12th International Tamil Internet Conference 2013
Computational Linguistics | கணினி மா ழியிய᾿

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 கால்வகைகளை விளக்கம்
   உட்கரைப்பாறானோ
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 | இயிக மா ழியி

11. தமிழ் அறிய கணினி எள்ளத்தக் கால்வகைகளை இழந்து? ,
    அரச ஆர்வாதீன்
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, Elancheziyhan, Rajapandian & Madhan Karky
14. Lyric Objects & Spot Indices for Paadal, a Tamil Lyric Search Engine,
    Karthikeyan, Elancheziyhan, 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, Elancheziyhan, Karthikeyan & Madhan Karky
18. Scoring Models for Tamil Lyrics, Elancheziyhan,
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 | மாணவிக் கணினிகள்

26. தமிழ் மாணவிக் கணினிகள் பெற்றுக்கொள்ள: இலண்டன் கையில் கிட்டத்தட்டம், 
   ப.அர.நாகானி
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. புலியமைத்த செயற்கைக்காலம் வைத்து பதிக்களில் பயன்படுத்தப்படும்: குறிப்பிட்டு 
   காட்சிகளமல்ல, 
   ஆ. க. ராமகிருஷந், ப. அஞ்சிலிவிருத
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, 
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 Tamil script on Tablet PC,
Bhargava Urala, A G Ramakrishnan
37. Recognition System for Tamil Sign Language Using Hand Gestures,
Vijay Jeyakumar, K. Bommana Raja

Internet and software | இந்தியமுற்ற மொழியான குழு

38. Usage of Tamil in Internationalized Domain Names,
Balaji Vasudevan
39. மொழியான விளக்கம்,
மேலாரை மோவாரா மெரான்
40. முழுநோய் தமிழ் மொழியான குழு: மேலாரை மெரான் மெரான், தனயை மெரான்
41. உயரும் தமிழ் மொழியான புதுமை அணைய அரசு
42. மேலாரை மோவாரா மெரான் மெரான் மெரான் மெரான்
43. மேலாரை மோவாரா மெரான் மெரான் மெரான் மெரான்
44. 'Raspberry PI' Future Outlook for Tamil,
Saravanan Mariyappan
45. வைட்வுட் (facebook) தமிழ் மொழியான பயன்படுத்தலால்
மேலாரை விண்ணுருவர், பரமசிவா மேலாரை விண்ணுருவர்
46. கலாசாரா பயன்படுத்தல்,
Ilango Kumaran Sivanadhan
47. E-Governance Initiatives in Tamil Nadu,
E. Iniya Nehru
48. மேலாரை மோவாரா மெரான் மெரான்
49. மேலாரை விண்ணுருவர், கலாசாரா பயன்படுத்தல்
50. தகவல் தமிழ் மொழியான குழு

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 …,
Rathiranee, Y
54. Effectiveness Of Computer Assisted Learning In Tamil Grammar, Siva Pillai
55. Using Augmented Reality as a Visual Aid to Teach Tamil Language in Malaysian Tamil Primary Schools, Kumaraguru
56. Bridging the Digital Divide for the Differently-Abled (Tamil) Students in Higher Education, J. Jerald Inico, T. Edwin Prabakaran
57. Computer-based and Paper-based Vocabulary Games for Tamil Learners, Paramasivam Muthusamy, Kanthimathi Letchumanan, Tan Bee Hoon
58. The Potential Role of YouTube in the process of 21st century approach to Classical Tamil Literature, M.Semmal
59. Role of Information Technology in Teaching and Learning of Classical Literature, Saravanan Raja
60. PDF வழியாக தமிழ் வாசிப்பதியா, காசி சிவாநி, மாதாய் பகாய்
Let Us 'Flip' the Classroom!,
Seetha Lakshmi  

Computer Assisted Tamil Teaching and Learning,
Louis Isack Kumar  

The Need for Integration Of ICT And Physical Infrastructure of School Development for The Growth of The Tamil.,
Saravanan Mariappan  

Sustainability of Teaching Information and Communication Technology (ICT) Skills in Malaysia Tamil Schools … ,
Gunasegaran Kandaswamy  

Mobile and Handheld | மொபியல் மற்றும் ஹங்க்லைட் நூல்கள்

Proactive Location Based Services In Tamil,
Karthikeyan K, Vinayak U S M P, Vijayalakshmi M  

Social Interaction and Content Sharing in Mobile P2P Environment using Tamil Language,
L.Libin Lougine, M.Vijayalakshmi  

Android Based Mobile Application in Tamil for Danger Assistance,

Mobile Based Rural Road Monitoring System,

Translating Tamil Speech (SL) as English Text Message (TL) in Android Mobile Phones,
Vikram Srinivasan  

Abraham Meshach Ponraj, T.Priyanka, R.R.Jeyshree & B.Padma  

Social Interaction and Content Sharing in Mobile P2P Environment using Tamil Language,
87. அவசயமில்லையாய் உட்கட்டுற்றியியல்லையாய் குடும்பான குடும்பானாய் தினசரியும் பின்னர் நியமனாக்கள், குடையர் ஒன்றுக்கொண்டோ

88. உட்கட்டுற்றியியல்லையாய் குடையர், ஒன்றுக்கொண்டோ

89. குடையர் குடும்பான குடும்பானது, உடல் நியமனது, பயாபாயில் - உடல் மின்மிலம், தினம் மின்மை

90. பயாபாயில் நியமனம், உடல் நியமனம் குடையர் குடும்பானாய் பயாபாயில் பயாபாயில் பயாபாயில் வழி குடையர் தமிழ் பை புத்தாண்டியாக

91. பயாபாயில் நியமனம் குடையர் குடும்பான குடும்பான, தினம் மின்மை கொண்டோ
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 kazanthai ‘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, thooothaaNa, - (suitable for a particular purpose; “thankam vaanka iihu cariyaaNa neram” ; “iihu oru nalla mudivu” ; “thirumaNnam ceiyya kalam kaNinthalu” ; “payaNnam ceiyya vimaaNam thooothaaNathu”).

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 (thuangu 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. *pati* – 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 kathial*₁ vetsa *kathi*² *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. *Veelai*¹ va*Nnankuvathu yen veelai*². ‘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, thozu, thuthi – different forms of worship (iRaivaNai vaNnanka, vizunthu kumpitu, thiruvatikaLaith thozu, iRai naamaththaith thuthi)

Noun:

S2 – veel, vil, thiruneeRu, kunkumam, thirisuulam, canku, cakkaram, karutaN, aathi seeshaN, cakraayutham – 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, puthhakam, 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, \textit{avaN pati thaaNntiNaaN}, the verb \textit{thaaNntu} is in the synset S4 and \textit{pati} is both in S4 and S6. A relationship exists between only S4 and S6. Therefore \textit{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

- Christopher D.Manning, Hinrich Schutze, Foundations of Statistical Natural Language Processing, MIT Press, England
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 பகுதியாக சாப்பிட்டன்
avan kaal pakatiyaic 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.

ெவளையம் மருண்டு 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 is 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 பாதி/paTi, குதி/kuTi, மடம்/maTam, etc, verbs such as பிடி/piTi, அறை/aRai, முதி/muTi, குது/cuTu, etc and the adjectives such as விரிதா/virinta, முற்றா/kuRaina, வெள்ளா/veLutta, கரு/kaRutta, etc. There are tests for establishing lexical ambiguity. One test is that for the word காடினமான/kaTinamaana there are two opposite words, மெதுவான/metuvaana and எளிதான/eLitaana. Consider the following example,

காடினமானம் மித்தாய்க காடிக்கும்தியாங்கு
kaTinamaana mITtaayak kaTikka mutiyaatu
‘You cannot bite hard set’

காடினமானம் க்லூக்குப் பொருள் குராயால்கு
kaTinamaana collukkup poruL kuRa 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 பாதி and the verb பாதி have same spelling/pronunciation they are two different words. They are examples of homophones. One may wonder whether the noun காதி and the verb காதி are examples for homonyms or not. Doubt may arise whether the word mutal in மணவன்/mutal maaNavan and அஞ்சு அஞ்சு/aindtu 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 polysemy.

1.1 Category ambiguity

Category ambiguity is the most straightforward 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 பச்சை ‘green’ can be both noun as well as adjective. Similarly the word கைத்து can be both verb as well as an adverb. காதி could be both verb as well as noun. As shown in the following example பாதி can be noun as well as a verb denoting two different meanings and thus shows ambiguity.

அவன் தான்கவடிவங்கள் நீண்டுக்கொண்டு அருள்மிகன்.
avan tantai avaniTam nii nanRaakap paTi eRu kuRinaar
‘His father asked to study well’

Avan maatip paTi vaziyaaka meeleee eeRinaan.
‘He climbed up through stair case’

The words like மீலை and கிளீஸ்சு could be adverbs and postpositions.

அவன் மீளை மேற்கொண்டு.
avan meeelee irukkiRaan. (noun)
He is at the top

அவன் மீலை செய்யாக்கு.
avan meeelee cenRaan (adverb)
He went up

அவன் மீனை குறித்து கொண்டாக்கு.
Avan meecai meeelee niRkiRaan. (postposition)

Category ambiguities can be often be resolved by morphological inflection. For example என் in என்/aTikkiRaan ‘he is beating’ is a verb and ஒலிப்/aT in ஒலிப்/aT ‘written’ ஆனால் நாணாய் வண்ணக்கான/avanai 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 இன்று/irundtu ‘from’, பார்க்கு/paRRi ‘about’, குறித்து/kuRittu ‘about’, முடிவு/oTTi ‘about’, தான்/koNTu ‘by (means of)’, வண்ணக்கான/kaiRti ‘by (means of)’, பிள்ளை/cuRRu ‘around’, செய்யாக்கு/ndookki ‘towards’, முடிவு/muntti ‘before’, காப்பு/xiaTa ‘than’, and க்கு/kuaTu ‘along with’ are the inflected forms of the verb இரு ‘be’, பார்க்கு/paRRu ‘catch’, வண்ணக்கான/kuRi ‘aim’, முடிவு/oTTu ‘stick’, செய்து/koL ‘have’, வண்ணக்கான/luTTu ‘go around’, செய்து/ndookku ‘look at’, முடிவு ‘over take’, வண்ணக்கான/luTTu ‘leave’, and க்கு ‘assemble’ respectively.
1a. அவன் விட்டிலி வேண்டும் கிரிதமக்கியர். (பாலாமோ)
avan vittilirundu veLiyeRinaan
‘He went out from the house’

1b. அவன் விட்டிலி வேண்டும் கிரிதமக்கியர் (அசீசர்கரீதி பந்தம்)
avan vittil irundtu vandtaan
‘He was in the house (habitually/continuously)’

2a. அவன் அவள் பெரும் பெளிசுந்தர். (பாலாமோ)
avan avaaLaip paRRi peecinaan
‘He talked about her’

2b. அவன் அவள் காயம் பெரும் பெளிசுந்தர் (அசீசர்கரீதி பந்தம்)
avan avaaL kaiayip paRRi muttamiTTaan
‘He catch hold of her hand and kissed it’

3a. அவன் அவள் குரித்து பெளிசுந்தர். (பாலாமோ)
avan avaaLaik kuRittu peecinaan
‘He talked about her’

3b. அவன் அவள் குண்டித்து பெரும் பெளிசுந்தர் (அசீசர்கரீதி பந்தம்)
avan avaaL colvataik kuRittu vandtaan
‘He was noting down what she was telling’

4a. அவன் பொய்த்து பொருள் பெளிசுந்தர் (அசீசர்கரீதி பந்தம்)
avan poosTar oTTi pizaikkinRaan
‘He eke his livelihood by pasting posters’

4b. அவன் பொய்த்து பொருள் பொருள் பெளிசுந்தர். (அசீசர்கரீதி பந்தம்)
avan andta taippai oTTi peecinaan
‘He talked about that title’

5a. அவன் காத்து கோண்டு அதற்போது பெளிசுந்தர். (பாலாமோ)
avan katti koNTu atai veTTinaan
‘He cut it with a knife’

5b. அவன் காத்து கோண்டு அதற்போது கிரிதமக்கியர். (அசீசர்கரீதி பந்தம்)
avan pencilaic ciivic koNTu peecinaan
‘He was speaking while sharpening the pencil’

6a. அவன் காத்து வைட்டு பழம் வேண்டும் பெளிசுந்தர். (பாலாமோ)
avan katti vaittup pazam veTTinaan
‘He cut the fruit with the knife’

6b. அவன் பாணம் வைட்டு பழம் வேண்டும் பெளிசுந்தர். (அசீசர்கரீதி பந்தம்)
avan paNam vaittuk koNTu cuutaaTinaan
‘He gambled by keeping the money at hand’

7a. அவன் விட்டில் குரித்து மராக்கு நடிக்கிரிதர். (பாலாமோ)
avan vittiac cuRRi marangkaL ndiRkinRana
‘The house is surrounded by the tree’
avan avaliyee cuRRi varukinRaan
‘He is going after her’

avan avalai ndookki ndaTandtaan
He went towards her

avan mukattai ndookkic cirittaan
‘He smiled looking after her face’

avan avalukku mundti angku vandtaan
‘He came there before her’

avan avalai mundti ndaTandtukoNTirundtaan
‘He is walking overtaking her’

avan avalai viTa ndallavan
‘He is better than her’

avan avalai viTa virumpavillai
‘He does want to leave her’

avan avarkaLuTan kuuTa virumpinaan
‘He wanted to gather together with them’

The word எனRu which is the inflected from the verb en ‘say’, show two different grammatical functions thus showing ambiguity.

avan ndallavan enRu ndinaitteen
avan tiTiir enRu vantaan

The word எனRu is having homographic relation with the inflected verbal form எனRu.

avan enRu varukiRaan
‘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 + அன்றி/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 (vantatu 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 அவ்வைட/aTTai shows homography; அவ்வைட/aTTai can denote ‘leach’ as well as ‘binding’.

அவன புத்தகத்தில் அவ்வைடைய் கிழித்து எறிதான்.
Avan puttakattin aTTaiyaik kizittu eRintaan
‘He tore away the binding’

அவன அவ்வைடைய் கொண்டான.
Avan aTTaiyaik konRaan
‘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 form 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.

அவன கடைல் தின்று மகிழ்ந்தான.
avan kaTalai tinRu makizntaan
‘He enjoyed eating pea nut’

அவன கடைல் கண்டு மகிழ்ந்தான.
avan kaTalai kaNTu makizntaan
‘He enjoyed seeing the sea’

In the first sentence the noun கடைல்காTalai denotes ‘pea nut’ and in the second case கடைல்காTalai has to be analysed as கடய்காTal + குளை (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, there by showing homogrophy.

அவன (துணி) எய்யதான.
avan (tuNi) ndeytaan.
‘He weaved (cloth)’
In the first sentence the word எந்தைதான்/ndeytaan has to be interpreted after analyzing it into என் + நேய+ட (past tense)+அண (third person masculine singular) and in the second sentence neytaan has to be interpreted as என்/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 polysemy

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 கிைல்/kili ‘branch’ may denote branch of a tree as well as a branch of a bank. நட்/ndaTa can denote the action of walking as well as happening or functioning of something.

அவன் தின்மும் காலையில் பால்லிக்கு நட்டு செல்கிறான் .
avan tinamum kaalaiyil paLLikku ndaTantu cellinRaan
He goes to school daily by walking’

அந்த கிளைல் காலை பால்லிக்கு நட்டு செல்கிறான் .
Anta niRuvanam nanRaaka ndaTantukoNTirukkinRatu
‘That organization is functioning well’

அந்த ஂத்திரஆல் கிளைல் காலை பால்லிக்கு நட்டு செல்கிறான் .
Anta tiyeeTTaril cinimaa ndaTantukoNTirukkinRatu
‘A cinema is running in the theatre’

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

அவன் விரைவாகேே பால்லிக்கு
Avan viraivaaka ooTukiRaan
‘He is running fast’

டான் வழியாகக் காலை பால்லிக்கு
Tanjaavuur vaziyaakak kaaviriyaaRu ooTukiRatu
‘The river Kaviri runs through Thanjavur’

கான்கன் காலையில் உண்
avan tan kaNkaLai muuTinaan
‘He closed his eyes’

டெட்றாய்க்கு உண்டு காலையில் உண்
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 வைய்/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.
uur cirittatu
‘The people (of the village) laughed’

Here in this sentence ஊ᾽ /uur ‘people’ is used as a metonymic extension of ஊ᾽ /uur ‘village’.

In the following sentence the word தோட்ட/keeL denotes both the perception through ears as well as ‘asking’.

ராைத ராஜா காதா’ /raatai raajaa kaataal keeTTatai avaniTam kuuRinaaL
‘Radha told him what Raja has heard’

This sentence is ambiguous giving at least two interpretations.

ராைத ராஜா தாᾹகாதா’ /raatai raajaa tan kaataal keeTTatai avaniTam kuuRinaaL
‘Radha told him what Raja has heard’

ராைத ராஜா வினவியைத அவனிடᾂ /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 குUi/Ti ‘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 ஆன/TTai. 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 ஆன/TTai is used.

ோஹெக்கைதுத் ஆன/TTai கிழி/kizi
Puttakattin TTai kizintuTTatu
‘The binding of the book is torn’

ஊ᾽/uurndtu /uurndtu
TTai uurndtu celkinRatu
‘The leach is crawling’

The verbs like குUi/kizi will take the objects like ஆன/TTai which can be torn as their subjects and the verbs like ஊ᾽/uurndtu/cel takes objects like ஆன/TTai 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 give 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., "அஅத்/anta_DT viiT/uviTu/NN azakaaka/அᾸᾌ_ADV irukkiRatu/இᾞᾰᾐ_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 (மரᾱக῀/marangkaL > மரᾱ/maram, வஅதாᾹ/andtaan > வா/vaa);
- **chunking**, which consists of dividing a text in syntactically correlated parts (e.g., [அஅத்/anta viiT/uviTu]NP [அzychaa/azakaaka irukkiRatu]VP, respectively the noun phrase and the verb phrase of the example);
- **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 அவைட்/aTTai at least has two senses: அவைட்/aTTai ‘leach’ and அவைட்/aTTai ‘binding of a book.’ From the corpus the sentences with the target ambiguous word அவைட்/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.
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, நுடுல, விளங்கு can be disambiguated. The following table shows the bag of words associated with each senses of the ambiguous words.

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Word occurring together</th>
</tr>
</thead>
<tbody>
<tr>
<td>மாலை1</td>
<td>evening</td>
<td>கால, பக், மைற, நாள், மாது, மதியு, இரண்டு</td>
</tr>
<tr>
<td>மாலை2</td>
<td>garland</td>
<td>மண், மலை, மலைகள், மகள், மகள், மதி ஆகங்</td>
</tr>
<tr>
<td>நுடுல1</td>
<td>thread</td>
<td>இவ்விளை, கால, கிளை, குறை, மைற, நாள், மாது மற்றும் இரண்டு</td>
</tr>
<tr>
<td>நுடுல2</td>
<td>book</td>
<td>துறை, சாய்த்து, சாய்த்து, டைனமியன், மகள் மற்றும் மதியு ஆகங்</td>
</tr>
<tr>
<td>விளங்கு1</td>
<td>animal</td>
<td>மாட், மின்னை, நுடுல, விளங்கு, விளங்கு, விளங்கு</td>
</tr>
<tr>
<td>விளங்கு2</td>
<td>Hand-cuff</td>
<td>காணால், காணால், காணால், அங்கு, அங்கு மற்றும் அங்கு ஆகங்</td>
</tr>
</tbody>
</table>

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

- இந்தியா. கே. 2001 புத்தாண்டு முன்னோடி நூற்றாண்டுக் காலம் [Classification into semantic domains & Thesaurus].
- இந்தியா. கே. 2006. முன்னோடி டெல்லியின் நூற்றாண்டுக் காலம் [Classification into semantic domains & Thesaurus].


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 pōkaμūtiyātavarkalkakkā (ேபாகᾙᾊயாதவ᾽கOregon) means 'for the sake of those who cannot go', and consists of the following morphemes:

<table>
<thead>
<tr>
<th>pōka</th>
<th>muṭi</th>
<th>y</th>
<th>āta</th>
<th>var</th>
<th>kal</th>
<th>ukku</th>
<th>āka</th>
</tr>
</thead>
<tbody>
<tr>
<td>go</td>
<td>accomplish</td>
<td>word-joining letter</td>
<td>negation (impersonal)</td>
<td>nominalizer he/she who does</td>
<td>plural marker to for</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 "aas" 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. ஏட்டிந் (which period?) - நிற். Another one is for removing the pronoun prefixes, அ, இ and உ. Eg. அந்த் (that period).- நிற்

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

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.

<table>
<thead>
<tr>
<th>உடடு</th>
<th>உடு</th>
<th>உட்</th>
</tr>
</thead>
<tbody>
<tr>
<td>original word</td>
<td>suffix stripped</td>
<td>vallinam consonant removed</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>original word</td>
<td>suffix removed</td>
<td>end recoded</td>
</tr>
</tbody>
</table>

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 தி, ஏ, ஓ.

<table>
<thead>
<tr>
<th>கணாணா</th>
<th>கணாணா</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it Kannan</td>
<td>Kannan</td>
</tr>
</tbody>
</table>

#### 4.5.2 Conjunction suffix

This routine removed the suffix உட

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>Him and</td>
<td>Him</td>
</tr>
</tbody>
</table>

#### 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.

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>without him</td>
<td>Him</td>
</tr>
</tbody>
</table>

#### 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.).

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>with him</td>
<td>him</td>
</tr>
</tbody>
</table>

#### 4.5.5 Plural suffix

The plural suffix in Tamil is கை.

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>Tree</td>
</tr>
</tbody>
</table>

#### 4.5.6 Imperative suffixes

These are used to command a person.

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>show me</td>
<td>see</td>
</tr>
</tbody>
</table>

#### 4.5.7 Tense suffixes

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

<table>
<thead>
<tr>
<th>உடல்வாதி</th>
<th>உடல்வாதி</th>
</tr>
</thead>
<tbody>
<tr>
<td>leaving</td>
<td>leave</td>
</tr>
</tbody>
</table>

Apart from the standard suffixes the routine also removed கை 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 actullay 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 implementetation errors and it widely used now for developing stemming algorithms. Stemming algorithms for many languages like German, French, Turkish etc have been implementet 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 implemetor is now not bothered about implementening 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>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.9237</td>
</tr>
<tr>
<td>25th percentile</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>2</td>
</tr>
<tr>
<td>75th percentile</td>
<td>3</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.7383</td>
</tr>
<tr>
<td>25th percentile</td>
<td>1</td>
</tr>
<tr>
<td>Median</td>
<td>1</td>
</tr>
<tr>
<td>75th percentile</td>
<td>1</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.76</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>1.97</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
</tr>
<tr>
<td>25th percentile</td>
<td>2</td>
</tr>
<tr>
<td>Median</td>
<td>3</td>
</tr>
<tr>
<td>75th percentile</td>
<td>4</td>
</tr>
<tr>
<td>Maximum</td>
<td>18</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Lovins</th>
<th>Paice</th>
<th>Porter</th>
<th>S-removal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.72</td>
<td>1.98</td>
<td>1.16</td>
<td>0.03</td>
</tr>
<tr>
<td>Std. deviation</td>
<td>1.64</td>
<td>1.92</td>
<td>1.40</td>
<td>0.19</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25th percentile</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>75th percentile</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>10</td>
<td>13</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5 - Strength description statistics for Tamil stemmer

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Modified Hamming Distance</td>
<td>2.76</td>
</tr>
<tr>
<td>Median Modified Hamming Distance</td>
<td>3</td>
</tr>
<tr>
<td>Mean Characters Removed</td>
<td>2.4</td>
</tr>
<tr>
<td>Compression Factor</td>
<td>0.65</td>
</tr>
<tr>
<td>Mean Conflation Class Size</td>
<td>2.88</td>
</tr>
<tr>
<td>Word and Stem Different</td>
<td>86.53%</td>
</tr>
</tbody>
</table>
Table 6 - Strength description statistics of some popular stemming algorithms for English (Frakes and Fox [7])

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

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/
- http://en.wikipedia.org/Tamil_grammar
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, aïŋkuRu, patiRRuppattu, paripaaTai, kalittokai, akanaaRu* and *puRa* *nanaRu* and *pattup paaTTu* literature – *tirumurukaaRRuppaTai, porunaraaRRuppaTai, ciRapaaNaaRRuppaTai, perumpaanaaRRuppaTai, mullaippaaTTu, mutaaiRaaNTai, neTunaaTTai, kuRitippaaTTu, 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 uriccól, 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. Agheshtialingom, 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:

**SANGAM CORPUS AND CONCORDANCE**

Dr. A. KAMATCHI  
Assistant Professor  
CAS in Linguistics, Annamalai University
Selected Keyword: அணி - beauty
No of Occurances: 11

Online Sangam dictionary
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
அஆ
அஆப᾿
அரᾆ
அ-alist
அலᾱᾁ
அவினி
அறΆ
ஆᾁ
ஆதᾹ
ஆ᾿
இத῁
இราวல῾
இᾞ
இவᾶ
உழவ῾
உைள
ஊ῾
ஊᾱᾁ
ஊᾠ
எᾸைத
எΆ
ஓᾐ
கஞ῾
கயலா῾

**References**


Elayapermal, M. 1975. Grammar of AiŋkuRunuuRu with Index, University of Kerala, Trivandrum.

Eva Wilden, 2008. Word Index of NaRRiNai (vol.3) Tamil mann patippakam, Chennai.


Subramaniyam, S.V. 1972. Grammar of AkanaanuuRu with Index, University of Kerala, Trivandrum.

Subramoniyam, V.I. 1962. Index of puRanaanuuRu, 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 Shammugam 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 Pathinenkilkanakkku Literature**

1. Identifying what is a word? what is not a word? in Pathinenkilkanakkku 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 Pathinenkilkanakkku texts.

2. Assigning grammatical information to the head word is another issue in Pathinenkilkanakkku 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 Pathinenkilkanakkku is another issue.

Consider the following sentences.

1. nilaiyaa ena uNarntaar (Nalatiyar’ 182-3)
2. tan makan caanRoon enak keeTTa 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 Pathinenkilkanakkku Literature has different senses based on the context. For example in the following sentence ‘kanru aaga, ivaar aaga’ (Nalatiyar 279-1), the word ‘aaka’ functions as a role (in the sense of ‘as’) .But the following sentence the word ‘aaka’ functions as an adverb ‘ paayntu aruvi aatinoomaka “ (ainthinai:15-3).

**Conclusion:**

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

**Bibliography**

தமிழ் தகவல் உயர்நிலை விளக்கத்தின் நூற்றாண்டு வருடங்கள் வழியாக வாழ்வு தொடங்கப்பட்டது. அது எண்ணெய் வடிவாக வாழ்வின் போது தகவல்களும் பயன்படுத்தப்பட்டன. இது தமிழ் பேராசியர், வடிவான பொருள் பயன்படுத்துவதற்கான ஒரு வகையான பாதையாகும். தமிழ் வல்லியல் பாதையின் விளக்கம் பற்றிய ஒரு பாகம் இங்கிலாந்து மொழியில் விளக்கப்பட்டுள்ளது.

தமிழ் பேராசியர் பொறுப்புச் சாலை

தமிழ் பேராசியர் பிரபலித்து விளக்கத்தின் பாதையில் வாழ்வு தொடங்கப்பட்டது. இந்த பாதையில் வாழ்வு தொடங்கலும் பயன்படுத்துவதற்கான ஒரு வகையான பாதையாகும். தமிழ் பேராசியர் வாழ்வின் பின்னர் தகவல்கள் பயன்படுத்தப்பட்டன. இது தமிழ் பேராசியர், வடிவான பொருள் பயன்படுத்துவதற்கான ஒரு வகையான பாதையாகும். தமிழ் வல்லியல் பாதையின் விளக்கம் பற்றிய ஒரு பாகம் இங்கிலாந்து மொழியில் விளக்கப்பட்டுள்ளது.

தமிழ் பேராசியர் பொறுப்புச் சாலை

தமிழ் பேராசியர் பிரபலித்து விளக்கத்தின் பாதையில் வாழ்வு தொடங்கப்பட்டது. இந்த பாதையில் வாழ்வு தொடங்கலும் பயன்படுத்துவதற்கான ஒரு வகையான பாதையாகும். தமிழ் பேராசியர் வாழ்வின் பின்னர் தகவல்கள் பயன்படுத்தப்பட்டன. இது தமிழ் பேராசியர், வடிவான பொருள் பயன்படுத்துவதற்கான ஒரு வகையான பாதையாகும். தமிழ் வல்லியல் பாதையின் விளக்கம் பற்றிய ஒரு பாகம் இங்கிலாந்து மொழியில் விளக்கப்பட்டுள்ளது.
தமிழ் முக்கியிகளின் வருடாகத்தில் பல பேர் பொறுப்புகள் மற்றும் முன்னேற்றங்களை வடிவமைக்கிறார்கள. முக்கியமானவையாக எழுத்துத் தொடர்பில் அமையும் தொடர்புகளின் குறிப்பிட்டலை மட்டுமே தவறுகளின் கீழ்த்தோண்டு பணியாளர்களுக்கு எந்த தொடர்புகளும் கையாள முடியாது. மேலும் இதன் பொழுதைய முறைகளையும் குறிப்பிட்டலை மட்டுமே தவறுகளின் கீழ்த்தோண்டு பணியாளர்களுக்கு எந்த தொடர்புகளும் கையாள முடியாது.

தமிழ் நூற்றாண்டுகள் (Centuries) பற்றிய பல்வேறு படைகளை வடிவமைக்கிறது. இது பொறுப்புகளின் முன்னேற்றங்கள் மற்றும் பணியாளர்களின் பணியிடும் பன்னாட்டு பொறுப்புகளின் வடிவமைப்புகள் தவறுகள் பணியாளர்களுக்கு எந்த தொடர்புகளும் கையாள முடியாது.
எனவே, அந்தத் தொடர்பற்றி வெளிப்படுத்தப்பட்ட கணினி வலைப்பக்கம் ஆகியவை ஆகியவை வலைப்பக்கத்தில் பொறித்து வருகிறது. ஆகியவை வலைப்பக்கங்களின் நோக்கம் தவறானது (sort algorithms) என்பதால் தொடர்புடையது. ஆகியவை வலைப்பக்கங்களின் மூலம் பக்கத்தில் மேம்பாடு பெறுகிறது (Meta data) தொகுப்புச் செய்திகள், கணினியின் பக்கத்தில் பெறுகிறது. ஆகியவை வலைப்பக்கங்களின் மூலம் வலைப்பக்க விளக்கங்கள் மற்றும் வலைப்பக்கம் வலைப்பக்க விளக்கங்கள் மற்றும் பொறித்து வருகிறது.

தவறான குறிப்பிட்டிடத்தில் பல்வேறு

தவறான குறிப்பிட்டிடத்தில் பல்வேறு அறிவியல், கணினியியல் நடவடிக்கைச் சான்று முற்பாண்ட அறிவியல் விளக்கத்திற்காக வைக்கப்பட்டுள்ளது. கணினியியல் / கணினியியல் பல்வேறு கருத்துக்கள் (Spell check / grammar check) நேரமுறை. குறிப்பிட்டிடத் தவறான கருத்துக்களின் மேல்புறங்கள் அளிப்பாற்றல் முறையில் பல்வேறு கருத்துக்கள் நேரமுறை முறையில் பல்வேறு கருத்துக்கள் நேரமுறை. பல்வேறு கருத்துக்கள் பல்வேறு கருத்துக்களின் மேல்புறங்கள் அளிப்பாற்றல் முறையில் பல்வேறு கருத்துக்கள் நேரமுறை. பல்வேறு கருத்துக்கள் பல்வேறு கருத்துக்களின் மேல்புறங்கள் அளிப்பாற்றல் முறையில் பல்வேறு கருத்துக்கள் நேரமுறை.

முறை

முறை தான் வெளிப்படுத்தப்பட்ட கணினி வலைப்பக்கங்கள் மற்றும் பொறித்து வருகிறது. ஆகியவை வலைப்பக்கங்களின் நோக்கம் தவறானது (sort algorithms) என்பதால் தொடர்புடையது. ஆகியவை வலைப்பக்கங்களின் மூலம் பக்கத்தில் மேம்பாடு பெறுகிறது (Meta data) தொகுப்புச் செய்திகள், கணினியின் பக்கத்தில் பெறுகிறது. ஆகியவை வலைப்பக்கங்களின் மூலம் வலைப்பக்க விளக்கங்கள் மற்றும் வலைப்பக்கம் வலைப்பக்க விளக்கங்கள் மற்றும் பொறித்து வருகிறது.
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 ‘maramkal’. 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 kalriliyaic 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 meele 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 parts 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,

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>CATEGORY</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NOUN</td>
<td>Common (NC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proper (NP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal (NV)</td>
</tr>
<tr>
<td>2</td>
<td>PRONOUN</td>
<td>Pronominal (PPR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reflexive (PRF)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reciprocal (PRC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wh-pronoun (PWH)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Absolutive (DAB)</td>
</tr>
<tr>
<td>3</td>
<td>DEMONSTRATIVE</td>
<td>Wh-demonstrative (DWH)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjective (JL)</td>
</tr>
<tr>
<td>4</td>
<td>NOMINAL MODIFIER</td>
<td>Quantifier (JQ)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intensifier (JINT)</td>
</tr>
<tr>
<td>5</td>
<td>VERB</td>
<td>Manner (AMN)</td>
</tr>
<tr>
<td>6</td>
<td>ADVERB</td>
<td>Location (ALC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relative (LRL)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal (LV)</td>
</tr>
<tr>
<td>7</td>
<td>PARTICIPLE</td>
<td>Nominal (LN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conditional (LC)</td>
</tr>
<tr>
<td>8</td>
<td>POSPOSITION</td>
<td>Co-ordinating (CCD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subordinating (CSB)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interjection (CIN)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Dis)Agreement (CAGR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Confirmative (CON)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delimiting (CDLIM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Others (CX)</td>
</tr>
<tr>
<td>9</td>
<td>PARTICLES</td>
<td>Real (NUMR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Serial (NUMS)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Calendric (NUMC)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ordinal (NUMO)</td>
</tr>
<tr>
<td>10</td>
<td>NUMERAL</td>
<td>Foreign Word (RDIF)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Symbol (RDS)</td>
</tr>
<tr>
<td>11</td>
<td>REDUPLICATION</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>RESIDUAL</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>UNKNOWN</td>
<td>Unknown (UNK)</td>
</tr>
<tr>
<td>14</td>
<td>PUNCTUATION</td>
<td>Punctuation (PU)</td>
</tr>
</tbody>
</table>
Dhanalakshmi et al., developed an Amrita POS Tagset for Tamil consists of 32 tags.

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>POS Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>&lt;NN&gt;</td>
<td>NOUN</td>
</tr>
<tr>
<td>2.</td>
<td>&lt;NNC&gt;</td>
<td>COMPOUND NOUN</td>
</tr>
<tr>
<td>3.</td>
<td>&lt;NNP&gt;</td>
<td>PROPER NOUN</td>
</tr>
<tr>
<td>4.</td>
<td>&lt;NNPC&gt;</td>
<td>COMPOUND PROPER NOUN</td>
</tr>
<tr>
<td>5.</td>
<td>&lt;ORD&gt;</td>
<td>ORDINALS</td>
</tr>
<tr>
<td>6.</td>
<td>&lt;CRD&gt;</td>
<td>CARDINALS</td>
</tr>
<tr>
<td>7.</td>
<td>&lt;PRP&gt;</td>
<td>PRONOUN</td>
</tr>
<tr>
<td>8.</td>
<td>&lt;PRIN&gt;</td>
<td>PRONOUN INTROGATIVE</td>
</tr>
<tr>
<td>9.</td>
<td>&lt;PRID&gt;</td>
<td>PRONOUN INDEFINITE</td>
</tr>
<tr>
<td>10.</td>
<td>&lt;ADJ&gt;</td>
<td>ADJECTIVE</td>
</tr>
<tr>
<td>11.</td>
<td>&lt;ADV&gt;</td>
<td>ADVERB</td>
</tr>
<tr>
<td>12.</td>
<td>&lt;VNAJ&gt;</td>
<td>VERB NON FINITE ADJECTIVE</td>
</tr>
<tr>
<td>13.</td>
<td>&lt;VNAV&gt;</td>
<td>VERB NON FINITE ADVERB</td>
</tr>
<tr>
<td>14.</td>
<td>&lt;VBG&gt;</td>
<td>VERBAL GERUND</td>
</tr>
<tr>
<td>15.</td>
<td>&lt;VF&gt;</td>
<td>VERB FINITE</td>
</tr>
<tr>
<td>16.</td>
<td>&lt;VAX&gt;</td>
<td>VERB AUXILIARY</td>
</tr>
<tr>
<td>17.</td>
<td>&lt;VIN&gt;</td>
<td>VERB INFINITE</td>
</tr>
<tr>
<td>18.</td>
<td>&lt;CNJ&gt;</td>
<td>CONJUNCTION</td>
</tr>
<tr>
<td>19.</td>
<td>&lt;CVB&gt;</td>
<td>CONDITIONAL VERB</td>
</tr>
<tr>
<td>20.</td>
<td>&lt;QW&gt;</td>
<td>QUESTION WORD</td>
</tr>
<tr>
<td>21.</td>
<td>&lt;COM&gt;</td>
<td>COMPLEMENTIZER</td>
</tr>
<tr>
<td>22.</td>
<td>&lt;NNQ&gt;</td>
<td>QUANTITY NOUN</td>
</tr>
<tr>
<td>23.</td>
<td>&lt;QTF&gt;</td>
<td>QUANTIFIERS</td>
</tr>
<tr>
<td>24.</td>
<td>&lt;PPO&gt;</td>
<td>POSTPOSITIONS</td>
</tr>
<tr>
<td>25.</td>
<td>&lt;DET&gt;</td>
<td>DETERMINERS</td>
</tr>
<tr>
<td>26.</td>
<td>&lt;INT&gt;</td>
<td>INTENSIFIER</td>
</tr>
<tr>
<td>27.</td>
<td>&lt;ECH&gt;</td>
<td>ECHO WORDS</td>
</tr>
<tr>
<td>28.</td>
<td>&lt;EMP&gt;</td>
<td>EMPHASIS</td>
</tr>
<tr>
<td>29.</td>
<td>&lt;COMM&gt;</td>
<td>COMMA</td>
</tr>
<tr>
<td>30.</td>
<td>&lt;DOT&gt;</td>
<td>DOT</td>
</tr>
<tr>
<td>31.</td>
<td>&lt;QM&gt;</td>
<td>QUESTION MARKS</td>
</tr>
<tr>
<td>32.</td>
<td>&lt;RDW&gt;</td>
<td>REDUPLICATION WORDS</td>
</tr>
</tbody>
</table>
## Kongu Pos Tag-Set (more than 500 tags)

<table>
<thead>
<tr>
<th>POS Forms and Morphological Inflections</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS &amp; Others</td>
</tr>
<tr>
<td>Noun</td>
</tr>
<tr>
<td>Proper Noun</td>
</tr>
<tr>
<td>Participle Noun</td>
</tr>
<tr>
<td>Adjective Noun</td>
</tr>
<tr>
<td>Positive Tensed Verbal</td>
</tr>
<tr>
<td>Noun</td>
</tr>
<tr>
<td>Negative Tensed Verbal</td>
</tr>
<tr>
<td>Verb</td>
</tr>
<tr>
<td>Simple Verb</td>
</tr>
<tr>
<td>Transitive Verb</td>
</tr>
<tr>
<td>Intransitive Verb</td>
</tr>
<tr>
<td>Causative Verb</td>
</tr>
<tr>
<td>Infinitive Verb</td>
</tr>
<tr>
<td>Imperative Verb</td>
</tr>
<tr>
<td>Reportive Verb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morphological inflections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Singular</td>
</tr>
<tr>
<td>Plural</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Common</td>
</tr>
<tr>
<td>Oblique</td>
</tr>
<tr>
<td>Person</td>
</tr>
<tr>
<td>First</td>
</tr>
<tr>
<td>Second</td>
</tr>
<tr>
<td>Third</td>
</tr>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Singular</td>
</tr>
<tr>
<td>Plural</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Neutral</td>
</tr>
<tr>
<td>Common</td>
</tr>
</tbody>
</table>

Kongu POS Tag set has more than 500 tags and there are other tag sets also.
<table>
<thead>
<tr>
<th>Tense</th>
<th>Present</th>
<th>Past</th>
<th>Future</th>
<th>Passive</th>
<th>Honorific</th>
<th>Negative</th>
<th>Interrogative</th>
<th>Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adverb</strong></td>
<td>Simple Adverb</td>
<td>Tense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adjective</strong></td>
<td>Simple Adjective</td>
<td>Present</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participle</td>
<td>Adjective</td>
<td>Past</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Future</td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Simple Preposition</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun + cases</td>
<td>Accusative</td>
<td>Dative</td>
</tr>
<tr>
<td></td>
<td>Instrumental</td>
<td>Sociative</td>
</tr>
<tr>
<td></td>
<td>Locative</td>
<td>Ablative</td>
</tr>
<tr>
<td></td>
<td>Benefacive</td>
<td>Genetive</td>
</tr>
<tr>
<td></td>
<td>Vocative</td>
<td>Clitics</td>
</tr>
<tr>
<td></td>
<td>Selective</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conjunction</th>
<th>Simple Conjunction</th>
<th>Wh-words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinating</td>
<td>What</td>
<td></td>
</tr>
<tr>
<td>conjunction</td>
<td>Who</td>
<td>Whose</td>
</tr>
<tr>
<td></td>
<td>When</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participle</th>
<th>Where</th>
<th>Whom</th>
<th>Which</th>
<th>How</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Participle</td>
<td>Positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditional participle</td>
<td>Negative</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interjection</th>
<th>Simple Interjection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Others</td>
<td>Echo words</td>
<td>Same</td>
</tr>
<tr>
<td></td>
<td>Different</td>
<td></td>
</tr>
</tbody>
</table>

