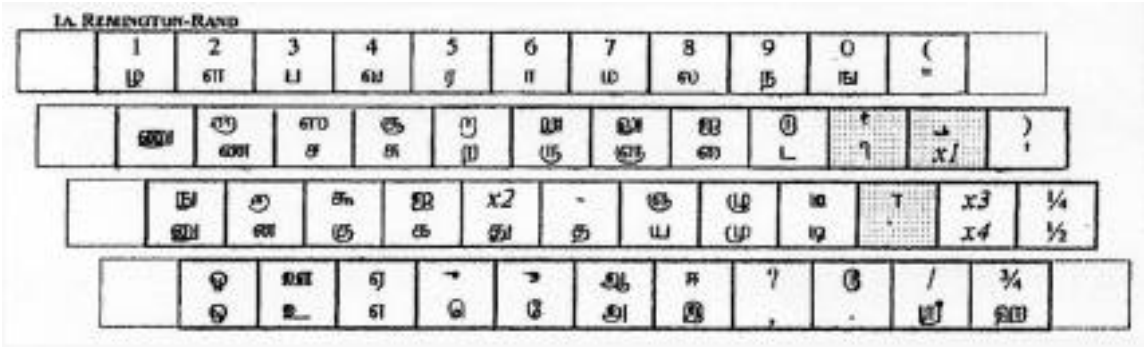
Ergonomic considerations in the design of the Tamil keyboard layout

S Sivasegaram

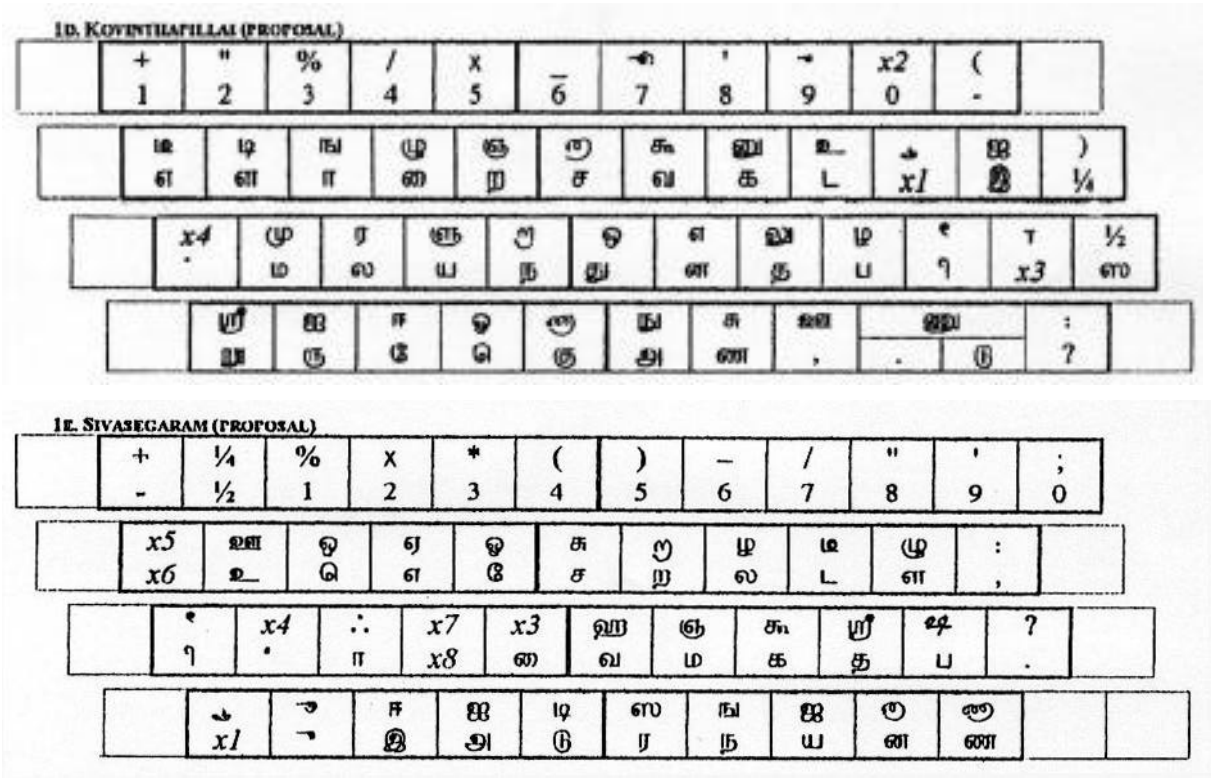
Faculty of Engineering, University of Peradeniya, Sri Lanka

