamil Speak Easy

தமிழ் கற்றல்



EduTamil

தமிழ் -ஆங்கிலம் (தங்கிஸ்) முறையில் தட்டச்சு



Tanglish - Type In Tamil

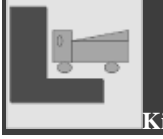
தமிழ் அரிச்சுவடி



Learn Tamil Alphabets



The Kids school (Tamil)



Kids Basics Flash Cards Tamil



Tamil Learning for Kids

தொடக்கநிலையில் தமிழ் கற்கும் பயனாளர்களுக்கு இக்குறுஞ்செயலிகள் மிகப் பயனாக இருக்கும். எளிமையாக மொழி கற்றுக்கொள்ளும் வகையில் வண்ணப் படங்களும், ஒலிப்படங்களும் இவற்றில் கொடுக்கப்பட்டுள்ளன. பயனாளர்கள் தாம் விரும்பிய வண்ணங்களிலும் வடிவங்களிலும் யதமிழ் எழுத்துக்களை எழுதிப் பழகலாம்தமிழ் கற்பிக்கும் குறுஞ்செயலிகளுள் இதுவும் ஒன்று. தொடக்க நிலையில் தமிழ் கற்க விரும்பும் பயனாளர்கள் இக்குறுஞ்செயலியினைப் பயன்படுத்தலாம். தொடக்கநிலையில் தமிழ் கற்போருக்கு இந்தக் குறுஞ்செயலிகள் முதன்மைப் பங்கு ஆற்றுகின்றன.

ஆ) தமிழ் தட்டச்சுக் குறுஞ்செயலிகள்

செல்லினம்



Sellinam

தமிழா



ThamiZha! -Tamil Visai

செல்லினம் 2003 ஆண்டு தை முதல் நாள் தொடங்கி இலவயமாக வழங்கப்பட்டு வருகின்றது. முரசு அஞ்சல் அல்லது தமிழ் 99 விசைப்பலகை முறையைப் பயன்படுத்தி எளிமையாகத் தமிழில் தட்டச்சு செய்யலாம். தமிழில் தட்டச்சு செய்வதற்கு உதவும் மற்றொரு வடிவம் தமிழா ஆகும். இணையக் கடைகளில்லிதுவும் இலவயமாகக் கிடைக்கின்றது. பயனாளர்களின் கையடக்கத் தொலைப்பேசியில் இவற்றைப் பதிவிறக்கம் செய்துகொண்டால், தமிழ்கிலத்தில் தட்டச்சு செய்யும் போக்கு காலப்போக்கில் மறைந்துவிடும்.

இ) தேடுபொறி குறுஞ்செயலிகள்
தமிழ் உலாவி



தமிழ் தொடர்பான வலைத்தளங்களுக்குச் சென்று, எழுத்துரு சிக்கலால் விழி பிதிங்கி நிற்கும்போது, உடனே நமது சிக்கலைத் தீர்க்கும் வல்லமை 'தமிழ் உலாவி'க்கு உண்டு. யூனிகோட் எழுத்து வடிவத்தைக் கொண்ட பத்திகளை மட்டுமே இந்த உலாவியைக் கொண்டு படிக்க முடியும் என்பதைக் கவனத்தில் கொள்ள வேண்டும்.

ஈ) தமிழ் எழுத்துரு தரும் குறுஞ்செயலிகள்

தமிழ் கீ போர்ட் யூனிகோட்



தமிழ் யூனிகோட்



தமிழ் யூனிகோட் ஃபொர்ட் எஸ்.எம்.எஸ்



தமிழ் கீபோர்ட்



எழுத்தாணி



Ezhuthani - Tamil Keyboard

இந்த குறுஞ்செயலியினை டிவிட்டர், பேஸ்புக், இமெயில் போன்றவற்றில் பயன்படுத்தலாம். (android app helps you to write tamil words easy and fast. if you write "thamiz" and that will change it into "தமிழ்". Simply Type english in keyboard. it is useful Tamil Unicode converter. You can make the tamil text posted in whatsapp, Twitter, Facebook, SMS and Email. Easy to copy and paste and does not need to be online to type)

உ) குறுஞ்செயலி தமிழ் அகராதிகள்

தமிழ் டிக்ஸனரி ஆஃப்லைன்



Tamil Dictionary Offline

தமிழ் டிக்ஸனரி ஆன்லைன்



Online Tamil Dictionary

தமிழ் டிக்ஸனரி : இங்லிஸ் தமிழ், தமிழ் இங்லிஸ்



Tamil Dictionary E-T & T-E

மேற்கண்ட அகராதிகள் தமிழரின் கையடக்கத் தொலைப்பேசிகளில் இருந்தால் மிகவும் பயனாக இருக்கும். அதிலும் தமிழ் தொடர்பான பணிகளில் அமர்ந்திருக்கும் கல்வியாளர்கள், செய்தியாளர்கள், மாணவர்கள் ஆகியோருக்கு உற்ற நண்பனாக இருக்கும். ஆப்லைன் அகராதி, இணைய வசதி இல்லாமல் நமது கையடக்கத் தொலைப்பேசியில் இயங்கும் வல்லமை கொண்டது.