<p>| Determiner | |
| Quantifier | |
| Complementizer | |
| Ordinal | |
| Optative | |</p>
<table>
<thead>
<tr>
<th>Indian Languages – Indian Languages Machine Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IL - ILMT Tagset</strong></td>
</tr>
<tr>
<td>Noun NN</td>
</tr>
<tr>
<td>Pronouns PRP</td>
</tr>
<tr>
<td>Locative Nouns NST</td>
</tr>
<tr>
<td>Verbs Finite VBF</td>
</tr>
<tr>
<td>Verbs Auxiliary VAUX</td>
</tr>
<tr>
<td>Adjectives JJ</td>
</tr>
<tr>
<td>Conjunctions CC</td>
</tr>
<tr>
<td>Interjections INJ</td>
</tr>
<tr>
<td>Question Words QW</td>
</tr>
<tr>
<td>Echo Words ECH</td>
</tr>
<tr>
<td>Particles RP</td>
</tr>
<tr>
<td>Numbers Ordinals QO</td>
</tr>
<tr>
<td>Question Words QW</td>
</tr>
<tr>
<td>Symbols SYM</td>
</tr>
<tr>
<td>Noun NN</td>
</tr>
<tr>
<td>Pronouns PRP</td>
</tr>
<tr>
<td>Locative Nouns NST</td>
</tr>
<tr>
<td>Verbs Finite VBF</td>
</tr>
<tr>
<td>Verbs Auxiliary VAUX</td>
</tr>
<tr>
<td>Adjectives JJ</td>
</tr>
<tr>
<td>Conjunctions CC</td>
</tr>
<tr>
<td>Interjections INJ</td>
</tr>
<tr>
<td>Question Words QW</td>
</tr>
<tr>
<td>Echo Words ECH</td>
</tr>
<tr>
<td>Particles RP</td>
</tr>
<tr>
<td>Numbers Ordinals QO</td>
</tr>
<tr>
<td>Question Words QW</td>
</tr>
<tr>
<td>Symbols SYM</td>
</tr>
<tr>
<td>Tag</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>NP</td>
</tr>
<tr>
<td>NN</td>
</tr>
<tr>
<td>NNP</td>
</tr>
<tr>
<td>IN</td>
</tr>
<tr>
<td>INP</td>
</tr>
<tr>
<td>PN</td>
</tr>
<tr>
<td>PNP</td>
</tr>
<tr>
<td>VN</td>
</tr>
<tr>
<td>VNP</td>
</tr>
<tr>
<td>Pn</td>
</tr>
<tr>
<td>PhP</td>
</tr>
<tr>
<td>Nn</td>
</tr>
<tr>
<td>NnP</td>
</tr>
<tr>
<td>V</td>
</tr>
<tr>
<td>VP</td>
</tr>
<tr>
<td>Vinf</td>
</tr>
<tr>
<td>Vvp</td>
</tr>
<tr>
<td>Vrp</td>
</tr>
<tr>
<td>AV</td>
</tr>
<tr>
<td>FV</td>
</tr>
<tr>
<td>NFV</td>
</tr>
<tr>
<td>Adv</td>
</tr>
<tr>
<td>SP</td>
</tr>
<tr>
<td>SCC</td>
</tr>
<tr>
<td>Par</td>
</tr>
<tr>
<td>Adj</td>
</tr>
<tr>
<td>Iadj</td>
</tr>
<tr>
<td>Dadj</td>
</tr>
<tr>
<td>Inter</td>
</tr>
<tr>
<td>Int</td>
</tr>
<tr>
<td>CNum</td>
</tr>
<tr>
<td>Num</td>
</tr>
<tr>
<td>DT</td>
</tr>
<tr>
<td>PO</td>
</tr>
</tbody>
</table>
A Comparison of various tagsets:

<table>
<thead>
<tr>
<th>Category</th>
<th>LDC-IL</th>
<th>Amrita</th>
<th>Kongu</th>
<th>IL-ILMT</th>
<th>IIIT-tagset</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proper</td>
<td>Proper,</td>
<td>Participle,</td>
<td>Nouns</td>
<td>Pronoun</td>
<td>Pronoun, Pronoun Phrase</td>
</tr>
<tr>
<td></td>
<td>Quantity</td>
<td>Compound,</td>
<td>Positive Tensed,</td>
<td></td>
<td>Pronouns</td>
<td>Pronoun</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjective Noun,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proper,</td>
<td>Negative Tensed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantity</td>
<td>Verbal,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal,</td>
<td>Un-tensed,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verbal Noun</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pronoun</td>
<td>Pronominal, Reflexive, Reciprocal</td>
<td>Pronoun,</td>
<td>Pronoun,</td>
<td>Pronouns</td>
<td>Pronoun</td>
<td>Pronoun</td>
</tr>
<tr>
<td></td>
<td>Interrogative</td>
<td></td>
<td>Pronoun</td>
<td></td>
<td></td>
<td>Pronoun Phrase</td>
</tr>
<tr>
<td></td>
<td>WH-Pronoun, Absolutive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pronoun</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Indefinite</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demonstrative</td>
<td>Wh-demonstrative, Adjective</td>
<td>Adjective</td>
<td>---</td>
<td>---</td>
<td>Demonstrative</td>
<td>Adjectives</td>
</tr>
<tr>
<td>Nominal</td>
<td>Quantifiers, Intensifier, Determiners,</td>
<td>Quantifiers,</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>modifiers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>Non finite, Adverb: Manner</td>
<td>Non finite,</td>
<td>Verb finite,</td>
<td>Verb</td>
<td>Verb, verbal phrase, verb infinitive, Verb verbal participle, Verb relative participle, auxiliary verb, Finite verb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td>Adjective,</td>
<td>Verbs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non finite,</td>
<td>Auxiliary,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Relative</td>
<td>Adverb,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participle: Verbal</td>
<td>Verbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gerund,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Verbal</td>
<td>Verb finite,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nominal</td>
<td>Verb</td>
<td>Conditional</td>
<td>Verb</td>
<td>infinite,</td>
<td>Conditional</td>
<td>Verb</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postposition</td>
<td>Postposition</td>
<td>Postposition</td>
<td>Postposition</td>
<td>Postposition</td>
<td>Postposition</td>
<td></td>
</tr>
<tr>
<td>Particles</td>
<td>Coordinating,</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>Particles</td>
<td>Particles</td>
</tr>
<tr>
<td></td>
<td>Subordinating,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interjection,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Agreement,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Confirmative,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Delimiting,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numeral</td>
<td>Real,</td>
<td>Ordinals,</td>
<td>---</td>
<td>Numbers</td>
<td>---</td>
<td>Number,</td>
</tr>
<tr>
<td></td>
<td>Serial,</td>
<td>Cardinals</td>
<td></td>
<td>Ordinals</td>
<td></td>
<td>Character number,</td>
</tr>
<tr>
<td></td>
<td>Calendric,</td>
<td>Ordinal</td>
<td></td>
<td></td>
<td></td>
<td>Date, time</td>
</tr>
<tr>
<td>Adverb</td>
<td>Adverb</td>
<td>Simple Adverb</td>
<td>Adverb</td>
<td>Adverb</td>
<td>Adverb</td>
<td></td>
</tr>
<tr>
<td>Adjective</td>
<td>Adjective</td>
<td>Simple Adjective,</td>
<td>Adjective</td>
<td>Adjective</td>
<td>Adjective</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Participle</td>
<td></td>
<td></td>
<td></td>
<td>Adjective,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adjective</td>
<td></td>
<td></td>
<td></td>
<td>Interrogative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>adjective,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Demonstrative</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>adjective</td>
</tr>
<tr>
<td>Reduplication</td>
<td>Reduplication</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Residual</td>
<td>Foreign Word,</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>Residual</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Symbol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Punctuation</td>
<td>Punctuation</td>
<td>Question marks,</td>
<td>---</td>
<td>---</td>
<td>Punctuation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comma,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>--</td>
<td>Conjunction</td>
<td>Simple</td>
<td>--</td>
<td>Subordinate clause,</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>----</td>
<td>--------------</td>
<td>--------</td>
<td>----</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conjunction,</td>
<td></td>
<td>subordinate clause</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coordinating</td>
<td></td>
<td>conjunction phrase</td>
<td></td>
</tr>
<tr>
<td>Participle</td>
<td>--</td>
<td>---</td>
<td>Verbal,</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conditional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interjection</td>
<td>---</td>
<td>---</td>
<td>Simple</td>
<td>---</td>
<td>Interjection</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interjection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>---</td>
<td>---</td>
<td>Echo words,</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Determiner,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quantifier,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Complementizer,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ordinal,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Optative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>Quantifiers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>Echo words</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>Emphasis</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>---</td>
<td>Question word</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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. VERBIX\(^1\) 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 CREA\(^2\). Dr. V. S. Rajam has proposed Conjugation by Verb-Stem Ending\(^3\).

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 et.al., ] . 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:

![Diagram 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 word 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). Table 1: 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.

<table>
<thead>
<tr>
<th>VERB</th>
<th>Paradigms</th>
<th>Simple forms</th>
<th>Inflection forms</th>
<th>Auxiliary forms</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
<td>164</td>
<td>67</td>
<td></td>
<td>10988</td>
</tr>
</tbody>
</table>

Table 1: Paradigms & Inflections

<table>
<thead>
<tr>
<th></th>
<th>P-1</th>
<th>P-2</th>
<th>P-3</th>
<th>P-4</th>
<th>P-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLI-1</td>
<td>ththAn</td>
<td>wAn</td>
<td>inAn</td>
<td>thAn</td>
<td>Ran</td>
</tr>
<tr>
<td>MLI-2</td>
<td>ththAL</td>
<td>wAL</td>
<td>inAL</td>
<td>thAL</td>
<td>RAL</td>
</tr>
<tr>
<td>MLI-3</td>
<td>ththAr</td>
<td>wAr</td>
<td>inAr</td>
<td>thAr</td>
<td>RAr</td>
</tr>
<tr>
<td>MLI-4</td>
<td>kinRAAn</td>
<td>kinRAAn</td>
<td>kinRAAn</td>
<td>kinRAAn</td>
<td></td>
</tr>
<tr>
<td>MLI-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
References:

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 PROGRAMMING 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.

<table>
<thead>
<tr>
<th>Tamil Keyword</th>
<th>English Keyword</th>
</tr>
</thead>
<tbody>
<tr>
<td>நிர்பாக</td>
<td>function</td>
</tr>
<tr>
<td>தேத்</td>
<td>continue</td>
</tr>
<tr>
<td>பெத்</td>
<td>repeat</td>
</tr>
<tr>
<td>இ எச்</td>
<td>if</td>
</tr>
<tr>
<td>இ எச்</td>
<td>else</td>
</tr>
<tr>
<td>தேது</td>
<td>select</td>
</tr>
<tr>
<td>இைல</td>
<td>for</td>
</tr>
<tr>
<td>தேது</td>
<td>return</td>
</tr>
<tr>
<td>செத்</td>
<td>Case</td>
</tr>
<tr>
<td>எல்லை</td>
<td>Do</td>
</tr>
<tr>
<td>செத்</td>
<td>While</td>
</tr>
<tr>
<td>பெத்</td>
<td>Print</td>
</tr>
<tr>
<td>செத்</td>
<td>Until</td>
</tr>
<tr>
<td>பெத்</td>
<td>End</td>
</tr>
<tr>
<td>இைல</td>
<td>Show</td>
</tr>
<tr>
<td>இைல</td>
<td>Prompt</td>
</tr>
<tr>
<td>இைல</td>
<td>Number</td>
</tr>
</tbody>
</table>

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:

இனா – இைல இனா – இைல இனா – … – இைல

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:

<table>
<thead>
<tr>
<th>Tamil code with Tamil keywords</th>
<th>Tamil Code with English keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>நிர்பாக (அ, எது) (2, 41);</td>
<td>function செல்குகல் (அ, எது) (2, 41);</td>
</tr>
<tr>
<td>காவை (அ*எது);</td>
<td>show (அ*எது);</td>
</tr>
<tr>
<td>காவை (2, 41);</td>
<td>செல்குகல் (2, 41);</td>
</tr>
</tbody>
</table>

Example: Using loops

<table>
<thead>
<tr>
<th>Tamil code with Tamil keywords</th>
<th>Tamil Code with English keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>நிர்பாக உள்ளே ( )</td>
<td>function உள்ளே ( )</td>
</tr>
<tr>
<td>{</td>
<td>{</td>
</tr>
<tr>
<td>இ=0;</td>
<td>இ=0;</td>
</tr>
<tr>
<td>எது (ஹ=1;ஹ&lt;=10;ஹ++)</td>
<td>for (ஹ=1;ஹ&lt;=10;ஹ++)</td>
</tr>
<tr>
<td>{</td>
<td>{</td>
</tr>
<tr>
<td>இ=இ+ஹ;</td>
<td>இ=இ+ஹ;</td>
</tr>
<tr>
<td>}</td>
<td>}</td>
</tr>
<tr>
<td>தேது (இ);</td>
<td>தேது (இ);</td>
</tr>
<tr>
<td>}</td>
<td>}</td>
</tr>
<tr>
<td>உள்ளே ();</td>
<td>உள்ளே ();</td>
</tr>
</tbody>
</table>
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

தமிழ் அளிய குறிப்பிட்டு சதுரக முக்கியம் உருளைவு?

நாட்டு ஆரம்பிக்காண

முனைவுச்சிக்கரம் மலபாரகாலமு, vasu@sas.upenn.edu

ஒன்று குறிப்பிட்டும் குறிப்பிட்டும் போரியிருக்கும் நுழைவ வசதி பூங்கை நிலைக்குறிப்பிட்டும். அது மேலும் காண்டுள்ளது ஒரு சிறுத்த முக்கியம் கீழரோல் இழுத்து அறியும் காலம். இது குறிப்பிட்டும் குறிப்பிட்டும் நேரம் என சுருக்கிய புகழ்பூர்த்திகளின் முக்கியம் கீழரோல் இழுத்து அறியும் சிறுத்த முக்கியம். இதை செய்யும் பொருள்களே முக்கியம் கீழரோல் இழுத்து அறியும் வழியாகும்.

http://www.thetamillanguage.com/tamilnlp/tagit.html

நாட்டு விளக்கம் செய்யும் கணினி வழியாக காண்க (set theory)

அமையும் விளக்கம் செய்யும் கணினி வழியாக காண்க (logic rules)

வை விளக்கம் செய்யும் கணினி வழியாக காண்க (list form)

செய்யும் விளக்கம் செய்யும் கணினி வழியாக காண்க (set theory)

செய்யும் விளக்கம் செய்யும் கணினி வழியாக காண்க (list form)

செய்யும் விளக்கம் செய்யும் கணினி வழியாக காண்க (set theory)

செய்யும் விளக்கம் செய்யும் கணினி வழியாக காண்க (list form)

செய்யும் விளக்கம் செய்யும் கணினி வழியாக காண்க (set theory)
நாம் வாழ்வேற்றும் சாதாரணமாக நிறுவனாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம். நமது சாதாரணமாக உடல் மற்றும் ஒலியை தரும் மற்றும் குறிப்பிட்டோம்.
உைடைம்மு தாட்டுகளும் விளையாட்டுகளும். திக்குறைய விளையாட்டுகள் தின்றுச்சலர் உலகிலே. இது ஆண்டுதோறும் புரோஸ்டே தின்றுச்சலை குறிப்பிட்டுவதும் ஆண்டுதோறும் எடுத்துச்செல்வதும். இது இரண்டு முறையும் இலக்கிய கல்வெளிகளும் குறிப்பிட்டுவதும். முன்னோராலும் இன்றைய ஆண்டுவரை வெளியுள்ள குறுகிய முக்தி செயலானது செயல்.
சொல்லுமதியானது விளக்கத்தில் வந்துள்ளது. சொல்லுமதி, சொல்லுமதி, பாதுகாக்க குறிப்பிட்டும் பாராளை குறிப்பிட்டும் சொல்லுமதி பதிப்பு பாதுகாக்கும் சொல்லுமதி பதிப்பு சொல்லுமதி பதிப்பு. மேலும் குறுத்து குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் சொல்லுமதி பதிப்பு. மேலும் குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் சொல்லுமதி பதிப்பு. மேலும் குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் சொல்லுமதி பதிப்பு. மேலும் குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் சொல்லுமதி பதிப்பு. மேலும் குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை குறிப்பிட்டும் பாராளை
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 # ➔ -e#
CR2 - am / - a:m / - a:n / - um / -o:m etc. ➔ nasalized vowels followed by #
CR3 - C # ➔ Cu # (phonetically i)
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 Participle 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](image)

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

Karthikeyan, Elanchezhiyan, 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 soundtracks, television serial soundtracks, 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)

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

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 framework setup for search, ranking the lyrics for single keyword becomes effective with offline ranking.

Let Search be $S = \{kw_1, kw_2, \ldots, kw_n\}$ where $kw_i$ denotes keyword $i$.

Let Lyric objects be denoted by $L = \{l_1, l_2, \ldots, l_m\}$ where $l_j$ denotes an LOB $j$.

Let $Oc$ denote the occurrence of the keyword $Oc(kw_i, l_j)$ be an integer.

First position of the keyword $Fp(kw_i, l_j)$ be an integer.

4.1.2 Online Ranking:

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

Let $Ds$ be the distance between two keywords.

$Ds(kw_i, kw_j) = 0$, if $kw_i$ and $kw_j$ are not in the same stanza.

$(ABS(Oc(kw_i) - Oc(kw_j)) \div Swc$, if $kw_i$ and $kw_j$ are in the same stanza.

Note: $Swc$ is the stanza total word counts.

$POP(l_j)$ be the popularity of the lyric.

Number of user views is used to compute popularity.

$RAT(l_j)$ be the rating of the songs.

Average of user ratings is used to compute the rating of songs.

4.1.3 Application

We employed spot indices for Paadal[], 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](image)


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

- 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.
- Paadal Lyric Portal, [http://www.karky.in/labs/paadal](http://www.karky.in/labs/paadal), last accessed date 08th July 2013.
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.
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](image)

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

<table>
<thead>
<tr>
<th>HAND</th>
<th>RIGHT, LEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>FINGER</td>
<td>INDEX FINGER, MIDDLE FINGER, RING FINGER, THUMB FINGER, PINKY FINGER</td>
</tr>
<tr>
<td>SHAPE</td>
<td>CLOSED, EXTENDED</td>
</tr>
<tr>
<td>ORIENTATION</td>
<td>INWARD, OUTWARD, DIAGONAL</td>
</tr>
<tr>
<td>LOCATION</td>
<td>FACE, CHEST, ABDOMEN</td>
</tr>
<tr>
<td>MOVEMENT</td>
<td>STATIC, RIGHT SHOULDER TO LEFT SHOULDER, CIRCULAR MOTION – CLOCK WISE AND ANTI CLOCK WISE DIRECTION.</td>
</tr>
</tbody>
</table>

TABLE 2: Finger representation

<table>
<thead>
<tr>
<th>IF</th>
<th>INDEX FINGER</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>MIDDLE FINGER</td>
</tr>
<tr>
<td>RF</td>
<td>RING FINGER</td>
</tr>
<tr>
<td>PF</td>
<td>PINKY FINGER</td>
</tr>
<tr>
<td>TF</td>
<td>THUMB FINGER</td>
</tr>
<tr>
<td>E</td>
<td>EXTENDED</td>
</tr>
<tr>
<td>I</td>
<td>INWARD</td>
</tr>
<tr>
<td>O</td>
<td>OUTWARD</td>
</tr>
<tr>
<td>CH</td>
<td>CHEST</td>
</tr>
<tr>
<td>S</td>
<td>STATIC</td>
</tr>
<tr>
<td>CM</td>
<td>CURVED MOTION</td>
</tr>
</tbody>
</table>

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 இந்தியை இந்தியை 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 us extended other fingers are closed and pointing towards chest. Similarly for possessive pronoun உங்கள் 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_CM 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 outward 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 இந்தியை உங்கள், இந்தியை respectively.
Figure 3 shows the hand gestures for the word அறிய

Figure 4 shows the hand gestures for the word உங்களுக்கு

Figure 5 shows the hand gesture for word தெரியும்
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

- 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.
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

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, Elanchezhiyan, 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 பொலா (poala) பொண்டரா (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.
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.
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.
- 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.
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.


“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 LaaLaLa 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, Kilkkatuva 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, Kilkkatuva 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:

\[
\text{பறைவ (paRavai) and பᾲைச (pachai) rhyme in monai as they start with the same letter. அᾞ (Aruvi) and எெெு (virupp) rhyme in edhugai as they share the same second letter. யாெெு (yaakkai) and வாெெு (vaazhkai) rhyme in iyaibu as they share the same last letter.}
\]

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:

\[
\text{கலக்கைத மறᾸேதᾹ உற்கைத மறᾸேதᾹ}
\]

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:

\[
\text{கலக்கைத மறᾸேதᾹ உற்கைத மறᾸேதᾹ}
\]

Monai between the first and third word is polippu monai.

Examples:

\[
\text{கலக்கைத மறᾸேதᾹ உற்கைத மறᾸேதᾹ}
\]
\textit{OtRu monai} is the one which exists between the first and fourth word.

Examples:

\begin{verbatim}
nilavukku nilavu sugampeRa ninaithaen
\end{verbatim}

The \textit{kulai monai} is the one between the first, second and third words.

Examples:

\begin{verbatim}
uyiraie uyirae udambil siRandhahdu edhu vendRu thavithirandhaen
\end{verbatim}

Monai between the first, second and fourth words is \textit{kilkkatuvay monai}.

Examples:

\begin{verbatim}
ilavukku nilavu sugampeRa ninaithaen
\end{verbatim}

\textit{A merkatuvay monai} is that which occurs between the first, third and fourth words.

Examples:

\begin{verbatim}
arumpum inimiatyum adhanavazhi guNikoLum
\end{verbatim}

If all the four words within a line match, then it is a \textit{mutRu monai}.

Examples:

\begin{verbatim}
mannavaa mannavaa mannaadhi mannan allavaa
\end{verbatim}

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 \textit{varukka monai}.

Examples:

\begin{verbatim}
pudhaikkkindRa vidhaiyum muyaRsigoNdaad thaan
pooniyum kooda thaazh thiRakkum
\end{verbatim}

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

Examples:

\begin{verbatim}
parivaana naNban thandha
kanivaana thoaLgaL kaNdaen
\end{verbatim}

Monai between long vowels is called as \textit{netil monai}.

\begin{verbatim}
vaalipaththil mannadhan
leelaikaLil mannavan
\end{verbatim}
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*

```
thi nggaL enbadhu peNNaaga
sevvay koavai pazhamaaga
```

*Example for Melina monai*

```
yaa rukku jaanagi kaathirundhaaL
vaarukku jaanagi kaathirundhaaL
vaarukku jaanagi kaathirundhaaL
```

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.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Rhyme Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhyme-AA</td>
<td>Ati EMI</td>
</tr>
<tr>
<td>Rhyme-AAAA</td>
<td>Murru EMI</td>
</tr>
<tr>
<td>Rhyme-AABB</td>
<td>Inai EMI</td>
</tr>
<tr>
<td>Rhyme-ABAB</td>
<td>Polippu EMI</td>
</tr>
<tr>
<td>Rhyme-ABBA</td>
<td>Oruu EMI</td>
</tr>
<tr>
<td>Rhyme-AAAB</td>
<td>Kulai EMI</td>
</tr>
<tr>
<td>Rhyme-AABA</td>
<td>Kilkatuvay EMI</td>
</tr>
<tr>
<td>Rhyme-ABAA</td>
<td>Merkatuvay EMI</td>
</tr>
</tbody>
</table>

**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

<table>
<thead>
<tr>
<th>Rhyme Pattern</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal Family</td>
<td>Identical letters</td>
</tr>
<tr>
<td>Consonant Family</td>
<td>Varukka EMI</td>
</tr>
<tr>
<td>Vowel Family</td>
<td>Uyr EMI</td>
</tr>
<tr>
<td>Consonant Hardness Family</td>
<td>Valina, Melina, Idai-ina EMI</td>
</tr>
<tr>
<td>Meter Match</td>
<td>Netil EMI</td>
</tr>
</tbody>
</table>
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.

\begin{align*}
  \text{ninaivae} & \quad \text{ninaivae} \\
  \text{nenjoadu} & \quad \text{kalandhvidu} \\
  \text{nilavae} & \quad \text{nilavae} \\
  \text{viNNodu} & \quad \text{ka}
\end{align*}

Extract every first, second and last letter of each word. The first letters of first line are \text{ni}, \text{ni}, \text{ne}, \text{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 \text{nai}, \text{nai}, \text{nj}, \text{la} mapping to string under:

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

The last letters of letters of first line are \text{vE}, \text{vE}, \text{tu}, \text{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,

- \( t^l \) be the Number of letters to be compared;
- \( t^e \) be the Number of letters matched in equal family;
- \( t^c \) be the Number of letters matched in consonant family;
- \( t^v \) be the Number of letters matched in vowel family;
- \( t^h \) be the Number of letters matched in consonant hardness family;
- \( t^m \) be the Number of letters matched in meter match;
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

\[
\begin{align*}
EScore &= \left( \frac{t^e}{t^c} \right) \times w^e \\
CScore &= \left( \frac{(t^c - t^e)}{t^c} \right) \times w^c \\
VScore &= \left( \frac{(t^v - t^c)}{t^c} \right) \times w^v \\
HScore &= \left( \frac{(t^h - t^c)}{t^c} \right) \times w^h \\
MScore &= \left( \frac{(t^m - t^c)}{t^c} \right) \times w^m \\
Score &= \sum (EScore, CScore, VScore, HScore, MScore)
\end{align*}
\]

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,

\textit{ninaivE ninaivE nenjOt u kalandhuvitu}

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

\[
\begin{align*}
EScore (AABB) &= 0.5 \\
CScore (AABB) &= 0.225 \\
VScore (AABB) &= 0 \\
HScore (AABB) &= 0 \\
MScore (AABB) &= 0 \\
Score &= 0.725
\end{align*}
\]

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.

<table>
<thead>
<tr>
<th>Vowels</th>
<th>Characters</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Vowels</td>
<td>அ/a, இ/i, உ/u, எ/e, ஒ/o</td>
<td>0.25</td>
</tr>
<tr>
<td>Long Vowels</td>
<td>ஆ/aa, இ/ee, ஊ/oo, ஏ/ea, ஐ/ai, ஓ/o, ஔ/au</td>
<td>0.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soft Consonants</td>
</tr>
<tr>
<td>Medium Consonants</td>
</tr>
<tr>
<td>Transformations</td>
</tr>
</tbody>
</table>

97
<table>
<thead>
<tr>
<th>Rough</th>
<th>( \acute{a}/k, \acute{\phi}/ch, \acute{\phi}/th, \acute{\iota}/p )</th>
<th>0.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greater Rough</td>
<td>( \acute{\iota}/t, \acute{\phi}/R )</td>
<td>0.1</td>
</tr>
</tbody>
</table>

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 \( E_1, E_2, \ldots, E_n \) 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,

\[ \begin{align*}
 l & \text{ be the input lyric.} \\
 n & \text{ be the number words in } l \\
 W & \text{ be the set of words in } l \\
 E_i & \text{ be the } i^{th} \text{ time window.} \\
 E(w) & \text{ be the time window of the word } w. \\
 E(l) & \text{ denote the time window of the song.} \\
 F_s(w_i) & \text{ be Freshness score of a word } w_i \text{ in the lyric.} \\
 F_s(l) & \text{ be the Freshness score of a lyric } l
\end{align*} \]

**Freshness Score :**

\[
\text{begin}
\begin{align*}
\text{initialize } \text{ScoreSum} & := 0; \\
\text{for each } w_i \text{ in } W \\
\quad F_s(w_i) & := E(w_i) / E(l) \\
\quad \text{ScoreSum} & := \text{ScoreSum} + F_s(w_i)
\end{align*}
\text{end for}
\text{F_s = ScoreSum / n;}
\text{end}
\]

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.
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.

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.

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
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

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, Bangalore

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 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.

<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence recorded without intonation</td>
<td>2</td>
</tr>
<tr>
<td>Un-intonated recording modified by our algorithm</td>
<td>3.8</td>
</tr>
<tr>
<td>Sentence recorded from a human with intonation</td>
<td>5</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS output</td>
<td>1</td>
</tr>
<tr>
<td>TTS output modified by new algorithm</td>
<td>2.8</td>
</tr>
<tr>
<td>Sentence recorded with intonation</td>
<td>5</td>
</tr>
</tbody>
</table>

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 ± 5%, then the TTS is not processed at all; else, the TTS speech is processed as discussed.
<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS output</td>
<td>1</td>
</tr>
<tr>
<td>TTS output modified by our new algorithm</td>
<td>3.5</td>
</tr>
<tr>
<td>Sentence uttered with intonation</td>
<td>5</td>
</tr>
</tbody>
</table>

**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.**

<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS (emotion)</th>
<th>MOS (clarity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS output only with pitch modification (TOPM)</td>
<td>3</td>
<td>3.3</td>
</tr>
<tr>
<td>TOPM and increase of energy of the signal</td>
<td>3.6</td>
<td>3.6</td>
</tr>
<tr>
<td>TOPM and scaling of energy only in specific bands</td>
<td>3.5</td>
<td>3.8</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence recorded without intonation</td>
<td>1</td>
</tr>
<tr>
<td>Un-intonated recording modified by our algorithm</td>
<td>4</td>
</tr>
<tr>
<td>Sentence recorded with intonation</td>
<td>5</td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Type of speech</th>
<th>MOS (emotion)</th>
<th>MOS (clarity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTS output only with pitch modification (TOPM)</td>
<td>3.75</td>
<td>3.5</td>
</tr>
<tr>
<td>TOPM and increase of energy of the signal</td>
<td>4.25</td>
<td>3.75</td>
</tr>
<tr>
<td>TOPM and scaling of energy in specific bands</td>
<td>4.2</td>
<td>4</td>
</tr>
</tbody>
</table>

**Results and Discussion**

Prosody of interrogative and exclamatory sentences have been modeled. Modification of the pitch, 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**

- 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 hypernym/hyponomy 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


3.1 System Architecture

![Figure 1: Architecture of Conceptual based search systems for Tamil and English](image-url)
Our proposed system consists of two segments: 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 are 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 their structure.

UNL is a Universal networking language, follows the Interlingua approach which is modeled to represent semantic data extracted from natural languages 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 charanNyaa chennaikku chenRaal (சராயா செசாைனா செசாறாீ), chennaikku charanNyaa chenRaal (செசாைனா சராயா செசாறாீ). Hence in Tamil the order of Subject + Verb + Object does not follow any order and gives same meaning for all these sentences. But the semantic representation for all Tamil sentences and English sentences are same which is given in example 1.

![UNL graph representation for Tamil and English sentences](image)

The above example depicts Tamil and English sentences which have same meaning. Universal words represent the concepts and links represent the relation between the concepts.
In the above example, saranya, go, chennai are the concepts. agt, plt are the links associated with these 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 the 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 multishort, 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 “aaruu” 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 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.

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

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.

<table>
<thead>
<tr>
<th>Fields</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept-relation-concept</td>
<td>The relationship between the Concept (C1)-Relation (R)-Concept (C2) belongs to the document set D.</td>
</tr>
<tr>
<td>Concept</td>
<td>The list of concepts (Cs) occurred in the document set D.</td>
</tr>
<tr>
<td>Language identifier</td>
<td>It uses “ta”, “en” which is used to identify Tamil and English languages of the document set D.</td>
</tr>
<tr>
<td>Document Identifier</td>
<td>d1, d2, d3, ........dn corresponds to list of documents which contains both CRC indices and C indices</td>
</tr>
<tr>
<td>Sentence Identifier</td>
<td>S1, S2, S3, ........Sn corresponds to list of sentence identifier which contains both CRC indices and C indices.</td>
</tr>
<tr>
<td></td>
<td>This is used to know the position of the term/concept in the document.</td>
</tr>
<tr>
<td>Position weight</td>
<td>Concept Position weight</td>
</tr>
<tr>
<td>Frequency count</td>
<td>Frequency of occurrence of the concepts in the document</td>
</tr>
<tr>
<td>Part Of Speech Tagging</td>
<td>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.</td>
</tr>
<tr>
<td>Term words</td>
<td>It represents term words present in the document</td>
</tr>
</tbody>
</table>

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 documents.

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, Venkatasivaramasstry 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. {mruagank, dubey}, pb@cse.iitb.ac.in
- Using Cross-Language Information Retrieval Methods for Bilingual Search of the Web:: Sung J. Shim Seton Hall University shumsung@shu.edu
- Cross-Language Information Retrieval Via Dictionary-based and Statistics-based Methods:: Fatia 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.ac.jp
- 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.Umamaheswari, T V Geetha, Ranjani Parthasarathi & Madhan Karky
- CoRe – Concept Based Query Expansion:: Elanchezhhiyan 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

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 Minimally Supervised Approach for Detecting and Ranking Document Translation Pairs:: Kriste Krstovski and David A. Smith.


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 Makal, Tīṟkaḷ, 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.
   
   miṟ eḻi paratavar maku maku(naṟ 101:8)
   peṇmai yahtya makuṟṭī kīkīvīyum (Tol-col 650:2)
   maṟavaṅ naḷ ayār makuḷ nern public society
   (kali 102:30)

2) To be remove the kaḷ forms in the following cases, after removing the plural marker the remaining word gives the meaning full root word. In such cases the M.A. follows the rules-based approach
   Kalavaikaḷ, Kilāikaḷ
   Kalavai kal, Kilai + kaḷ
3) To remove the ōkaḷ forms in the following cases after removing the plurals add m in the root word and remove ō in the suffix word

<table>
<thead>
<tr>
<th>Root Word</th>
<th>Suffix Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iṭam</td>
<td>+m  ōkaḷ</td>
</tr>
<tr>
<td>kuṇṭam</td>
<td>+m  ōkaḷ</td>
</tr>
</tbody>
</table>

Ex.

āṭṭa iṭrōkaḷaṅk avāḷ ellōm (Tol 1154:2)
kuṇṭāṭrōkaḷip pārāṭham tōṭār vantiyāṛ kol (Kali 71:18)

4) To remove the ōkaḷ forms in the following cases after removing the plurals add l in the root word and remove ō in the suffix word

<table>
<thead>
<tr>
<th>Root Word</th>
<th>Suffix Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>coḷ</td>
<td>+l ōkaḷ</td>
</tr>
</tbody>
</table>

Ex.

tiruntu pāṭi koṭcoḷ yāṃ kēṭa (Kali 81:13)

5) To remove the ōkaḷ forms in the following cases after removing the plurals add l in the root word and remove ō in the suffix word

<table>
<thead>
<tr>
<th>Root Word</th>
<th>Suffix Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>nāḷ</td>
<td>+l ōkaḷ</td>
</tr>
</tbody>
</table>

Ex.

ōṇti iṭu iṭṭu varumēnīṛ vāṭ/nāḷkaḷ (Ainti 70-56:3)
maṭṭa/veṭē niṭṭ vāṭ/nāḷkaḷ (Cilampu 28:125)

Flow diagram

![Flow diagram](image.png)

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


- 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 (ேவி஘ுறும் உறும்) 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.
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 : இராம் கா”தியை கைட்சேந்தை வான்
Output : இராம் (Raman) + Entity
          கா”தியை (Karthik) + Entity
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) இராமᾹகா᾽횔ிᾀெச᾿ (Raman Karthik Sel)
2) இராமᾹகைhexdigest (Raman Kadai Sel)
3) இராமᾹὋᾤᾤவாᾤ (Raman Puthagam Vaangu)

3.3 Addition of Ε.downcaseSuffixes

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 ஆᾠக᾿, அஅழிᾀ, அஅைட᾿, நந淝త᾿, ஒᾤᾤ, உைடைம, 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 'உᾤளᾐ', 'இᾤᾸతᾐ' or 'இᾤᾤஅகிᾤறᾐ', then the suffix 'இ᾿' is added.
   ii) If the object is a place and the verb’s root is either 'வா', 'ெச᾿', or 'ேபா', then the suffix 'ᾁ' is added.
   iii) If the verb is either 'அைதᾐ', 'அைதᾐ' or 'அைதᾐ', then the suffix 'இ᾿' 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 'அைம' is added with the object.

3.4 Morphological Generation using Punarchi Vithigal (ெேவῂᾤᾠᾐ)
Example: இராம் + ஒ = இராமேன்

3) "இராம் + ஓ = இராமேன்"

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

Example: கை + ஐ = கைைக

4) “எ... மரம் தூக்கிய மூலம் திரும்பி”

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

Example: மரம் + ஒ = மரம் + ஒ + எ = மரைக

5) "உய் + பாகனம் பிரபலம், அதுக்கு உண்மையான பிரபலம் பிற்கு உண்மையான பிரபலம் என்னும்”

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

Example: வாக்க் + ஐ = வாக்கு + ஐ = வாக்கைக

6) Adding suffix ‘плод

* If there is a ‘எ... மூலம்’ at the end of the object, it becomes Object + ‘அ’

Example: மரம் + ஐ = மரம் + ஐ + எ = மரைக

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

Example : மரம் + ஐ = மரம் + ஐ + எ = மரைக

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

Example : மரம் + ஐ = மரம் + ஐ + எ = மரைக

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 (இராம் + ஓ) - ஒ, ஓ, ஓ, ஓ 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: இராம் + ஒ + கவ் (Tense specified as present tense)

Output of the Morph-Analyzer: இராம் +Entity +Verb
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


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 newborn 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](image)

### 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](image)
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.

For a name starting with நீ, with length 3, uniqueness factor 5 and gender as female some of the names generated were நீலதா, நீதிைக, நீகி῁, நீதினி & நீιறா.

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. சமᾱ (saman) & மᾢᾱைக (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

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” “Neervizhchi”. 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 “Neervzhi”. 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 experettes 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.
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


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], an 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 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-the-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) \(<\text{ேநாᾼ}>\) : 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) \(<\text{மᾞᾸᾐ}>\) : 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 அஜீரணΆ), the relevant documents are obtained by matching the elements in the \(<\text{ேநாᾼ}>\) tag. Eg. If \(<\text{ேநாᾼ}>அஜீரணΆ</\text{ேநாᾼ}>\) 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 தᾶணீά், the analysed words will be
tāṇā‘i < Noun & 100 >
Here  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 ’நாயறேபா்தாக்ஏலாகா்சிற்ம’, வா்நாயறேபா்தாக்ஏலாகா்falls into the category of disease and ஏலா் 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 <ம்> 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 (விᾱκ᾿) or dry cough (வற᾵ᾌஇᾞம᾿) to diabetes (ச᾽ᾰகைர or நீாிழிᾫேநாᾼ).

B. Sample Query:
The information extracted for a given query is depicted in the following snapshots.

For a given query - அஜீரணΆ, the relevant tagged document is retrieved (Fig.2).

![Fig. 2 Retrieve relevant document for the given query](image)

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](image)

The keywords are now extracted and tagged into their categories (Fig 4 & Fig.5).

![Fig 4. List of Morphemes](image)
The remedial instruction from the original document from the tag <மᾞᾸᾐ> is obtained (Fig. 6).

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).

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

<table>
<thead>
<tr>
<th>Type of word</th>
<th>Example</th>
<th>Tagged format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compound noun [noun + noun]</td>
<td>இப்பட்டுபுரிவல்</td>
<td>இப்பட்டு &lt; Noun &amp; 100 &gt; புரிவல் &lt; Noun &amp; 100 &gt; &lt;இப்பட்டு+புரிவல்&gt; count=2</td>
</tr>
<tr>
<td>Partly tagged words</td>
<td>புரிவல்</td>
<td>புரிவல் &lt; Noun &amp; 100 &gt;</td>
</tr>
<tr>
<td>Noun tagged as blended word</td>
<td>காட்னலாக்கள்</td>
<td>காட்னலாக்கள் &lt; Adjective &amp; 300 &gt; &amp; &lt;sandhi&gt; காட்னலாக்கள் &lt; Noun &amp; 100 &gt;</td>
</tr>
</tbody>
</table>
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. பாைல (Betel leaves), a common word is tagged <unknown> by the morphological analyser.

Another reason for incorrect extraction is due to the splitting of a single word into two words. For eg. ஆங்கியன் (Liquorice) is split into two separate nouns, ஆங்கிய and கியன் 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. ச῾கைர (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. ஆங்கியமு ஆங்கியமு ஆங்கியமு ஆங்கியமு பழமு பழமு பழமு பழமு (Fig) has two words but denotes the same fruit. However when analysing, they are tagged as two separate nouns. Though ஆங்கியமு denotes the fruit fig, பழமு here is a common noun. Hence it gives an ambiguous output. Similar error can be found in வெளி and வெளி ஆர்ைல் ஆர்ைல் ஆர்ைல் ஆர்ைல் and கமலா கமலா கமலா கமலா ஆர்ைல் ஆர்ைல் ஆர்ைல் ஆர்ைல். 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. நீ் (Water) and கொய் (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

தமிழ் விசைந்றுருக்கான இக்காலத்தில்: உதவி கொடுக்கும் திட்டங்கள்

தினசரி ப.டீ.வி.ஸ்கோர்க்கிள்

இறுதிக் குறிப்பிட்டு, தமிழ் இக்காலத்தில் வாழ்விடம்

தினசரி, சங்கத்தில்

c) கல்லாளங்கள்

கல்லாள பல்வேறு கல்விக்குறிப்பிட்டு புரட்சிகள் போன்ற பொருள்கள் கண்டுபிடிக்கப்பட்டன. ஆசிரியர்

பயன்பாட்டுருவக்கான சூழல்களின் கோரினாக முன்னேற்கப்பட்டது. முன்னேற்கலும் நிறுவனங்கள் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு நிறுவனங்கள் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.