The beginning

The Tamil typewriter came into use near the middle of the last century. Two makes, Remington-Rand and Olympia (**Figs 1a,b**), based on the machine used for the Roman script entered the market, with distinct keyboard layouts and the mechanism modified to accommodate certain diacritical devices on "dead keys". Neither keyboard layout seemed to be scientifically designed. Subsequently, the layout was modified so that the Tamil characters were located along the second, third and fourth rows of the keyboard, with the four Grantha consonant symbols occupying top row uppercase positions. This layout New Remington, **Fig 1c**, after Siromoney (1963), was standard until the arrival of electric typewriting, which allowed greater flexibility in keyboard layout.



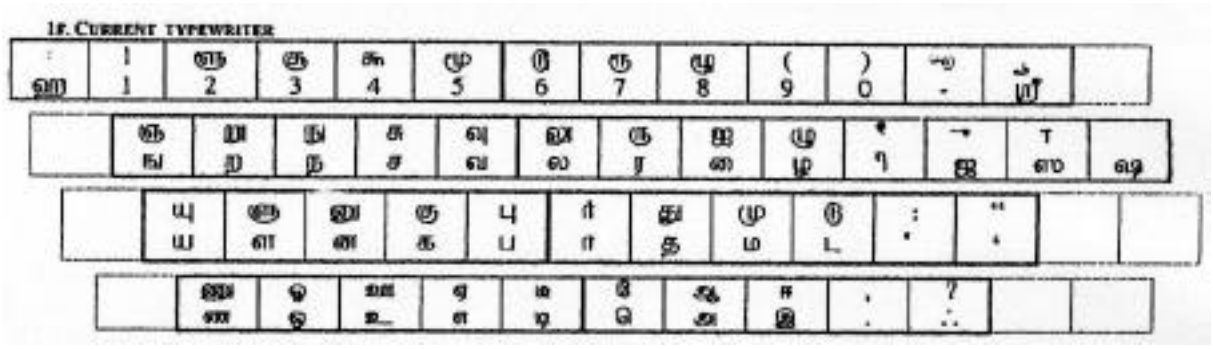
Ergonomic considerations prompted Kovinthapillai (1971) to propose the layout of **Fig. 1d,** based on an even loading of the fingers; and Sivasegaram (1976, 1979) that of **Fig. 1e**, based on load according to the capability of each finger as determined by Hoke (cited in Barnes, 1964), and considerations of sequence of operation of the hands, compactness of the keyboard and ease of mastery. The proposals, like any other without state patronage, were doomed to fail since the cost of changeover to the manufacturer had to be offset by financial benefits and the state was the single biggest consumer of the product.



An interesting result of the studies of Kovinthapillai and Sivasegaram, based on widely different sets of text, was that the overall frequency distributions for the main symbols and diacritical devices did not show a statistically significant difference. A recent study by Cheran, Vaseeharan and Cheran, based on sample prose writings from the Project Madurai database, also showed essentially the same pattern. Texts with a highly Sanskritised or anglicised style could lead to a difference in distribution, with a moderately higher frequency for the Grantha characters. It is unlikely, however, that the distribution for the traditional characters will differ significantly from that for 'purer' Tamil.