ஊ) தமிழ்ச் சிறுகதைகள், நாவல்கள், கவிதைகள் தரும் குறுஞ்செயலிகள்

டாக்டர் மு. வ சிறுகதைகள்



Mu Va Tamil Short Stories

மகாகவி பாரதியார் சிறுகதைகள்



Bharathi - Tamil Stories

கல்கியின் நாவல்கள்



Ponniyin Selvan (Tamil)



Dhiyaga Boomi in tamil - kalki



Magudapathi by Kalki in Tamil



Mohini Theevu - Kalki in Tamil



Kalki Short Stories 3 - Tamil



Dr. CN Annadurai Tamil Stories



Pudhumai Pithan Tamil Stories



Thirukural E-Book - Tamil



Thirukkural with meanings



Ra Karthikesu in Tamil Stories



Bharathiyar Paattu (Tamil)



Madhole - Tamil Kavithaikal

Madhole - Tamil Kavithaikal :



Life Quotes- Tamil Ponmozhi



Uvamaigal



Swami Vivekananda-Tamil Quotes

2. இசை



Old Tamil Movies



Watch Tamil Movies



Tamil Short Films



Tamil TV Serials and Shows



AR Rahman Tube - Tamil



Watch Tamil Live TV - Free

Watch Tamil Live TV - Free



♥♥ Tamil Movie Song Lyrics ♥♥



Tamil New Hit Songs Ringtones



TAMIL Radio



Tamil Radio FM



Famous Tamil Rhymes



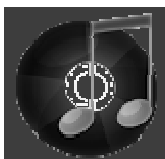
Tamil Karaoke



Tamil Nursery Rhymes



Tamil Nursery Songs /Game Kids



Tamil Video Songs (HD)



Tamil Songs



Thirumanthiram



Thevaaram - A book of Devotion



Thiruvagasam Part-1



Islam Tamil



Tamil Bible Ad Free



Tamil Crorepati Quiz Game

3. நாடகம்.



Dinakaran - Tamil News



Google tamil news



Wikitrack Tamil

Saaranga Infotech Wikitrack Tamil is an android application that helps you to track the updates, changes on Tamil Wikipedia.



India Today Tamil



AdaDerana Tamil



Dinamalar for Tablets



Tamil NewsPapers



Tamil Cinema News & Events



Tamil News Alerts



Tamil Seithigal



Tamil Naal Kaati



Tamil SMS



TV Tamil Open Directory



Facebook - Tamil Posting



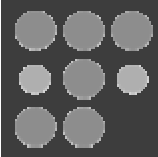
i LYRICS Tamil Songs



All Tamil



Tamil Radio and News



Tamil Quiz



Tamil Comedy & Punch Dialogues



Tamil Songs Downloader



Tamil Tāy Vālttu



Learn Univerce in TAMIL



500+ Tamil Jokes



Hangman Tamil Game - Classic



Tamil Baby Names



Tamil business Directory



Tamil Marriage Match

மொபைல் இயங்குதளத்தில் (os) தற்போது ஆன்ட்ராய்டு (Android) சக்கைப்போடு போட்டு வருகிறது. இதன் எளிமை மற்றும் வசதிகள் காரணமாக தற்போது இதனுடைய பயனாளர்கள் எண்ணிக்கை பல மடங்கு அதிகரித்து வருகிறது. இதே வேகத்தில் சென்றால் 2015 ல் அசைக்க முடியாத இடத்தில் இருக்கும் என்று வல்லுனர்கள் கணித்துள்ளார்கள். இது ஒரு சுதந்திர மென்பொருள் (Open Source) என்பதும் இது மிக பிரபலமானதிற்கு முக்கிய காரணமாகும். இவ்வளவு சிறப்புகள் இருந்தாலும் தமிழ் மொழியில் போதுமான குறுஞ்செயலிகள் இல்லை என்றே கூறவேண்டும். இது தமிழ் ரசிகர்களை வருத்தத்தில் ஆழ்த்துகிறது. ஆப்பிள் நிறுவனத்துடன் ஒப்பிடும்போது குறைகள் அதிகமாகவே தென்படுகின்றன. கூகுள் விரைவில் இந்த வசதியை தரும் என்று நம்புவோமாக.