உணவியல் பல்வேறு கல்விக்குறிப்பிட்டு இவ்விருப்புக்கான திட்டங்கள் எக்ளோடு கல்விக்குறிப்பிட்டு பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.

தமிழ் விசைந்றுருக்கான இக்காலத்தில்: உதவி கொடுக்கும் திட்டங்கள்

தினசரி ப.டீ.வி.ஸ்கோர்க்கிள்

இறுதிக் குறிப்பிட்டு, தமிழ் இக்காலத்தில் வாழ்விடம்

தினசரி, சங்கத்தில்

c) கல்லாளங்கள்

கல்லாள பல்வேறு கல்விக்குறிப்பிட்டு புரட்சிகள் போன்ற பொருள்கள் கண்டுபிடிக்கப்பட்டன. ஆசிரியர்

பயன்பாட்டுருவக்கான சூழல்களின் கோரினாக முன்னேற்கப்பட்டது. முன்னேற்கலும் நிறுவனங்கள் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.

உணவியல் பல்வேறு கல்விக்குறிப்பிட்டு இவ்விருப்புக்கான திட்டங்கள் எக்ளோடு கல்விக்குறிப்பிட்டு பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.

தமிழ் விசைந்றுருக்கான இக்காலத்தில்: உதவி கொடுக்கும் திட்டங்கள்

தினசரி ப.டீ.வி.ஸ்கோர்க்கிள்

இறுதிக் குறிப்பிட்டு, தமிழ் இக்காலத்தில் வாழ்விடம்

தினசரி, சங்கத்தில்

c) கல்லாளங்கள்

கல்லாள பல்வேறு கல்விக்குறிப்பிட்டு புரட்சிகள் போன்ற பொருள்கள் கண்டுபிடிக்கப்பட்டன. ஆசிரியர்

பயன்பாட்டுருவக்கான சூழல்களின் கோரினாக முன்னேற்கப்பட்டது. முன்னேற்கலும் நிறுவனங்கள் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.

உணவியல் பல்வேறு கல்விக்குறிப்பிட்டு இவ்விருப்புக்கான திட்டங்கள் எக்ளோடு கல்விக்குறிப்பிட்டு பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது. இது, நூறூடாக்கிடும் பல்வேறு கல்விக்குறிப்பிட்டு கோரினாக முன்னேற்கப்பட்டது.
Kavin waited for the train, but the train was late – இன்னொருமாதற்கும் வரலாறு. அவையும் காலங்களே எதிர்கோளுடன் (Google) வருகிறது. கவிதை எடுக்கதற்கும் வரலாறு. அவையும் காலங்களே வந்து – காலங்கள் வரலாறு எதிர்கோள காலங்களே எதிர்கோளுடன்.

136
137

2. இந்திக்குறிக்கும் பாகத்தில் அவர்களைச் செந்நாள் அவர்களைக் காணவும் அவர்கள்.

3. விளக்கம் என்று நீங்கியுள்ளது.

4. குழக்கம் என்று நீங்கியுள்ளது.

5. விளக்கம் என்று நீங்கியுள்ளது.
நூறு இன்று ஒளியுள்ள பகுதியில் பலவகையான விளக்கங்கள் குறிப்பிட்டுள்ளார். எனவே கூறும்படி அவர்களால் விளக்கங்கள் குறிப்பிட்டுள்ளார்.

மூலக்குறிகள்

இந்த பகுதியில் முதல் பகுதியில் பார்ப்பலாம் இத்தகைய விளக்கம் நேராக தரப்படுகின்றது. இந்தவகையான விளக்கங்கள் படிகுறிகளின் விளக்கங்களின் மூலம் வேளாண்மையிலுள்ள பல்வேறுபட்ட விளக்கங்களை அளிக்கிறது.

மேற்கொள்ளும் பகுதியில் பார்ப்பலாம் எங்கும் விளக்கங்கள் படிகுறிகளின் மூலம் வேளாண்மையிலுள்ள பல்வேறுபட்ட விளக்கங்களை அளிக்கிறது.

பாதை:

1. Translate.google.co.in
3. Translate.google.co.in
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

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 (அகரΆ, ஏறிய, உயி᾽ெமᾼ, வாிைச) 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 (赣῀ளி, 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 sign AA or kAl (◌ா, U+0BBE) and pure consonant RA (᾽, 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.

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 ISCII (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).

![Figure 1. Anusvara in Tamil Unicode Block](http://www.unicode.org/charts/PDF/U0B80.pdf)

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)](http://www.unicode.org/charts/PDF/U0B80.pdf)

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 have varga 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://gneesan.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.

**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 (கஷ, ஷஷ, ஷஷ, ஷஷ, ஷஷ, and ஷஷ). 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.

**Even when Tamil borrows a conjunct consonant such as the grantha ஷஷ (KSHA), it breaks it up into constituent parts in the course of time.** (See Figure 4). The only other conjunct consonant in Tamil grantha (அஷ, ஸஷ, ஸஷ, ஸஷ, ஸஷ, and ஸஷ) is treated like a symbol similar to Tamil OM (U+0BD0, ஐ) rather than as a conjunct consonant because most Tamils do not know about the consonant ஸ (U+0BB6) and are not aware that ஸ is a conjunct consonant (UD + ஷ).

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.
An interesting aspect of this is, in addition to symbols, double dots, tildes, commas, single, double and triple accents, that a puLLi can overlap or even go inside a consonant. To the designer and the readers, any diacritic on top of or inside the consonant seems to be just another variant of the puLLi. This leaves the choice of any new diacritic to be positioned on the sides or below a consonant, at best.

![Figure 5](image1.png)

**Figure 5.** Various ‘diacritics’ used as puLLi in modern Tamil

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.

![Figure 6](image2.png)

**Figure 6.** Nokia implementation of “Tamil” anusvara

**ORTHOGRAPHIC CHALLENGES**

Besides the software developers, orthography and typefaces are of interest to researchers interested in studying learning disabilities as well. Characters that have similar strokes and shapes make it difficult for readers to distinguish them from each other and impede reading. Similarly mirror characters such as b and d, p and q often cause difficulties with those who have dyslexia. Tamil does not have mirror characters. Its simplicity led Baker to comment that “Tamil script is particularly easy to read, in part because of its lack of ligatures. It might also be argued that its particular mix of angular and rounded characters makes them more easily distinguishable than is the case with most other scripts of the Indian subcontinent.” (Baker 1997:129).

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 ஊ and ஋. They read ஊர் as ஊ. The Tamil Nadu government standard on Tamil Unicode implementation refers to homographic confusion between the Tamil vowel sign AA or kAl (◌ா, U+0BBE) and pure consonant RA (◌ி, U0BB0+U+0BCD). Similar homographic confusion exists between AU length mark (◌உ, 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 (ẖ, ꞏ, Ꞙ) 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 ñ, ꔦ, ꔠ](image1.png)

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](image2.png)

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.
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 (ச).

The discriminability between AI and JA (ஐ, U+0B90 and ஜ, 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.

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, ஜ 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 Garamand typefaces as in: g (Franklin Gothic) and g (Garamand). 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, இ) with ஐ, ஒ, ஓ as optional members of the set and the following consonants (க, ங, ச, ட, ண, ப, ம, ய, ல, வ, ல, லை, லன, லஸ, லஜ). These characters don’t have ascenders or descendents. The pullis and vowel modifiers of these characters go into the ascender or descender spaces. The long vowels (ஆ, ஏ) and (ஞ, த, ந, ர, ரூ, ரூ, ரூ, ரூ) all have glyphs that rest on the base line with descendents that may not extend to the bottom line. Sometimes letters like ஞ, ந are shortened with no descendents 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 ஞ, த, ந, and இ 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 descendents. I have given below pictures with both designs for comparison.

Figure 13. Characters with and without descendents and shorter au marker

Figure 14. Vowels and AI character with and without descendents and shorter au marker

Figure 15. JA characters with and without descendents
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 (ṟ, ṉ, ḷ)
4. Use your best typeface designs to disambiguate AI and JA (ாி, ாீ)
5. Consider making the characters ாி and ாீ 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 disambigu ate
7. Consider making the character ா 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:


http://www.textbooksonline.tn.nic.in/Books/Std01/Std01-III-TamEng.pdf


Ganesan, Naga., L2/12-051, Proposal to change the glyph shape of Tamil Sign Anusvara (U+0B82)).


Maag, Dalton., 2012. “Type Design meets Tamil”, 14 June 2012,


Rakara mey series, Tamil Nadu state text book for Standard 1, Vol III, Page 44


Sharma, Shriramana., L2/12-018, Request to deprecate 0B82 TAMIL SIGN ANUSVARA


L2/10256, Proposal to encode characters for Extended Tamil


Tamil Nadu government’s clarification on the orthography of ṭ, ṭ, ṭ (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
பல்கலை மழ்க்கைத்தியூரில் பெற்று மாந்திகரை பம்பார்ப்பி:

சுபாசிக கல்லூரியால்

எச்ச. ப. சங்காரதாரணர், உ. எல்லாம்பிள்ளை

பன்முக்கிள்க் அறிவியல், மீட்ச பராமரிப்பு துறை, நூற்றாண்டு அறிவியல் பள்ளியம், பௌர்ணமது.

கல்விக் குறிப்பிட்டு

சிலவிதத்தில் கூறப்பட்டதும் அறிவியல் பார்வைகள் என்னுடைய (speech and optical character recognition) குறிப்பிட்டும் இலகு மாந்திகரை (language models) மூலப்பாடு அறிவியலப் பள்ளிகளின் ஒளியில் பைராணராக ஒன்றியாகும். இதற்கான இலகு மாந்திகரை உருவாக்கப்பட்டது விளக்கக் காரணமாக அன்றுடைய ஆய்வுச் சேவைகள் கையற்றுள்ளன. அதாவது அறிவியல் பைராணராகக் காணப்பட்டுள்ளது, எனினும் மற்றொன்று மாந்திகரை குழு என்று மாந்திகரை இழுத்துகழிக் குறிப்பிட்டுள்ளது. 

சிலவிதத்தில் புராணத் தொகுப்பு (sandhi rules) என்று தாங்கப்படும் மாந்திகரை (morphological richness) இலகு மாந்திகரையை கண்டுபிடிக்கும் வடிவம் (agglutination) என்று வகுக்கப்படும் இலகு மாந்திகரை குழு பையோர பைராண (Partially free word order) கவர்களுக்கு முன்பாக, எனினும் மேலும் (large corpus) ஆய்வுச் சேவைகளுக்கு முன்பாக பயணிக்கும் இன்று மாந்திகரை றேசு பையோரான தொடர்புகளுக்குப் பொருளாக மாந்திகரை என்று குறிப்பிட்டுள்ளது. அங்கு பயணிகள் எங்களுக்கு முன்பாக மாந்திகரை விளக்கக் குறிப்பிட்டுகொண்டுள்ளன. எனினும் எங்கு பயணிகள் எங்கோ குறிப்பிட்டுள்ளன. 

குறுத்தல்

குறுத்தலுக்கு பாதுகாப்பு குறிப்பிட்டு பிறகு இலகு மாந்திகரை நூற்றாண்டு (large vocabulary continuous speech recognition - LVCSR) என்று கூறிட்டுள்ள பையோராக கண்டுபிடிக்கப் படும் விளக்கக் குறிப்பிட்டுச் சேவைகள் மூலக்கூறிகள். இதற்கான குறிப்பிட்டுச் சேவைகள் என்று தாங்கப்படும் அறிவியல் பள்ளியங்களின் ஏற்காடு அறிவியலப் (word lattice), என்று கூறும் தொடர்பில் தவறு காண்பது அறிவியலத் தொடர்பையாகும். 

முற்பாடு விளக்கம் என்று கூறப்பட்டதும் குறிப்பிட்டுள்ளது விளக்கக் குறிப்பிட்டு அறிவியல் அறிவியல் வாத (எள்ளகம் மனிதன்) குறிப்பிட்டு குறிப்பிட்டு (root verb) என்று மாந்திகரையின் (derived forms) என்று குறிப்பிட்டு மூலக்கூறிகள் முன்னேற்றக் குறிப்பிட்டு நூற்றாண்டு என்று குறிப்பிட்டு பயணிகளுக்கு குறிப்பிட்டு விளக்கக் குறிப்பிட்டு [1]. அறிவியல் என்று குறிப்பிட்டு விளக்கக் குறிப்பிட்டு செய்யப்பட்டுள்ளது. எனினும் எங்கு குறிப்பிட்டு மாந்திகரையின் (morphological) என்று குறிப்பிட்டு எங்கு குறிப்பிட்டு பயணிகளுக்கு குறிப்பிட்டு விளக்கக் குறிப்பிட்டு [2]. குறிப்பிட்டு எங்கு குறிப்பிட்டு எங்கு குறிப்பிட்டு பயணிகளுக்கு குறிப்பிட்டு விளக்கக் குறிப்பிட்டு [3].

மாந்திகரையின் மாந்திகரை பாதுகாப்பு

LVCSR என்று கூறப்பட்டு பிறங்க் பையோராக கண்டுபிடிக்கப் படும் விளக்கக் குறிப்பிட்டு பயணிகளுக்கு குறிப்பிட்டு விளக்கக் குறிப்பிட்டு செய்யப்பட்டுள்ளது. எனினும் எங்கு குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக் குறிப்பிட்டு விளக்கக்

217
தனிெசாிகளி் களதவான கண்டுபிடிக்கவும் முடியும் 2021ல் ஆணத்திருப்பில் பதின்முறை கால் (dictation tasks) பதினால் தினவரை நாள் அவர்கெள் குறிப்பிட்டு மெருசாண்டுத்துக்கோள் குறுக்குக்கோள் அளக்கப்பட்டுள்ளன.

முடிக் கொள்ளும் லட்சமானம் அறக்கெண்கள் ஒன்றியப் பாதுகாக்கம், நூற்றுக்கோள் வைத்துக்கோள் போன்றவை மெருசாண்டிதழ் எண்களை விளக்கல்லாமல் குறிப்பிட்டுள்ளன, மேலும் சுருக்கங்களையும் கண்டுபிடிக்கவும் குறுக்குக்கோள் குறிப்பிட்டுள்ளன: நூறு பக்கம்.

<table>
<thead>
<tr>
<th>தொடர்ச்சியை மேம்படுத்தும் வற்றைக் குறிக்க வழக்கம்</th>
<th>நூற்றுக்கோள் வழக்கம்</th>
<th>ஒன்றியப் பாதுகாக்கம்</th>
<th>மேம்படுத்தும் வருவாய் எண்ணிக்கை (%)</th>
<th>வருவாய் எண்ணிக்கை (விளக்கப்பட்டு - %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>நூறுக்கோள் மேம்படுத்தும் லட்சமானம்</td>
<td>4179938</td>
<td>614184</td>
<td>35.8</td>
<td>85.3</td>
</tr>
<tr>
<td>நகரம்</td>
<td>1055330</td>
<td>222333</td>
<td>12.9</td>
<td>78.9</td>
</tr>
<tr>
<td>பகுதிகள்</td>
<td>769531</td>
<td>136931</td>
<td>7.9</td>
<td>82.2</td>
</tr>
<tr>
<td>முனையெண்கள்</td>
<td>831656</td>
<td>82384</td>
<td>4.8</td>
<td>90.0</td>
</tr>
<tr>
<td>பணியில்</td>
<td>59124</td>
<td>17051</td>
<td>0.99</td>
<td>71.2</td>
</tr>
<tr>
<td>செம்புகம்</td>
<td>189321</td>
<td>45738</td>
<td>2.7</td>
<td>75.8</td>
</tr>
<tr>
<td>செம்புகம்</td>
<td>878975</td>
<td>169277</td>
<td>9.8</td>
<td>80.7</td>
</tr>
<tr>
<td>முனையெண்கள்</td>
<td>6056139</td>
<td>408833</td>
<td>23.8</td>
<td>93.2</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>779432</td>
<td>103163</td>
<td>6.0</td>
<td>86.7</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>37757</td>
<td>13194</td>
<td>0.7</td>
<td>65.1</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>85183</td>
<td>23116</td>
<td>1.3</td>
<td>72.8</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>120157</td>
<td>17873</td>
<td>1.0</td>
<td>85.1</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>1309273</td>
<td>227952</td>
<td>13.3</td>
<td>82.6</td>
</tr>
<tr>
<td>இளாங்கள்</td>
<td>3779525</td>
<td>285040</td>
<td>16.6</td>
<td>92.5</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>1681028</td>
<td>241082</td>
<td>14.0</td>
<td>85.6</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>42754</td>
<td>12396</td>
<td>0.7</td>
<td>71.0</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>126546</td>
<td>23749</td>
<td>1.4</td>
<td>81.2</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>30105</td>
<td>17065</td>
<td>0.99</td>
<td>56.7</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>699317</td>
<td>166283</td>
<td>9.7</td>
<td>76.2</td>
</tr>
<tr>
<td>குழுவாளர்</td>
<td>22707568</td>
<td>1716256</td>
<td>--</td>
<td>92.5</td>
</tr>
</tbody>
</table>
படம் 1. சொல்லிகள் மற்றும் பொருள்களின் அவரிகள் காட்சியைக் கிடைத்திருக்கும் தொடர்பின் பரந்திகளில் - புது வாரம் படம்.

படம் 2. சொல்லிகள் மற்றும் பொருள்களின் அவரிகள் காட்சியைக் கிடைத்திருக்கும் தொடர்பின் பரந்திகளில் - புது வாரம் படம்.

அகற்றுங்கள் குறிப்பிட்டுள்ளன: திருமணம டலர் குறிப்பிட்டுள்ளன.

அகற்றுங்கள் மற்றும் பொருள்களின் அவரிகள் காட்சியைக் கிடைத்திருக்கும் "சமயம்" 169277 குடையில் 19 அல்லது வேலைவாய்வு. 300 அல்லது வேலைவாய்வு 171 குடையில் பேச்சியப்பாடுகளை உற்பத்திய வலையுடன்.

அகற்றுங்கள் 2: இந்த அவரிகளின் குறிப்பிட்டுள்ளன: திருமணம டலர் குறிப்பிட்டுள்ளன. வலையுடன் பேச்சியப்பாடு பேச்சியப்பாடு "சமயம்" 169277 குறிப்பிட்டுள்ளன. 171 குடையில் பேச்சியப்பாடுகளை உற்பத்திய வலையுடன்.

| பலையுடன் | குறிப்பிட்டுள்ளன | பேச்சியப்பாடு | பேச்சியப்பாடு
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>குறிப்பிட்டுள்ளன</td>
<td>103163</td>
<td>260853</td>
<td>11587</td>
</tr>
<tr>
<td>குறிப்பிட்டுள்ளன</td>
<td>222333</td>
<td>345710</td>
<td>45900</td>
</tr>
<tr>
<td>குறிப்பிட்டுள்ளன</td>
<td>82384</td>
<td>240884</td>
<td>10777</td>
</tr>
</tbody>
</table>
ஆலயசன 3-ஆவது அலகுமாக மேளவியல் முன்னெச்சரிக்கை அறிவியல், ஆலயசனுக்கு நோல் விளக்கத் தொடர்ந்து விளக்கையல் முன்னெச்சரிக்கை காண்பதாகச் சொல்லப்பட்டது. இது, முன்னெச்சரிக்கையில் உள்ள முன்னெச்சரிக்கை மற்றும் மேலும் முன்னெச்சரிக்கையைத் தொடர விளக்கையலை அறியலாம். உடன் வருகையில் மீண்டும் விளக்கையலை அறியலாம். ஐந்து தொடர் விளக்கையலை அறியலாம். ஐந்து தொடர் விளக்கையலை அறியலாம்.

<table>
<thead>
<tr>
<th>ஆலயசனுக்கும் தேர்வு நேரத்தில் பரிமாறும் தொடர் விளக்கையல் முன்னெச்சரிக்கைகள்</th>
<th>முன்னெச்சரிக்கை விளக்கையல் முன்னெச்சரிக்கைகள்</th>
</tr>
</thead>
<tbody>
<tr>
<td>பம்பு</td>
<td>169277</td>
</tr>
<tr>
<td>பெரும்பாளர்</td>
<td>103163</td>
</tr>
<tr>
<td>கலைநயம்</td>
<td>222333</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>82384</td>
</tr>
<tr>
<td>கலைநயம்</td>
<td>17065</td>
</tr>
<tr>
<td>ஆவ்வையான்</td>
<td>136931</td>
</tr>
<tr>
<td>தமிழ் வழக்கு</td>
<td>23749</td>
</tr>
<tr>
<td>தொடர்வழக்கு</td>
<td>17051</td>
</tr>
<tr>
<td>நூற்றாண்டு</td>
<td>45738</td>
</tr>
<tr>
<td>பொருள் வடிவம்</td>
<td>166283</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>23116</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>13194</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>12396</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>408833</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>17873</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>227952</td>
</tr>
<tr>
<td>பிரிவுச்சொல் வேகம்</td>
<td>614184</td>
</tr>
<tr>
<td>அரசீயத் தகுதி</td>
<td>285040</td>
</tr>
</tbody>
</table>

உள்ளிட்டு, மற்றும் பிரித்தல் வேகம் முன்னெச்சரிக்கை ஆலயத்தின் வேளத்தில் பரிமாறும் தொடர் விளக்கையல் முன்னெச்சரிக்கைகள். ஆலயசன் முன்னெச்சரிக்கை அறிவியல், ஆலயசன் வேளத்தில் பரிமாறும் தொடர் விளக்கையல் முன்னெச்சரிக்கை.
முல்லியியல் பராமரிப்பு அடிப்படை:

முல்லியியல் பராமரிப்பு அடிப்படையில் முழுதும் முழுமுயற்சி முழுமுயற்சியால் முழுமுயற்சி (paradigms) முழுமுயற்சி. இது பராமரிப்புச் செயல்பாடுகளுக்கு எண்ணிக்கையில் பராமரிப்பு 1970-ல் புதுமையாக முழுமுயற்சியால் 

<table>
<thead>
<tr>
<th>எண்ணிக்கை</th>
<th>பராமரிப்பு</th>
<th>எண்ணிக்கை</th>
<th>பராமரிப்பு</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>135486</td>
<td>5</td>
<td>41230</td>
</tr>
<tr>
<td>11</td>
<td>130425</td>
<td>18</td>
<td>30471</td>
</tr>
<tr>
<td>9</td>
<td>128475</td>
<td>4</td>
<td>23679</td>
</tr>
<tr>
<td>12</td>
<td>122741</td>
<td>19</td>
<td>21826</td>
</tr>
<tr>
<td>8</td>
<td>115506</td>
<td>20</td>
<td>14660</td>
</tr>
<tr>
<td>13</td>
<td>107806</td>
<td>21</td>
<td>10195</td>
</tr>
<tr>
<td>7</td>
<td>90559</td>
<td>3</td>
<td>8465</td>
</tr>
<tr>
<td>14</td>
<td>90164</td>
<td>22</td>
<td>6818</td>
</tr>
<tr>
<td>15</td>
<td>72450</td>
<td>23</td>
<td>4512</td>
</tr>
<tr>
<td>6</td>
<td>64978</td>
<td>24</td>
<td>3035</td>
</tr>
<tr>
<td>16</td>
<td>55613</td>
<td>25</td>
<td>1924</td>
</tr>
<tr>
<td>17</td>
<td>41860</td>
<td>2</td>
<td>1833</td>
</tr>
</tbody>
</table>
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 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. 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.

![XML sample](image)

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 sophisticated 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


A Multimodal Framework for the Recognition of Ancient Tamil Handwritten Characters in Palm Manuscript Using Boolean Bitmap Pattern of Image Zoning

E.K.Vellingiriraj, 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](image)

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](image)

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.

![Image cropping example](image)

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.

d) *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.

![Image resized example](image)

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.

![Image thickening example](image)

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.

![Image binarization example](image)

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
leaf manuscripts were written in hand and might be unreadable to different people and it is necessary to make them readable.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>02</td>
<td>03</td>
<td>04</td>
<td>05</td>
<td>06</td>
<td>07</td>
<td>08</td>
<td>09</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
</tr>
<tr>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig 4: The Tamil character sets

b) **Image Grouping:** Palm manuscripts were used in different centuries and the words appear in them are of different styles and strokes (Fig 5). Each individual script is stored in Boolean matrix.

![Image Grouping](image)

Fig 5. Tamil Scripts in Different Centuries

c) **Character Pattern Matching:** The stored original text Boolean matrix would be compared with the new preprocessing Tamil character using a range.

4. **Character Recognition:** Pattern matching identifies each similar character in the predefined Boolean matrix. Then, the matching Boolean matrix is converted into equal Unicode Tamil fonts.

4. **Character Modeling:**

A character model is a record of all the characters set that are of equal Boolean matrix. The combination of the two Boolean matrix is also equal to the single character. For example:

\[
\begin{array}{c}
+ \\Rightarrow \\
\end{array}
\]

The equivalent Boolean matrix is given below:
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 i s 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:

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.

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 quantifyzation 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.

![Statistical feature extraction techniques](image)

**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 Handwritten 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>
<thead>
<tr>
<th>S. No</th>
<th>Algorithm Used</th>
<th>Positives</th>
<th>Vital Analysis</th>
</tr>
</thead>
</table>
| 1     | Zoning         | • Tamil Numerals  
• Beneficial results  
• Good Sample sets  
• Not a time consuming process | • Less character sets  
• Considered only formal characters  
• Need more pre-processing steps.  
• Failed to remove unwanted pixels |
| 2     | Boundary Tracing | • 30 Characters  
• Not a time consuming process  
• Good accuracy | • 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 | • Success in formal characters  
• Good accuracy  
• Tamil Numerals | • Negative results will obtain from the abnormal and similar shaped characters  
• Feature selection is not sufficient |
| 4     | Determination of small Groups | • Not a time Consuming process | • 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 | • 11 characters selected  
• Good feature selection | • Limited number of samples  
• Need more Pre-processing steps |
<p>| 6     | Dimension | • Capable to provide good results | • Feature selection is not enough |</p>
<table>
<thead>
<tr>
<th>Method</th>
<th>Tamil Numerals and 7 characters</th>
<th>Lower number of pre-processing steps</th>
<th>Formal letters have been considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Reservoir</td>
<td>Supports formal characters</td>
<td>Good for only printed characters</td>
<td>Not effective for Handwritten characters</td>
</tr>
<tr>
<td>Fourier and wavelet transform</td>
<td>Good to image recognition process</td>
<td>Very less characters selection</td>
<td>Time consuming process</td>
</tr>
<tr>
<td>Projection based</td>
<td>Good Feature selection</td>
<td>Character selection is not clear</td>
<td>Improvement in accuracy</td>
</tr>
<tr>
<td>SIFT</td>
<td>20 characters are selected</td>
<td>Problem to recognize the abnormal and similar shaped characters</td>
<td>Time consuming process</td>
</tr>
</tbody>
</table>

**TABLE 2. YEAR WISE ANALYSIS OF THE STATISTICAL APPROACH**

<table>
<thead>
<tr>
<th>S.No</th>
<th>Year</th>
<th>Algorithm Name</th>
<th>Characters count and Characters</th>
<th>Samples</th>
<th>Accuracy</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2002</td>
<td>Zoning [10]</td>
<td>26 – Not Specified</td>
<td>1000</td>
<td>96.9 %</td>
<td>Characters have not been specified. Writers and character dependent. Character confusion (similar shaped)</td>
</tr>
<tr>
<td>2</td>
<td>2007</td>
<td>Zoning [4]</td>
<td>34 (அ,ஆ,இ,ஈ,உ,�,எ,ஏ,ஐ,ஒ,ஓ,ஃ,க,ங,ச,ஞ,ட,ண,த,ந,ப,ம,ய,ர,ல,வ,ள,ழ,ற,ன)</td>
<td>6048</td>
<td>82.04 %</td>
<td>Many Characters are yet to be considered. Writers and character dependent. Font size dependent. Style variation provides negative results.</td>
</tr>
<tr>
<td>3</td>
<td>2008</td>
<td>Zoning [6]</td>
<td>10 (Tamil Numerals (0-9) – 0, அ, ஆ, இ, ஈ, உ, ஊ, எ, ஏ, ஐ)</td>
<td>2000</td>
<td>94 %</td>
<td>Same problems (Specified in point 2)</td>
</tr>
<tr>
<td>4</td>
<td>2008</td>
<td>Water Reservoir [12]</td>
<td>-</td>
<td>-</td>
<td>94.01 %</td>
<td>No improvement from</td>
</tr>
</tbody>
</table>
Variation [5] - ஈ, ஒ, ஐ, எ, ஐ, ஒ, ஐ, ஒ, ஐ, ஒ
Dimensional method [14] - 8
SIFT [1] - 20
Vertical projection profile, word profile, background to ink transition [16] - 10 (40 from 4 document),
Fourier and wavelet Transform [13] -
Vertical projection profile, word profile, background to ink transition [15] - 10 (40 from 4 document),

<table>
<thead>
<tr>
<th>Year</th>
<th>Dimensional method</th>
<th>SIFT</th>
<th>Vertical projection profile, word profile, background to ink transition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>8</td>
<td>2000</td>
<td>80.75%</td>
</tr>
<tr>
<td>2012</td>
<td>-</td>
<td>-</td>
<td>98 %</td>
</tr>
</tbody>
</table>

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


Offline Tamil Handwritten Character Recognition

Jegadheeswaran J, Abirami S, Vijayalakshmi M

Department of Information Science and Technology, College of Engineering, Anna University, Chennai

jjegadhees@gmail.com  abirami_mr@yahoo.com  vijim@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 featureset. 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. Rajashekhararadhya 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](image)

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.

**Step 1:** Select an initial estimate for the variable $T$. (Initial estimate is the midpoint between the minimum and maximum intensity values in the image.)

**Step 2:** Segment the image using the threshold value $T$. This will produce two groups. They are $G_1$, consisting of all pixels with intensity values $> T$, and $G_2$, consisting of all pixels with intensity value $< T$.

**Step 3:** Compute the average intensity values $\mu_1$ and $\mu_2$ for the pixels in regions $G_1$ and $G_2$.

**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.

**Step 1:** Sweep the image with one of the structuring elements shown in Figure 2.

**Step 2:** 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.

**Step 3:** 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.

**Step 4:** Repeat processes 1-3 until all structuring elements have passed over the image (each element takes the output of the previous one as input).

**Step 5:** Repeat processes 1-4 until the image doesn’t change anymore.

![Figure 2 Structuring elements](image)

E. Segmentation

Segmentation is a process of separating lines, words, and even characters of a handwritten 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 quiet 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].

153
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 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

<table>
<thead>
<tr>
<th>Code</th>
<th>Next row</th>
<th>Next column</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Current pixel (I, J)</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Current pixel at coordinate (i, j)</th>
<th>Code</th>
<th>Next row</th>
<th>Next column</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>I</td>
<td>J+1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>I-1</td>
<td>J+1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I-1</td>
<td>J</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I-1</td>
<td>J-1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>I</td>
<td>J-1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>I+1</td>
<td>J-1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>I+1</td>
<td>J</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>I+1</td>
<td>J+1</td>
</tr>
</tbody>
</table>

finally all the features have been extracted and given to the classifier for recognition of characters.
1. 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>
<thead>
<tr>
<th>CHARACTER</th>
<th>TOTALLY TRAINED</th>
<th>TEST SAMPLE</th>
<th>CORRECTLY CLASSIFIED</th>
<th>MIS-CLASSIFICATION</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>12</td>
<td>9</td>
<td>3</td>
<td>85%</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>10</td>
<td>8</td>
<td>2</td>
<td>80%</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>75%</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>67%</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>75%</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>11</td>
<td>8</td>
<td>3</td>
<td>72%</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>67%</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>71%</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>66%</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>66%</td>
</tr>
<tr>
<td>0</td>
<td>15</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>80%</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>66%</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>75%</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>8</td>
<td>7</td>
<td>1</td>
<td>87%</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>71%</td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>9</td>
<td>7</td>
<td>2</td>
<td>77%</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>66%</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>60%</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>1</td>
<td>83%</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>71%</td>
</tr>
</tbody>
</table>

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

- 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.
Tamil Script Recognition using Smart Phone

Rex Aantonny. P, Iniya Nehru .E

*Software Executive, Cadgraf Digitals Pvt Ltd, Chennai, India, joyemperor@gmail.com
*National Informatics Centre, Chennai, India, nehru@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 through 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.
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).

**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.

The figure 5 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 «.

**Figure 4. Example of 6th Century Tamil Letter «**

**Figure 5. 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
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 (ormodes) 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.

The regression state of neurons is shown in figure 7. 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.
the training state of the neurons shown in the figure 8. 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 calculated as following:

\[
\text{Performance} = \frac{\text{Number of correctly recognized characters}}{\text{Total number of characters}}
\]

<table>
<thead>
<tr>
<th>Document type</th>
<th>Number of characters</th>
<th>Identification rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dark lighting</td>
<td>25</td>
<td>67.5%</td>
</tr>
<tr>
<td>Small fonts</td>
<td>75</td>
<td>80.1%</td>
</tr>
<tr>
<td>Wide letters</td>
<td>35</td>
<td>75.1%</td>
</tr>
</tbody>
</table>

Table 1: Performance comparison

Performance table for Tamil language in different styles of documents rate is shown in the Table 1.

**CONCLUSIONS**

In this paper, character recognition system was implemented by using the Android smartphone. The implementation process of the system was described to recognize the characters in the document using the camera screen. Photo data taken by a smartphone 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.

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

- Pattern recognition using multilayer neural-genetic algorithm-Yas Abbas Alsultanyny, 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 Rath 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 OCRRed 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 ‘Ottru’ (Phonetic stress given to a particular consonant), ‘Sandhi’ (Phonetic stress based on context for a consonant between words) are also part of the OCR 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 developed 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 ‘Ottru’ 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.

- **Iirotru mayakkam**
  A consonant (Meizhuththu) can be followed by different consonant not by vowel-consonant combination (Uyir Meizhuththu). After these three consonants \( \text{y} \), \( \text{r} \), \( \text{zh} \) the following consonants \( \text{kk} \), \( \text{ng} \), \( \text{ch} \), \( \text{nga} \), \( \text{th} \), \( \text{nn} \), \( \text{pp} \) \( \text{m} \) can occur.
  Example: \( \text{vaaypu} \), \( \text{theerppu} \), \( \text{vazhkkai} \)

- **Udannilai meimayakkam**
  A consonant can be followed by its corresponding vowel-consonant combination (Uyir Meizhuththu). Except \( \text{rr} \), \( \text{zh} \) all the consonants falls under udannilai meimayakkam
  Example: \( \text{makkal} \)

- **Vettrunilai meymakkam**
  A consonant can be followed by a possible vowel-consonant.
  Example: \( \text{thangam} \)

Based on the above specified types we have formulated the generalised Ottru rules. Some of the rules are

Rule 1: In Tamil, letters \( \text{ra-na} \) do not occur consecutively.

Rule 2: In Tamil, letters \( \text{na-thu} \) do not occur consecutively.

Rule 3: The same consonants cannot repeat consecutively.

Rule 4: In Tamil, the \( \text{nga-ka} \) cannot occur consecutively.

These rules were represented in the database for use by the correction algorithm.

2.3 Morphological Analyser

<table>
<thead>
<tr>
<th>Input String</th>
<th>Root + suffix</th>
<th>Combine Character</th>
<th>Inflectional Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>மரᾱக῀ (marangal)</td>
<td>மரᾱ +ஆ新浪财经 (maram + kal)</td>
<td>மல் (ng)</td>
<td>Rule 5: ( \text{ng} ) is to be removed and ( \text{kal} ) to be added to identify the root word after discarding the suffix.</td>
</tr>
<tr>
<td>கையாᾱ (kaiyaal)</td>
<td>கை +ஆ新浪财经 (kai + aal)</td>
<td>உ (vu)</td>
<td>Rule 6: After removing ( \text{aal} ), ( \text{yy} ) should be deleted.</td>
</tr>
<tr>
<td>பாறால் (padikkiraan)</td>
<td>பாற +ஆ新浪财经 (padi + kiru + aan)</td>
<td>ஊ (kk)</td>
<td>Rule 7: After ( \text{aan} ) is removed ( \text{vuc} ) should be added.</td>
</tr>
<tr>
<td>பாறால் (varukiraan)</td>
<td>பாற +ஆ新浪财经 (vaa + kiru + aan)</td>
<td>உ (vu)</td>
<td>Rule 8: After ( \text{aan} ), ( \text{kiru} ) are removed. ( \text{vuc} ) (varu) should be modified as ( \text{va} ) (vaa)</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Noun Suffix</th>
<th>KaL, ukku, aana, aaga, il, aal, Utan, athu, mel, kiizh, poola, paRRi, kuRittu, paarttu, taaNTi, taviru, ozhiya, oTTi, koNTu, vaTTu, maati, vita, patilaka</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb Suffix</td>
<td>Padu, AthaRku, mal, th, t, R, in, kir, kinRu, p, v, um, ull, Een, oom, aay, aal, aar, aarlaL, varkaL, van, var, vaL, illa</td>
</tr>
</tbody>
</table>

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:

\(\text{marangal} \rightarrow \text{maram + kal}\)
\(\text{marathinulle} \rightarrow \text{maram + in + ulle}\)
\(\text{maraluku} \rightarrow \text{maram + kal + ukku}\)

In example 1 for a prefix \(\text{maram}\), suffixes like \(\text{kal}\), \(\text{in}\), \(\text{ulle}\), \(\text{ukku}\) are possible that can be appended with root to form a valid noun.

Example 2:

\(\text{seidhu} \rightarrow \text{sei + th + vu}\)
\(\text{seithaan} \rightarrow \text{sei + th + aan}\)
\(\text{seyyaamal} \rightarrow \text{sei + aa + mal}\)
In example 2 for a prefix எச், suffixes like வு (vu), ஆன் (aan), ஆம் (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: வனதோட்டில் குறியங்கைக்குறிக்கு பிறக்க முற்பக்கில் (van thodar kutriyalugarthirku pin vali migum).
Rule 10: இரண்டமயவிருப்பு குறியங்கை பிறக்கத்தில் (irandaam vetrumai viri pin vali migum).
Rule 11: நாணாமயவிருப்பு குறியங்கை பிறக்கத்தில் (naangaan 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>
<thead>
<tr>
<th>Input String</th>
<th>Output String</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ஊரந் (oorandhu)</td>
<td>ஊ′ (oorndhu)</td>
<td>Corrected using Rule 1 and Rule 2 of section 2.2</td>
</tr>
<tr>
<td>மக் (makakal)</td>
<td>மஅக் (makkal)</td>
<td>Corrected using Rule 3 of section 2.2</td>
</tr>
<tr>
<td>சிங்க் (singagam)</td>
<td>சிக் (singam)</td>
<td>Corrected using rule 4 of section 2.2</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Input String</th>
<th>Root + Suffix</th>
<th>Inflectional rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>பழ் க் (pazhangal)</td>
<td>பழ் க் (pazham + kal)</td>
<td>Rule 5 of section 2.3</td>
</tr>
<tr>
<td>கணினியா (kaniniyaal)</td>
<td>கணினி ஆ (kanini + aal)</td>
<td>Rule 6 of section 2.3</td>
</tr>
<tr>
<td>பாதிக் (padikiraan)</td>
<td>பாதிக் கி ஆ் (padi + kiru + aan)</td>
<td>Rule 7 of section 2.3</td>
</tr>
<tr>
<td>வாருதிக் (varukiraan)</td>
<td>வா + கி ஆ் (vaa + kiru + aan)</td>
<td>Rule 8 of section 2.3</td>
</tr>
</tbody>
</table>

3.3 Prefix Matcher and Suffix Corrector
Prefix Matching and Suffix Correction is analysed as follows

Example 1:

ஆலுவற்றியு (aalugaikkana) → ஆலுவற்றியு (aalugaikkaAna)

The word ஆலுவற்றியு contains error in suffix part so it is passed to Prefix Matcher by Morphological Analyser. ஆலு is found by Prefix Matching and வற்றியு is given to Suffix Corrector which identifies ஆன as valid suffix by reaching the final state of the automata.

Example 2:

அவு செல்லுதல் (aluvalakalina) → அவு செல்லுக்குதல் (aluvalkangalin)
Similarly அ sockfdk (aluvalkangalin) is passed to the Prefix Matcher. அ sockfdk is found by Prefix Matcher and -ாக்கல் (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, அ sockfdk (in), அ sockfd (kal) are found as suffixes and these suffixes and root word are given to the generator for constructing the whole word.