The current typewriter keyboard layout (**Fig. 1f**) retaining many of the features of the New Remington keyboard was not optimised for load distribution between fingers. It has, however, features that are beneficial to learners of touch-typing and to an increasing proportion of users with no experience in touch-typing. The design was subject to the restriction that the keyboard had to accommodate, besides the main vowel and consonant symbols and the 'aaytham', the four Grantha characters and the cluster *srii,* diacritical device for the pure consonant, devices for the *aa*, *i* and *ii* series, two sets of *u* and *uu* devices as used with the consonants *p, y* etc. and with the Grantha letters, the device for lengthening the vowel in *Nu, thu, nu* etc., the devices for the *e, ee* and *ai* series, 14 vowel consonants of the *u* series and 8 of the *uu* series, as well as *di* and *dii.* It should also be noted that designs preceding the script reform of the late 1980s had to accommodate three additional vowel consonants in the *u* series

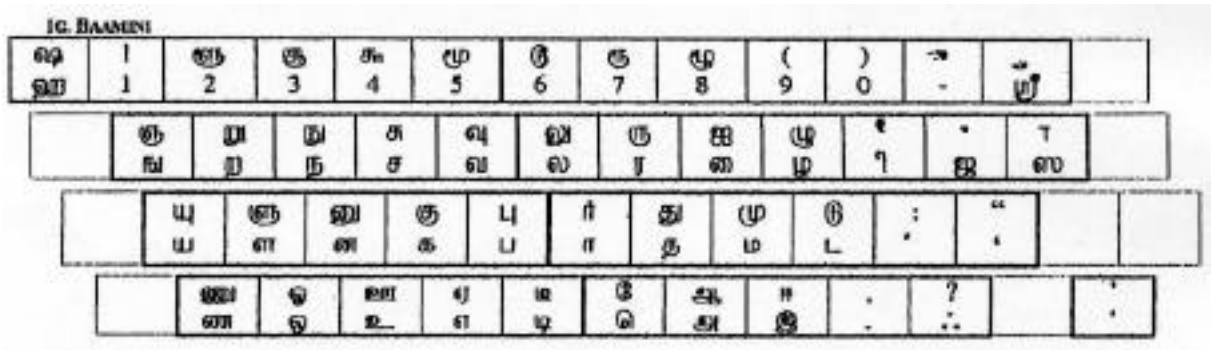
and an additional device for the *ai* series, and that all designs excluded at least one symbol that was not in common use.



The layout also sacrificed a degree of compactness in the interest of consistency, to allocate the upper case of all but two consonant symbols to the corresponding *u* vowel consonant. As in the case of its predecessors, attention to the load distribution between fingers was poor as was that to the sequence of use of the two hands.

Developments since exposure to electronic communication

The enthusiasm of the Tamil diaspora for adopting Tamil to the electronic media also had the negative effect of individuals who took the initiative to develop fonts for Tamil word-processing choosing to make "personalised improvements" to the layout.



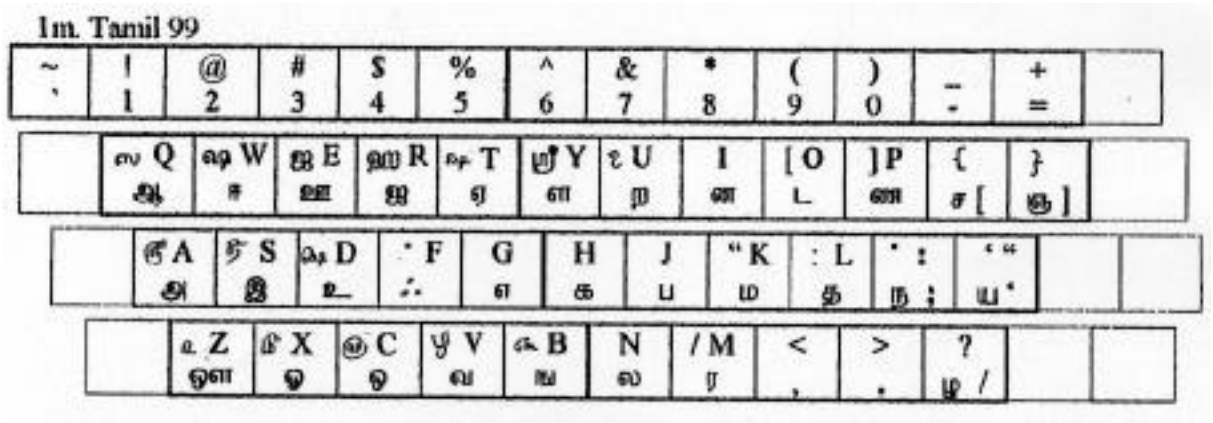
While the majority of font systems (like the one in **Fig. 1g**) conformed to the current typewriter keyboard, with differences affecting only the changes to the location of certain keys, several deviated from it in the location of some of the characters falling outside the three main rows. These font systems, by and large, are now not in general use. There have also been designs that differed more important ways, for purposes which will be commented on in the paragraphs that follow. The net result, however, was that documents prepared with one set of fonts could not always be correctly reproduced with another. The gravity of this problem has been recognised so that publishers insist on preparation of documents produced either with a true font system conforming to the standard typewriter keyboard or with a word-processing software of their choice.