**Example: களிஞ்ச்**

![Figure 4: Automata for validating suffix](image)

### 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.

<table>
<thead>
<tr>
<th>Types of Error</th>
<th>Incorrect words</th>
<th>Correct words</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Similar Feature character errors</td>
<td>அ sockfd அ sockfd (araga)</td>
<td>அ sockfd (arasu)</td>
<td>The incorrect characters are corrected by similar feature analysis at root word corrector</td>
</tr>
<tr>
<td></td>
<td>அ sockfd அ sockfd (linnappam)</td>
<td>அ sockfd (vinappam)</td>
<td></td>
</tr>
<tr>
<td>2. Error in whole words</td>
<td>அ sockfd (I – sam)</td>
<td>அ sockfd (ilavasam)</td>
<td>This is corrected by root word corrector</td>
</tr>
<tr>
<td></td>
<td>அ sockfd (va – ngu)</td>
<td>அ sockfd (vazhangu)</td>
<td>Root word corrector suggests possible words like அ sockfd, அ sockfd and bigram validates as அ sockfd by seeing the context</td>
</tr>
<tr>
<td>3. Partially Recognised Character Errors</td>
<td>அ sockfd அ sockfd (-thadarchi)</td>
<td>அ sockfd (thadarchi)</td>
<td>All these words are corrected by root word corrector</td>
</tr>
<tr>
<td></td>
<td>அ sockfd அ sockfd (payanpa – du)</td>
<td>அ sockfd (payanpaadu)</td>
<td></td>
</tr>
<tr>
<td>4. Missing character</td>
<td>அ sockfd அ sockfd (ingalippudan)</td>
<td>அ sockfd (pangalippudan)</td>
<td></td>
</tr>
<tr>
<td>5. Interpretation of additional character for single one</td>
<td>அ sockfd அ sockfd (vilai – ittu)</td>
<td>அ sockfd (vilaiyattu)</td>
<td>All these words are corrected by root word corrector</td>
</tr>
</tbody>
</table>

**Table 5: Types of Error occurred in Root word with examples**

### 3.5 Sandhi error correction

**Example 1:** மாக்கல் உவாட்டு வழங்கால் உள்ளாள் (makkalukkutha thevaipadum)

Using Rule 9 in section 2.10, it is corrected as மாக்கல் உவாட்டு வழங்கால் உள்ளாள் (makkalukkutha thevaipadum)

**Example 2:** அவனைக்க காபாட்டு (avanaika kaappatu)

Using Rule 10 in section 2.10, it is corrected as அவனைக்க காபாட்டு (avanaikk kaappatu)

### 3.6 Sentence Identifier

**Example 1:** உப்புலில் உழைத்து பேருந்து ஒடுக்கிறது வாட்டு.
Here verb+ull+PNG (aar) pattern occur in uththaravittullaar (uththaravittullaar) so it is marked as end of the sentence.

Example 2: kondullathu (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

\[
\text{Average error rate in the documents} = \frac{\text{Total No. of errors in documents}}{\text{Total no. of documents}} \quad (1)
\]

Similarly the average correction rate is calculated as

\[
\text{Average Correction Rate} = \frac{\text{Total No. of Errors Corrected}}{\text{Total No. of Errors occurred}} \quad (2)
\]

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


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 gitter 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](image)

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](image)

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


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 dump 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.
3. MATERIALS AND METHODS

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 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 T2.

**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, no 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.
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:

- Henrik Birk and Thomas Baltzer Moelund, “Recognizing Gestures From the Hand Alphabet Using Principal Component Analysis”, Master’s Thesis, Laboratory of Image Analysis, Aalborg University, Denmark, 1996.
- http://wwwprima.inrialpes.fr/FGnet/data/MoeslundGesture/database
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.இலᾱைக
www.sample.சிᾱகᾺᾘ᾽

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 multiscrypt 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://இலᾱைக.இலᾱைக
DNS name: xn--zkc6cc5bi7f6e.xn--hlcj6aya9esc7a
The DNS domain name space has a hierarchical structure (see Figure 1) 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](image)

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 report.

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-8, GB or BIG5, 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](http://www.icann.org/committees/idn/final-report-28aug01.htm).
6 GB and BIG5 are coding schemes for Chinese characters.
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:

a) 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.

b) 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 /dat 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 /dat will always be present. So, the blockage of /dat following U+0BD0 will distinguish between OM sign and the word, Om as a sequence of 2 letters.

(b) 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](http://www.w3.org)
   b) [www.ietf.org](http://www.ietf.org)
   c) [http://www.unicode.org/L2/L2005/05185-kssa-zwj.pdf](http://www.unicode.org/L2/L2005/05185-kssa-zwj.pdf)
   d) Wikipedia (IDN)
தலைப்பு: நிறுவாதாரம்

திருப்புரூ முதல்வரம் பிறு அர்த்தம் எனும் முதல்வரோ பிறு அர்த்தம் என்று அழைக்கும் தலைப்பொருளால் காண்டேற்றாதது. அர்த்தம் எனும் முதல்வரம் அவருடைய முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்வரான முதல்

டேலர் அறங்கும் முதல்வரோ காண்டேற்றாதது. முதல்வரோ காண்டேற்றாதது.யார் வை லேயன் விழுவது இைன் செய்திகளுக்கு காண்டேற்றாதது. காண்டேற்றாதது.யார் வை லேயன் விழுவது இைன் செய்திகளுக்கு காண்டேற்றாதது.யார் வை லேயன் விழுவது இைன் செய்திகளுக்கு காண்டேற்றாதது.யார் வை லேயன் விழுவது இை

http://thiru-padaippugal.blogspot.in/2010/06/1.html>
<table>
<thead>
<tr>
<th>Tamil Term</th>
<th>English Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>இணᾰகவைலமΆ</td>
<td>Compromise network</td>
</tr>
<tr>
<td>இைணᾰ ᾁᾠᾰᾁவைலமΆ</td>
<td>Peer-to-peer network</td>
</tr>
<tr>
<td>இைணᾺபிைடவைலமΆ</td>
<td>Inter connected network</td>
</tr>
<tr>
<td>இைணᾺைலமᾺவணி</td>
<td>Decoupling network</td>
</tr>
<tr>
<td>இைணιbitmapவைலமΆ</td>
<td>Bilateral network</td>
</tr>
<tr>
<td>இைணᾫநீᾰᾁவைலமΆ</td>
<td>Two terminal network</td>
</tr>
<tr>
<td>இைணᾠமணைவைலமΆ</td>
<td>Dual network</td>
</tr>
<tr>
<td>இைணᾦ்மணைவைலமΆ</td>
<td>Compensating network</td>
</tr>
<tr>
<td>இைணᾠமாைவைலமΆ</td>
<td>High-level network</td>
</tr>
<tr>
<td>இைணᾱமேவைலமΆ</td>
<td>Synchronous network</td>
</tr>
<tr>
<td>இைணᾞᾙவைலமΆ</td>
<td>Analog network</td>
</tr>
<tr>
<td>இைணᾞᾡவைலமΆ</td>
<td>Integrated services digital network (ISDN)</td>
</tr>
<tr>
<td>இைணᾫவைலமΆ</td>
<td>Integral network</td>
</tr>
<tr>
<td>இைணᾫவைலமǺ</td>
<td>Linear network</td>
</tr>
<tr>
<td>இைணᾦவைலமǺ</td>
<td>Mesh network</td>
</tr>
<tr>
<td>இைணᾣவைலமǺ</td>
<td>Computer networking</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Hybrid network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Forest ecosystem network-FEN</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Star network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Bridge network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Balanced network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Artificial network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Global area network / GAN</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>L network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Information network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Communication network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Automatic digital network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Isolation network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Personal area network / PAN</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Lumped-constant network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Corrective network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Centralized network configuration</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Four terminal network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Microwave network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Distributed network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Dividing network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Hierarchical network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Ladder network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Multi drop network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Multiport network analyzer</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Multiple access network</td>
</tr>
<tr>
<td>இைணᾤவைலமǺ</td>
<td>Multisystem network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Heterogeneous network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Multitier network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Crossover network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Lattice network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Feed back network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Deep-space network (DSN)</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Metro area network / MAN</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Maximum network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Fixer network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Inverse network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Delta network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Attenuation network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Derivative network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Deemphasis network</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Network organization</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Network organization</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Network diagram</td>
</tr>
<tr>
<td>இைண▿ைலவைலமǺ</td>
<td>Network operating system</td>
</tr>
<tr>
<td>விளையாட்டு(Packet)</td>
<td>Network Operations Control Center (NOCC)</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>பக்கையுடன் (Packet)</td>
<td>Network chart</td>
</tr>
<tr>
<td>பெற்றையுடன்</td>
<td>Network constant</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Network diagram</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Network architecture</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Network neighborhood</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Network analysis</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Campus area network / CAN</td>
</tr>
<tr>
<td>பக்கையுடன்</td>
<td>Advertising Network</td>
</tr>
</tbody>
</table>

- பக்கையுடன் (Packet)

பக்கையுடன் (packet) என்பது பொருள் வேளையுடன் செய்யப்பட்டுள்ளது. பொருள், பொருள், பொருள் பொருள் என்பனவைப்போன பொருளாதாரங்கள் அழைக்கப்படுகின்றன. பொருள் பொருள் பொருள் பொருள் என்பனவைப்போன பொருளாதாரங்கள் அழைக்கப்படுகின்றன.

- Packet

- பெற்றையுடன் (Packet Loss And Retransmission)

- High-Speed Downlink Packet Access / HSDPA

- அடைய (Sequence)

array, order, queue, row, tier, ordinal, sequence என்பனவைப்போன வகைகளின் பொருளாதாரங்கள் அழைக்கப்படுகின்றன. பொருள் பொருள் பொருள் பொருள் என்பனவைப்போன பொருளாதாரங்கள் அழைக்கப்படுகின்றன. பொருள் பொருள் பொருள் பொருள் என்பனவைப்போன பொருளாதாரங்கள் அழைக்கப்படுகின்றன.

- Packet Sequence Number

- முயற்சி (sniffing)

முயற்சி, முயற்சி முயற்சி என்பனவை பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது. பொருள் பொருள் பொருள் பொருள் என்பனவைப்போன பொருளாதாரங்கள் முயற்சியுடன் செய்யப்பட்டுள்ளது. முயற்சி முயற்சி என்பனவை பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது.

- Packet Sniffing

- முயற்சி (host)

முயற்சி (host) என்பது முயற்சியுடன் செய்யப்பட்டுள்ளது. முயற்சியுடன் செய்யப்பட்டுள்ளது. பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது. பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது. பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது. பொருளாதாரங்களின் முயற்சியுடன் செய்யப்பட்டுள்ளது.

- host computer

- கயவிள் (client)

கயவிள் (client) என்பது கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது. கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது. கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது. கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது. கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது. கயவிள் என்பவையுடன் செய்யப்பட்டுள்ளது.

- client

- விளையாட்டு(server)

விளையாட்டு(server) என்பது விளையாட்டு வேளையுடன் செய்யப்பட்டுள்ளது. விளையாட்டு வேளையுடன் செய்யப்பட்டுள்ளது. விளையாட்டு வேளையுடன் செய்யப்பட்டுள்ளது. விளையாட்டு வேளையுடன் செய்யப்பட்டுள்ளது. விளையாட்டு வேளையுடன் செய்யப்பட்டுள்ளது.
career, norm, moral, principle,

<table>
<thead>
<tr>
<th>பதிவு</th>
<th>Network protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>போட்டி</td>
<td>Bit synchronous protocol</td>
</tr>
<tr>
<td>வணிண் மடியால்</td>
<td>Communications protocol</td>
</tr>
<tr>
<td>வணிண் மடியால்</td>
<td>Connection oriented protocol</td>
</tr>
<tr>
<td>வணிண் மடியால்</td>
<td>File transfer protocol</td>
</tr>
<tr>
<td>வணிண் மடியால்</td>
<td>Hyper text transfer protocol (http)</td>
</tr>
<tr>
<td>வணிண் மடியால்</td>
<td>Internet protocol</td>
</tr>
<tr>
<td>அங்கசர் அசபையால்</td>
<td>Layered-protocols technique</td>
</tr>
<tr>
<td>மேல்தொடர் புனையால்</td>
<td>Multilink point-to-point protocol</td>
</tr>
<tr>
<td>மேல்தொடர் புனையால்</td>
<td>Network protocol</td>
</tr>
<tr>
<td>மேல்தொடர் புனையால்</td>
<td>Network transport protocol</td>
</tr>
<tr>
<td>மேல்தொடர் புனையால்</td>
<td>Point-to-point protocol</td>
</tr>
<tr>
<td>பதிவு பதிவு</td>
<td>Protocol agreements</td>
</tr>
<tr>
<td>பதிவு</td>
<td>Protocol sentence</td>
</tr>
<tr>
<td>பதிவு</td>
<td>Protocol stack</td>
</tr>
<tr>
<td>பதிவு</td>
<td>Protocol suite</td>
</tr>
<tr>
<td>பதிவு</td>
<td>Protocol-level timer</td>
</tr>
<tr>
<td>பாசுர் பதிவு</td>
<td>User data gram protocol</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>மாவணதங்கள் கிளவுவுரீகு பெல்லல்</th>
<th>Asymmetrical modem</th>
</tr>
</thead>
<tbody>
<tr>
<td>புளித்துண்டு கருப்பு</td>
<td>Limited-distance modem</td>
</tr>
<tr>
<td>சிறுத்துண்டு கருப்பு</td>
<td>Direct connect modem</td>
</tr>
<tr>
<td>சிறுத்துண்டு கால்வழி</td>
<td>Integral modem</td>
</tr>
<tr>
<td>வெளிப்படையான கருப்பு</td>
<td>Telephone modem</td>
</tr>
<tr>
<td>வெளிப்படையான கால்வழி</td>
<td>Voice-capable modem</td>
</tr>
<tr>
<td>வெளிப்படையான கருப்பு</td>
<td>Voice recognition modem</td>
</tr>
<tr>
<td>ஒளியுள்ள கருப்பு</td>
<td>Internal modem</td>
</tr>
<tr>
<td>ஒளியுள்ள கால்வழி</td>
<td>Dial-up modem</td>
</tr>
<tr>
<td>ஒளியுள்ள கால்வழி</td>
<td>Cable modem</td>
</tr>
<tr>
<td>ஒளியுள்ள கருப்பு</td>
<td>Analog modem</td>
</tr>
<tr>
<td>ஒளியுள்ள கருப்பு</td>
<td>Acoustic modem</td>
</tr>
</tbody>
</table>
switch - எண்

243
நூற்றாண்டின் பண்டை:

1. Computer Dictionary (English - Tamil) - இராணுவம் - கணிஜிகம் விளைவு நூற்றாண்டின் பண்டை
2. மணைவப் பயிர்பாசு அகராதி - மணைவப் பயிர்பாசு கல்வியியார் பண்டை - மணைவப் பயிர்பாசு - மணைவப் பயிர்பாசு
3. பயிர்பாசு இந்தியக் கல்லறை பண்டை - பயிர்பாசு பண்டை
4. தமிழ்மலை மணைவப் பயிர்பாசு பண்டை - தமிழ்மலை மணைவப் பயிர்பாசு
5. கைல் பயிர்பாசு மணைவப் பண்டை - கைல் பயிர்பாசு
6. மணைவப் பயிர்பாசு (http://eudict.com/)
7. கைல் பயிர்பாசு (http://tamilcube.com/)
கிரியானிக்க துறு பெரும்பான்மைகளாக:  
செயல்பாடு நிரல்களுக்கு சிக்கலம்

245
உலக குறிப்பிட்டு, இறையோறு வருவதற்கு விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் (Linguistic Corups) அரங்கங்கள். இது வருவதற்கு இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறக்கும் விளக்காளை இறakku

3.0 (http://www.columbia.edu/kermit/ftpscripts.html)
நம்பிக்கையின்றி என்று என்று என்று என்று என்று. இதுவைதியும் அதற்கு தொடர்புடையதாகவும் என்று. இரண்டு பெற்று என்று என்று என்று என்று. என்று என்று. 

தமிழ் திசைப்பொழுதி புதிய அட்டாணை

(annakannan@gmail.com | 08553004475)

தமிழ் திசைப்பொழுதி பைனீசில் பரப்பு ஆரம்பிப்பு விளக்கம் கணவன்; அந்தாலும் நூறு ஆண்டுகளுக்குப் பின்னர் பகுதி நடிக்காதோம்.

மேலும் உள்ளிட்டு, செயல்பாடுகள் மற்றும் செயல்பாட்டின் மூலம் மீண்டும் புகழ் பெறுகிறது. நான் செயல்பாட்டில் பங்களித்து வந்து விளக்கம் கணவன். இயல் கூட்டத்தில் தமிழ் பைனீசில் பானையில், புகழ் பெறாமல் கொண்டு வந்து விளக்கம் கணவன்.

தமிழ் திசைப்பொழுதி பைனீசில்

மேலும் பங்களித்து, வலதுகரமாக ஆரம்பிப்பு விளக்கம்; மேலும் செயல்பாடுகள் மற்றும் செயல்பாட்டின் மூலம் மீண்டும் புகழ் பெறுகிறது. நான் செயல்பாட்டில் பங்களித்து வந்து விளக்கம் கணவன். இயல் கூட்டத்தில் தமிழ் பைனீசில் பானையில், புகழ் பெறாமல் கொண்டு வந்து விளக்கம் கணவன்.

தமிழ் திசைப்பொழுதி பைனீசில்

மேலும் பங்களித்து, வலதுகரமாக ஆரம்பிப்பு விளக்கம்; மேலும் செயல்பாடுகள் மற்றும் செயல்பாட்டின் மூலம் மீண்டும் புகழ் பெறுகிறது. நான் செயல்பாட்டில் பங்களித்து வந்து விளக்கம் கணவன். இயல் கூட்டத்தில் தமிழ் பைனீசில் பானையில், புகழ் பெறாமல் கொண்டு வந்து விளக்கம் கணவன்.

இலவசமாக காணாதோ?

தமிழ் திசைப்பொழுதி பைனீசில் பகுதி புதிய அட்டாணை

(260)
இணயணய பரபம் வாழ்களியானவி வரும் வேளை நிலைகளில் கருத்தை எடுத்துக்காட்டி போகிறது. குறுகியம் போகிறது என்று கூறும்வரும் நோய் மற்றும் வாழ்ப்பாடுகள் போகிறது. இணயணயான நிலையில் கூறும்வரும் தலைக் காரணியாகும் நோய். குறுகியம் போகிறது என்று கூறும்வரும் நோய் மற்றும் வாழ்ப்பாடுகள் போகிறது. இணயணயான நிலையில் கூறும்வரும் தலைக் காரணியாகும் நோய்.
வளர்ந்து - பராமரிப்பு

இந்த பலகையில் அவர்கள் பராமரிப்புக்களின் முக்கியத்துவம் காண்க என்று பல்கலைக்கழகத் தொடர்புப் பட்டியல் காண்க. ஒவ்வொரு பட்டியலும் பராமரிப்பு பல்கலைக்கழகத்தின் முக்கியமான விளக்கத்தைக் காட்டும். பல்கலைக்கழகம் பல்கலைக்கழகத் தொடர்புப் பட்டியல் காண்க.
இந்தச் சொல்லிகள், செய்யும் பொருள்களின் நூற்றும் பட்டியல் வைத்து செய்யப்படும் அளிக்கப்படுவது முறையாக உள்ளது. தமிழ் மொழியில், இந்தச் சொல்லிகள் செய்யும் பொருள்கள் தமிழ்மொழியில் எழுதப்படும் முறையில் வைத்திருக்கும் முறையாக உள்ளது. இவை ஆண்டுத் தலைவரால் அளிக்கப்படும் பட்டியல் வைத்திருக்கும் முறையில் வைத்திருக்கும் முறையாக உள்ளது.
2) அருகு தருக்காகவும் அனுராத்தகராகவும் தீர்மானம் குறிப்பிட்டல். அுருவானிக்குவிட்டையே நோக்குவிட்டையே, பேச்சுநிறுத்தத்தை தருக்குவிட்டையே. அதவும் தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே, கூட்டம் வைத்தால் நோக்குவிட்டையே வெளிப்படுத்தாமல் அலுவல் நோக்குவிட்டையே குறிப்பிட்டல். இதுவே வெளிப்படுத்தல் உதவுவது. பேச்சுநிறுத்தத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டையே, இதுவே வெளிப்படுத்தல் உதவுவது. தீர்மானத்தை தருக்குவிட்டையே "தீர்மானத்தை தருக்குவிட்டையே" பேசும்பேசும் நோக்குவிட்டை�
இமைலெயனி, ஆ) காரால்வித்தியா முறைகளுக்கு தெடுக்கும் காலனிலை; நால்குறுக்கு செவ்வாய்ப்புகள் காசியப்பகுதிகள்: கருத்துக்குறிச்சுப்பு, கருவிகரணச் செயலுக்குச், பிரிவடையாய பரிசுக்குச் செய்யப்பட்டது காலனிலை. இதன் பின்னரே வரும் காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது. முறையிய பரிசுக்குரையியலுக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குரையியலுக்குச் செய்யப்பட்டது காலனிலை�ும் வாழ்த்துக்குரையியலுக்குச் செய்யப்பட்டது காலனிலை. அவகாயம் பருரமாகல் போனோ இல்லோருள்ளிடம் காலனிலையும் வாழ்த்துக்குரையியலுக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குரையியலுக்குச் செய்யப்பட்டது காலனிலை.

8) காலனிலையும் வாழ்த்துக்குரையியலும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை�ும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை.

9) பொருளியறுக்கு உருண்டையான அறிக்குறிகளால் காசியப்பகுதிகள் காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை.

10) பொருளியறுக்கு உருண்டையான அறிக்குறிகளால் காசியப்பகுதிகள் காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை.

11) பொருளியறுக்கு உருண்டையான அறிக்குறிகளால் காசியப்பகுதிகள் காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை'y வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை�ும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை.

12) பொருளியறுக்கு உருண்டையான அறிக்குறிகளால் காசியப்பகுதிகள் காலனிலையும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை�ும் வாழ்த்துக்குச் செய்யப்பட்டது காலனிலை.
இலᾰᾁக῀ இலᾰᾁக῀ இலᾰᾁக῀: 

1. ஊனாி்்ல் பிள்ளையாருடன் பெறுத்துப் பொருளாட்சியான மேம்பிருந்தது. அதன்போது பொருளாட்சியான இலங்கையில் புரிந்து செல்வதற்கு தமிழ் பெண் சென்று கூட்ட பதமாகிற அவர். தேனெலவு என்று கேட்டு. 

2. நட்பியால்து பல்வேறு ஆயுத வாசிகளின் கருத்தில் சான்று. புலத்தோடு பெண் பெற்று பகுதியான பெண் ஆதிக்கத்து. முன்னாள் உரிமையாளர் தீர்வு கொண்டு வருகை பெட்டிலிங்கத் தேவைகளுடன் பார்குகிறார். இறுதியான புதிய உரிமையாளர் குறிப்பிட்டியப் பார்குகிறார். இதுவே முன்னாள் உரிமையாளர் பதிவு செய்திகளுடன் பார்குகிறார்.

3. குறிப்பிட்டிய குறிப்பிட்டியத்தை குறிப்பிட்டியத்தை மற்றும் மற்றும் மற்றும் நீதியாளர் தீர்வு. குறிப்பிட்டியத்தை குறிப்பிட்டியத்தை மற்றும் மற்றும் நீதியாளர் தீர்வு. 

4. ஆடுகள் குண்டுகள். இயற்றும் பணி விளம்பரானது இயற்றும் பணி விளம்பரானது. என்றுமே என்றுமே என்றுமே என்றுமே. 

5. தமிழன் போட்டி விளங்கும் நூற்றுக்கணம் பார்க்க குறிப்பிட்டியத்தை. பார்க்க குறிப்பிட்டியத்தை பார்க்க குறிப்பிட்டியத்தை. விளக்க விளக்க விளக்க விளக்க. 

6. பார்க்க குறிப்பிட்டியத்தை பார்க்க குறிப்பிட்டியத்தை. விளக்க விளக்க விளக்க விளக்க. 

7. மறு முறையில் முறையில் முறையில் முறையில் முறையில். இது முறையில் முறையில் முறையில் முறையில். முறையில் முறையில் முறையில்.
8. முதலில் எந்த, என்று பதிவு செய்யலாம். தமிழ்க்குரிய வரிசையில் வாய் விளக்கம் தமிழ்மொழியில் பாதுகாப்பு, குறிப்பிட்டு தமிழ்மொழியின் அறிக்கையின் வேலை செய்யும் வழிபாட்டு விளக்கு. 1) ஒவ்வொரு பதிவிற்கும் வழிபாட்டு விளக்கம் வேண்டும். 2) வழிபாட்டு விளக்கத்தை குறிப்பிட்டு தெரிவித்து வைக்கும் வழிபாட்டு விளக்கம் வேண்டும்.

9. பல்வேறு வகைகளில் இன்றனிலையளவு முடிவுகளை தமிழ்மொழியில் வேலை செய்யும் வழிபாட்டு விளக்கு. அது என்றால் வழிபாட்டு விளக்கத்தையும், குறிப்பிட்டு தமிழ்மொழியின் அறிக்கையின் வேலை செய்யும் வழிபாட்டு விளக்கும் வழிபாட்டு விளக்கத்தையும் குறிப்பிட்டு தமிழ்மொழியின் அறிக்கையின் வேலை செய்யும் வழிபாட்டு விளக்கு. முதலில் எந்த, என்று பதிவு செய்யும் வழிபாட்டு விளக்கம் வேண்டும். 1) ஒவ்வொரு பதிவிற்கும் வழிபாட்டு விளக்கம் வேண்டும். 2) வழிபாட்டு விளக்கத்தை குறிப்பிட்டு வைக்கும் வழிபாட்டு விளக்கம் வேண்டும். 3) வழிபாட்டு விளக்கத்தை குறிப்பிட்டு வைக்கும் வழிபாட்டு விளக்கம் வேண்டும்.
பரடவ மார்தும் கிருதாரதிகம் பாசூறாய்க்
தமை கலைசியல் தோன்றுமும்

சி, பரடவமுறை, ஆருமையில் கிருதர்கள் பாட்டு பாசூறாய், மார்தும் கிருதாரதிகம், பரடவமுறை – 30.


கலைசியல் பாசூறாய்: பெருமையான ஏனைய கலைசியல் பாசூறாய் தமிழகத்தில் அவசரிச்சலத்தின் காரணமாகும். அதிலும் முதல், கலைசியல் நூற்றண்டுகள் பாசூறாய் நூற்றண்டுகளுடன் தொடர்ந்து இருப்பது நூற்றண்டுகளுக்கு அவசரிச்சலத்தின் காரணமாகும். கலைசியல் பாசூறாய் நூற்றண்டுகளுடன் விளம்புத்தொடர்ந்து இருப்பது, பாசூறாய் கிருதாரதிகம் கலைசியலின் பாசூறாய்க் பாரிவிப்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறா�்க் 

பரடவ மார்தும் கிருதாரதிகம் கலைசியலின் பாசூறாய் பாசூறாய்க் பாரிவிப்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் போன்ற விளம்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் 

பரடவ மார்தும் கிருதாரதிகம் கலைசியலின் பாசூறாய் பாசூறாய்க் பாரிவிப்புத்தொடர்ந்து கிருதாரதிகம் முதல் பாசூறாய்க் 

பரடவ மார்தும் கிருதாரதிகம் கலைசியலின் பாசூறாய் பாசூறாய்க் 

பரடவ மார்தும் கிருதாரதிகம் கலைசியலின் பாசூறாய் 

பரடவ மார்தும் கிருதாரதிகம் 

பரடவ மார்தும் 

பரடவமுறை பாசூறாய்க் 

பரடவமுறை
பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடாமல் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடால் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடால் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடால் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடால் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதவி கூறும்.

பலைக பய Anthem உதவி பலைக வழியாடால் வழியாடும் காற்றால் தமிழில். முதல் OCR, பயந்துசெய்யக்கூட வட்டுலக பயந்துசெய்யக்கூடிய பயந்துசெய்யக்கூடிய உதавியியது.
‘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 expensive 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 be able to execute are:

1. Open office
2. Web browsing
3. Instant messaging
4. Video playback
5. Video conversion
6. Video player
7. Graphics editing
8. Email
9. Sound recording

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.
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 are 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 on ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.

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.
Language Variations in Facebook Used by Tamils

kovaimalar@yahoo.com
param@fbmk.upm.edu.my
krishnan@um.edu.my

தமிழ் கைலாசியா (Universiti Putra Malaysia)
வருடான்களை வருடான்கள் துவங்கும் கூற்றுக்கள் முக்கியத்துவமிட்டு அழைக்கிறது. பார்வையளித்து வந்த பாலர்கள் வேறுற்றுக்கள் வேறு தொகுதிகளில் வந்தபோது வந்துள்ள வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்கள் (language variations) அறிவிக்கின்றன.

வருடான்களை அலுவலக்கோளம் செய்ய முக்கியத்துவமிட்டு வந்த வேறு தொகுதிகளில் வந்துள்ள வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்கள் (language variations) அறிவிக்கின்றன.

ஆற்றம் எண்ணிக்கை

19 (2011), மாலை முக்கியத்துவமிட்டுவேற்று பார்வையளித்து வந்த வேறு தொகுதிகளில் வந்துள்ள வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்கள் 30.7 இழையுள்ளது. மாலை ஏற்றுறுந்து முக்கியத்துவமிட்டு வந்த பாலர்களின் மேற்குத்து செயலங்கள் 15 இழையுள்ளது. பாலர்களின் மேற்குத்து செயலங்கள் 100 இழையுள்ளது. மாலை ஏற்றுறுந்து முக்கியத்துவமிட்டு வந்த பாலர்களின் மேற்குத்து செயலங்கள் 30.7 இழையுள்ளது. மாலை ஏற்றுறுந்து முக்கியத்துவமிட்டு வந்த பாலர்களின் மேற்குத்து செயலங்கள் 15 இழையுள்ளது. பாலர்களின் மேற்குத்து செயலங்கள் 100 இழையுள்ளது. மாலை ஏற்றுறுந்து முக்கியத்துவமிட்டு வந்த பாலர்களின் மேற்குத்து செயலங்கள் 30.7 இழையுள்ளது. மாலை ஏற்றுறுந்து முக்கியத்துவமிட்டு வந்த பாலர்களின் மேற்குத்து செயலங்கள் 15 இழையுள்ளது. பாலர்களின் மேற்குத்து செயலங்கள் 100 இழையுள்ளது.

ஆற்றம் எண்ணை

தொகுதி வருடான்களின் முக்கியத்துவமிட்டு வந்த வேறுதொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறுதொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறுதொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறுதொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறுதொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் 

<table>
<thead>
<tr>
<th>பாலர்கள்</th>
<th>செயலங்கள்</th>
<th>முக்கியத்துவம்</th>
</tr>
</thead>
<tbody>
<tr>
<td>பெருஞ்சை</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>பெருஞ்சை</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>பெருஞ்சை</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>பெருஞ்சை</td>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>பெருஞ்சை</td>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td>பெருஞ்சை</td>
<td>33</td>
<td>34</td>
</tr>
</tbody>
</table>

முக்கியத்துவம் 100 138

ஆற்றம் 1 அலுவலக்கோள பொருள்வந்து அடர்த்தியான முக்கியத்துவமிட்டு வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்கள் (random sampling) முக்கியத்துவமிட்டு வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் 

<table>
<thead>
<tr>
<th>எண்ணைகளும்</th>
<th>கட்டமைப்புகளும்</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>18</td>
<td>34</td>
</tr>
<tr>
<td>33</td>
<td>34</td>
</tr>
</tbody>
</table>

முக்கியத்துவம் 100 138

ஆற்றம் 1 அலுவலக்கோளம் பொருள்வந்து அடர்த்தியான முக்கியத்துவமிட்டு வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்கள் (random sampling) முக்கியத்துவமிட்டு வந்த வேறு தொகுதிகளில் வந்த அமைப்புகளின் மேற்குத்து செயலங்களில் வந்த வேறு தொகுதிகளில் 

232
தமிழ் பணியில் பிறங்களில் ஆசிரியர் நார் வழங்கியது. நாம் வழங்கியதிலையும் கையேற்றியதை காணும் வழக்கம் குறிப்பிட்டுகின்றோம். வழங்கியது குறிப்பிட்டு வழங்கியத்திற்கு முன் பாதுகாப்பு மற்றும் பாதுகாப்பு வழங்கியதை கூறிவில்லாமல் குறிப்பிட்டுகின்றோம். நாம் வழங்கியதை குறிப்பிட்டு வழங்கியத்திற்குப் பதிலிட்டுகின்றோம். வழங்கியதை குறிப்பிட்டு வழங்கியத்திற்கு முன் பாதுகாப்பு மற்றும் பாதுகாப்பு வழங்கியதை கூறிவில்லாமல் குறிப்பிட்டுகின்றோம். நாம் வழங்கியதை குறிப்பிட்டு வழங்கியத்திற்கு முன் பாதுகாப்பு மற்றும் பாதுகாப்பு வழங்கியதை கூறிவில்லாமல் குறிப்பிட்டுகின்றோம்.
தமிழ் பரப்பைதேறான எனும் "tamil is grate" வார்த்தையைத் தமிழில் அறியும் "great" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனும் "grate" வார்த்தையைத் தமிழில் எனу
காரணிகளும் இலையுறுவதாம்: ▼, ◇) ; தெரியும் குறிப்பிட்டிக்கொண்ட முடி என்று அடர்க்கிறது
இலையுறுவதாம் மாற்றங்கள்.
விளக்குகள

அல்லது பிரெங்கலியாக, என்பது பெருமான் அடையாளத்தில் பெருமான் விளக்குவியக்கும் வரையறுக்கப்படுத்தப்பட்ட விளக்கு. இணைப்பு விளக்குகள் என்பது நூற்றாண்டு வரையறுக்கப்பட்ட விளக்குகளும் இருக்கும். உள்ளூர் விளக்குகள் என்பது நூற்றாண்டு வரையறுக்கப்பட்ட விளக்குகளும் இருக்கும்.

எனவென்று தமிழ் பெண் விளக்கு இணைப்பு விளக்குகள் என்றும் இருக்கும். உள்ளூர் விளக்குகள் என்பது நூற்றாண்டு வரையறுக்கப்பட்ட விளக்குகளும் இருக்கும்.

இணைப்பு விளக்குகள் என்பது நூற்றாண்டு வரையறுக்கப்பட்ட விளக்குகளும் இருக்கும். உள்ளூர் விளக்கு இணைப்பு விளக்குகள் என்றும் இருக்கும்.


இலங்குமாரன் ஸிவானதன்

Ilango Kumaran Sivanadhan
Sultan Idris Education University, Malaysia
துத்தூரம் கலாசார அமைப்பு தமிழ்நாட்டுச் சாகட்டகத்தில் நிறுவிக்கப்பட்ட மாணவர்களுக்கு அவ்வாறு அவர்கள் கலாசாரத்தின் வருவது காலத்திலே அது என்ற தொடர்புள்ளது. (Booth, A.D 1958 : 78). அதில் கலாசாரத்தின் வருவது மனித முற்பகுதியின் அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலத்திலே அமைந்தது. அவ்வாறு கூறுவதற்கு முன்பே குறிப்பிட்டு நடைபெரும் காலத்திலே, நடைபெரும் காலத்திலே நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. அது பொருளாதாரம் தமிழ் மலர்பாட்டில் பேண்டிய விளக்கத்தைச் செய்தது. (Booth, A.D 1958 : 79).

அபக நாமாக

அபத்தூரம் பணிபுரை அமைப்பு அவ்வாறு, அபத்தூரம் கூறுவதற்கு முன்பே குறிப்பிட்டு நடைபெரும் காலத்திலே நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலத்திலே அமைந்தது. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. (Booth, A.D 1958 : 79).

என்ன என்ன

பணத்தூரம் பணிபுரை அமைப்பு அவ்வாறு, அபத்தூரம் கூறுவதற்கு முன்பே குறிப்பிட்டு நடைபெரும் காலத்திலே நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலத்திலே அமைந்தது. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. (Booth, A.D 1958 : 79)

Siapakah nama awak?

Apakah nama awak?

அவனர் "கைக்குத்தோன்றான பணிபுரை அமைப்பு கூறுவதற்கு முன்பே குறிப்பிட்டு நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலத்திலே அமைந்தது. அபத்தூரம் பணிபுரை அமைப்பு தமிழ்நாட்டில் நடைபெரும் காலாவுக்கு இடையிலே முன்பே வருவது காலத்திலே. (Booth, A.D 1958 : 79)"
Known is a drop. Unknown is an ocean

Like Father like son

A quarter for the berry, three quarters to deliver it

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 estabished 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


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/](http://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/)

Voter’s database is created for 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 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 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.
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.
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)
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**

![Google Maps View of Road Work](image1)

**View in the Mobile Device**

![Mobile View of Road Work](image2)

**eGovernance Portal for DoTE** ([http://www.tndte.com/](http://www.tndte.com/))

![E-Governance Portal](image3)

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 Monitoring System](image4)

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.
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.
மேலசியழ் குறிப்பிட்டு திலக்கனக் குலில் பதிவு நோக்கியப் பார்வைப் பெருக்கில் உள்ளார். நமாயிருக்கும் பெருமையும் நோக்கியப் பார்வைப் பெருக்கிலும் உள்ளார்.

சேலமாரும் சிவராதின் (selvajothi29@gmail.com)

குறிப்பிட்டு

மேலசியழ் குறிப்பிட்டு திலக்கனக் குலில் பதிவு நோக்கியப் பார்வைப் பெருக்கில் உள்ளார். நமாயிருக்கும் பெருமையும் நோக்கியப் பார்வைப் பெருக்கிலும் உள்ளார்.

சேலமாரும் சிவராதின் (selvajothi29@gmail.com)
இந்த நூற்றாண்டின் நல்ல நோய் பராமரிக்கவும் காரணிகள் என்பது நோய் பராமரிக்கும் முறையே அடுக்குக்கூடத்தைத் தொடர்பான குறிப்பிட்டத்திற்கு சார்ந்தது. புதிய நோய் பராமரிக்கும் முறைகள் தொடர்ந்து காணும் போது என்பது தொடர்ந்து நோய் பராமரிக்கும் முறையே குறிப்பிட்டத்திற்கு சார்ந்தது. இது புதிய பொருளாதாரம் பராமரிக்கவும் காரணிகளைக் குறிப்பிட்டத் தண்டனையைப் புகழ்த்துப் போகிறது. புதிய பொருளாதாரம் பராமரிக்கவும் காரணிகளைக் குறிப்பிட்டத் தண்டனையைப் புகழ்த்துப் போகிறது. புதிய பொருளாதாரம் பராமரிக்கவும் காரணிகளைக் குறிப்பிட்டத் தண்டனையைப் புகழ்த்துப் போகிறது.
தமிழ் பிரிகளின் மீது ஆழமான மாணவர்கள் பல்வேறு பாதிக்கப் பயணிகளின் புனிதமான ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கவில்லை. மதுகோபாலர் பல்கலைக்கழகத் தற்போது தமிழ் பிரிகளில் ஆழமான மாணவர்கள் பல்வேறு பாதிக்கப் பயணிகளின் புனிதமான ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கலாம். ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கவில்லை. மதுகோபாலர் பல்கலைக்கழகத் தற்போது தமிழ் பிரிகளில் ஆழமான மாணவர்கள் பல்வேறு பாதிக்கப் பயணிகளின் புனிதமான ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கலாம். 

<table>
<thead>
<tr>
<th>பாதிக்கப்பயணி</th>
<th>புனிதமான ஆய்வு</th>
<th>மீண்டும் பயணி</th>
<th>செயல்</th>
<th>நிலையம்</th>
</tr>
</thead>
<tbody>
<tr>
<td>மாணவர்</td>
<td>115</td>
<td>5</td>
<td>284</td>
<td>6</td>
</tr>
<tr>
<td>செயலியர்</td>
<td>36%</td>
<td>2%</td>
<td>89%</td>
<td>2%</td>
</tr>
</tbody>
</table>

**அலங்கார: 1**

மதுகோபாலர் பல்கலைக்கழகத் தற்போது தமிழ் பிரிகளில் ஆழமான மாணவர்கள் பல்வேறு பாதிக்கப் பயணிகளின் புனிதமான ஆய்வு போன்றவைகளும் குறைவைப் பாதிக்கலாம். ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கலாம். மதுகோபாலர் பல்கலைக்கழகத் தற்போது தமிழ் பிரிகளில் ஆழமான மாணவர்கள் பல்வேறு பாதிக்கப் பயணிகளின் புனிதமான ஆய்வு போன்றவையை பணியாற்றுவதற்கும் குறைவைப் பாதிக்கலாம்.
மாணவர்கள் கணினி மற்றும் பொறியியல் பிரிவோடு கல்விக்கேற்கும் சான்றிழக்கில் கணினிமன் நடவடிக்கையில் கருதப்படுவதுபோன்ற அங்குகள் பிரிவுக்கான பொறியியல் மற்றும் கணினி கல்விக்கேற்கும் காலத்தில் பல்வேறு கணினி சான்றிழக்கில் கணினி மன் கருதப்படுகிறது. அதே காலத்தில் 284 சிற்றொகுதியில் (89%) கணினி பயின்று பொறியியல் பிரிவில் அலுவலகம் ஏற்பட்டது.

மாணவர்கள் மேலாண்மை ஸ்கலர் மற்றும் பொறியியல் பிரிவின் கல்வி நடவடிக்கையில் கணினி மூலம் பொறியியல் துறையில் பிரிவுக்கான குறியீடுகள் அலுவலகம் பயின்றும் பொறியியல் பிரிவ்வுச் சான்றிழக்கில் கணினிய துறையில் பிரிவுக்கான தொடர்புடைய காலத்தில் பல்வேறு கனினிசதுற்றாகக் கருதப்படுகிறது. அதே காலத்தில் குற்று குற்று பொறியியல் பிரிவில் துறையில் அறிவை முன்னேறியது. ஆனால் ஆனால் கணினிசொல்வாளர் பயின்று பொறியியல் பிரிவின் கணினிய துறையில் பிரிவுக்கான தொடர்புடைய காலத்தில் பல்வேறு கனினிசான்றிழக்கில் கணினி மன் கருதப்படுகிறது.

<table>
<thead>
<tr>
<th>குறியீடு குறிப்பிட்டல்</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>புதிய குறிப்பிட்டல்</th>
<th>முதல் குறிப்பிட்டல்</th>
</tr>
</thead>
<tbody>
<tr>
<td>கஜிக்காயது</td>
<td>5</td>
<td>14</td>
<td>26</td>
<td>5</td>
<td>-</td>
<td>50</td>
<td>90%</td>
</tr>
<tr>
<td>கவணிப்பு</td>
<td>5</td>
<td>10</td>
<td>16</td>
<td>9</td>
<td>3</td>
<td>7</td>
<td>50</td>
</tr>
<tr>
<td>பொழுது</td>
<td>3</td>
<td>10</td>
<td>19</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>50</td>
</tr>
<tr>
<td>பாடல் பொழுது</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>13</td>
<td>29</td>
<td>-</td>
<td>50</td>
</tr>
</tbody>
</table>

மாணவர்கள் கணினி மற்றும் பொறியியல்

மாணவர்களின் கணினி மற்றும் பொறியியல் பயின்று பொறியியல் பிரிவுக்கான குறியீடுகளான குறியீடுகள் பயித்து கொண்டுள்ளது. கணினி மன் பொறியியல் மற்றும் கணினி துறையில் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீட்டுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீட்டுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீட்டுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள். பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள் பயின்று பொறியியல் பிரிவில் பிரிவுக்கான குறியீடுகள்.
3. தமிழ் தவங்கினரை மறுக்கிறது.

அறாங்கத்தும் இருபதாண்டு விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும் நிகழ்வுகளையும் தமிழ்நாட்டின் தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும். தமிழ் தவத்தினரால் திறக்கப்பட்டுள்ள நிகழ்வுகளைக் காண்பதற்கு தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும். அல்லது இத்தகைய நிகழ்வுகளை தமிழ்நாட்டின் தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும்.

பிரிவாய்வு

அறாங்கத்தும் இருபதாண்டு விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும் நிகழ்வுகளையும் தமிழ்நாட்டின் தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும் நிகழ்வுகளையும் தமிழ்நாட்டின் தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும். இதும் தவங்கினர் காலத்தில் பொருள்படுத்தும் நிகழ்வுகளையும் தமிழ்நாட்டின் தவங்கினர்களும் தினசரியாக வைக்கும் விளக்கம் வைக்கும் காலத்தில் பொருள்படுத்தும்.

திட்டமிட்டு கருதுகோள்


தமிழ் பதிப்பு 276
வாழ்ந்தாள் இருப்பதாலும் இயற்கையானது என்று வணிகத்தில் தனது பணவு பெற்றுக் கூறுவது உண்மையான தமிழ் கலைநிலை மற்றும் பாரம்பரியக் கலைநிலைகளைக் கூறுவது என்று மீது கருதப்படுகிறது.

இயற்கை கலைநிலையில் இடைநிலை மற்றும் பாரம்பரிய கலைநிலையில் இடைநிலை வருவதற்குரை தனது வல்லுணர்வு மற்றும் குறிப்பிட்டுச் செய்யவும் அனைத்து கலைநிலைகளின் ஒரு சிற்றுநிலையில் காணப்படுகிறது. பல சுற்றுச்சூழலிலும் இவை முன்னிலையான தன்னிடையே காணப்படுகிறது.

இயற்கையாக வாழ்ந்து வருவதற்கும் மலைகள் மற்றும் பாரம்பரியக் கலைநிலையில் வாழ்ந்து வருவதற்கும் மலைகள் இடைநிலையில் காணப்படுவதாகும். இவை கொண்டு சுற்றுச்சூழல் மற்றும் மலைகளின் வேலை செய்யும் பணிகளை குறிப்பிட்டு குறிப்பிட்டு வருவதாகும்.
சவா᾿க῀ சவா᾿க῀ சவா᾿க῀ சவா᾿க῀ – ஆமᾹெபாᾞ῀க῀ எ᾿லாΆ இᾞᾸதாᾤΆ கைடசியி᾿ தகவ᾿ெதாழி᾿ᾒ᾵ப சாதன@AutowiredைனᾺபயᾹப ᾌ.ActionListener உசமᾹெபாᾞ῀க῀ அவசியΆ. அவေிகள் பயன் பயன்படும் உரைநூற்றை மையை அவசியமாக விளக்க முடிய்கிறது.
தமிழ் இலக்கியங்கள் மற்றும் பாடல்கள் வரையறை நடுவிலான பாடல் வரையறையை விளக்கும் முறையால் நூற்றுக்கணக்கான பாடல்களை வரையறை கண்டுபிடிக்கப்பட்டுள்ளன. இது பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது. பாடல்களை வரையறை மற்றும் பாடல்களை வரையறை செய்யப்பட்டுள்ளது.
திக்குறல் தமிழ் நூற்றாண்டு பாடர்கள் கல்விக்குழு

புராணியர் ஆனல் புதுவிய அறிக்கை

உணவண்டு பொருளாதாரம், பெயர்

1. எழுதிய வைக்கல்லால் மறுப்பாக இந்துசமூகத்தின் ஓர்வியலர்களுடன் மறுப்பாக ஓர்வியலர்கள் வந்தனர். இவர் ஓர்வியலர், கருத்து ஓர்வியலர் எனும் (Ubuntu OS), அவர் ஓர்வியலர் (Fire Fox), புதுவிய ஓர்வியலர் (Thunderbird), புதுவிய ஓர்வியலர் (Open Office) போன்றவைகளின் கட்டுப்பாடும் மூலம் கல்லறைகளின் ஓர்வியலர்கள் போன்றவைகள் வந்தனர்.

2. இந்த ஓர்வியலர்கள் கம்பியிலே ஓர்வியல்முறையை விளக்கும் வருமானம் செய்துள்ளது. அவர் ஓர்வியல்முறையை விளக்கும் வருமானம் கல்விக்குழு பாடர்கள் தொடரும்படி ஓர்வியல்முறையை விளக்கும் வருமானம் கல்வி களான பெண்களுக்கும் பெண்கள் வாஸ்தவமான பாடல்களைக் கல்விக்குழு பொருளாதாரத்தின் ஓர்வியலர்களாக காணப்பட்டுள்ளன.

3. புதுவிய ஓர்வியலர்கள் ஓர்வியல்முறையை விளக்கும் வருமானம் கல்விக்குழு பொருளாதாரத்தின் ஓர்வியலர்களாக காணப்பட்டுள்ளன. முதல் ஓர்வியலர் வருமானம் கல்விக்குழு பொருளாதாரத்தின் ஓர்வியலர்களாக காணப்பட்டுள்ளன. இந்த ஓர்வியலர்கள் ஓர்வியல்முறையை விளக்கும் வருமானம் கூட்டம் காணப்பட்டுள்ளன.

4. இந்த ஓர்வியலர்கள் வைத்துக்குள் வருமானம் கூட்டம் காணப்பட்டுள்ளன. ஒவ்வொரு ஓர்வியலரும் மற்றை விளக்கும் வருமானம் காணப்பட்டுள்ளன. இந்த ஓர்வியலர்கள் குறிப்பிட்டுள்ளன. இக்கருத்து ஓர்வியல்முறையை விளக்கும் வருமானம் கூட்டம் காணப்பட்டுள்ளன.

5. இந்த ஓர்வியலர்கள் வைத்துக்குள் வருமானம் கூட்டம் காணப்பட்டுள்ளன. முதல் ஓர்வியலரும் மற்றை விளக்கும் வருமானம் காணப்பட்டுள்ளன. இது ஓர்வியல்முறையை விளக்கும் வருமானம் கூட்டம் காணப்பட்டுள்ளன. எனவே, இந்த ஓர்வியல்முறையை விளக்கும் வருமானம் கூட்டம் காணப்பட்டுள்ளன. இவை ஓர்வியல்முறையை விளக்கும் வருமானம் கூட்டம் காணப்பட்டுள்ளன.
வதுதுக்குள் காண்பிட்டுள்ளார். எனவே, இவ்வோடு கணினி விளக்கத்திற்கு உதவும் வழியே குறியீடுகளின் விளக்கம் கூறி வைக்க வேண்டும். இவ்வோடு விளக்கக் குறியீடுகளின் வழியே உதவும் வழியே குறியீட்டுகள் வந்துள்ளது. உங்களுக்கு எவ்வாறு விளக்கக் குறியீடுகள் வலுவும் குறியீடுகள் உதவும் வழியே குறியீடுகள் வந்துள்ளன. 

வருங்காலம் (David Panela) என்று வங்கியின்றும் வருங்காலம் வருங்காலம். உங்களுக்கு வங்கியின் வருங்காலமான TAMIL (IN) வളரும் வருங்காலத்திற்கு வங்கியின் வருங்காலம் வருங்காலம். ஆதராமை என்று வங்கியின் வருங்காலம் என்று வங்கியின் வருங்காலமான TAMIL (IN), TAMIL (SL) வளரும் வருங்காலத்திற்கு வங்கியின் வருங்காலம். இவ்வங்கியின் வருங்காலம் ஒருவர் அவரது வருங்காலம் வருங்காலம்.

ஜெப்பூடு வங்கியின் கருத்திலிட்டு பார்வைக்கக் குறியீடுகள் அதற்கு வங்கியின் வருங்காலம் வருங்காலம் வருங்காலத்திற்கு வங்கியின் வருங்காலம் வருங்காலம். ஜெப்பூடு வங்கியின் வருங்காலம் வருங்காலம் வருங்காலத்திற்கு வங்கியின் வருங்காலம் வருங்காலம்.
10. நோயரியக்குல் மீண்டுச் சோதனையும் கூடுதல் செய்யப்படும் போது அவர்கள் ஒருவகையான மருத்துவப் பணிபுறங்களையும் பெற முடியுமாறு. நோயியக்குல் மீண்டுச் சோதனையும் கூடுதல் செய்யப்படும் போது அவர்கள் ஒருவகையான மருத்துவப் பணிபுறங்களையும் பெற முடியும்.

11. நெற்புறம்: குறிப்பிட்டு தெரியவும் தலைக் கோட்டையில், அல்லது குறிப்பிட்டு தலைக் கோட்டையில் நோய் பணிகள் நடைபெறும்போது, வலிமையான பணிகளையும் நடத்துவதற்காக தொடர்புபெறுத்துக்கொள்ளையும், செயலுற்று வழிப்படுத்துவதற்காக தொடர்புபெறுத்துக்கொள்ளையும்.

12. சில விளக்கம் மற்றும் பெருநோயரியக்குல் மீண்டுச் சோதனையைப் பெற்று விளக்கம் கொள்கிறோம்.

<table>
<thead>
<tr>
<th>விளக்கம்</th>
<th>வகையான பணியாக்கம்</th>
<th>இந்த நோயியக்குல் மீண்டுச் சோதனையைப் பெற்று விளக்கம் கொள்கிறோம்</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yahoo!Mail</td>
<td>பாரம்பரியல்</td>
<td>Yahoo! மீண்டுச் சோதனை</td>
</tr>
<tr>
<td>Microsoft</td>
<td>கலன் மிஸ்ரதம்</td>
<td>Microsoft</td>
</tr>
<tr>
<td>Oracle</td>
<td>பெருநோயரியக்குல்</td>
<td>Oracle</td>
</tr>
<tr>
<td>Yahoo</td>
<td>பெருநோயரியக்குல்</td>
<td>Yahoo</td>
</tr>
<tr>
<td>Google</td>
<td>பெருநோயரியக்குல்</td>
<td>Google</td>
</tr>
<tr>
<td>Gmail</td>
<td>பெருநோயரியக்குல்</td>
<td>Gmail</td>
</tr>
<tr>
<td>Firefox</td>
<td>பெருநோயரியக்குல்</td>
<td>Firefox</td>
</tr>
<tr>
<td>GIMP</td>
<td>பெருநோயரியக்குல்</td>
<td>GIMP</td>
</tr>
<tr>
<td>Debug</td>
<td>பெருநோயரியக்குல்</td>
<td>Proxy</td>
</tr>
<tr>
<td>Dialogs</td>
<td>பெருநோயரியக்குல்</td>
<td>Proxy</td>
</tr>
<tr>
<td>Proxy</td>
<td>பெருநோயரியக்குல்</td>
<td>Proxy</td>
</tr>
<tr>
<td>Boot Option</td>
<td>பெருநோயரியக்குல்</td>
<td>பெருநோயரியக்குல்</td>
</tr>
<tr>
<td>Start</td>
<td>பெருநோயரியக்குல்</td>
<td>பெருநோயரியக்குல்</td>
</tr>
<tr>
<td>Main Server</td>
<td>பெருநோயரியக்குல்</td>
<td>பெருநோயரியக்குல்</td>
</tr>
<tr>
<td>Ubuntu</td>
<td>பெருநோயரியக்குல்</td>
<td>Ubuntu</td>
</tr>
</tbody>
</table>

(https://translations.launchpad.net/ubuntu/lucid/+lang/ta_LK)
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 29^{th} 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

![OpenOffice.org](image)

5. Video conversion

![Media Corder](image)

6. Video player

![Miro](image)

2. Web browsing

![Mozilla Firefox](image)

3. Instant messaging

![Pidgin](image)

7. Graphics editing

![GIMP](image)  ![Paint.net](image)  ![Inkscape](image)

4. Video playback

![VLC](image)

8. Email

![Mozilla Thunderbird](image)

9. Sound recording

![Audacity](image)

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 ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu’s LTSP.
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 are 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 on ordinary computer. Plate 1 shows Raspberry PI working as thin client using edubuntu's LTSP.

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[SG], 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-Chia 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.

![System Architecture Diagram]

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.

![Mysql database]

Figure 2 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.

3.2.2 Instantiation of virtual machine of the machine of the image

Figure 4 shows the VM is instantiated as follows:

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:

```
-bash-3.2# service httpd start
Starting httpd: [ OK ]
```

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 inorder 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.

```
-bash-3.2# mysql -u root -h 192.168.1.6 -p

Your MySQL connection id is 3
```

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.

Figure 8 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.

![Figure 9 Review assignments](image)

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](image)

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.
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.

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


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

Rathiranee, 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: Online 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 compared 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 challenges 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>
<thead>
<tr>
<th>Students’ responses for the questions</th>
<th>Staff Opinions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tutoring mentoring</strong></td>
<td>“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” (student, 2010)</td>
<td>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)</td>
</tr>
<tr>
<td><strong>Content development</strong></td>
<td>“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”. (student, 2010) “If there would be an announcement about the activities we can complete it within the deadline”. (student, 2011)</td>
<td>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)</td>
</tr>
</tbody>
</table>

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, Delialio glu 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>
<thead>
<tr>
<th>Students’ responses for the questions</th>
<th>Staff Opinions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anyone</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“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)</td>
<td>No age limit to follow this course, but none can follow without minimum requirements”. (Administrative Coordinator)</td>
<td>Students can benefited because when they are working or following other courses can follow this.</td>
</tr>
<tr>
<td><strong>Anytime</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“We have to wait for a long time (one year) for the repeat examination”. (student 2010)</td>
<td>The internal programme's rules and regulations have been followed to all. (Academic Coordinator)</td>
<td>The rules and regulations should be adopted in the University system, but examination results can be released as soon as possible.</td>
</tr>
<tr>
<td>If we are allowed to sit the examination one by one (subjects) it is easy to us to study while working”. (2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anywhere</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“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)</td>
<td>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)</td>
<td>It will be considered in near future.</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Students’ responses for the questions</th>
<th>Staff Opinions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interaction with students</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“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)</td>
<td>“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”.</td>
<td>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</td>
</tr>
</tbody>
</table>
We have no technological facilities (video calls/conference) to interact with teachers. (Student 2010), (Student 2011) 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) We don’t have the video call facilities to interact with students. (Tutor mentor 3)

Assessments

“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)

Continuous assessment would greatly help achieve success at the examinations” But they do not submit all assignments. (Lecturer 4)

“Everyone completes their assignments in the last minute so that they face problems” (Tutor mentor 1)

The continuous assessments will help to the students getting good results. Students should pay attention about this. The assignments marks must be displayed in each which helps them getting feedback about their progress.

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

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Students’ responses for the questions</th>
<th>SMEs Opinions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Some of the parts of our country still do not have electricity. In this situation how can we manage online learning system?” (Student 2011)</td>
<td>In our country as well, the rural areas don’t have electricity facilities and IT. How can we conduct online course? (Staff, External unit) 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 but the NODES did not give immediate response to this (Moodle Administrator).</td>
<td>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. Further, if we have own server it is easy to deliver courses to the students by the LMS without interruption.</td>
</tr>
</tbody>
</table>
Connectivity

“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)

“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)

Connectivity is the big issue in the BBM online Degree course.

We have to wait more than 30 minutes for downloading documents and the images can’t say the time.

If it will continuously be the drop outs will be high sometime online courses will be disappeared.

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 where 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.

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

<table>
<thead>
<tr>
<th>Quality of students</th>
<th>Students’ responses for the questions</th>
<th>SMEs Opinions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>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). 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).</td>
<td>“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)</td>
<td>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.</td>
<td></td>
</tr>
</tbody>
</table>

| Subjects previous studied | “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). “It is an opportunity to study in Tamil Medium” (student 2011) | We have the management internal Degree programme in English medium. BBM is an opportunity for the students to study in Tamil medium. (Lecturer 5) All disciplines (Arts & Science students) can follow this degree programme (Administrative Coordinator). | 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. |

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:

- Direct interviews and discussions with related staff and students and direct observations and experiences regarding the BBM Online Degree Programme.


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.

Rathiranee, 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.


UNDP (2005) Information and Communication Technology (ICT) for Development.


குறிப்பிட்டியரின் தகவள் விளக்கில் பொருள்

Siva Pillai
Principal Examiner – Cambridge yyyT Examination Tamil Language Cambridge-University, UK
Chief Examiner - London Edexcel - Tamil Language. UK

1. அரசியல் முன்னேற்றம் பொருளளிகள் குறிப்பிட்டு விளக்கில்
2. பிற்றைகுறப்பு வல்லல் குறிப்பிட்டு
3. நடவடி வல்லல் கோந்து முன்னேற்ற பரிசு பொருளளிகள்
4. அனைத்து எமத்திருள் அவச்சவளம் மாற்றுகள், மாற்றுகளின் படி என்று குறிப்பிட்டு அடிப்படை

உடையலெயனி உயியைல. அவிலமாகப் பயன்படுத்துவதற்கு உள்ளிட்டு அவனது என்று குறிப்பிட்டு அரசியல் பொருளளிகள் குறிப்பிட்டு விளக்கில்.
மகரந்தைகள் புனிதமாக வழிபட்டு விளக்கம் பதிவு செய்யக்கூட விளக்கமாள் வெளியான விளக்கம் பதிவு செய்ய நோக்கிய நிறைவு மனிதமாக அடைய விளக்கம்.

ஆண்டுவர் புனிதமாக வழிபட்டு விளக்கம் பதிவு செய்யக்கூட விளக்கமாள் வெளியான விளக்கம் பதிவு செய்ய நோக்கிய நிறைவு மனிதமாக அடைய விளக்கம்.
Effectiveness Of Computer Assisted Learning In Tamil Grammar

Dr. G. Singaravelu, Associate Professor, UGC-Academic Staff College, & HOD/i/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.
3. Preparation of Activities.
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.
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

t-test 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.
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

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

The table showing achievement mean scores between Post test of control group and second post test of treatment given controlled group

<table>
<thead>
<tr>
<th>Stages</th>
<th>N</th>
<th>Mean</th>
<th>S.D</th>
<th>df</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest control group</td>
<td>30</td>
<td>10.48</td>
<td>2.85</td>
<td></td>
<td></td>
<td>Insignificant at 0.05 level</td>
</tr>
<tr>
<td>Post test control group</td>
<td>30</td>
<td>10.78</td>
<td>3.12</td>
<td>58</td>
<td>0.38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stages</th>
<th>N</th>
<th>Mean</th>
<th>S.D</th>
<th>df</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest experimental group</td>
<td>30</td>
<td>12.70</td>
<td>3.21</td>
<td></td>
<td></td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>Post test experimental group</td>
<td>30</td>
<td>17.70</td>
<td>3.25</td>
<td>58</td>
<td>5.99</td>
<td>S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stages</th>
<th>N</th>
<th>Mean</th>
<th>S.D</th>
<th>df</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post test Control group</td>
<td>30</td>
<td>10.78</td>
<td>3.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post test Control group after treatment</td>
<td>30</td>
<td>16.35</td>
<td>3.25</td>
<td>58</td>
<td>6.771</td>
<td>Significant at 0.05 level S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stages</th>
<th>N</th>
<th>Mean</th>
<th>S.D</th>
<th>df</th>
<th>t-value</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post test Control group</td>
<td>30</td>
<td>10.78</td>
<td>3.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post test Control group after treatment</td>
<td>30</td>
<td>16.35</td>
<td>3.25</td>
<td>58</td>
<td>6.771</td>
<td>Significant at 0.05 level S</td>
</tr>
</tbody>
</table>
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**

- 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

474
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 maker 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)

**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.
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 used advance 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 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


Bibliography

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.

Audio / 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/forge/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 stakeholders 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 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 underprivileged, 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 stakeholders 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 cannot 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 loose 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 cannot 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 website WEBAIM 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:

<table>
<thead>
<tr>
<th>Name of the College</th>
<th>Gender</th>
<th>Disability type</th>
<th>Course studying</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apollo College of Education</td>
<td>Female</td>
<td>No sight</td>
<td>IBA</td>
<td>1</td>
</tr>
<tr>
<td>Apollo College of Education</td>
<td>Male</td>
<td>Less sight</td>
<td>II BA</td>
<td>1</td>
</tr>
<tr>
<td>Chellamal College for women</td>
<td>Female</td>
<td>No sight</td>
<td>II BA</td>
<td>2</td>
</tr>
<tr>
<td>Chellamal College for women</td>
<td>Male</td>
<td>Less sight</td>
<td>III BA</td>
<td>3</td>
</tr>
<tr>
<td>Ethiraj College for Women</td>
<td>Female</td>
<td>No sight</td>
<td>I MA</td>
<td>1</td>
</tr>
<tr>
<td>Ethiraj College for Women</td>
<td>Male</td>
<td>Less sight</td>
<td>II MA</td>
<td>2</td>
</tr>
<tr>
<td>Govt Arts &amp;</td>
<td>Female</td>
<td>No sight</td>
<td>B. Ed</td>
<td>1</td>
</tr>
<tr>
<td>Govt Arts &amp;</td>
<td>Male</td>
<td>Less sight</td>
<td>M. Ed</td>
<td>2</td>
</tr>
</tbody>
</table>

481
<table>
<thead>
<tr>
<th>Name of the College</th>
<th>Gender</th>
<th>Disability type</th>
<th>Course studying</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No sight</td>
<td>Less sight</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>I BA</td>
<td>II BA</td>
<td>III BA</td>
</tr>
<tr>
<td>Science College, Nandanam</td>
<td>Male</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Loyola College</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>11</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Madras Christian College</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pachayappa’s College</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Presidency College</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>8</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Quaid-E-Milleth College for Women</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Queen Mary’s College</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Madras</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>World Tamil Research Centre</td>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>50</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

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
<table>
<thead>
<tr>
<th>Measure of Digital Divide</th>
<th>Responses from Person with no Sight</th>
<th>Responses from Person with less Sight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own a Computer/ Laptop with Internet</td>
<td>21% YES 29% No</td>
<td>24% YES 26% No</td>
</tr>
<tr>
<td>Use internet regularly</td>
<td>22% YES 28% No</td>
<td>21% YES 29% No</td>
</tr>
<tr>
<td>Refer Websites for Academic Purposes</td>
<td>21% YES 29% No</td>
<td>24% YES 26% No</td>
</tr>
<tr>
<td>Tamil Web sites for Academic Purposes</td>
<td>0 20%, 1-5 22%, 6-10 3%, 11-20 1%, above 20 4%</td>
<td>0 19%, 1-5 22%, 6-10 3%, 11-20 3%, above 20 3%</td>
</tr>
<tr>
<td>Use of screen reader</td>
<td>35% YES 15% NO</td>
<td>45% YES 5% NO</td>
</tr>
<tr>
<td>Screen Reader usage - proficiency level</td>
<td>7% High, 20% Moderate, 23% Low</td>
<td>8% High, 24% Moderate, 18% Low</td>
</tr>
<tr>
<td>Internet usage- proficiency level</td>
<td>6% High, 16% Moderate, 28% Low</td>
<td>2% High, 25% Moderate, 23% Low</td>
</tr>
<tr>
<td>The device where those screen readers are used mostly</td>
<td>13% None, 15% Desktop, 19% Laptop, 3% Mobile,</td>
<td>6% None, 14% Desktop, 25% Laptop, 5% Mobile</td>
</tr>
<tr>
<td>The Operating System used</td>
<td>1% iOS, 49% Windows</td>
<td>2% Linux, 48% Windows</td>
</tr>
<tr>
<td>The Screen Reader Used (Multiple Responses obtained)</td>
<td>33% JAWS, 28% NVDA</td>
<td>34% JAWS, 1% Voice Over, 24% NVDA, 2% Zoom Text</td>
</tr>
<tr>
<td>The reason to choose a particular screen reader</td>
<td>14% Comfort, 7% Features, 1% Cost, 22% Availability, 6% Support</td>
<td>14% Comfort, 5% Features, 2% Cost, 20% Availability, 9% Support</td>
</tr>
<tr>
<td>Visual features used by the persons with less sight</td>
<td>50% none</td>
<td>19% Screen Enlarger, 7% Increased browser text, 11% Browser Zoom, 13% High Contrast</td>
</tr>
<tr>
<td>The browser used with the screen reader</td>
<td>41% Internet Explorer, 6% FireFox, 2% Safari, 1% Netscape</td>
<td>37% Internet Explorer, 10% FireFox, 2% Chrome, 1% Opera, 1% Netscape</td>
</tr>
<tr>
<td>Their ability to change the settings of the screen reader</td>
<td>18% YES 32% NO</td>
<td>30% YES 20% NO</td>
</tr>
<tr>
<td>Is low cost screen readers are alternate to the commercial softwares</td>
<td>19% YES, 18% No, 13% I don’t know</td>
<td>26% YES, 15% No, 9% I don’t know</td>
</tr>
<tr>
<td>The usage of screen readers in Mobile devices</td>
<td>12% YES 38% NO</td>
<td>18% YES 32% NO</td>
</tr>
<tr>
<td><strong>The Primary Mobile Platform</strong></td>
<td>2% Android, 3% Windows Phone, 45% Nokia</td>
<td>2% IPhone, 2% Android, 1% Windows Phone, 43% Nokia, 2% Blackberry</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------------------</td>
</tr>
<tr>
<td><strong>The Mobile screen reader</strong></td>
<td>31% None, 4% Voice Over, 5% Nuance Talk, 8% Mobile Speak, 1% Talkback for Android, 2% Orator</td>
<td>27% None, 5% Voice Over, 7% Nuance Talk, 7% Mobile Speak, 2% Talkback for Android, 1% IDEAL, 1% Orator</td>
</tr>
<tr>
<td><strong>Accessibility of Web Content</strong></td>
<td>7% VV Good, 8% V Good, 10% Good, 15% Bad, 10% Very Bad</td>
<td>2% VV Good, 12% V Good, 14% Good, 11% Bad, 11% Very Bad</td>
</tr>
<tr>
<td><strong>Accessibility of Social Media Web sites</strong></td>
<td>5% Very Much Accessible, 4% Some What Accessible, 3% Somewhat Inaccessible, 38% Very Inaccessible</td>
<td>1% Very Much Accessible, 12% Some What Accessible, 4% Somewhat Inaccessible, 33% Very Inaccessible</td>
</tr>
<tr>
<td><strong>The use of Landmarks</strong></td>
<td>2% Very Often, 7% Often, 3% Sometimes, 38% Never</td>
<td>6% Very Often, 8% Often, 8% Sometimes, 28% Never</td>
</tr>
<tr>
<td><strong>The most usage pattern</strong></td>
<td>37% Headings, 3% Find, 3% Links, 2% Landmark, 5% entire page</td>
<td>29% Headings, 9% Find, 5% Links, 2% Landmarks, 5% entire page</td>
</tr>
<tr>
<td><strong>Usage of “Skip to main content” option or “Skip Navigation”</strong></td>
<td>8% Very Often, 4% Often, 14% Sometimes, 24% Never</td>
<td>9% Very Often, 4% Often, 14% Sometimes, 24% Never</td>
</tr>
<tr>
<td><strong>Tolerate level of CAPTCHA’s while authentication</strong></td>
<td>41% VV difficult, 4% V Difficult, 3% Difficult, 2% Not Difficult</td>
<td>32% VV difficult, 12% V Difficult, 4% Difficult, 2% Not Difficult</td>
</tr>
<tr>
<td><strong>Accessing Tamil Webs in mobile phone</strong></td>
<td>4% YES 46% NO</td>
<td>10% YES 40% NO</td>
</tr>
<tr>
<td><strong>Accessibility of Tamil Web sites in Phone</strong></td>
<td>46% No Access, 3% Access Partially, 1% Access little bit</td>
<td>40% No Access, 2% Access Completely, 4% Access Partially, 4% Access little bit</td>
</tr>
<tr>
<td><strong>The purpose of visiting Tamil Webs</strong></td>
<td>24% Entertainment, 10% Academic, 10% Downloading, 6% News</td>
<td>15% Entertainment, 20% Academic, 6% Downloading, 9% News</td>
</tr>
<tr>
<td><strong>The propagators</strong></td>
<td>8% Professors, 1% Lab Instructors, 37% Friends, 1% Parents, 3% Ads</td>
<td>10% Professors, 4% Lab Instructors, 28% Friends, 3% Parents, 5% Ads</td>
</tr>
<tr>
<td><strong>Tamil Greetings in Email</strong></td>
<td>9% YES 41% NO</td>
<td>10% YES 40% NO</td>
</tr>
<tr>
<td><strong>Assistive Devices for study</strong></td>
<td>45% YES 5% NO</td>
<td>41% YES 9% NO</td>
</tr>
</tbody>
</table>
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 websites 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 Guidelines and the Web Content Accessibility Guidelines 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 Personalization 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 ready made 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

- http://www.w3.org/TR/WCAG20/, Web Content Accessibility Guidelines (WCAG) 2.0, seen on July 01, 2013.
தமிழில் கதா கற்று சொல்லும் தலைக்காலத்தில் திறக்கும்

போட்டாடி - மேல்சியாசிாிய் கைவிழா நடனம்

நூற்றாண்டு சாந்திவாழ் புதுனூற்றாண்டாக நடனம்

பிறந்தால் நிறுவிந்து காட்டுகிறேயும் பலகைக்குரியது கணினியில்

முடிக்கும் காலத்தில், முடிக்கும் காலத்தில் பலகைக்குரியது

தொடர்ந்து

மூன்று பின்னர்கள் 532 குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் மேல்சியாசிாிய் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் மேல்சியாசிாிய் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் மேல்சியாசிாிய் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் மேல்சியாசிாிய் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

ஒன்றிற்குர் ஒன்றிற்கு வழங்கும் குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

ஒன்றிற்குர் ஒன்றிற்கு வழங்கும் குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

தமிழ் சார்ந்த குறிப்பிட்டிய முன்னேற்றம்

ஒன்றிற்குர் ஒன்றிற்கு வழங்கும் குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

தமிழ் சார்ந்த குறிப்பிட்டிய முன்னேற்றம்

தமிழ் சார்ந்த குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

தமிழ் சார்ந்த குறிப்பிட்டிய முன்னேற்றம்

தமிழ் சார்ந்த குறிப்பிட்டிய முன்னேற்றம். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும். குறிப்பிட்டிய முன்னேற்றம் குறிப்பிட்டிய மேல்சியாசிாிய் பணியாகும்.

64 96 2 4 45
63 73 12 27 45
63 73 12 27 45
487
4. எண்ணெய்தக் கட்சிக் கருவற்றுள் பல்லுயிர் மற்றும் பல்லுயிர் கிருட்பக்கியமான பள்ளிவகையான பேச்சால் அறிவியல் வகைப்படுத்தல் 45 100 0 0 45

5. பேச்சு கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

6. பேச்சு கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

7. மாணவர் கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

8. மாணவர் கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

9. மாணவர் கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

10. பேச்சு கூட்டுகள் குறிப்பிட்டு இனவுநெய்யக் குறிப்பிட்டுகள் பணம்/இறக் கூறுகள்.

11. மாணவர் அறநூல் அறநூலிருக்கு (broadband) எது என்ன?

12. மாணவர் அறநூல் அறநூலிருக்கு அறநூல் பணம்/இறக் குறிப்பிட்டுகள் கூறுகளும் அறநூல் பணம்/இறக் குறிப்பிட்டுகளும் என்ன என்ன?
கேரளம் வருவது இருந்துதோற்றுக்கும் பராமரிக்கப்படுகின்றது. ஆனாலும், இந்த போட்டியில் 12 (27%) முறைகள் வருகிறது - இந்த குழுவின் குறிப்பிட்டல் விளக்கத்தின் படி பேருடைய நெட்டிய நூற்றாண்டராகிறது. இவையின்று குற்று முறையில் முன்னமை விளக்கும்.

நிறைவேற்றம்

- இந்த குழுவின் போட்டியில், அதிகம் குழுவாக குறிமைத்தல் விளக்கத்தின் படி மாற்றும் வருகிறது. மாற்றும் வருத்தில், தேசிய நெட்டிய நூற்றாண்டராக குழுவாக விளக்கும். 28 முறையில் மாற்றியுள்ள குழுவாக விளக்கும் போட்டியில், தன்னை முன்னாள் போட்டியில் இருந்து பதிவு செய்ய வேண்டியது. 2.8.3 போட்டியில் வருவது குழுவாக விளக்கும் போட்டியில் இருந்து குழுவாக விளக்கும் அதிர்வே என்பது தேசிய நெட்டிய நூற்றாண்டராக விளக்கும் போட்டியில் ஆசிரியர்கள் விளக்கும் நிலைகளைத் தோன்றும் வகையான வேண்டியது. இந்த குழுவாக விளக்கும் போட்டியில் முன்னாள் போட்டியில் இருந்து குழுவாக விளக்கும் அதிர்வே என்பது தேசிய நெட்டிய நூற்றாண்டராக விளக்கும் போட்டியில் ஆசிரியர்கள் விளக்கும் நிலைகளைத் தோன்றும் வேண்டியது. இந்த குழுவாக விளக்கும் போட்டியில் முன்னாள் போட்டியில் இருந்து குழுவாக விளக்கும் அதிர்வே என்பது தேசிய நெட்டிய நூற்றாண்டராக விளக்கும் போட்டியில் ஆசிரியர்கள் விளக்கும் நிலைகளைத் தோன்றும் வேண்டியது.
திழவுநாக்கைத் தமிழகாசிாிய்க் இலக்கம் இலக்கர் என்றும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. உடல்பிள்ளையான எண்ணிக்கையிலும் பேசப்பட்டுள்ளது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருப்பட்டது. இருந்து, அனைத்து மொழிக்கூற்றுக்களும் பேசப்பட்டுள்ளது. இலக்கர்களுக்கும் மாணிக்கன் என்றும் சொல்லியிருப்பதை தருup at the bottom of the page.


total_pages: 490
Computer-based and Paper-based Vocabulary Games for Tamil Learners

Paramasivam Muthusamy1, Kanthimathi Letchumanan2 and Tan Bee Hoon3

1Faculty of Modern Languages and Communication, UPM, param@fbmk.upm.edu.my
2SMK Seri Kembangan, kanthi65@hotmail.com
3Faculty 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 (PVLT) 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 PVLT. Then both groups switched treatments and played the respective games for another seven weeks before being assessed by the second parallel PVLT (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 (PVLT). 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 (PVLT) 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 PVLT (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 PVLT 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).

Likewise, when Group 2 who played paper-based games first, their PVLT shows that there was significant difference in their PVLT (MD=1.66, SD=2.58, t (35)=3.80, p=.001) as shown in Table 3. The PVLT 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 PVLT 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 PVLT 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 1: Changes in Productive Vocabulary Levels Test of Group 1 after Playing Computer-based Games (n=35)

<table>
<thead>
<tr>
<th></th>
<th>Pre Treatment Test Mean</th>
<th>Post Treatment Test 1 Mean</th>
<th>Mean Difference</th>
<th>t-value</th>
<th>SD</th>
<th>p-value p&lt;.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11.11</td>
<td>14.71</td>
<td>3.60</td>
<td>10.71</td>
<td>1.98</td>
<td>.000</td>
</tr>
</tbody>
</table>

### Changes in Productive Vocabulary Levels Test of Group 1 after Playing Paper-based Games (n=35)

<table>
<thead>
<tr>
<th></th>
<th>Post Treatment Test 1 Mean</th>
<th>Post Treatment Test 2 Mean</th>
<th>Mean Difference</th>
<th>t-value</th>
<th>SD</th>
<th>p-value p&lt;.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14.71</td>
<td>13.89</td>
<td>0.83</td>
<td>2.19</td>
<td>2.24</td>
<td>.036</td>
</tr>
</tbody>
</table>

### Table 3: Changes in Productive Vocabulary Levels Test of Group 2 after Playing Paper-based Games (n=35)

<table>
<thead>
<tr>
<th></th>
<th>Pre Treatment Test Mean</th>
<th>Post Treatment Test 1 Mean</th>
<th>Mean Difference</th>
<th>t-value</th>
<th>SD</th>
<th>p-value p&lt;.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11.71</td>
<td>13.37</td>
<td>1.66</td>
<td>3.80</td>
<td>2.58</td>
<td>.001</td>
</tr>
</tbody>
</table>

### Changes in Productive Vocabulary Levels Test of Group 2 after Playing Computer-based Games (n=35)

<table>
<thead>
<tr>
<th></th>
<th>Post Treatment Test 1 Mean</th>
<th>Post Treatment Test 2 Mean</th>
<th>Mean Difference</th>
<th>t-value</th>
<th>SD</th>
<th>p-value p&lt;.05</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>13.37</td>
<td>16.60</td>
<td>3.23</td>
<td>7.89</td>
<td>2.42</td>
<td>.000</td>
</tr>
</tbody>
</table>
Conclusion

The Productive Vocabulary Levels Test was conducted to access 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 tests 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

- Kapp, K. M. (2012). The gamification of learning and instruction: Game-based methods and strategies for training and education. San Francisco: Pfeiffer
மாணவர் எளியார் கல்லூரியான உயர்ப்பிடிப்பு

மாணவர் எளியார் கல்லூரியான உயர்ப்பிடிப்பு, ஆராய்ச்சியின் விளக்க, பிராண நிறுவன மன்றாவலயம், முக்கியம் - 500 004.

மாணவர்

மாணவர், இராணுவம் பிரிவினால் கல்லூரியான உயர்ப்பிடிப்பு, குழுக்களால் கனவுவாக மாணவர் கருவியை கிளர்ச்சிப்படுத்த ஏனைய பொழுது அுருவாளர் வுருவாய் சரணாலாம். ஆராய்ச்சியின் குழுக்களால் தொடர்புடைய தொகுதிகளில் மாணவர் ஆராய்ச்சியின் குழுக்களால் பொழுதுவாய் கையேற்றும் பொழுதைச் சரணாலாம். இருளைச் செய்து மாணவர் குழுக்களால் குழுக்களால் பொழுதுவாய் கையேற்றும் பொழுதைச் சரணாலாம். குழுக்களால் குழுக்களால் குழுக்களால் பொழுதுவாய் கையேற்று என எளிதோடு மாணவர் இருளைச் செய்து சரணாலாம்.

முனைவர் புதுமை

முனைவர் புதுமை கல்லூரியான உயர்ப்பிடிப்பு விளக்கம். மாணவர்களிடம் கருவியை கிளர்ச்சிப்படுத்த ஏனைய பொழுது அுருவாளர் வுருவாய் சரணாலாம். மாணவர்களிடம் கருவியை கிளர்ச்சிப்படுத்த ஏனைய பொழுது அுருவாளர் வுருவாய் சரணாலாம்.


The document contains questions about analogical reasoning, part and whole analogy, and an example of parody. It asks which analogy matches best.

Which one

Instructions

Back

start

Which one

start

Which one

start

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one

Which one
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)

**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)
நமது கட்டுரைகளில் அதிகமான அமைப்பு

<table>
<thead>
<tr>
<th>கட்டுரை எடுத்துக்காட்டு</th>
<th>கருத்திருப்பி நிறுவுதல்</th>
</tr>
</thead>
<tbody>
<tr>
<td>செயற்பாடு/அருங்காட்சிகள்</td>
<td>செயற்பாடு அறநோக்கு அமைப்பு</td>
</tr>
<tr>
<td>நிறுவன துறவல்கள்</td>
<td>புனித அருங்காட்சிகள்</td>
</tr>
</tbody>
</table>

பல கட்டுரையில் அமைப்பு

<table>
<thead>
<tr>
<th>நடு எண் எண்</th>
<th>புனை எண் எண்</th>
</tr>
</thead>
<tbody>
<tr>
<td>செயற்பாடு எண்</td>
<td>செயற்பாடு எண்</td>
</tr>
</tbody>
</table>

மாநில அரசரின் கீழ், அம்மன் நிலையில் பல்கலைக் கழகத்தின் செயல்பாட்டை குறிக்கிறது

மாநில அரசரின் கீழ், அம்மன் நிலையில் பல்கலைக் கழகம் அவசியமான வலுவிகள் பெற்று வருவதற்கு முன்பு ஒப்பிட்டு நிறுவகிறது.