The word-processing software, Anjal pioneered the use of the Roman keyboard and represents an alphabetic system of writing, as opposed to the syllabic systems deriving from the Brahmi script. It particularly benefits those without access to Tamil keyboards and, very importantly, averts problems arising from the diversity in keyboard layout.

The design of **Fig 1h** was inspired by that of Anjal but the keyboard was functionally similar to that of **Fig. 1g**. Some subsequent keyboard designs (**Figs 1j, k**) combined the features of true font systems with the alphabetic system to economise on keyboard space and to improve speed of typing: for example, the entire *i, ii, u,*and *uu* series of vowel consonants could be formed by typing the consonant symbol followed by the pure vowel symbol. Notably, these keyboard layouts are specific to word processing software packages. Nevertheless, each hybrid system, despite its benefit of economy, suffers the disadvantage of incompatibility with true font systems, as well as several other hybrid systems.

The Tamil-99 keyboard is, really, a successor to Anjal, with Tamil vowel and consonant symbols taking the place of Roman symbols and, occasionally, certain combinations of Roman symbols. While much thought has gone into electronic data storage, communication and handling, the design of the keyboard combines the weaknesses of the Anjal and the true font keyboards.

Some software packages such as Kamban 3.0 (see layout in **Fig. 1l**) appear to follow the logic of Tamil-99 (**Fig. 1m**), but with no benefit except the arrangement of the vowel and consonant in their alphabetic sequence, something that an experienced typist will not recognise as a significant merit.



A critical assessment

Keyboard layout design has continued to improve since the first typewriter. Although the current typewriter keyboard is lacking in important respects, especially the loading of fingers according to capability and alternation of the use of hands, it is easy to master and to use. It would have been good if this layout served as the reference for all subsequent keyboards. However, changes in strategy for word processing militate against the use of that layout in software-specific keyboard design, although not in true-font word processing, which is usable with the more universal word-processing software packages.

The unfortunate situation in Tamil word processing is that a regular user could be compelled to be familiar with at least three systems for Tamil word processing, while facilities do not seem to exist for conversion from one to another and, even if developed, could be expensive to the user. Thus, it will become necessary to prepare afresh a large volume of documents held as electronic files.

It is worth noting here that the standard Sinhala typewriter keyboard layout, which is also used in word processing, is commendable with respect to the two main ergonomic requirements of loading of fingers according to capability and alternation of strokes between the two hands (Jayatillake and Sivasegaram, 1975). It is also relatively easy to master, despite the larger number of symbols than in Tamil.

Word processing strategies adopting features of Anjal have the advantage of economy as they eliminate the need to accommodate on the keyboard a large number of vowel consonants, especially of the *u, uu* series. However, extending the logic of Anjal to vowel consonants of the *e, ee, ai, o, oo,* and *au* series entails a new problem, in view of which many word processing software packages have opted for hybrid systems that retain the devices for the *e, ee,* and *ai* series (which are also used to generate the *o, oo,* and *au* series). This is in deference to the methodology followed in writing and as a consequence in typing, namely that the diacritical devices is added either after (pure consonant, *i, ii, u,* and *uu* series), before (*e, ee,* and *ai* series) or both before and after (*o, oo,* and *au* series) the consonant symbol.