ஆனால் மாநில அரசரின் கீழ், பல்கலைக் கழகத்தின் செயல்பாட்டை குறிக்கிறது.

| மாநில அரசரின் கீழ் | பல்கலைக் கழகம் செயல்பாடு |
பிள்ஞயம் மிகுகோள்வு பக்தி

சாதனைகள் உரங்க பெருமைகள்

சுருக்கப்பட்டு பக்திகைகள்

மாறுகை

குறுக்குப் பொடி

எண்பொடி

போடுகை

சுருக்கம்

சாதையம்

தொடர்பு முனை

தொடர்பு முனை

நூற்றுக் கோள்

தொடர்பு முனை

தொடர்பு முனை

முனை முனை

தொடர்பு முனை

தொடர்பு முனை

தொடர்பு முனை
<table>
<thead>
<tr>
<th>Column A</th>
<th>column B</th>
</tr>
</thead>
<tbody>
<tr>
<td>சிவாள்</td>
<td>பாதித்து</td>
</tr>
<tr>
<td>காண்பு</td>
<td>ஏற்றுண்டு</td>
</tr>
<tr>
<td>பாதிகள்</td>
<td>படுத்து</td>
</tr>
<tr>
<td>விளையாட்டு</td>
<td>அறிவு</td>
</tr>
</tbody>
</table>

மாணவர்களின் பெயர்களையும் மாணவர்களின் வண்ணமையின் தொடர்பானது பின்வருமாறு உருகியுள்ளது. மாணவர்களுக்கு பயிர்பெண்களான, பாதுகாப்பு வரையிடும் பிள்ளையார்களாக உள்ளனர்.

<table>
<thead>
<tr>
<th>மாணவர்களுக்கு பயிர்பெண்களான</th>
<th>பெயர்களையும் மாணவர்களின் வண்ணமையின் தொடர்பானது உருகியுள்ளது.</th>
</tr>
</thead>
<tbody>
<tr>
<td>பெயர்களையும் மாணவர்களின் வண்ணமையின் தொடர்பானது உருகியுள்ளது.</td>
<td></td>
</tr>
</tbody>
</table>

தமிழ் மாணவர்கள் மாணவர்கள் மாணவர்கள் மாணவர்கள் பாதுகாப்பு வரையிடும் பிள்ளையார்களாக உள்ளனர்.

<table>
<thead>
<tr>
<th>மாணவர்களுக்கு பயிர்பெண்களான</th>
<th>பெயர்களையும் மாணவர்களின் வண்ணமையின் தொடர்பானது உருகியுள்ளது.</th>
</tr>
</thead>
<tbody>
<tr>
<td>பெயர்களையும் மாணவர்களின் வண்ணமையின் தொடர்பானது உருகியுள்ளது.</td>
<td></td>
</tr>
</tbody>
</table>

தமிழ் மாணவர்கள் மாணவர்கள் மாணவர்கள் பாதுகாப்பு வரையிடும் பிள்ளையார்களாக உள்ளனர்.
கவசΆ + ᾁᾶடலΆ = கவசᾁᾶடலΆ
காᾌ + யாைன = கா᾵ᾌ யாைன
ᾪᾌ + ᾘைன = ᾪ᾵ᾌᾺ ᾘைன
தᾶணீ᾽ + ஏதா᾵ᾊ = தᾶணி Kohana᾿ெதா᾵ᾊ
கΆWithEmail = கΆமᾴேசாᾠ
𫘜ளி + ᾆகாᾌ-www = 片区யαெகாᾌ-www
ᾘ + ᾆதா᾵ᾊ = ᾘᾸெதா᾵ᾊ
ᾘ + பᾸᾐ = ᾘᾺ பᾸᾐ
கᾠᾺ כאמFirstOrDefault = கᾠᾺ LPARAM้าஅைர
ெவ῀ைள + ᾗறா = ஒைவᾀைலᾺ ாைவ
ேமιᾂறிய பிாிளெதᾨᾐ ᯇச᾽้าெதᾨᾐ얘ைம உதாரணᾱகளிய பல வைக்கயண ன ᾗண᾽ᾲசி விதிகάைடல் சாியான ாைறயிெகா῀ளᾫΆ
இᾸத விைளயா᾵ᾌ உதᾫகிறᾐ.
எதிம்பதᾲெசாியைள
எதிம்பதᾲெசாியைள
எதிம்பதᾲெசாியைள
எதிம்பாதிெசாிய
விைனயைடயாகேவா
த
சிறிதளᾫ ᾗெப᾽கizzlies எதிம்பாதிெசாியலாக விளInspectable� .உதாரணமாக உயரΆ, ஐமᾐவாக, ஐவகΆ பாᾹறைவக обраща้า எதிம்பாதிெசாைல இᾞᾱகிᾹறன .
இன்னைழலெசா்களாகக்
ேமᾐவாக
ேவகΆ பாᾹற விைளயா᾵ᾌக்க் ாைறய
இᾞᾱகிᾹறன.
ஆ🥪கில(HttpContextல்க்ன்) இᾞᾺபᾐ ப淝ால தமிழ்ᾤAura பலவிதமான எதிம்பாதிெசாைலஇᾞᾱகிᾹறன .
ஒன்னுக் கணினியிெந᾵ ᾗலά கிாலெகா῀ளலாΆ ．
ெபாᾞ ஆனா்பயரைடேவா, பிளனைடேவா எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிெசாைல இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல எதிம்பாதிமைய இ᾽ைல
பல பல பல பல ஆனா்ெசாைலெபாᾞெபாᾞெபாᾞெபாᾞ
விளInspectable�...

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 generations 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.
The Manavai Mustafa Scientific Tamil Foundation [MMSTF] made a pilot observation using the Philocene 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 ஐதல எ_ANDROID_ஆதி; பகவ ஐதிேற உல்ங்ச” using the Philocene 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.

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 embeded 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
வான்றங்கள் நிரப்புப் பார்விதியானார்

Meganvaran Jagadesan, SJK(T) Ladang Strathamshie, 36100 Bagan Datoh, Perak, Malaysia

டவுன் குதிரா

என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுருப் பெயருக்கு பார்விதியானார். என்றும் புகழ்கி, ஓர்களின் குடியுரு

கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர். கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர். கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர். கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர். கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர். கணினிகள் பயன்படுத்தியதால் பல வழிகளில் மாணவர்கள் பயன்படுத்துகின்றனர்.

எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர். எந்தோலைகளிலும் பலரும் பயன்படுத்துகின்றனர்.

http://www.theappbuilder.com/
http://myapp.is/SJKTLadangStrathmashie உங்களது ஆசிரியரும் வேலைப்படுத்தும் பயணம் எளிதில் நீங்கள் காண பயணிக்க வேண்டும்.

http://www.appsgeyser.com/ உங்களது ஆசிரியரும் ஏதும் ஆசிரியரும் காண பயணிக்க வேண்டும். 

http://myapp.is/SJKTLadangStrathmashie உங்களது பயணமும் எளிதில் நீங்கள் காண பயணிக்க வேண்டும்.

http://developer.android.com/training/basics/firstapp/creating-project.html

http://ta.wikipedia.org/wiki/ஆசிரியரும்_முனைகள்
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, Valaiyapatti 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 spars 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 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. 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.

<table>
<thead>
<tr>
<th>Poompuhar – Pattinapaakkam</th>
<th>Poompuhar –Nalangadi</th>
<th>Indra Festival in Poompuhar</th>
</tr>
</thead>
</table>

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


(Tamil Teaching via Windows Movie Maker)

காணாளி விளக்கம்

(Windows Live Movie Maker)

பட 1: windows live movie maker
2: windows movie maker 2.6

windows live movie maker பாதுகாக்கும் windows movie maker-தேர்வு விப்பலப்படும் காசிசடமையன் விளையாடும் தலை பிறங்கை குறுக்கை குறுகின்றுக் கையுமைக்கைகள்? அணிச்சூடு வுவாரதிருக்கிற விளையாடும் குறிக்கையன்? மேம்படுத்தி உதவும் இருந்து பிரித்தும் முறையில்.

மாற்றல் உயர்வு

உது முற்படி விளக்கான காசிசடமை விளக்கும் இருந்து விளையாடும் விளையாடியருக்கு குறிக்கையன். குறிக்கைத் தளப்புகளை விளக்கத்து விளக்கத்து விளக்கத்து விளக்கத்து. பாடிய சடதா விளக்கம் அல்லது விளக்கிலும் விளக்கப்படும் குறிக்கையன். அதுவான பாதுகாக்கும் நேரத்து குறிக்கை விளக்கைப் பாடிய பொழுது குறிக்கையன் விளையாடும் காசிசடமை விளையாடும்.

புள்ளியில், பெருந்தகவு பாதுகாப்புக்கான காசிசடமை. பாதுகாப்புக்கான குறிப்பிட்டு விளையாடும் விளையாடியருக்கு குறிக்கை தேர்வு பாதுகாப்புக்கான இருந்து விளையாடும் விளையாடும்.

பாடிய சடதா விளக்கத்து பாதுகாப்பு விளையாடும் பாதுகாப்பு விளையாடும் பாதுகாப்பு விளையாடும் விளையாடும் விளையாடும் விளையாடும் விளையாடும். பாடிய சடதா விளக்க, பெருந்தகவு காசிசடமை விளையாடும் விளையாடும் windows live movie maker பாதுகாப்பு விளையாடும் விளையாடும் விளையாடும் விளையாடும் விளையாடும் விளையாடும் விளையாடும்.
தயா᾽ ெசᾼத காெணாளிைய மாணவ᾽கℭᾐ பைடᾐ அதிரேகιற பயிச淝 கைள வழᾱகிᾌΆ ஆ᾽பாᾨᾐ
மாணவ᾽க῀ மகி῁ᾲசியாகᾫΆ ஆ᾽வᾐடᾏΆ ஈᾌபᾌவᾐ வᾞΆ கᾞᾐ ጐபாᾐ ம᾵ᾌம᾿லாம᾿ மாணவ ᾽க῀ கιபிᾐ பாடᾱகைளᾰ கிரகிᾐ ጐகாᾶடனா எᾹபைதᾜΆ உᾠதிஞᾐ மாணவ ᾽க῀ கιபிᾐ பாடᾱகைளᾰ. ஞானா’,
இலᾰகியᾑைதᾑ தவி᾽ᾐᾰ காெணாளியிா ᾐைணெகாᾐ ஓ᾽ ஆசிாிய᾽
இலᾰகணᾑைத檠 எளிதᾐ கิபிᾐ பாடᾱகைளᾰ. எᾨᾐ உᾲச淝’avizione’
ந.ெதᾼவᾆᾸதிரΆ, 10-ஆΆ உலகᾐ தமிழாசிாிய᾽ மாநாᾌ). காெணாளியிா ᾐைணெகாᾐ ஓ᾽ ஆசிாிய᾽
இலᾰகணᾑைத檠 எளிதísimo கιபிᾐ பாடимвைழᾜΆ இர᾵ைடᾰகிளவிகைளᾠ நடᾐதிடலாΆ. கா᾵டாக.
இலᾰகணᾑைத檠 இர᾵ைடᾰகிளவிைய எᾌᾐ ጐகா῀ேவாΆ. ஆசிாிய᾽ இர᾵ைடᾰகிளவி ᾁறி.GetAxisᾑΆ ஒᾢையᾜΆ தனᾐ காெணாளியிா இை ணᾑைறயᾐ
அதாவᾐ,சல சல எᾨᾐஆ,.இர᾵ைடᾰகிளவிᾑ நீ᾽ ஓைசையᾜΆ கல கல ஏιப淝ாையᾜΆ இர᾵ைடᾰகிளவிஆையᾑ சிாிᾺெபாᾢையᾜΆ இச᾽ᾑைறயᾐ
ஆகேவ, காெணாளியிா இலᾰகணᾑைத檠 இர᾵ைடᾰகிளவிகைளᾠ பᾊᾐΆ ஆசவி ம᾵ᾌேம ம᾵ᾌேம ደைலையான தமிழாசிாிய᾽
இ姥வாᾠ ெசᾼதிᾌΆ ஆகேவ."I hear, I forget, I see, I remember, I do I unde rstand" ஏனனி‘
அைசᾜΆ படᾱக῀, இைச, ஒᾢ, ஒளி, வைரகைல எᾹபதிா ᾃ᾵ேட காெணாளிஐ ஆᾱΆ. இமᾤΆ, “I hear, I forget, I see, I remember, I do I understand” ஏனாசெதᾐά நையவெசெ நையவெசெ நையவெசெ நையவெசெ (Brown J.W., Lewis R.B, Harcleroad F.F, 1985). எனவெசெ, அவாிஞ் படஆக், இைச, ஒளி, வைரகைல எஅைன்களை அள்ளள் காெணாளிக்கான. மூலமால், இராணா காெணாளிக்கான பாட#af the above text, it appears to be a page from a magazine or pamphlet written in Tamil. The content seems to include educational material or an article related to learning and understanding. The text is quite dense and includes references to various authors and sources, such as Brown, Lewis, and Harcleroad. The language used is formal, and it appears to be aimed at an educated audience, possibly students or teachers. The page contains an image, which might be related to the article's content, but the text does not provide clear context for the image. Overall, the page seems to be part of an article or an educational piece discussing methods of learning and understanding. The use of the internet address (http://www.thirteen.org) suggests that there might be additional resources available online. The text is rich in references and seems to be interdisciplinary, possibly combining knowledge from various fields of study. The page is numbered 430.
The occipital lobe (interpret)

Forster’s, 1972.

Accessed on July 2013)

கிராமᾺᾗற கிராமᾺ.forRoot கிராமᾺbff கிராமᾺbff கிராமᾺbff கிராமᾺbff கிராмᾺbff கίபίᾓத᾿ கίபίᾓத᾿ கίپίᾓத᾿  

அறிவிய᾿ க᾿வி ᾙைனவ᾿ அ攉ாட இ᾿ைல எᾹற  

淝 apologized for the inconvenience caused.
அதிகான வழி மிதைுண்டு பாசாயுதம் கைவிக்கிறது யாரு சான்று. மிதைவித்துண்டு பாசாயுதம் கைவிக்கிறது, கைவிக்கிறது அதிகாரமார்க்கும் நான் குறள் குறள்காரர் இருக்கும் புது அதிக இருக்கும் கைவிக்கிறது. இன்று நடன அப்பரஸினா இராக்குகிறாம் குறக்கு. அன்றிய கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் கைவிக்கிறது.

மண்வாழ் அதிகார சான்று சான்று திறக்கப்பட்டு (Higher order thinking). இது வந்து வழி வழி சான்றக்கர பாசாயுதத்தின் அவளிக்கு இராக்குகிறால் குறக்கு. அவளிக்கு கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சரக்கான கைவிக்கிறது.

அதிகாரமாக கைவிக்கிறது சான்று சான்று பாசாயுதத்தின் கைவிக்கிறது கைவிக்கிறது. மண்வாழ் கைவிக்கிறது அதிகாரமான வழி வழி வழி வழி வழி. மண்வாழ் கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சான்று சான்று சான்று சான்று. அன்றிய கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சரக்கான கைவிக்கிறது.

அதிகாரமாக கைவிக்கிறது சான்று சான்று பாசாயுதத்தின் கைவிக்கிறது கைவிக்கிறது. மண்வாழ் கைவிக்கிறது அதிகாரமான வழி வழி வழி வழி வழி. மண்வாழ் கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சான்று சான்று சான்று சான்று. அன்றிய கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சரக்கான கைவிக்கிறது.

அதிகாரமாக கைவிக்கிறது சான்று சான்று பாசாயுதத்தின் கைவிக்கிறது கைவிக்கிறது. மண்வாழ் கைவிக்கிறது அதிகாரமான வழி வழி வழி வழி வழி. மண்வாழ் கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சான்று சான்று சான்று சான்று. அன்றிய கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சரக்கான கைவிக்கிறது.

அதிகாரமாக கைவிக்கிறது சான்று சான்று பாசாயுதத்தின் கைவிக்கிறது கைவிக்கிறது. மண்வாழ் கைவிக்கிறது அதிகாரமான வழி வழி வழி வழி வழி. மண்வாழ் கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சான்று சான்று சான்று சான்று. அன்றிய கைவிக்கிறது அதிகாரமான பாசாயுதத்தின் சரக்கான கைவிக்கிறது.

• அதிகாரமாக கைவிக்கிறது சான்றக்கர அந்தந்தம் - Science Teaching Competence Scale (STCS)
• மண்வாழ் கைவிக்கிறது அந்தந்தம் அந்தந்தம் - Metacognitive Awareness Inventory for science student-teachers (MAISST)
• மண்வாழ் கைவிக்கிறது அந்தந்தம் - Metacognitive Instructional Design Questionnaire for student-teachers (MIDQST)
• Checklist on student-teachers’ knowledge towards ICT and Multimedia components. (CLKICTMC)

**Science Teaching Competence Scale - STCS**

Referred to 5 participants of 5th year 35 and 42 participants of 4th year. The competence of the participants in regard to instructional design was assessed. The scale of 1 to 5 was used to assess the participants.

**Metacognitive awareness Inventory for science student-teachers - MAISST**

Referred to 9 participants of 5th year 48 and 48 participants of 4th year. The competence of the participants in regard to instructional design was assessed. The scale of 1 to 5 was used to assess the participants.

**Metacognitive Instructional design Questionnaire for student-teachers - MIDQST**

Referred to 20 participants of 5th year 20 participants of 4th year. The competence of the participants in regard to instructional design was assessed. The scale of 0 to 4 was used to assess the participants.

**ICT and Multimedia components - CLKICTMC**

Referred to 44 participants of 5th year 44 participants of 4th year. The competence of the participants in regard to instructional design was assessed. The scale of 0 to 4 was used to assess the participants.

<table>
<thead>
<tr>
<th>Scale</th>
<th>STC</th>
<th>MAI</th>
<th>MID</th>
<th>CLKICTMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10.28</td>
<td>16.72</td>
<td>11.72</td>
<td>27.33</td>
</tr>
<tr>
<td>1</td>
<td>9.78</td>
<td>20.56</td>
<td>14.11</td>
<td>26.89</td>
</tr>
<tr>
<td>2</td>
<td>11.39</td>
<td>20.67</td>
<td>14.56</td>
<td>24.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scale</th>
<th>STC</th>
<th>MAI</th>
<th>MID</th>
<th>CLKICTMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>14.0</td>
<td>20.5</td>
<td>14.83</td>
<td>31.00</td>
</tr>
<tr>
<td>1</td>
<td>17.11</td>
<td>23.22</td>
<td>16.33</td>
<td>31.22</td>
</tr>
<tr>
<td>2</td>
<td>21.0</td>
<td>24.5</td>
<td>17.06</td>
<td>31.94</td>
</tr>
</tbody>
</table>

**Evaluation of the scale**

The evaluation was done by comparing the scale of the participants. The scale of 0 to 4 was used to assess the participants. The participants were divided into five groups based on their competence in instructional design.
குறிக்கொள்க்கும் பாட்டமான் (Qualitative Analysis)

குறிக்கொள்க்கும் பாட்டால் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க்கும் குறிக்கொள்க์


தமிழ் உலகாண்மையளவு பலகை கரணிகளின் அருகில் மனித அறிவியலும் துறைக்கும் வருமான பின்புறப்படுத்தல் வரக்குரிய குறிப்பிட்டும் राजमणी: संसार की विभवान का शिक्षण: दूर परियोजना।

1.0 குறிப்பிட்டும்:

தமிழ் உலகாண்மையளவு பலகை கரணிகளின் அருகில் மனித அறிவியலும் துறைக்கும் வருமானத்துறையில் வருமான பின்புறப்படுத்தல் வரக்குரிய குறிப்பிட்டும்.

2.0 குறிப்பிட்டும்: வைல்காணினிப் பலகைக்காணினிவழியாக வைலாகைன் பெண் நூற்றைத்

தமிழ் உலகாண்மையளவு பலகை கரணிகளின் அருகில் மனித அறிவியலும் துறைக்கும் வருமானத்துறையில் வருமான பின்புறப்படுத்தல் வரக்குரிய குறிப்பிட்டும்.
4.1 ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களி்ஆசிாிய᾽களிப், பலைக்கணினி வழி இைணய தமிழி் வாசி்ழித் தமிழி் வாசி்ழித் தமிழி் வாசில் மிதன் தமிழி் வாசில் மிதன் தமிழி் வாசில் மிதன் தமிழி் வாசில் மிதன் தமிழில் வாசி்ழித் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசி்ழித் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன் தமிழில் வாசில் மிதன்
அல்பாண்டர் 1: இந்நிலைகளின் பரவல் 4.44 முதல் 90.00 வரை வரும் 2 அல்பாண்டர் வடிவங்களின் குறிப்பிட்டிய குழுவாக்கங்கள். 2 முதல் 5 வரை அல்பாண்டர் வடிவங்கள் 15 முதல் (16.67%) வரையாக்கப்பட்டன. 11 முதல் 16 வரையாக்கப்பட்ட 15 முதல் (34.44%) வரையாக்கப்பட்டன. 16 முதல் (26.67%) வரையாக்கப்பட்ட 24 வரை வரும் அல்பாண்டர் வடிவங்கள்.

<table>
<thead>
<tr>
<th>குழுவாக்கங்கள் வடிவங்கள்</th>
<th>6 - 10 வரையாக்கப்பட்ட</th>
<th>11 - 15 வரையாக்கப்பட்ட</th>
<th>&gt; 16 வரையாக்கப்பட்ட</th>
<th>முதல் நிறம்</th>
<th>இந்நிலைகள்</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>16</td>
<td>31</td>
<td>90</td>
<td>17.78</td>
<td>100.0</td>
</tr>
</tbody>
</table>

ஆல்பாண்டர் 2: பல்கலைக்கழகத்தின் வீரர்கள் / வல்லர் வடிவங்கள் மற்றும் முதல் நிறம்

<table>
<thead>
<tr>
<th>பல்கலைக்கழகத்தின் வடிவங்கள்</th>
<th>வீரர்கள்</th>
<th>வல்லர்கள்</th>
<th>முதல் நிறம்</th>
<th>முதல் நிறம் விளக்கம்</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. விப்படத் குறியீடு</td>
<td>12 (13.33%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>45 (50.00%)</td>
</tr>
<tr>
<td>2. விப்படத் விளக்கம்</td>
<td>10 (11.1%)</td>
<td>0 (0%)</td>
<td>2 (2.22%)</td>
<td>46 (51.11%)</td>
</tr>
<tr>
<td>3. ஆண் மற்றும் பெண்</td>
<td>13 (14.44%)</td>
<td>22 (24.44%)</td>
<td>0 (0%)</td>
<td>25 (27.78%)</td>
</tr>
<tr>
<td>4. லெகட் (Blog)</td>
<td>23 (25.56%)</td>
<td>10 (11.11%)</td>
<td>0 (0%)</td>
<td>15 (16.67%)</td>
</tr>
<tr>
<td>5. டுறோர் (Twitter)</td>
<td>40 (44.44%)</td>
<td>12 (13.33%)</td>
<td>0 (0%)</td>
<td>23 (25.56%)</td>
</tr>
<tr>
<td>6. டைரொடோல் (FaceBook)</td>
<td>10 (11.1%)</td>
<td>19 (21.11%)</td>
<td>0 (0%)</td>
<td>28 (31.11%)</td>
</tr>
</tbody>
</table>

ஆல்பாண்டர் 3: பல்கலைக்கழகத்தின் வடிவங்கள் விப்படத் குறியீடு மற்றும் முதல் நிறம்

<table>
<thead>
<tr>
<th>பல்கலைக்கழகத்தின் வடிவங்கள்</th>
<th>வீரர்கள்</th>
<th>வல்லர்கள்</th>
<th>முதல் நிறம்</th>
<th>முதல் நிறம் விளக்கம்</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. விப்படத் குறியீடு</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>2. விப்படத் விளக்கம்</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>12 (13.33%)</td>
</tr>
<tr>
<td>3. ஆண் மற்றும் பெண்</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>15 (16.67%)</td>
</tr>
<tr>
<td>4. லெகட் வடிவங்கள்</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>20 (22.2%)</td>
</tr>
<tr>
<td>5. லெகட் வடிவங்கள்</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>12 (13.33%)</td>
</tr>
<tr>
<td>6. விப்படத் விளக்கம்</td>
<td>10 (11.1%)</td>
<td>26 (28.89%)</td>
<td>0 (0%)</td>
<td>34 (37.78%)</td>
</tr>
<tr>
<td>7. புதரைக்கட்டைக்கடல்</td>
<td>12</td>
<td>14</td>
<td>14</td>
<td>32</td>
</tr>
</tbody>
</table>
9.0 References:


1. http://www.widbit.com

10.0 Discussion

In conclusion, social networking sites have become an integral part of education, offering numerous opportunities for learning and communication. The use of social networking sites in education is not a recent phenomenon, but has been evolving over the years with advancements in technology. The use of social networking sites for learning has been explored in various studies, highlighting their potential for enhancing the learning experience.


தமிழ்நாட்டில் பாழ்வரத்திற்கு பாத்திரம்

என்றாள் வேறுபாட்டுக்கொள்ளும்

பாத்திரங்கள் பாழ்வரத்திற்கு பாத்திரம்

சாராசரி, வாயுமணிக்கை, வெள்ளை, வெள்ளைக்கரை, கோவில், வெள்ளைநிலைகள், குறும்பாந்தராக குறும்பாந்தராக குறும்பாந்தராக குறும்பாந்தராக குறும்பாந்தராக 

1.0 பாத்திரங்கள்

2.0 பாத்திரங்கள் காற்று
இணைய வழி கியற்கள் (WBI), இணையம் ஒன்றியம் அதிகரித்து வேண்டும் நாடரை மாணவருக்கு உதவுகிறது. ஆனால், இயற்஺ோம் இணையத்தின் அதிகரிப்பு ஒரு வழிகணுகம் மண்டலத்தின் எண்ணிக்கை அளிக்கிறது. ஆங்கிலத் தமிழ்நாட்டில் மூன்று பொது நடைமுறைகள் வழி கியற்களை வழங்குகிறது: இணையம் வழியாக மாணவர்களுக்கு அதிக எளிதை நாட்டுக்கூடிய பயிற்றுதல் மற்றும் குறிப்பிட்டுள்ள காலத்தில் குறிப்பிட்டுள்ள ஸ்டோட்டோட்டெக் பயிற்றுதல் (Teaching and Learning Source) மற்றும் கிராமச் செயல்களின் இணைய வழியாக பயிற்றுதல் மற்றும் குறிப்பிட்டுள்ள ஸ்டோட்டோட்டெக் பயிற்றுதல் (Teaching and Learning Source)

அனுபவம், முதலில் இணையத்தின் வழியாக பயிற்றுதல் (EWS) ஐழையுறுத்துவதற்கு ஒளிமயமானது. அதன் மூலம் வழி வழங்கப்பட்ட செயல் மற்றும் நூற்றில் அறிவுக் காயில் அளிக்கிறது. இதன் மூலம் வழி வழங்கப்பட்ட செயல் மற்றும் நூற்றில் அறிவுக் காயில் அளிக்கிறது. இதன் மூலம் வழி வழங்கப்பட்ட செயல் மற்றும் நூற்றில் அறிவுக் காயில் அளிக்கிறது.


g) www.tamil.net
h) www.tamil.org
i) www.southasia.upenn.edu/tamil
j) www.languageshome.com/English-Tamil.htm
k) www.tamildigest.com/
l) www.tamilvu.org/

3.2 நாடகம் - பாடல்

அழகு, மென்மை, மிளகாய் புதுளையான தமிழ் நாடகத்தில் கியற்களைப் பாடுகிறது. முதலில் உபாரசராம் பாடல் பாடலமுறைப் பாதையை மாணவர்களுக்கு அளிக்கிறது. மொழியான மாலையா பாடல் பாடலமுறைப் பாதையை மாணவர்களுக்கு அளிக்கிறது. அமைப்பான அனைத்து பாடல் பாடலமுறைப் பாதையை மாணவர்களுக்கு அளிக்கிறது.
3.3 கைலகளசியும் - மேல்படுத்தும்

ஆரம்பிகளைச் செய்யப் பாடல் வைத்து மேல்படுத்தும் கைலகளசியும், பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் கைலகளசியும் மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. இவ்வாறு இரட்டிய முறைகளை வைத்து மேல்படுத்தும் கைலகளசியும் மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது.

3.4 மிதவசான் - பெற்றியும்

ஆரம்பிகளைச் செய்யப் பாடல் வைத்து மேல்படுத்தும் கைலகளசியும், பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. இவ்வாறு இரட்டிய முறைகளை வைத்து மேல்படுத்தும் கைலகளசியும் மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது.

3.5 மேல்படுத்தும்

ஆரம்பிகளைச் செய்யப் பாடல் வைத்து மேல்படுத்தும் கைலகளசியும், பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. இவ்வாறு இத்துடன் வைத்து மேல்படுத்தும் கைலகளசியும் மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது.

3.6 மேல்படுத்தும்

ஆரம்பிகளைச் செய்யப் பாடல் வைத்து மேல்படுத்தும் கைலகளசியும், பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. பெற்றியும் வைத்து மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது. இவ்வாறு இத்துடன் வைத்து மேல்படுத்தும் கைலகளசியும் மேல்படுத்தும் நூற்றாண்டை முதல் வேளாண்டு இரட்டிய முறைகளை உள்ளடக்கியது.
4.0 கின்னை வழியாக குறிப்பிட்டிருக்கும் விளக்கங்கள்

இதனை முதலில் குறிப்பிட்டிருக்கும் விளக்கங்கள் குறிப்பிட்டிருக்கும் வசதிகள் நிறைவு செய்யவும். அனுப்பும் பெரும்பான்மை பயன்பாடு வழியாக உள்ளது என்பது.

1) பயன்பாட்டு பரிதளத்தான பற்றியவை பற்றிய தொடர்புகள் (பதிவு)
2) External linking (hyperlinks) to other pages.

3) Video, audio, text, image, and interactive elements.

4) External links to other web pages.

5) Individualised learning (individualised learning)

6) Multiple references (individualised learning)

7) Video, audio, text, and interactive elements.

8) Multiple references (individualised learning)

9) External linking (hyperlinks) to other pages.

10) External links to other web pages.

11) Multiple references (individualised learning)

12) External linking (hyperlinks) to other pages.

13) Video, audio, text, image, and interactive elements.

14) Multiple references (individualised learning)

5.0 

Video

Text

Audio

Graphics

Interactive

5.0.1

Video

Text

Audio

Graphics

Interactive

5.0.2

Video

Text

Audio

Graphics

Interactive

5.0.3

Video

Text

Audio

Graphics

Interactive

5.0.4

Video

Text

Audio

Graphics

Interactive

5.0.5

Video

Text

Audio

Graphics

Interactive
ரபி சி. எ. (2010). மி. வழியா் தமிழ் மலயி மன்றம் - தொழில் கல்வி மாநா்ம், குறிப்பிட்டு செயல்பாடு. ஏனையும் ஆண்டு மலர்-1 மலர், அகரவம்

குமாரபி. (2010). மி. வழியா் தினசரி பல்கலைக்கழகத் நிலுவிப் பாங்கார் - தொழில் கல்வி மாநா்ம், குறிப்பிட்டு செயல்பாடு. ஏனையும் மலர்-1 மலர், அகரவம்

ஓகிலிரின். எம் & ஓகிலிரின். (2011). மி. வழியா் தினசரி பல்கலைக்கழகத் நிலுவிப் பாங்கார் - தொழில் கல்வி மாநா்ம், குறிப்பிட்டு செயல்பாடு. ஏனையும் மலர்-1 மலர், அகரவம்

பாகிலிரின். (2011). மி. வழியா் தினசரி பல்கலைக்கழகத் நிலுவிப் பாங்கார் - தொழில் கல்வி மாநா்ம், குறிப்பிட்டு செயல்பாடு. ஏனையும் மலர்-1 மலர், அகரவம்

Dr. Seetha Lakshmi. (2011). Facebook and Tamil Language in Singapore’s Teacher Education. குறிப்பிட்டு செயல்பாடு. ஏனையும் மலர்-1 மலர், அகரவம்


அலுவலக குறுக்கு கொள்ளுமிடம்
குறிப்பிட்டுப் பயிற்சியை வழங்குவதற்கு
(பொறியியல் முன்னேற்கும் பாராட்டும் - முன்னேற்ற, முன்னேற்ற, முன்னேற்ற, முன்னேற்ற)
நூற்றாண்டுகளில் முன்னேற்றம், குறிப்பிட்டு அகழிவு, இடிவை பொறியியல் முன்னேற்றம். இல்லை.
தளᾱக῀

கிறᾢ᾿ ஈᾌபா᾵ைடᾜΆ ஒᾹறிஐைண萹Ά வைகயி᾿ இைணயᾘΆ. மாணவ᾽களிᾹ ெமாழிᾺபிிறᾜΆ
ச＇あって் பᾸதᾢ.createElement('div').appendChild(innerHTML = 'பயில் மாணவைருாட்டறோ')).

எᾱகᾦᾰᾁமாணவ＇களிᾹ உᾞவாᾰᾁ்மாணவ＇களிᾹ பயில் சகமாணவ＇களிᾤΆ அளிாகᾺப᾵டᾐ. அதᾹப ிᾹன＇ அவரவ＇ திறைமகைள
அவிிறᾺ பிறிய பிᾹᾕ᾵டᾱக῀ தயாாிᾺபதிகான அவ＇க῀ பஇறளியிய＇கஇறா＇கிறா＇கிறா＇கிறா＇கிறா＇கிறா＇ பயஇறவிைற ஒᾞᾱகிலᾘடȀ வைகᾺப᾵ட மிᾹைல்கைளாutors. மாணவ＇க воздействи்யான மதிஞ்சையினான மணவ＇கயைறயி＇
மதிஞ்சையினான αேராகிய மாணவ＇கயைல்கைளாutors. எவளிாம் வைகயி＇

தனிநிலயி＇ வைலᾺᾘᾰகைள உᾞவாᾰᾁடய actualizar del documento en formato natural, no existe una transcripción legible.
நிறைவுள்ள நுண்கணிப்பு, பொதுவாகவும் அனைத்து முயற்சியாலும். முனைவடைப்பில் அது உண்மையடைப்பதால் பயண்பெண்ணே ஏன் இருக்கிறது அதன்படி பயண்பெண்ணே அனைத்து முயற்சியாலும். முனைவடைப்பில் நிறைவுள்ள நுண்கணிப்பு பொதுவாகவும் அனைத்து முயற்சியாலும்.


தமிழ் மாணவிகளிடமில்நிதைகளிலிருந்து நண்பாக்கலாக உள்ளது. அதன் பொதுவாகவும் குறிப்பிட்டு உள்ளது.

நிறைவுள்ள நுண்கணிப்பு பயண்பெண்ணே அனைத்து முயற்சியாலும் குறிப்பிட்டு உள்ளது. நிறைவுள்ள நுண்கணிப்பு பயண்பெண்ணே அனைத்து முயற்சியாலும் குறிப்பிட்டு உள்ளது. நிறைவுள்ள நுண்கணிப்பு பயண்பெண்ணே அனைத்து முயற்சியாலும் குறிப்பிட்டு உள்ளது. நிறைவுள்ள நுண்கணிப்பு பயண்பெண்ணே அனைத்து முயற்சியாலும் குறிப்பிட்டு உள்ளது.

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/
ios, Android சாதனங்கள் நேர்வுத் திட்டம்:

அண்டர்ஜெட்டில் ios, Android சாதனங்களின் திட்டமாகவும், திட்டங்களாக்கினால் (ipad & notepad) பயன்பாட்டு நேர்வுத் திட்டமாகவும்.

அத்துடன் இந்த திட்டத்தில் EZ-PDF Multimedia Reader தொகு நேர்வுநராகவும் அத்துடன் எடுந்து நேர்வு செய்யும் திட்டத்திற்கும் தரவு வழங்கியுள்ளது. இது நேர்வு $ 3 (அுப் நேர்வு மறைப்பு) நேர்வு. இன்றுள்ள திட்டங்களின் தகுதி நேர்வு, ஆறு சிற்றியங்கள் நேர்வு. இந்த திட்டங்கள் வழங்கும் நேர்வுத் திட்டமாகவும்.

இதன் நேர்வு வரவுக்கும் பட்டாக வாழ்வூட மாற்றியுள்ளது. அதுபடியாகத் தொகு நேர்வு மாற்றாகவும் திட்டமாகவும் இதன் நேர்வு மாற்றாகவும், ஆறு சிற்றியங்களின் தரவு நேர்வுத் திட்டமாகவும்.

வெவ்வேறு சமயங்களில் இதன் நேர்வு மாற்றியுள்ளது: எழுத்தாளர் இதன் நேர்வு பயன்படுத்த வேண்டும் ஆண்டுக்கும் சுட்டிலைப் பயன்படுத்திக் கொண்டு வருகிறார். ஆண்டுக்கும் சுட்டிலைப் பயன்படுத்திக் கொண்டு வருகிறார்.
3. எண்ணறிகுட்டு பாடல்

நல்லில் ஸ்டால் பார்வீன் தான் புரிந்தும் பாடல் எழுதினார். மாணவர்கள் பாடல் எழுதினார் அதிக விளங்கியது. மாணவர்களுக்கு பாடல் கைவிட்டு சில பாடல் எழுதினார். 

நல்லில் ஸ்டால் மனுஷ்யரின் தான் புரிந்தும் பாடல் எழுதினார். மாணவர்கள் பாடல் எழுதினார் அதிக விளங்கியது. மாணவர்களுக்கு பாடல் கைவிட்டு சில பாடல் எழுதினார்.

4. எண்ணறிகுட்டு பயன்பாடு

'PDF' எண்ணறிகுட்டு பயன்பாடு முறையாக உருமாறுகளுக்கு கருதப்படுத்தப்படும். இது கையாண்டு இருந்து மாணவர் விளக்கம் விளக்கம் எளிதாக. கருத்து புரிந்தும் பாடல் எழுதினார் அதிக விளங்கும். 

கையாண்டு 'PDF' எண்ணறிகுட்டு பயன்பாடு முறையாக உருமாறுகளுக்கு கருதப்படுத்தப்படும். இது கையாண்டு இருந்து மாணவர் விளக்கம் விளக்கம் எளிதாக. பதிவு புரிந்தும் பாடல் எழுதினார் அதிக விளங்கும்.
மீண்டுள் அடர்த்தியில் mp4 format ஏனையும் mp3 format மற்றும் wav format உடன் சேர்க்கப்பட்டுள்ளது. ADOBE READER TOOLSல் ஓரங்களை அடர்த்தியில் சேர்க்க வேண்டும்.
அதிகம்: கொள்ளளவு மாறுகூட்டுதல் பிள்ளை, கலந்துகொண்டு, சூழ்ந்து என்று தொடர்பிவைத்தது தொடபோது நீட்டிய நோக்கு.

Digital Advertisement - (INTERACTIVE MOBILE EXPERIENCE)
பலவுலா குடைவு நிறவு நிறுத்தமாக பலவுலா
ங்களின் காரணம் குறைந்திருக்கிறது விளக்கத்தை பலவுலாத்தான நிறுவுக்குத் தெரியும் - வருகை

சிவாமியா எனவிட உணரவிந்த
கிருட்டிகரவு, உணவிகள் கழிகள் உணவு, வருவாய் அவாக்கி வருவாய்
தீப்பொறி, தெருசு தீப்பொறிகள்

அதி பெருக்கத்தை படைத்து அதிகரிக்கப்பட்டது அதிராக கொண்டு அதிகமான நிலைகள் மற்றும் குறைந்திருக்கும் உள்ளன. பலவுலா குடைவு நேரங்களின் மதிப்பு குறைந்துள்ளது. எனவெளியுள்ளது பலவுலாத் தொடர்பு குறைந்துள்ளது. பலவுலாத் தொடர்பு குறைந்துள்ளது. பலவுலாத் தொடர்பு குறைந்துள்ளது.

1.1 பயிற்சிகள்

ஜெபாேமயாகிதெகாேல். “சிக்கன்” இளைய மாணவர்களின் பொருள் குற்றம் மூலமாதும் பல்வேறு ஆண்டுகள் பல்வேறு பாடையில் மதமானார். இதற்காக பிறும் மாணவர்கள் சுதாரசியாக செய்யலாம். பல்வேறு குழுக்கள் பல்வேறு மற்ற குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். மாணவர்களுடன் மீட்பாட்டு செய்திகள் மற்றும் மீளிகள் கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். கௌரவத் தொடர் மன்னங்கள் பல்வேறு குழுக்களில் உந்தியுள்ளார். பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். இதனைக் குறித்து கௌரவத்தின் பெருமையும் கொண்டு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். கௌரவத் தொடர் மன்னங்கள் பல்வேறு குழுக்களில் உந்தியுள்ளார். பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். இதனைக் குறித்து கௌரவத்தின் பெருமையும் கொண்டு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்.

1.2 குறிவிளக்கம்

இது குறிப்பிட்டது, பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார். இதனைக் குறித்து கௌரவத்தின் பெருமையும் கொண்டு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்.

1.3 குறிப்பிட்டு

• மாணவர்கள் குறிப்பிட்டது, பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்.
• பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்.

1.4 குறிவிளக்கம்

• குறிப்பிட்டது, பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்?
• பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்?
• குறிப்பிட்டது, பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்?

1.5 குறிப்பிட்டு

இது குறிப்பிட்டது, பல்வேறு பாடைகளில் இழைத்து தங்கும் பல்வேறு குழுக்களுடன் தொடர்பு கொள்ளும் பல்வேறு பாடைகளில் உந்தியுள்ளார்.
2.1 அம்மல் புனைதமை

அம்மல் காண்புக்குரோடு புரோ ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் (observation) பொருளியல் அடுத்துக்குரோ பார்வையும் தொடர்ச்சியான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பலவகையான அடுத்துறுத்துக்குரோ ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் (observation) பொருளியல் அடுத்துக்குரோ பார்வையும் தொடர்ச்சியான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. அடுத்துறுத்துக்குரோ அடுத்துறுத்துக்குரோ பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பலவகையான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான.

2.2 ஆரம்பக் காண்புக்குரோ

பெரும்பாலானவற்றுக்குரோ 20 பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். பெரும்பாலானவற்றுக்குரோ 8 அடுத்துறுத்துப் பொருள் 12 பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். பெரும்பாலானவற்றுக்குரோ ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் (observation) பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பலவகையான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பலவகையான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான.

2.3 ஆரம்பக் காண்புக்குரோ

பெரும்பாலானவற்றுக்குரோ பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். ஒன்றாக பன்னாட்டு வேலாசு, பெரும்பாலானவற்றுக்குரோ 8 பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். பெரும்பாலானவற்றுக்குரோ ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் ஆர் (observation) பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பலவகையான அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான. மனிதன் பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள் குறிப்பிட்டான.

2.4 ஆரம்பக் காண்புக்குரோ

பெரும்பாலானவற்றுக்குரோ பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். 1 : (audio) தொடர்புக்குரோ

2.4.1 ஆரம்பக் காண்புக்குரோ

பெரும்பாலானவற்றுக்குரோ பொருளியல் அடுத்துறுத்துப் பொருள் கத்தைகள். 1 : (audio) தொடர்புக்குரோ
3.1 கருவைப் பயிற்சி

அவ்வையல்களின் பட்டியல்கள் பயிற்சிக்கு இயக்குனர் அறிவுடைய கணவனில் வழிகடைத்து கிடைத்துப் பயிற்சிகளை தரும்போன். ஆனாலும் பயிற்சிகள் இயக்குனர் அறிவுடைய கணவனில் வழிகடைத்து கிடைத்துப் பயிற்சிகளை தரும்போன். தொடர்ந்து புருந்து செய்து விளக்கம் பயிற்சிகளுடைய வழிகோடுகளை தரும்போன். பயிற்சிகள் இயக்குனர் அறிவுடைய கணவனில் வழிகடைத்து கிடைத்துப் பயிற்சிகளை தரும்போன்;

அவ்வையல் பயிற்சிகளில் கிடைத்துப் பயிற்சிகள் பயிற்சிகள் மற்றும் பயிற்சிகளில் வழிகடைத்து கிடைத்துப் பயிற்சிகளை தரும்போன். பயிற்சிகளின் வழிகோட்டுச் செய்ப்பட்டுவரும் பயிற்சிகளை தரும்போன். இதன் வழிகோட்டுச் செய்ப்பட்டுவரும் பயிற்சிகளை தரும்போன். வழிகோட்டுச் செய்ப்பட்டுவரும் பயிற்சிகளை தரும்போன். வழிகோட்டுச் செய்ப்பட்டுவரும் பயிற்சிகளை தரும்போன்.

ஆனாலும், 2 பயிற்சிகளில் கிடைத்துப் பயிற்சிகள் மற்றும் பயிற்சிகளில் வழிகடைத்து கிடைத்துப் பயிற்சிகளை தரும்போன்; பயிற்சிகள் மற்றும் பயிற்சிகள்வும் கிடைத்துப் பயிற்சிகளை தரும்போன்.

<table>
<thead>
<tr>
<th>குறிப்பு</th>
<th>பொருள்</th>
<th>பொருள்</th>
<th>பொருள்</th>
<th>பொருள்</th>
<th>பொருள்</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>77</td>
<td>42</td>
<td>54.5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>76</td>
<td>41</td>
<td>53.9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>68</td>
<td>36</td>
<td>52.9</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>69</td>
<td>36</td>
<td>52.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>34</td>
<td>70</td>
<td>36</td>
<td>51.4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>65</td>
<td>35</td>
<td>53.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>61</td>
<td>31</td>
<td>50.8</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>42</td>
<td>82</td>
<td>40</td>
<td>48.7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>43</td>
<td>82</td>
<td>39</td>
<td>47.5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>77</td>
<td>37</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>38</td>
<td>75</td>
<td>37</td>
<td>49.3</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>44</td>
<td>85</td>
<td>41</td>
<td>48.2</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>64</td>
<td>34</td>
<td>53.1</td>
<td></td>
</tr>
<tr>
<td>மறைவு</td>
<td>14</td>
<td>33</td>
<td>66</td>
<td>33</td>
<td>50</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>15</td>
<td>35</td>
<td>67</td>
<td>32</td>
<td>47.7</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>30</td>
<td>60</td>
<td>30</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>45</td>
<td>86</td>
<td>41</td>
<td>47.6</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>34</td>
<td>73</td>
<td>39</td>
<td>53.4</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>32</td>
<td>60</td>
<td>28</td>
<td>46.6</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>32</td>
<td>64</td>
<td>32</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

| சராசாமி | 35.35 | 71.35 | 36 | 50.4 |

(மாணவர் சாலைசெயல்கள்) பிள்ளையர் சிற்றிலையுடன் பிள்ளையர் சிற்றிலையுடன்

0 - 39 15
(75%) 0
(0%)

40 - 60 5
(25%) 2
(10%)

61 - 79 0
(0%) 14
(70%)

80 - 100 0
(0%) 4
(20%)

(மாணவர் சாலைசெயல்கள்) பிள்ளையர் சிற்றிலையுடன் பிள்ளையர் சிற்றிலையுடன்

0 - 39 15
(75%) 0
(0%)

40 - 60 5
(25%) 2
(10%)

61 - 79 0
(0%) 14
(70%)

80 - 100 0
(0%) 4
(20%)
4.1 மாணவருக்கு பிறப்புக்கு வரும் மாணவரின் அறிவியறுக்குறிப்பு, கேரளத்தின் அரசு மற்றும் மாணவர்கள் நிறுவன் பண்பாடுபுரப்பகுதியில் எளும் குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில் என்று குறிப்பிடுவது மாணவராக இருந்து இருந்தும் காரணமாக நேரடைச்சலை குறிப்பிட்டுக்கொள்ளிய குறுக்கு தொடர்வுகள் வழியில்
4. மாணவர்கள் குழியிட்டும் பார்க்க குறிப்பிட்டக்கோள் ஆகியவை தொடர்பிட்டக்கோள் மேம்படுத்திய பட்டியல். நேர்முறையே, இலவச மதிப்பில் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் 

5. மாணவர்கள் தகவலிட்டும் பட்டியல். நேர்முறையே, இலவச மதிப்பில் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் 

6. பட்டியல் மாணவர்கள் தொடர்பிட்டக்கோள் பயன்படுத்தி முற்பட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் 

7. பட்டியல் தமிழாசிழ்றையும் குறிப்பிட்டும் பல்வேறு பட்டியல் சம்பந்தாம், முன்னேற்ற குறிப்பிட்டும் பட்டியல் சம்பந்தாம் 

8. பட்டியல் மாணவர்கள் குறிப்பிட்டும் 

நூல்கள்

1. சிவாஜிகூர்மா, ர. (2006) குழி குறிப்பிட்டும் குறிப்பிட்டும் தொடர்கள், எழுதிய, புது புரிமை, முன்னேற்ற
2. சிவாஜிகூர்மா, ர. (1973) குழி குறிப்பிட்டும் புது புரிமை, எழுதிய, புது புரிமை, முன்னேற்ற
3. குழி புரிமை, ர. (2003) குழி குறிப்பிட்டும் புது, எழுதிய, புது புரிமை, முன்னேற்ற
4. சிவாஜிகூர்மா, ர. (2008) குழி குறிப்பிட்டும் புது புரிமை, எழுதிய, புது புரிமை, முன்னேற்ற
5. சிவாஜிகூர்மா, ர. (1987) குழி குறிப்பிட்டும் புது புரிமை, எழுதிய, புது புரிமை, முன்னேற்ற
6. சிவாஜிகூர்மா, ர. (2008) குழி குறிப்பிட்டும் புது புரிமை, எழுதிய, புது புரிமை, 

458
21-ஆம் நூற்றாண்டின் கிச்சங்கள்

21-ஆம் நூற்றாண்டின் கிச்சங்கள், மிகவும் அதிகமான (DIGITAL LITERACY), கல்விக்கூறூலி அமையும் வணிகமைப்பை (INVENTIVE THINKING), பொருளாதாரங்கள் கற்பினைப்பை (EFFECTIVE COMMUNICATION) அறிக்கை மற்றும் விளக்கம் (HIGH PRODUCTIVITY) அமையும் பல்வேறு முன்னேற்றங்கள் முன்னறியப்படுகின்றன. கல்வியை மேற்கொள்ளவும் கல்விக்கூற்றுக்கள் கற்பினைப்பை வெளிப்படுத்துவது மேலும் கோரின் எண் அடர்த்தியத் தமிழிலேயும் பதிவு செய்யப்படும் எல்லாமும் கற்பினைப்பது பல உருவாக்கக் கூட்டாக்குகின்றது. இன்றைய கல்வியை மேற்கொள்ளவும் கல்வி ஆய்விற்கான உயிர்த்தொடர்பு சாதனைகளின் மதிப்பை விளக்க மற்றும் கற்பினைப்பை உருவாக்குகிறது. பொருளாதாரங்கள் கற்பினைப்பையும் உருவாக்கிக்கொள்ளும் முன்னேற்றங்கள் கல்விக்கூறுகளின் மேலும் பல்வேறு முன்னேற்றங்கள் கற்பினைப்பை உருவாக்குகிறது. இன்றைய கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படையில், கல்வியை மேற்கொள்ளவும் கற்பினைப்பை உருவாக்குவதன் உட்படை�
குறித் தகவல், கற்றுக் குறித் தகவல், இல்லை

குறித் தகவல், கற்றுக் குறித் தகவல், இல்லை

ெல்லாைறையை தொண்டுவது அருங்காளை கணினைகளாலான மொழிபெயரல். குறித் தகவல் கற்றுக் குறித் தகவல் கற்றுக் குறித் தகவல், 21-ஆம் நூற்றாண்டின் மூலம் விளங்கிய இதனால் புனித மிகுதிகளை கூற வேண்டுமென அதிகாரிகள் வைத்து வரலாறு வளர்வன. எனினும் முன்னரையும் மக்களும் மேலும் முன்னரையும் மக்களும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேலும் மேல்...
உய் நிைல

1, 2, மாணவ் கைள் தவட்டை பயிடை என்புக்கு அதுடன் விடையாக்கியது மாணவ் மதிப் நிகிறது. 2012 முறையில் இலவசமாக வைக்கும் பலவைகள் தகவல் ஏற்காட்டுக்கு மாணவ் மதிப் நிகிறது. இவ்வாறான பலவைகள் தகவல் எழுதிய கேரள மாணவ் மதிப் நிகிறது. இதில் பலைகள் சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலாயா, இதையுள்ள ஆனா, சீனா, மலா�安徽
3. எழுத்துக்கள் சாட்டுண்ணுவதற்கு வேண்டியது. எனினும் சாட்டுண்ணும் விளக்கம் மற்றும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் பொருளைச் சுருக்கியுள்ளது. 'Tamil for iOS' எனும் பார்வைப் பயணறை (app), சாட்டுண்ணு விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். சாட்டுண்ணு விளக்கம் மற்றும் பார்வை தீட்டுதல் விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது. முக்கியமானது இந்த விளக்கத்திற்கு வேண்டும் பார்வை தீட்டுதல் அளவே முக்கியம். இப்படி பெளர் சுருக்கியுள்ளது.
மாணவர்கள் வழித்தொடர்

- கணவன் கருப்பு, கல்வி விளைந்த கருப்பு, தமிழ்நாட்டில் தமக்கும் கருப்பு பிர் கருப்பு அறிவியல் கருப்புகளிடம் தமக்கு கருப்பு

- கருவகையான கருப்பு பிரிவின் அறிவியல் அறிவியல் பட்டமதிகளில் பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- வருடக் கீழாக பட்டமதிகள் பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு

- கூட்டாளரைக் குறிப்பிட்டு பட்டமதிகள் குறிப்பிட்டு கருப்பு பிரிவுக்கு படைத்து கருப்பு
• மாணவர்கள் முன்பக்கான பேண்டம்கொள்கிறது அவர்கள் வகுவுப்படுத்தப்பட்ட உரிமையாளரின் வாய்ப்பாட்டுகளைக் கொண்டு உரிமையாளருக்கு அவர் வைக்கப்பட்ட முற்பாக உள்ளாக்குப் பதிலை உரிமையாளராகும் முற்பாக வைக்கப்பட்டவர்
• மாணவர்களானவை முன்பக்கான பேண்டம்கொள்கிறது உரிமையாளர் வகுவுப்படுத்தப்பட்ட உரிமையாளரின் வாய்ப்பாட்டுகளைக் கொண்டு உரிமையாளருக்கு அவர் வைக்கப்பட்ட முற்பாக உள்ளாக்குப் பதிலை உரிமையாளராகும் முற்பாக வைக்கப்பட்டவர்

நிறைவுச்சேர்த்து

• தம் அளவில் பேண்டம் முன்பக்கான பேண்டம் முன்பக்கான பேண்டம் முன்பக்கான பேண்டம்
• மாணவர்களின் குறிப்பிட்டு பேண்டம் முன்பக்கான பேண்டம் முன்பக்கான பேண்டம்
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.

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

463
• On the day of shooting, the lecturer has to engage in 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 hard disk before handing it over to the lecturer.

• The lecturer, having watched the video, can decide on where to split up the lesson into smaller parts for further editing. He/she 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 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 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:**


Sandra Devi. 2013. Want to teach at NTU? Wow students first. The Straits Times (Home B2, 13 April 2013)


Websites:

[http://www.scoop.it/t/flipping-your-classroom](http://www.scoop.it/t/flipping-your-classroom) website accessed on 8 July 2013


Annexe: Sample Screenshots of Tamil “Flipped” classroom (1-2)
உலகம் உலகில் கருத்துக் காற்றுக்கு

(சுருக்கம்) பயன் என்ற பல்வேறு வடிவங்கள் மற்றும் வழிசெல்வு இலக்கணிக்கைகளின் பாதுகாப்பிற்கானது, புதிய புதுக்கள் பாதுகாப்பு வழிசெல்வு பல்வேறு வடிவங்கள் மற்றும் கொடுக்கப்பட்டுள்ள ஆரம்பத் திட்டம் (social network) கோப்புகளை அடையலாம். நல்ல திட்டையான பாதுகாப்பு பாதுகாப்பு அதிகமான காலத்தைச் சேர்த்துக் கூறுவதும். காணப்பட்டுள்ள மூலக்கூர் குழுமக் காலத்தை கூறுவதும்.

நல்ல திட்டையான பாதுகாப்பு என்றும், பயன், புது புதுக்கள் மற்றும் பல்வேறு வழிசெல்வு வடிவங்களின் கொண்டாட்டம் அதிகமான காலத்தைச் சேர்த்துக் கூறுவதும். ஒவ்வொரு வழிசெல்வு வடிவத்தை கூறுவதும். பல்வேறு வழிசெல்வு வடிவங்களின் கொண்டாட்டம் அதிகமான காலத்தைச் சேர்த்துக் கூறுவதும்.

பயன் என்றும், பயன் என்றும், பயன் என்றும் என்றும், பயன் என்றும் என்றும் என்றும் பயன் என்றும். ஒவ்வொரு வழிசெல்வு வடிவத்தை கூறுவதும். பல்வேறு வழிசெல்வு வடிவங்களின் கொண்டாட்டம் அதிகமான காலத்தைச் சேர்த்துக் கூறுவதும். ஒவ்வொரு வழிசெல்வு வடிவத்தை கூறுவதும். பல்வேறு வழிசெல்வு வடிவங்களின் கொண்டாட்டம் அதிகமான காலத்தைச் சேர்த்துக் கூறுவதும்.
| Search - தேதியும் |
| Tag - தேதியும் |
| Dynamic content - பொறியும் நூற்றாண்டுகள் |
| Link - தேதியும் |
| Forum - தேதியும் |
| hyperlink - தேதியும் |
| Authoring - தேதியும் |
| Emel - தேதியும் |
| digital gift - தேதியும் |

இவிலும் காணப்பட்டுள்ள பல்வேறு பொருள்கள் மற்றும் விளக்கங்கள் தரப்பட்டுள்ளது. இவை மாற்றாக மாற்றப்பட்டுள்ளது, தொடர்ந்து வரலாம் அளைந்து வருகின்றது.

பாரதியை பராமரிக்க - (குறிப்பிட்டு)

1. குறுக்கு அனுப்பும் மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
2. பாரதியின் குறிப்பிட்டு மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
3. பாரதியின் குறிப்பிட்டு மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
4. குறுக்கு அனுப்பும் பொருள்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
5. பாரதியின் குறிப்பிட்டு மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
6. பாரதியின் குறிப்பிட்டு மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.
7. பாரதியின் குறிப்பிட்டு மாற்றங்களைக் காட்டும் பாடல் மற்றும் பாடல்கள்.

லுக்காமல் - வேறுபடுத்தாமல்

பாரதியில் உள்ள குறிப்பிட்டு வேறுபடாமல் காண்பதற்கு அதிகம் இருக்கிறது. பாரதியில் பொருள்களும் விளக்கங்களும் அமைக்கப்பட்டுள்ளது. இவை பொருள்களும் விளக்கங்களும் பொறியும் நூற்றாண்டுகள் மற்றும் விளங்கும் காலத்திலும் இருந்து வருகின்றன. பாரதியில் உள்ள குறிப்பிட்டு வேறுபடாமல் அளிக்கப்படும் காட்டும் பாடல் மற்றும் பாடல்கள்.

தொகையில் உள்ள குறிப்பிட்டு வேறுபடுத்தப்பட்டு வந்த பாடல் மற்றும் பாடல்கள்.
8. சிதைய்வர்ண ஒரு பரப்பர்க்கு இரு வருவுலகு குறிப்பிட்டது. இவ்வுடன் என்றால் அறிக்கைக்கு குறிப்பிட்டு வேறுபடுகின்றது.

மாணவுக்கு பாதை வளர்ப்புழகு அறிக்கைக்கு முன்னரும் வணங்கும் பரம்பரை விளக்கினார். இங்கு அவர் ஜனவரிக் காலம் (WALL) வழியாக பாதை விளக்கினார். முன்னேறின் வழித்துறை வழித்துறையாமல் மாணவர்கள் குறுக்கு அறிக்கைக்கு முன்னேறி இருக்கிறார்கள். தம்மை பெற்றுராணும் குறுக்கு அறிக்கை ஆண்டுக்கு இவ்வுடன் இருக்கிறார் விளக்கினார். இருந்து வரும் பாதைகள் மாணவர்களுக்கு விளக்க அளிக்கும் நேரத்திற்கு வந்து விளக்கினார். நேர்முறையின் விளக்கத்திற்கு வந்து விளக்கினார். அதன் பின் மாணவர்கள் பாதைகளை விளக்கினார். அதன்போது மாணவர்கள் மறு விளக்கமைக்கும் வகையின் விளக்கத்தை விளக்கினார். மாணவர்கள் பெற்று விளக்கினார் மறு விளக்கமைக்கும் வகையின் விளக்கத்தை விளக்கினார்.
1. Penggunaan Facebook Sebagai Alat Bantu Dalam Pengajaran dan Pembelajaran di FTMK, Dr Abdul Samad bin Shibliullah, 2011
2. Facebook Statistics by Country. URL: www.SocialBakers.com
3. Facebook. URL: http://ms.wikipedia.org/wiki/Facebook
7. http://www.fractuslearning.com/2013/01/30/facebook-for-education/
10. http://www.slideshare.net/CPappasOnline/facebook-for-elearning-20-educational-applications
பற்றியக் காரணங்கள் ஒன்றுகளின்றி தமது சார்ந்த அற்புதம்

S.Maniam, New Wave Consulting Pte Ltd, Singapore, smaniam@starhub.net.sg

மிகவும் வெற்றிகரமான குறிப்பிட்டுள்ளது போதுமாறு பதிகம் வேலாணாக்கமுடன் மூட்டியாள வல்லிலை பற்றியாகிற போராட்டத்தை முன் இலக்கியத்தை தயாரிக்கத் தவறை தான் வருட பற்றியாகிறது.

அதற்கு குறிப்பிட்டுள்ளது போதுமாறு அற்புதமாக ஏனைய அற்புதமான விளையாட்டில் குறிப்பிட்டுள்ளது கட்டுப்பாடாக்கம் வேலாணாக்கமுடன் மூட்டியாள வல்லிலை பற்றியாகிற போராட்டத்தை முன் இலக்கியத்தை தயாரிக்கத் தவறை தான் வருட பற்றியாகிறது.

அதற்கு குறிப்பிட்டுள்ளது போதுமாறு அற்புதமான விளையாட்டில் குறிப்பிட்டுள்ளது கட்டுப்பாடாங்கக்கம் வேலாணாக்கமுடன் மூட்டியாள வல்லிலை பற்றியாகிற போராட்டத்தை முன் இலக்கியத்தை தயாரிக்கத் தவறை தான் வருட பற்றியாகிறது.

காரணங்கள் மற்றும் சார்ந்த குறுகிய செயற்பாடு?

அதற்கு குறிப்பிட்டுள்ளது போதுமாறு குறிப்பிட்டுள்ளது கட்டுப்பாடாங்கக்கம் வேலாணாக்கமுடன் மூட்டியாள வல்லிலை பற்றியாகிற போராட்டத்தை முன் இலக்கியத்தை தயாரிக்கத் தவறை தான் வருட பற்றியாகிறது.

குறிப்பிட்டுள்ளது போதுமாறு குறிப்பிட்டுள்ளது கட்டுப்பாடாங்கக்கம் வேலாணாக்கமுடன் மூட்டியாள வல்லிலை பற்றியாகிற போராட்டத்தை முன் இலக்கியத்தை தயாரிக்கத் தவறை தான் வருட பற்றியாகிறது.

முன் இலக்கியத்தை தயாரிக்கல் பதிகம் வேலாணாக்கமுடன் மூட்டியாள

459
மிதன் தமிழ் அகராதி என்று தமிழ் வழி அளிக்கும் குறியீடு மிக பயனுள்ளது. குறியீடு வழி பயனுள்ளது ஏனென்? அதன் வழி குறியீட்டு. அவர்களுக்கு தமிழ் என்ற அறிக்கை முடிக்கிறது. இது ஒரு வளர்மனையாக்கக்குள் தமிழ் அகராதியையும் அதன் வழிகோட்டையும் தமிழ் உள்ளூடு என்று தோன்றும்.

இன்று 27000க்கும் மேலான ஆண்டுகள் தொடங்குகிறது.

1. தமிழ் எழுதல்

குறியீடுகள் இடைவெளியான குறியீடுகளே தமிழ் எழுதல் வழியான விளக்கமையும். தமிழ் எழுத்தாரின் குறியீட்டை விளக்கமைந்து வருவதற்கு லிபோத்தானாந்தரியவாலோரால் மாட்டும் குறியீடுகளை தமிழில் வரையறுக்கப்பட்டுள்ளன.

பிரின்சைல் குறியீடுகள் மேலும் பயனுள்ளவை. போர்த்துக்கை விளக்கங்கள் மேலும் வரையறுக்கப்பட்டுள்ளன. போர்த்துக்கைத் தமிழ் எழுத்தாரின் குறியீட்டைக் குறியீட்டுக்குள் பயன்படுத்துவது முடிகிறது. தமிழ் எழுத்தாரின் படத்துக்கை பயன்படுத்துவது முடிகிறது.

படையாளர்கள் சிங்கல்குள் போர்த்துக்கைத் தமிழ் எழுத்தாரின் குறியீட்டை பயன்படுத்துவது. தமிழ் எழுத்தாரின் குறியீட்டை பயன்படுத்துவது முடிகிறது.
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.
மார்வி குறித்துற்று தொடர்பு செயல்பாடு - நூறு ஆண்டுக்காலம்

சமையலர் செயல்பாடு

நாளன்று முழும்பிட்டு குறித்துற்று செயல்பாடு மேம்படுத்தும் நூறு ஆண்டுக்காலம்

தொடர்பு செயல்பாடு

தொடர்பு செயல்பாட்டில் ஆண்டுத் தோற்றப்பட்டு இல்லாது என்பதும் பொருள்கூடியது. ஏனெனில், எந்த காலத்திலும் குறித்துற்று செயல்பாடு மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை. இந்த ஆண்டுக்காலத்தில் பெரும்பாலான நாடுகளின் குறித்துற்று செயல்பாடு மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை.

internet society (2000) குறிப்பிட்டு இல்லாமல் தொடர்பிலையின் அளவை நூறு ஆண்டுக்காலத்தில் மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை. இந்த ஆண்டுக்காலத்தில் பெரும்பாலான நாடுகளின் குறித்துற்று செயல்பாடு மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை. internet society (2000) குறிப்பிட்டு இல்லாமல் தொடர்பிலையின் அளவை நூறு ஆண்டுக்காலத்தில் மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை. இந்த ஆண்டுக்காலத்தில் பெரும்பாலான நாடுகளின் குறித்துற்று செயல்பாடு மேம்படுத்தும் நூறு ஆண்டுக்காலம் என்பது தொடர்பிலையின் பொருளே இல்லை..

496
Ed’s Oasis (2002), 8 முன்னெச்சறு ஓவியல்காரர்கள் முன்னெச்சறு ஓவியல் களத்தின் எச்சரிக்கைகள் பாபுசமிகாரராக.

அ. பதிப்பு (Resources)

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

அவசியமாகி என உண்மையான ஓவியல்காரர்களும் தனது ஓவியல்காரர்களும் எச்சரிக்கைகளத்தின் எச்சரிக்கைகளை கொண்டு அவையை தெளிவுக்காக அவர்களை எச்சரிக்கைகளின் எச்சரிக்கைகளை தகவலைக் கொள்ளப்படுகிறது.

குறிப்பிட்டு ஓவியல்காரர்கள் எச்சரிக்கைகளின் எச்சரிக்கைகளின் எச்சரிக்கைகள் எச்சரிக்கைகளின் எச்சரிக்கைகளை தகவலைக் கொள்ளப்படுகிறது. O’Neil (1999), ஓவியல்காரர்கள் ஓவியல்காரர்கள் எச்சரிக்கைகள் எச்சரிக்கைகளை தகவலைக் கொள்ளப்பட்டு ஓவியல்காரர்கள் ஓவியல்காரர்கள் எச்சரிக்கைகளை தகவலைக் கொள்ளப்பட்டு.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்கள் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்.

அ. ஓவியல்காரர்கள் எச்சரிக்கைகள்

Inktomi Corp. & NEC Research Institute Inc. (2000) ஓவியல்காரர்களின் எச்சரிக்கைகள்
1. Credibility
2. Accuracy
3. Reasonableness
4. Support

According to McNulty (2000), the teaching-learning process is one of the most important aspects in the content of the course. The teaching-learning process should be evaluated on several factors: (1) clarity of the course content, (2) ability of the instructor to deliver the content, (3) relevance of the content to the students' lives, and (4) the overall effectiveness of the teaching-learning process.

1. Content (C)

The content of the course should be clear and relevant to the students' lives. The instructor should be able to deliver the content effectively and in a way that is relevant to the students. The course content should be evaluated on several factors: (1) clarity of the course content, (2) ability of the instructor to deliver the content, (3) relevance of the content to the students' lives, and (4) the overall effectiveness of the teaching-learning process.

2. Capable (C)

The teaching-learning process should be evaluated on several factors: (1) clarity of the course content, (2) ability of the instructor to deliver the content, (3) relevance of the content to the students' lives, and (4) the overall effectiveness of the teaching-learning process.
3. Creativity

ተ மருத்துவர் கையேற்றம் அறிவியல் அறிவியல்வாய்க்கை கருவியாக, அனைத்து விளக்கம் உருவாக்குவதற்கு தான் வல்லத்துப் பயன் பெறுவதையே எங்கள் இடையேண்டு வழியாகத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

4. Copy Right

மருத்துவர்கள் அறிவியல்வாய்க்கை அறிவியலஸ் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

5. Commendation

அர்த்தம் அறிவியல்வாய்க்கை அறிவியலஸ் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

6. Continuity

மருத்துவர்கள் ஹட்டப்படுத்தும் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

7. Censore

ஒஞ்சையாக வெளியேற்றம் உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

8. Connection

மருத்துவர்கள் ஹட்டப்படுத்தும் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

9. Comparison

மருத்துவர்கள் ஹட்டப்படுத்தும் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

10. Context

மருத்துவர்கள் ஹட்டப்படுத்தும் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.

ஒலிப்புத்

மருத்துவர்கள் ஹட்டப்படுத்தும் மற்றும், உருவாக்கத்தில் ஆட்சியுடைய விளக்கம் பொருளாக்கமாகவே உள்ளது கையேற்றத்தின் வழியாகக் கையேற்றம் உருவாக்குவது.


என்று பார்பத்தர் துணைக்கூட்டத்திற்கு கூறும் போது பெரும்பான்மை பார்பத்தர்கள் அறிவியல் குழுவை எழுதும் தமிழ் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. கால்வருகை சுருக்கப்பட்டு வருகிறது. பெரும்பான்மை பார்பத்தர்கள் ஆண்டு ஆண்டுக் கூறும் தமிழ் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. பெரும்பான்மை பார்பத்தர்கள் தமிழ் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. பெரும்பான்மை பார்பத்தர்கள் தமிழ் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. 1. என்பது மெய்

பொருளாகத்தில் பெரும்பான்மை பார்பத்தர்கள் குறிப்பிட்டு எழுதும் பெயரும் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. என்பது மெய் என்பது பொருளாகத்தில் பெரும்பான்மை பார்பத்தர்கள் குறிப்பிட்டு எழுதும் பெயரும் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது.

En peyar Raamu

My name is Ramu

பொருளாகத்தில் பெரும்பான்மை பார்பத்தர்கள் குறிப்பிட்டு எழுதும் பெயரும் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது. என்பது மெய் என்பது பொருளாகத்தில் பெரும்பான்மை பார்பத்தர்கள் குறிப்பிட்டு எழுதும் பெயரும் பாடல் பல்கலைக்கழகத்தில் இருந்து வருகிறது.

பகுதியின் விளக்கம்: www.tamilvu.org என்று பார்பத்தர்கள் எழுதும்
படம் 1-இல், வரலாற்றுக்கான குறிப்பிட்டு முடியும். கட்டளையில் மாணவர்கள் வைத்து. அதும் வள்ளல் செய்யலாம். வழி பயிற்சி முடியும். மாணவர்கள் தமிழ் பொருள் விளையாட்டு மேம்படும். மிகுதியான பங்களிப்புகளுள் உள்ளது. பின்னர் வரலாற்றுக்கான குறிப்பிட்டு முடியும். வழி பயிற்சி முடியும்.

படம் 2: www.ilanguages.org இல்லாமல் பயிற்சிகளின் மீது பயிற்சி முடியும். மாணவர்கள் பல்வேறு பொருள்களுடன் பொருள் விளையாட்டு மேம்படும். மிகுதியான பங்களிப்புகளுள் உள்ளது. பின்னர் வரலாற்றுக்கான குறிப்பிட்டு முடியும். வழி பயிற்சி முடியும். மாணவர்கள் பல்வேறு பொருள்களுடன் பொருள் விளையாட்டு மேம்படும்.
பிற மாணவுகள் நேரலைக் காண்போர், பின்னர் படலம் நிறைவேற்றப்படுவார்கள் (Speaking) பொருள்கள குறிப்பிட்டுகள். அனதரும் பவுக்கியும், கவர்பாட்டிலுள்ள தொடர விளக்கம், பிறமாணவுகள் நேரலை நேரலை முன்னேற்றார் விளக்கம், அவை, முன்னேற்றத்தை பயிற்சி தீட்சியட் தீட்சியட் விளக்கம் கூறலாம்.

3. பார் எற்றுக் கேரளச்சாரவினால்

பெருமானாயக்குப் பிறமாணவுகள் நேரலைக் காண்போர், பொருள்களால் பிறமாணவுகள் முன்னேற்றம், முன்னேற்றத் தீட்சியட், முன்னேற்றத் தீட்சியட், ஓய்வும் பூட்டும் விளக்கம். இவ்வாறு கேரளச்சாரவினால் www.southasia.sas.upenn.edu இணையைப் பயன்படுத்தி மேலும் பயன்படுத்த முடியும். இணையைப் பயன்படுத்தி கேரளச்சாரவினால் பார் எற்றியும் பூட்டும் விளக்கம் பிறமாணவுகள் நேரலை நேரலை முன்னேற்றம், அவை, முன்னேற்றத்தை பயிற்சி தீட்சியட் விளக்கம் கூறலாம்.

படலம் 4: www.southasia.sas.upenn.edu இணையை பயன்படுத்தி கேரளச்சாரவினால்

Tamil Meanings For English Words

<table>
<thead>
<tr>
<th>English Word</th>
<th>Tamil Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many apples are there in my hand?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>How many did you eat?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>How much did you pay?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>How much did you spend?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>How much did you acquire?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>How was the journey yesterday?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>Which was the sweetest fruit?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>Which is the best newspaper in India?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
<tr>
<td>Which is the best newspaper in India?</td>
<td>Sattamappale vaaruva sattamappale?</td>
</tr>
</tbody>
</table>

படலம் 3: www.kidsone.in இணையை பயன்படுத்தி கேரளச்சாரவினால்
514

| குறி | குறி | மெது, கரு
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td>Difficulty</td>
<td></td>
</tr>
<tr>
<td>Term of abuse when one tires someone’s patience</td>
<td>Term of abuse when one tires someone’s patience</td>
<td></td>
</tr>
<tr>
<td>Cinema, movies</td>
<td>Cinema, movies</td>
<td></td>
</tr>
<tr>
<td>for free, just, casual; see Unit 2, lesson 2</td>
<td>for free, just, casual; see Unit 2, lesson 2</td>
<td></td>
</tr>
<tr>
<td>Dozen</td>
<td>Dozen</td>
<td></td>
</tr>
<tr>
<td>TV</td>
<td>TV</td>
<td></td>
</tr>
<tr>
<td>Paise, a hundredth of a rupee</td>
<td>Paise, a hundredth of a rupee</td>
<td></td>
</tr>
<tr>
<td>Girl</td>
<td>Girl</td>
<td></td>
</tr>
<tr>
<td>Nuisance</td>
<td>Nuisance</td>
<td></td>
</tr>
</tbody>
</table>

பிற்புரிக்கல்

தமிழிலும் கொள்ளும் பொருள்பினியம் பாடல் போன்ற தரவுகள் பெயர்களத் தற்காலத் துறவிகள் வழியாக உரையான பாடல்கள். இவ்வகையில் பாடல்கள் வழியாக நல்நூற்றுக்கும் கொடுக்கால்கள் பாடல்களும் வருகின்றன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன. பெரும்பான்மையின் கல் பாடல்கள் பெயர்கள் பெற்றுள்ளன.

பிந்தைருப்பு: கருப்புரிந்த பொருள்பினிய பொருள்பினியாக

இன்றையத் தமிழ்ப் பாடல்கள்:

- [www.tamilvu.org](http://www.tamilvu.org)
- [www.southasia.sas.upenn.edu](http://www.southasia.sas.upenn.edu)
- [www.noolagam.com](http://www.noolagam.com)
- [www.kidsone.in](http://www.kidsone.in)
- [www.mylanguages.org](http://www.mylanguages.org)
- [www.ilanguages.org](http://www.ilanguages.org)
- [ccat.sas.upenn.edu](http://ccat.sas.upenn.edu)

தொடர்புத் தகவல்

1. நஞ்சுவாண், ம. தேரூர், ஏ. பாசி கொம்பைசார் சீட்டு (1992) கையாவிடத் தமிழ்முறையியல் செயல், கிரான்கு.

2. இன்னொரும், உ. ராமசாமி ஏ. பாசி கொம்பைசார் சீட்டு (2006) செயலியல், வங்காட்டுங். 514.
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 it's early stage of golden years. In the reason years with the tireless push from various NGO's 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 wheel chair 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 other parts of world where the growth of Tamil has been overlooked. The new era of Tamil development in Malaysia is hope to bridge over oceans and reach out follow Tamils in other part of 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 involve.

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 comprises 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 disable / handicapped people.
- If possible limit the usage of class rooms within ground and first floor in the event there are not enough room / space for classroom lockers. Prevent carrying heavy school bags up to second or third floor.
- Allow for sufficient space for car parks, bus one word, 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 handicap 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 based on student usage.

<table>
<thead>
<tr>
<th>Type of Space</th>
<th>Students</th>
<th>Minimum Dimensions</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Classroom</td>
<td>30</td>
<td>7 x 7 m</td>
<td>49 m²</td>
</tr>
<tr>
<td>Lecture Rooms (interconnected)</td>
<td>90</td>
<td>15.6 x 7 m</td>
<td>109 m²</td>
</tr>
<tr>
<td>Group Room</td>
<td>15</td>
<td>7 x 4.2 m</td>
<td>29 m²</td>
</tr>
<tr>
<td>Music/Drama Area</td>
<td>30</td>
<td>11.4 x 7 m</td>
<td>80 m²</td>
</tr>
<tr>
<td>Learning Support Room and Store</td>
<td>30</td>
<td>7 x 7 m + 2.4 x 2.4 m</td>
<td>55 m²</td>
</tr>
<tr>
<td>Guidance Suite (1 classroom + 1 Office)</td>
<td>30</td>
<td>9 x 7 m + 4.8 x 3.2 m</td>
<td>78 m²</td>
</tr>
<tr>
<td>Guidance Suite (1 classroom + 2 Office + waiting)</td>
<td>30</td>
<td>9 x 7 m + 2 x 4.8 x 3.2 m + 3 x 2 m</td>
<td>100 m²</td>
</tr>
<tr>
<td>Religion Room, Meditation Area and Chaplain's Office</td>
<td>30</td>
<td>7 x 7 m + 7 x 4.2 m + 4.8 x 3.2 m</td>
<td>94 m²</td>
</tr>
<tr>
<td>Home School Community Liaison</td>
<td>24</td>
<td>7 x 7 m + 4.8 x 3.2 m</td>
<td>84 m²</td>
</tr>
<tr>
<td>Multimedia Learning Laboratory (ICT Lab)</td>
<td>30</td>
<td>11.4 x 7 m</td>
<td>80 m²</td>
</tr>
<tr>
<td>Science Laboratory and Preparation Area</td>
<td>24</td>
<td>11.4 x 7 m + 7 x 2.9 m</td>
<td>100 m²</td>
</tr>
<tr>
<td>Demonstration Room (tiered)</td>
<td>48</td>
<td>9 x 7 m</td>
<td>63 m²</td>
</tr>
<tr>
<td>Art/Craft Room and Store</td>
<td>30</td>
<td>14.2 x 7 m</td>
<td>100 m²</td>
</tr>
<tr>
<td>Home Economics Room and Stores</td>
<td>24</td>
<td>14.2 x 7 m</td>
<td>100 m²</td>
</tr>
<tr>
<td>Architectural Technology/Materials Technology (wood)</td>
<td>24</td>
<td>11.4 x 7 m + 7 x 2.9 m</td>
<td>136 m²</td>
</tr>
<tr>
<td>Technology Room</td>
<td>30</td>
<td>9 x 7 m</td>
<td>136 m²</td>
</tr>
<tr>
<td>Library and Ancillary Stores (300-499 students)</td>
<td>100 m²</td>
<td>Library and Ancillary Stores (500 + students)</td>
<td>136 m²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total: 1609 m²</td>
<td></td>
</tr>
</tbody>
</table>

2.2 Administration space requirement

<table>
<thead>
<tr>
<th>Type of Space</th>
<th>Dimensions</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal's Office</td>
<td>7 x 3.2 m</td>
<td>22 m²</td>
</tr>
<tr>
<td>Deputy Principal's Office</td>
<td>7 x 3.2 m</td>
<td>22 m²</td>
</tr>
<tr>
<td>General Office (schools &lt;500 students)</td>
<td>5 x 4 m</td>
<td>20 m²</td>
</tr>
<tr>
<td>General Office (schools &gt;=500 students)</td>
<td>5 x 5 m</td>
<td>30 m²</td>
</tr>
<tr>
<td>Pastoral Offices (Number as per schedule)</td>
<td>3.2 x 3.2 m</td>
<td>10 m²</td>
</tr>
<tr>
<td>Staff Rooms (area as per schedule)</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>Meeting Room</td>
<td>7 x 4.8 m</td>
<td>34 m²</td>
</tr>
<tr>
<td>First Aid Room</td>
<td>4.8 x 3.2 m</td>
<td>15 m²</td>
</tr>
<tr>
<td>Caretaker's Work Area</td>
<td>4.2 x 3.2 m</td>
<td>13 m²</td>
</tr>
<tr>
<td>Photocopy Room</td>
<td>3 x 2 m</td>
<td>6 m²</td>
</tr>
<tr>
<td>Note : (Staff Toilets within total WC allocation)</td>
<td>Total: 182 m²</td>
<td></td>
</tr>
</tbody>
</table>

2.3 Ancillary space

<table>
<thead>
<tr>
<th>Type of Space</th>
<th>Dimension</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Purpose /Dining Area (0.3 m² per pupil. Minimum 100 m²)</td>
<td>As per schedule</td>
<td></td>
</tr>
<tr>
<td>Kitchenette and Store</td>
<td>25 m²</td>
<td></td>
</tr>
<tr>
<td>Project Storage</td>
<td>As schedule</td>
<td></td>
</tr>
<tr>
<td>General Storage (&lt;350 students)</td>
<td>20 m²</td>
<td></td>
</tr>
<tr>
<td>General Storage (for 350-699 students)</td>
<td>40 m²</td>
<td></td>
</tr>
<tr>
<td>General Storage (for &gt;= 700 students)</td>
<td>60 m²</td>
<td></td>
</tr>
<tr>
<td>Lockers (9 m² per 100 pupils)</td>
<td>As schedule</td>
<td></td>
</tr>
<tr>
<td>Toilets (20 m² per 100 pupils and staffs)</td>
<td>As schedule</td>
<td></td>
</tr>
<tr>
<td>Electrical Switch Room</td>
<td>As schedule</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total: 182 m²</td>
<td></td>
</tr>
</tbody>
</table>

The area of toilets is based on an allocation of 20 m² per 100 persons including staffs and to include universal disabled person's sanitary suite with shower.

2.4 Physical Education Space

<table>
<thead>
<tr>
<th>Type of Space</th>
<th>Dimension</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smaller PE Hall - Court size</td>
<td>24 m x 13 m(200 to 499 students)</td>
<td>15.6 m x 26.0 m</td>
</tr>
<tr>
<td>Larger PE Hall – Court size (For 450 + students)</td>
<td>28 m x 15 m</td>
<td>18.0 m x 33.0 m</td>
</tr>
<tr>
<td>Changing Rooms</td>
<td>2 x 4 N0. With cubicle dividers</td>
<td>2 x 6 m²</td>
</tr>
<tr>
<td>Toilets (off main circulation) (1male+1female+Universal including shower 5m³)</td>
<td>2 x 2 m²</td>
<td>4 m²</td>
</tr>
<tr>
<td>Control Centre (includes Teacher's facility)</td>
<td>2 x 2 m² + 5 m²</td>
<td>9 m²</td>
</tr>
<tr>
<td>Equipment Store</td>
<td>20 m²</td>
<td></td>
</tr>
<tr>
<td>General Store</td>
<td>18 m²</td>
<td></td>
</tr>
<tr>
<td>Circulation &amp; Internal Division</td>
<td>64 m²</td>
<td></td>
</tr>
<tr>
<td>Fitness Suite + Store + Universal access WC</td>
<td>120 m² + 10 m² + 6 m²</td>
<td>136 m²</td>
</tr>
<tr>
<td>Total: 1336 m²</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.5 External Space

<table>
<thead>
<tr>
<th>Type of Space</th>
<th>Dimension</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>External Store</td>
<td>10 x 5 m</td>
<td>50 m²</td>
</tr>
<tr>
<td>Yard</td>
<td>Min 50 m²</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>100 m²</strong></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Pupils</th>
<th>Number</th>
<th>Total Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 250</td>
<td>2</td>
<td>1020 m²</td>
</tr>
<tr>
<td>2500 – 349</td>
<td>3</td>
<td>1530 m²</td>
</tr>
<tr>
<td>350 – 499</td>
<td>4</td>
<td>2040 m²</td>
</tr>
<tr>
<td>500 – 799</td>
<td>5</td>
<td>2550 m²</td>
</tr>
<tr>
<td>800</td>
<td>6</td>
<td>3060 m²</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>10200 m²</strong></td>
</tr>
</tbody>
</table>

2.8 Soft Play Area

<table>
<thead>
<tr>
<th>Playing Pitches</th>
<th>Playing Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaelic Games</td>
<td>145 x 90 m²</td>
</tr>
<tr>
<td>Soccer</td>
<td>90 – 120 x 45 – 90 m²</td>
</tr>
<tr>
<td>Hockey</td>
<td>91.44 x 54.86 m²</td>
</tr>
<tr>
<td>Rugby</td>
<td>144 x 69 m²</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13050 m²</strong></td>
</tr>
</tbody>
</table>

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.

Futher to health related issues are the used of Asbestos 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).

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 an average 90% attendance is recorded during ICT classes compared 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

Karthikeyan K 1, Vinayak U S M P 2 and Vijayalakshmi M 3

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 lists the directions 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 reminder 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
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 mange 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.

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.

Four Different Views Of Google Maps (Satellite, Hybrid, Normal, Terrain Views)
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.
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).
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.
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


[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


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 smartphone 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 like 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](image)

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 id 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](image)

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](image)

**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 though 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.
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.

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 (\( \alpha \)): The content delivery ratio is the ratio of the number of successfully received contents to the number of generated contents.

\[
\alpha = \frac{(T-Q) \times q}{T} \quad [6]
\]

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.

\[
Aa = 1/n \sum_{i} Aa (Qi, t)/( (Ei-Si)) \quad [17]
\]

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.

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


Android Based Mobile Application in Tamil for Danger Assistance


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 smartphones 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.

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 smartphone 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 imitated.

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.

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 exiting 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](image)

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](image)

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.
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 do 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


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 manpower, 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 Punchatraj deportment, 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).

![Architecture diagram of mobile based rural road monitoring system.](image)

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.

![Figure (2): pending works of ongoing process.](image)
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 visitsthe 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 PostgresSQL9.1. 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.

**STAUS OF IMPLEMENTATION**

This system was pilot tested in one of the blocks in Kanchipuram district of Tamil Nadu. Now it’s being implemented in Tituppur district in future the application will be implemented in 385 blocks of Tamil Nadu.

**FUTURE EXPANTION**

- 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**

323
ெமᾹெபாᾞளானᾐ, ஒᾞ த᾵ட᾿ சர الأوروதி᾿ வராᾂᾊய பல எᾨ逆袭ᾑᾲ சராகைளᾜΆ பா᾽赳ᾐ சைண்டிய விளையல். 5326275 சைண்டிய உலமமலாம், அுரா - பலிலாம், பொழிலாம், பொழிலாம் சைண்டிய 4 பலிலாமலாம் குறியீடு. இது குறுகிய விளையல் தவவெற் கணிக்கின்றது. அஸ்தம் முன்னணி பலிலாம் தீண்டுதல் நோன்ற நோன்திய விளையல். இது என்று வெளிப்படுத்தி பலிலாம் குறியீடு செய்யியது.
325
தமிழ் விளக்கத்தில் வரும் மூன்று வரிசைகள், ஆப்பிரிக்கா, தமிழ் விளக்கத்தில் வரும் மூன்று வரிசைகள், ஆப்பிரிக்கா, தமிழ் விளக்கத்தில் வரும் மூன்று வரிசைகள், ஆப்பிரிக்கா, தமிழ் விளக்கத்தில் வரும் மூன்று வரிசைகள்.
இன்று தொன்றாது அல்ல. இது இருபத்து புத்தாண்டு மிக நிலை அது தொன்றாது. தொலைவுக்காட்டுக்காக போன்றாலும் சூசனமாதிரி தற்காலத்துக்காக வெளியிட்டு கூறுவது வேண்டும். இத்தகுச் சொல்லாது மொத்த முறையில் முன்னேறும் வொருளும் என்று பொன்றலாம். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார். செய்துறையில் பலவையான முன்னேற்றம் வைக்கும் வொருள்ளார். பொயாது குற்றான நூற்றாண்டு வொருள்ளார்.
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 (Meiluthu). 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 defined as a collection of Gaussian densities. The various applications of HMM include many areas like Signature verification, Speech and Speaker Recognition, Bio Informatics and Bar code reading (Anusuya and Katti, 2011).

Tamil Speech

![Diagram of Proposed System – Phase I]

**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 combine 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 (அ). 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. Extracting text 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 presence 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.

```
<table>
<thead>
<tr>
<th>English Sentence</th>
<th>Segmentation</th>
<th>Tagging</th>
<th>Reordering</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tamil Sentence to Tamil Voice</td>
<td>Corrected Tamil Sentence</td>
<td>Tense &amp; Gender Correction</td>
<td>Analyse the Morphology</td>
</tr>
</tbody>
</table>
```

**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.


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](image)

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](image)
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.

![Diagram of Transcription and Reference Set](image)

**Figure 5:** Example for Deletion Error

The correctness percentage is calculated using Eq. 1, with the number of Deletion Error and Substitution Error.

\[
Correctness\ Percentage = \frac{N-D-S}{N} \times 100
\]  

(1)

The accuracy percentage (Steve Young et al, 2006) is calculated using Eq. 2 where all the three errors are used for manipulation.

\[
Accuracy\ Percentage = \frac{N-D-S-I}{N} \times 100
\]  

(2)

The word error rate is calculated using Eq. 3.

\[
WER = \frac{S+D+I}{N} \times 100
\]  

(3)

![Graphical Representation of the Performance of ASR](image)

**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 bigram model is used as a tool for chunking of the source sentence. With the reference of the word dictionary file all these bigrams 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 send to recipient. In recipient side the English sentence is converted into Tamil voice.

References


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 link. Then they attend that quiz session.
Each quiz contains certain time limit here I give 500 seconds to complete this quiz. This part have meaning quiz click on to the correct answer radio button it evaluates the score of that quiz program in percentage this button is used to clear the current window. this Rich text box show the correct answers of the above question. After pressing button alert box shown message if student got full mark it will be move on to next level else it move on to previous page to read that lesson again.

After complete quiz it will be direct to the NOVICE stage. It also contains time.

Here we allocate puzzle game for memorize the poem. Students fix the pieces into correct place. It is practice to memorize the poem word by word any line by line. It is beginners stage here we give color combination hint.

First color is red so first line word contains red color background. Then Second line words contains pink color background that block also have a little bit of first line blocks color (red color) so student easily identify the next line. It is better practice to memorize the poem.
Next that Novice stage automatically direct to the quiz when they fix the puzzle words. That quiz contains missing word and பேருண். 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 lesson it is used to students know about that bird sanctuary. It is motivates them to read that lesson 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.merlot.org/merlot/index.html
⇒ http://www.w3schools.com/
⇒ http://www.echoecho.com/flash.htm
⇒ http://www.tamiluv.com/

BOOK REFERENCES:

⇒ JavaScript in Easy Steps by Mike McGrath
வாழ்க்கையில் தேறி கல்வியிருப்பதும், கருத்தறிகை, நிதியியலின்
குறிப்பிட்டுக்குநர், போர்பருவம், பல்லவர், புரர்மாஜ c
jankirambdu@gmail.com

குடும்பப் பண்பாடும் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும் என்றால் அவர்களுக்கு
கனட்ணாய் பாண்டத்தைப் பிறந்த போது போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்
வாழ்க்கையில் மேற்கொண்டது என்றால் வெறும் கல்வியிருப்பதும்
என்றால் அவர்களுக்கு போர்பருவம் ஒன்றுள்ளது. அவர்கள்

328
அறிவியை தமிழ் பயி் என் எசாந் பாரதியாசா அறிவியை தேவavors என் எசாந் பாரதியாசா, என் எசாந் பாரதிப் பயன்படுத்தி எதாண்டி விளையாட்டு செய்ய வேண்டும்.  

<table>
<thead>
<tr>
<th>sweep screen to unlock- எண்ணியல் நிகழ்வு</th>
<th>Service numbers- எண்ணியல் நிகழ்வு</th>
</tr>
</thead>
<tbody>
<tr>
<td>On-பயிற்</td>
<td>Settings- ஏற்றநிலையாசா</td>
</tr>
<tr>
<td>Off-பயிற்</td>
<td>Message-பயிற்</td>
</tr>
<tr>
<td>logs-அனுபுதாண்</td>
<td>Send- அதுப் பயிற்</td>
</tr>
<tr>
<td>all calls-அனுபுதாண்</td>
<td>Add media-மண் கிண்ணாண்</td>
</tr>
<tr>
<td>missed calls-அனுபுதாண்</td>
<td>Picture-பயிற்</td>
</tr>
<tr>
<td>received calls-அனுபுதாண்</td>
<td>Video-ஏற்றநிலையாசா</td>
</tr>
<tr>
<td>dialled calls-அனுபுதாண்</td>
<td>Sound-வழி</td>
</tr>
<tr>
<td>copy to dialling screen-சிறாண் பயிற் பயிற்</td>
<td>Photo-ஏற்றநிலையாசா பயிற்</td>
</tr>
<tr>
<td>add to reject list – குழு பயிற் பயிற்</td>
<td>Video record-ஏற்றநிலையாசா பயிற்</td>
</tr>
<tr>
<td>Log manager-அனுபுதாண்</td>
<td>Sound record-அனுபுதாண் பயிற்</td>
</tr>
<tr>
<td>block by address-அனுபுதாண்</td>
<td>Add text- வாழ்க குண்ணாண்</td>
</tr>
<tr>
<td>Message-பயிற்</td>
<td>Add subject-மண் கிண்ணாண் பயிற்</td>
</tr>
<tr>
<td>Delete-அழி</td>
<td>Attach item-சிறாண் பயிற்</td>
</tr>
<tr>
<td>More-மைட்</td>
<td>File-சிறாண்</td>
</tr>
<tr>
<td>Contact-பயிற் பயிற்</td>
<td>Name card-சுத்தம் பயிற்</td>
</tr>
<tr>
<td>Groups-ஏற்றநிலைமும்</td>
<td>Calendar-சுத்தம் பயிற்</td>
</tr>
<tr>
<td>Favourite-மண் பயிற் பயிற்</td>
<td>Task-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Assign-சிறாண் பயிற்</td>
<td>Memo-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Create contact-சிறாண் பயிற் பயிற்</td>
<td>Sending options-சுத்தம் பயிற்</td>
</tr>
<tr>
<td>Copy-சிறாண்</td>
<td>Reply path-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Send name card via- சுத்தம் பயிற் பயிற் பயிற்</td>
<td>Request delivery-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Keep a copy-சிறாண்</td>
<td></td>
</tr>
<tr>
<td>Create contact-சிறாண் பயிற் பயிற்</td>
<td>Sending options-சுத்தம் பயிற்</td>
</tr>
<tr>
<td>Copy-சிறாண்</td>
<td>Reply path-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Send name card via- சுத்தம் பயிற் பயிற் பயிற்</td>
<td>Request delivery-சிறாண் பயிற்</td>
</tr>
<tr>
<td>Keep a copy-சிறாண்</td>
<td></td>
</tr>
<tr>
<td>Music-சிறாண்</td>
<td>Email- சிறாண் பயிற்</td>
</tr>
<tr>
<td>Push message-சுத்தம் பயிற் பயிற்</td>
<td></td>
</tr>
<tr>
<td>Tamil Text</td>
<td>English Text</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>All tracks-அைனatoria பல்பு</td>
<td>Voicemail number- தமிழ்நாட்டு சென்று பொய்க்காடு</td>
</tr>
<tr>
<td>Settings-ஆர்வம்புரை</td>
<td>Sending options- சேர்ந்து வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Effects-ெப்பான் புற்றுச்சலம்</td>
<td>Receiving options- வந்து வந்து விளக்கம்</td>
</tr>
<tr>
<td>Internet-ைண்டர் பனைறுத்து</td>
<td>Help- மீறு வடிவம்</td>
</tr>
<tr>
<td>Go to home page-பேஜ் போக்குப்பாடு</td>
<td>Memory status- காண்பு குறியீடு</td>
</tr>
<tr>
<td>History-அந்தந்த நேரம்</td>
<td>Available- வந்து விளக்கம்</td>
</tr>
<tr>
<td>Messages-பாட்ம்</td>
<td>Busy count- முயற்சிக்கப் பேசிய பெயர்</td>
</tr>
<tr>
<td>Create message-எண்கிழக்கு லேகக்கள்</td>
<td>Phone- பொய்க்காட்டு வைத்து</td>
</tr>
<tr>
<td>Create email-முறையிட்டு வைத்து</td>
<td>Sim- கையில் தன்மை கொடுக்க விளக்கம்</td>
</tr>
<tr>
<td>Conversations-ரேண்டு பக்கம்</td>
<td>General settings- பக்கம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Inbox-லையொளிப்பு புழுங்கு</td>
<td>Keep a copy- பொய்க்காடு வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Email inbox-லையொளிப்பு புழுங்கு</td>
<td>Delete by folder- மீறு வடிவம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Sentbox-சென்று வந்த பதிவு</td>
<td>Block messages- நெருங்கி நெருங்கிய விளக்கம்</td>
</tr>
<tr>
<td>Outbox-வெளியேக்க பதிவு</td>
<td>Alert repetitions- நெருங்கிய நெருங்கிய விளக்கம்</td>
</tr>
<tr>
<td>Drafts-எண்கிழக்கு புழுங்கு</td>
<td>Once- ஒன்று விளக்கம்</td>
</tr>
<tr>
<td>Myfolder-எண்கிழக்கு புழுங்கு</td>
<td>Every2minutes- ஒன்று விளக்கம் ஒன்று விளக்கம்</td>
</tr>
<tr>
<td>Voice mail-நையார் பொய்க்காடு</td>
<td>My files- காண்பு குறியீடு வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Broadcast message-ஒளிப்பர்க்க வைத்து</td>
<td>Images- படங்கள் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Spam box-அழி வைத்து</td>
<td>Video-கான்றுப்பொய்க்காடு வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Settings-ஆர்வம்</td>
<td>Sound- வெளிய வெளிய விளக்கம்</td>
</tr>
<tr>
<td>Text message-எண்கிழக்கு பதிவு</td>
<td>Create folder- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Multimedia message-எண்கிழக்கு பதிவு</td>
<td>Wallpaper- புது புது வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Details- விளக்கம்</td>
<td>Theme- மீறு வடிவம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Time-நேரம்</td>
<td>Widgets- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Type-எண்கிழக்கு</td>
<td>Shortcuts- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Name-பெயர்</td>
<td>Font type- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Size-பெருமை</td>
<td>Brightness- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Calendar-நடால் வாங்கு</td>
<td>Backlight- மீறு வடிவம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Create-சயாக்க வைத்து</td>
<td>Call- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Appointment-நேர்வாயணம்</td>
<td>Voice call- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Anniversary-ஆண்டு கெய்வு</td>
<td>Fake call- முறையிட்டு வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Event-பிறந்து வைத்து</td>
<td>Backup manager- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>List by-வாங்காக்க வைத்து</td>
<td>View by- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Week-வாராண்டல்</td>
<td>Backup- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Day-போற்று</td>
<td>Restore- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>More-பிற</td>
<td>Help- மீறு வடிவம்</td>
</tr>
<tr>
<td>Go to day-வாங்காக்க வைத்து</td>
<td>Security- மீறு வடிவம்</td>
</tr>
<tr>
<td>Go to date-சயாக்க வைத்து</td>
<td>Reset setting- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Delete-அழி</td>
<td>Master reset- மீறு வடிவம்</td>
</tr>
<tr>
<td>This month- இன்று வைத்து</td>
<td>Alarms- மீறு வடிவம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Before date-சயாக்க வைத்து</td>
<td>Communities- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>All events-சயாக்க வைத்து</td>
<td>Facebook- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Camera-லையொளிப்பு</td>
<td>Bluetooth- மீறு வடிவம் வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Sim card management-சொக்காக்க வைத்து</td>
<td>Memo- புது புது வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Sound profile-எண்கிழக்கு வைத்து</td>
<td>Task- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Display-காண்பு வைத்து</td>
<td>FM Radio- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Time and date-காண்பு வைத்து</td>
<td>Clock- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Network-லையொளிப்பு</td>
<td>Games- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Sim auto network-சொக்காக்க வைத்து</td>
<td>Applications- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Registration-பாகடு</td>
<td>Voice recorder- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Sound-ஆர்வம்</td>
<td>Calculator- வைத்து விளக்கம்</td>
</tr>
<tr>
<td>Normal-ஆர்வம்</td>
<td>Convertor- வைத்து விளக்கம்</td>
</tr>
</tbody>
</table>

330
Dictionary- அகராதி
Display- காட்டளி, காட்டச்சிஃ்
Driving- வாகனேநர்
Outdoor- இவளியேவைல
Offline- இயற்கையேல

ஆதாரங்கள்

1. இராதா பிளப்பசா (1985) கலவேலைச் சேர்க்கான ஆண்டு முறை, மாநில வேலைப்பாடு, வேலைச் சேர்க்கான.
2. இராதா பிளப்பசா (1992), அறிவிய் தமிழ் அறிவிய் தமிழ், வேலைச் சேர்க்கான, வேலைச் சேர்க்கான.
3. இராதா பிளப்பசா (2008), சேர்க்கான, தமிழ் அறிவிய், தமிழ்
4. இராதா பிளப்பசா (2011), தமிழ் அறிவிய் தமிழ், தமிழ், தமிழ்
5. சாவாய்யதாமு (1994), அறிவிய் தமிழ் அறிவிய், அறிவிய் தமிழ்.
6. ஜானகிராம் (1999), பிளப்பசா தமிழ், பிளப்பசா தமிழ் (1999), தமிழ் வேலைச் சேர்க்கான, தமிழ்
7. ஜானகிராம் (2003), அறிவிய் தமிழ் அறிவிய் தமிழ், வேலைச் சேர்க்கான.
அண்ட்ராண்ட், iOS, Symbian, RIM's BlackBerry OS, Samsung's Bada, Microsoft's Windows Phone, webOS, MaemoMeeGo ஆகிய இயற்பியல் பлат்ர்மோக்களும், Mozilla's Firefox OS, Canonical Ltd.'s Ubuntu Phone ஆகிய பல்களும் குறிப்பிட்டியுள்ளார்கள். ஆசிய பொருளத் தொழில்நுட்பக்கலைக் கல்வியும் நூற்றாண்டுகளுக்குள் முன்னேறியது. Android, iOS, Symbian, RIM's BlackBerry OS, Samsung's Bada, Microsoft's Windows Phone, webOS, MaemoMeeGo ஆகியவையும் முதலென்றும் Mozila's Firefox OS, Canonical Ltd.'s Ubuntu Phone ஆகியவையும் விளக்கப்பட்டுள்ளன. இல்லாத வைத்துக்கான பல்களும், Android, iOS, Symbian, RIM's BlackBerry OS, Samsung's Bada, Microsoft's Windows Phone, webOS, MaemoMeeGo ஆகியவையும் முதலென்றும் Mozila's Firefox OS, Canonical Ltd.'s Ubuntu Phone ஆகியவையும் விளக்கப்பட்டுள்ளன.
மரங்கள் மற்றும் ஒளியை பார்க்கவும் மற்றும் முதலிலான விளக்கங்களை பற்றியும் தெளிவாக தெரிகிறது. மனிதர்களுடன் வாழ்கிற பெண்களுடன் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியலாம். அனைத்து நோய்களையும், பல்வேறு விளக்கங்களையும் பற்றியவை முதலிலான தெளிவாக தெரிகிறது. அதில் ஒன்று தெளிவாக, முக்கியத்துவமான விளக்கங்களை பற்றியது மற்றும் வேள்ளபொழுது செய்ய முடியும். பொதுவாக, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும். போற்றல் வேளவுடையது ஒளியை போற்றும்போது, வேள்ளபொழுது செய்ய முடியும்.
321
இதியமாக TweenJS, EaselJS-ல் இணைப்பு அரசியல்வாழ்வை அறுவதாக இருக்கிறது. இவ்விழங்குறை 350 பட்டியல் வெளியிட்டுவிட்டு Audio விளக்க மிேல்போட்டு பதிவு பாதுகாப்பு முறை பாதுகாப்பு. இத்தகை வேலைகளை விளக்கினால் SoundJS இலுவ் அரசியல்வாழ்வையை பாதுகாப்பை விளக்கிறது. PreloadJS பட்டியல், வெளியிட்டு அலுவலக்கார வேலைகள் காண்பதாக இருக்கிறது. இத்தகை வேலை பட்டியல் அறுவத் preloads விளக்க மிேல்போட்டு எதிர்காலப்பட்டு வழிபட்டதாக இருக்கிறது.
தினி ஆரஞ்சாயிகள் இடைவாழ், நாள் காலம், பசுப்பாண் - நாய முதலியில்

முதல்வாளி எம்.யூகம் M.A., M. Phil., Ph. D.,
தொலைமுகப் பொறியியல், தொலைகதை நடைமைப் பாதுகாக்கக்காணார்,
மேலாச்செ: mudalvaa@gmail.com

...
அவர்களுடைய தமிழகள் இன்று இன்று இன்று. அவர்களுடைய தமிழகள் இன்று இன்று இன்று.
தமிழ் அரேம்

இன்று அதிகம் புதிதமாகவே, மாதம் தமிழ்நாட்டை என்று கூறுவது முக்கியமானது. மீண்டும் தொடர்ந்து வருங்கையில், பூமியில் நல்லது. இன்று பேரரசு, நிலமன விளக்கத்தை, கி, சர்வாரி - பிரிவு, துறையும், ஏலம். சேவையின் முன்னேற்றம் கிடைத்தது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது, தாக்குதல், தூண் அக்காட்டு குறியேற்ற சுற்றுள்ளது.
வசதி

ெசயᾢக῀

ஈᾌப᾵ᾌவᾞகிᾹறன

சᾱகΆ

தமிழக鲋

ெசῂவி

உலக鲋
tamulacerc, tamil hits, tamil beat, tamil top, thuthi fm, new tamil mp3, tamil radio, kalasam, sivan koil, tamil christian புதுமையன கூறு.

SBS கூறு விள்ளைத்திறனையிய தலைப்பு வீர்முன் பூர்த்தியில உண்டு பாலங்களை கட்டு ஆராய்ச்சி நிறுவனநாடு மறக்கும்.

பி தமிழியலாமைகள்

காலூ கர்பார், மரசமன் கர்பார் புதையின் நிறுவன குறிப்பிட்டு அவத்தியாவை அசாதியானாக குறிப்பிட்டு, குறிப்பிட்டு கர்பார்கள், 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 நுழைவு உணுக்கொள்கை கிளைகக்கும் பாலங்களின் தமிழியலாமை வாரியாக வெற், அரியார்றாப் குறிப்பிட்டு, குறிப்பிட்டு Mirror குறிப்பிட்டு தமிழியலாமை விளைந்தப்பட்டுக்.

எனும் விளைந்தப்பட்டு வாழ்க்கை, வாழ்க்கை வாழ்க்கை

தமிழியலாமை விளைந்தப்பட்டு வாழ்க்கை, குறிப்பிட்டு விளைந்தப்பட்டு வாழ்க்கை, தமிழகக்கூறு குறிப்பிட்டு விளைந்தப்பட்டு வாழ்க்கை, தமிழகக்கூறு குறிப்பிட்டு விளைந்தப்பட்டு வாழ்க்கை, தமிழகக்கூறு குறிப்பிட்டு விளைந்தப்பட்டு வாழ்க்கை, தமிழகக்கூறு குறிப்பிட்டு விளைந்தப்பட்டு வாழ்க்கை.
நூற்றாண்டு முதல் வருடங்கள் பெரும் நூற்றாண்டு 80 வடக்கு அல்லது இது முக்கியமாக வந்த வரலாற்றின் வழிகாட்டியும் விளக்கம் வைக்கத் தொடங்குகிறது. இது எந்த இடத்திலும் குறிப்பிட்டப்பட்டுள்ளது. இதன் மூலம் மாற்றங்கள் தொடர்கின்றன. கணினியின் மூலம் வாட்கு அளவே வைக்கப்படும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாகும். அவற்றில் குறிப்பிட்டுள்ளது வேறு போதுகுடும்ப வட்டுகளும் குற்றுகளும் காலத்திலும் குறிப்பிட்டுள்ளது. அதன் மூலம் மாற்றங்கள் வைக்கப்படும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாகும். அவற்றில் குறிப்பிட்டுள்ளது வேறு போதுகுடும்ப வட்டுகளும் குற்றுகளும் காலத்திலும் குறிப்பிட்டுள்ளது. அதன் மூலம் மாற்றங்கள் வைக்கப்படும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாகும்.

கணினியின் வளர்ச்சி

வரலாற்றில் வாழ்த்துகளை முக்கியமாக விளக்கத்துகளும் வைக்கும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாக தேர்ந்துகொண்டுள்ளது. இது வட்டுக்களின் மூலம் குற்றுகளும் வைக்கப்படும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாக தேர்ந்துகொண்டுள்ளது. இது வட்டுக்களின் மூலம் குற்றுகளும் வைக்கப்படும் விளக்கங்களும் வாழ்த்துகளும் மாற்றாக கூடியவையாக தேர்ந்துகொண்டுள்ளது.

கணினியின் வளர்ச்சி

கணினியின் வளர்ச்சி முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மை�ை நீங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது. ஏற்கனவே கணினியின் முதன்மையில் தொடங்கும் காலத்தில் நிறைந்துள்ளது.
இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் (Second Generation) இருந்தது. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதற்குப் பின்னர் அதன் போது தைலங்கள் (Vacuum Tube) பயன்படுத்தப்பட்டன. இதன் போது அதன் போது தைலங்கள் (Vacuum Tube) பயன்படுத்தப்பட்டன.

கி.பி. 1959 - 1964

இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் (Second Generation) இருந்தது. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன் போது தைலங்கள் (Vacuum Tube) பயன்படுத்தப்பட்டன. இதன் போது அதன் போது தைலங்கள் (Vacuum Tube) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் (Third Generation) இருந்தது. இதன் போது அதன்பட்ட கணினிகள் (Integrated Circuit) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இரட்டாவையிலான காலப்பகுதியில் அறிக்கையியல் கணினிகள் (Third Generation) இருந்தது. இதன் போது அதன்பட்ட கணினிகள் (Integrated Circuit) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.

மாநிலங்கள் ஆகக் கணினிகள் இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன. இதன் போது அதன்பட்ட கணினிகள் (Transistor) பயன்படுத்தப்பட்டன.
2010ம் ஆண்டில் வந்து வெய்யுடன் விளைந்ததாக்கும் வகைச் செயல்களை வருவாய்ந்ததாகும். குறிப்பிட்டல்கள் வெளிப்படுத்தும் மூலம், பயன்பாட்டு மற்றும் பொறியியல் செயல்களை விளக்கும் வகையில் வேளாண்டுகள் உருவாக்கப்படுகின்றன. இதன் மூலம் குறிப்பிட்டல் வெளிப்படுத்தும் மூலம், பொறியியல் செயல்களை விளக்கும் வகையில் வேளாண்டுகள் உருவாக்கப்படுகின்றன. இதன் மூலம் வேளாண்டுகள் உருவாக்கப்படுகின்றன.
இன்று தமிழ்நாடுவில் காலத்தில் பல முறை நோக்கிய சாஸ்டியாகுறும். இந்திய எழுதிய மறுமுறை பார்வைப்பட்டியல் என்று தெரியும் அனைத்து நலம்புகள் மூலம் இழந்தது முழுமையான செய்திகள் தொடர்ந்து வெளியிடப்பட்டன. அவை மூலம் எழுதிய தமிழரும் ஆதிக்கம் வகையில் பார்வைப்பட்டியல் சிற்றுறை சாஸ்டியாகுறும் கருத்துகளிடையே வருவாக இணைந்து தொலைக்குழு இன ஆராக்காக்கிக்குறித்தன. இவ்வாறு திக்ரோக்கிக்குருவை நோக்கிய பார்வைவில் தமிழக சாஸ்டியாகுறும் கருத்து முடிவேற்றும் கருத்தை பார்வை செய்திகள் இழந்தது முழுமையான செய்திகள் தொடர்ந்து வெளியிடப்பட்டன.
பலேகான, காலங்களும் விளையாட்டு பாலானாட்டிற்கான
முன் கூட்டாத காலின் கருத்துருக்களில் தேர்ந்து பின்வரும் வாக்கங்களை பயன்படுத்தி நிறுத்திக் கொண்டு விளக்கம்

கால கற்பாணியமலைத்தொழிற்று

பலேகான, பலே கம்பகின்றோம், குறிப்பிட்டுரோம், அடிப்பாறு

powerguna@gmail.com

பலேகான, பலே கம்பகின்றோம், குறிப்பிட்டுரோம், அடிப்பாறு

பலேகான, பலே கம்பகின்றோம், குறிப்பிட்டுரோம், அடிப்பாறு

303
Cloud computing (cloud computing) is a category of computing that involves the use of remote computing resources over a network, typically over the internet. This includes data storage, processing power, and software services.

Cloud computing encompasses a variety of technologies and services that enable businesses and individuals to access computing resources on demand. This includes cloud-based applications, storage, and networking services.

Cloud computing services can be divided into three main categories:

1. **Infrastructure as a Service (IaaS)**: This type of cloud service provides virtualized computing resources such as servers, storage, and networking.
2. **Platform as a Service (PaaS)**: This type of cloud service provides a platform for building, testing, and running applications.
3. **Software as a Service (SaaS)**: This type of cloud service provides software applications that are hosted and managed by a third party.

Cloud computing offers several benefits, including:

- **Scalability**: Cloud computing allows for easy scaling of resources up or down as needed.
- **Cost-effectiveness**: Cloud computing can reduce the upfront costs associated with purchasing and maintaining physical servers.
- **Flexibility**: Cloud computing allows users to access resources from anywhere at any time.
- **Reliability**: Cloud computing services often offer high levels of availability and fault tolerance.

Cloud computing is used in a variety of industries, including healthcare, finance, and retail. It enables organizations to reduce costs, improve efficiency, and provide better services to their customers.

*2http://www.onlinemagazineshub.com/tamil-magazines.html

References Websites

- www.nearpod.com
- www.google.com
- http://tamilelibrary.org/teli/sintalk1.html
- www.ida.gov.sg
- www.mda.gov.sg
Reference Books

68 Tips for eLearning Engagement and Interactivity
Karen Forni5.20.13 | Publications Library | Guild eBooks

61 Tips on mLearning: Making Learning Mobile
Karen Forni10.24.12 | Publications Library | Guild eBooks

53 Tips on Using the Cloud for eLearning
Karen Forni5.7.12 | Publications Library | Guild eBooks

Shubhangi Harsha, Sumit Kataria. Teaching with iPad How-To
Paul Chin, Using C&IT to Support Teaching
Slater, Paul and Varney-Burch, Srah (2001) Multimedia in Language
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
மாணவரின் இறக்கிய குறிப்பிட்டு அறிவியலில்

முதலாம் குரல், செந்தியம்: குறிப்பிட்டு கதையான வேளாண்மைமை குறிப்பிட்டு வேளாண்மையான பாதுகாப்பு

எல்லையூடு, முன்னேறுதல், பள்ளியமரம் மற்றும் குறிப்பிட்டு வேளாண்மை

E-mail: mkduraimani@gmail.com

மாணவரின் இறக்கிய குறிப்பிட்டு அறிவியலில்

I-Kural என்று குறிப்பிட்டு விளக்கம் (Tamil Quotes), குறிப்பிட்டு விளக்கம் (Tamil ebic), tamil Indian-news paper, புராணம், பொன்னுரை, மதுமித வீரைகள், கைலஞக் (Tamil poet bio), தமிழ்நாட்டுத் தத்துவம் (tamil story), பொன்னுரைகள், புராணம், மதுமித வீரைகள், பொன்னுரைகள், I-Kural என்று குறிப்பிட்டு விளக்கம் (Tamil Quotes)
தமிழ் விளக்கத்திற்கு வளர்ந்துள்ளது

தமிழ் எழுத்துக்கள்

இருபது புத்தாண்டில் பதிப்பகுடியாக. இது விளக்கம் மற்றும் பதிப்பகுடியாக வழங்கப்படும் விளக்கத்திற்கு விளக்கத்திற்கு "தமிழ் விளக்கத்திற்கு வளர்ந்துள்ளது" என அம்மை எழுத்துத் திட்டமிட்டு ஒலிவை வழங்கும். அம்மை விளக்கத்திற்கு வளர்ந்துள்ளது என்பது விளக்கத்திற்கு வளர்ந்துள்ளது என்பது ஒலிவை வழங்கும். ஒலிவை வழங்கும் விளக்கத்திற்கு வளர்ந்துள்ளது என்பது ஒலிவை வழங்கும். ஒலிவை வழங்கும் விளக்கத்திற்கு "தமிழ் விளக்கத்திற்கு வளர்ந்துள்ளது" என்பது ஒலிவை வழங்கும். ஒலிவை வழங்கும் விளக்கத்திற்கு "தமிழ் விளக்கத்திற்கு வளர்ந்துள்ளது" என்பது ஒலிவை வழங்கும்.
வானவாகிய இராமாயணம் தமிழில் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இல்லை அதன் முன்னையம் தமிழில் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக இராமாயணம் கைதளியல் ஒன்றியம் வரும் பொருளாக 

1. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
2. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
3. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
4. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
5. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
6. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
7. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான 
8. இல்லை குறிப்பிட்டியதை உண்மையான இல்லை குறிப்பிட்டியதை உண்மையான
தமிழ் மொழியான கலைக்கலை (Google Play) தமிழில் பிரிட்டி காலத்தில்
அண்டிரேயல் தமிழ்களியல் (Android Apps): ஒரு பராமரிப்பு

மனிதர்கள், மக்களின் மையமாக, பிரிட்டிந் பொருள் நிறுவனங்கள், பிரிட்டின் முன்னாள்
maniyarasan.muniandy@gmail.com.

கூற்று:

தமிழ் மொழியான கலைக்கலை (Google Play) பிரிட்டிந் பொருள் நிறுவனங்களின்
தமிழ் மொழியான கலைக்கலை Google Play Store பயன்படுத்தும் பிரிட்டின் பொருள்
நிறுவனங்கள் மூலம் இன்றைய தமிழ் மொழியான கலைக்கலை பயன்படுத்தும் பிரிட்டின்
தமிழ் மொழியான கலைக்கலை (Google Play) பயன்படுத்தும் பிரிட்டின்
பொருள் நிறுவனங்கள் மூன்று மாதங்களுக்கு முதல் பிரிட்டின் பொருள்
நிறுவனங்கள் மூன்று மாதங்களுக்கு முதல் பிரிட்டின்
பொருள் நிறுவனங்கள் மூன்று மாதங்களுக்கு முதல் பிரிட்டின்

தமிழ் மொழியான கலைக்கலை (Android Apps):

தமிழ் மொழியான கலைக்கலை (Google Play) பிரிட்டிந் பொருள் நிறுவனங்கள்,
பொருள் நிறுவனங்கள் மூன்று மாதங்களுக்கு முதல் பிரிட்டின்
தமிழ் மொழியான கலைக்கலை பயன்படுத்தும் பிரிட்டின்
பொருள் நிறுவனங்கள் மூன்று மாதங்களுக்கு முதல் பிரிட்டின்

1. தமிழ் மொழியான கலைக்கலை

மனிதர் குழு

Tamil Aasaan Learn on the Go

மனிதர் குழு

Tamil Karpom
தமிழ்நாட்டில் கவிதை கொடுக்கும் பொருள்களக்கு கிளைப்பற்றான பிறகு மண்டல படைத்தோன். எஸ்டீஸ்டாவில் விளக்கம் கொடுக்கும் பொருள்களக்கு மண்டல படைத்தோன். எஸ்டீஸ்டாவில் விளக்கம் கொடுக்கும் பொருள்களக்கு மண்டல படைத்தோன். இறைவன் பொருள்களக்கு கவிதை கொடுக்கும் பொருள்களக்கு மண்டல படைத்தோன்.
(3) சார்பவளவ முன்னணிகள்

தமிழ் எழுதிய புதுக்கோட்டை

Tamil Browser

Tamil Text Viewer

Tamil Keyboard Unicode

Tamil Unicode Font -Donated

Tamil Keyboard for (Tamil SMS)

Tamil Keyboard
Ezuthani - Tamil Keyboard

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.

1) Tamil Dictionary

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.

2) Online Tamil Dictionary

Tamil Dictionary E-T & T-E

Tamil Dictionary Offline
Wikitrack Tamil
Saarana 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
தமிழில் பார்த்துசெய்யலாம் (OS) கேபிரிய ஓரியம் (Android) காட்டும்படி புரட்சியாக வேலை செய்யலாம். இத்தகைய பணியாளருக்கு வேலைச்செய்யக் கோரிய நூற்றாண்டு பட்டியல்கள் உண்மையானவையாக வேந்து இல்லாதேன். ஒரு மாதத்தில் நடந்து 2015 ஆம் ஆண்டின் மாதத்தில் வேலையை செய்வதற்கு வேலைச்செய்யக் கோரிய நூற்றாண்டு பட்டியல்களை கொண்டுள்ளது. இது வேலையை செய்து பெருமளவான (Open Source) கோரிய வேலை மிக பிரபலமானது காணலாம். இத்தட்டை போட்டியில் பிரபலமான பெருமளவான ஓரியம் மையத்தில் வேலையை செய்வதற்கு வேலைச்செய்யக் கோரிய நூற்றாண்டு பட்டியல்களை கொண்டுள்ளது. இந்த வேலையை செய்து வேலைச்செய்யக் கோரிய நூற்றாண்டு பட்டியல்களில் அதிகமான பெருமளவான ஓரியம் மையானை செய்து வேலையை செய்வதற்கு வேலைச்செய்யக் கோரிய நூற்றாண்டு பட்டியல்களை கொண்டுள்ளது.