Writing is based on geometric shapes and their sequences as recognised and retained in the mind, rather than a phonetically guided pattern, even in languages with an alphabetic script. (When writing, one does not go through a ritual of spelling words, but rather recalls words, or even phrases, as individual units comprising sequences of letters. This pattern applies to typewriting as well).

For example when one types the word "கொடுப்பு" (koduppu) what comes to mind is a sequence of symbols கொ, டு, ப், and பு, and the way in which each is formed, namely க followed by ெ and then ா; then ட modified in some way (in this case, by curtailing the letter to add a loop); then ப followed by a dot on top, and finally ப to which one adds a downward stroke. Each of the letters கொ, டு, ப் and பு would be dealt with equal ease, as a sequence or as a whole unit, as necessary. However, the writer or typist is not guided by the phonetic sequence (க் ெ ட உ ப் ப உ) to generate the written word.

While the phonetic sequences of the *aa, i, ii, u,* and *uu* series of vowel consonants agree with the graphic sequence in which the symbol and diacritical device occur, the sequences are in conflict in the case of the *e, ee, ai, o, oo,* and *au* series, besides the fact that the consonant symbol is the vowel consonant of the *a* series to which a device is added afterwards to convert it to the pure consonant.

This is a major flaw in systems that follow the logic of Anjal but with Tamil characters on the keyboard. Whereas Anjal conforms to the principle of stereotyping, by following the methodology for transliteration in the Roman script, adaptations such as in Tamil-99 and Kamban (Figs 1*l*,m) are in breach of that principle.

Tamil-99 has added to the problem by introducing linguistic features into the word processing software. Some of these, besides the larger issue pointed out in the preceding paragraphs, were pointed out by the writer at a preparatory seminar for the Tamil Inayam 2000 preparatory conference in 1999. The criticism was ignored at the time, as the Tamil 99 was, to most participants, *fait accompli,* while others had too little time to study the system in detail. Cheran et al. have since drawn attention to three major defects in their critique of Tamil-99 and proposal for change.

These and other defects are illustrated by the following examples:

To generate the pure consonant symbol, a dot is added after the consonant symbol (Rule 3) whereas the consonant symbol without modification is the vowel consonant of the *a* series (Rule 2). This is inconsistent with phonetic principles. To be consistent, the keyboard should carry the pure consonant form (f; etc.) rather than (f; etc.)

Typing consonant symbols in succession will result in the first being rendered a pure consonant (Rule 4): கக becomes க்க , சச becomes ச்ச, etc. Thus, for example, typing "அததன்" would yield "அத்தன்" and "பரரின்" would yield "பர்ரின்".

Rule 5 overrides Rule 4 so that typing the same consonant symbol in succession more than twice will not lead to a sequence of pure consonants so that ககக becomes க்க, and not க்க்;. A fourth symbol in succession would, however, yield க்கக். The result is that, for example, the sequence keyed in as "அடடடடஏ" could yield "அட்டட்டடே" instead of an intended "அடடடடே" or "அட்டடடடே" or "அட்டட்டடே" or "அடட்டடே" or "அடட்டட்டே" or "அடடட்டடே", placing a dot where it does not belong, and forcing one of seven equal possibilities on the user. Here, the cure to one problem has paved the way to a bigger problem.

Rule 7 legislates that, in a sequence comprising a nasal consonant symbol and the corresponding voiced consonant, the nasal will be interpreted as a pure consonant. For example, ஙக will become ங்க and ஞச will become ஞ்ச. The result is that, for example, the sequence "கஆணடஆ" will be rendered as "காண்டா" instead of the intended "காணடா" and the sequence "சரஇகமப" will be rendered "சரிகம்ப" rather than the intended "சரிகமப".

The corrective actions prescribed in the software are inadequate since they would require the user to be constantly alert to the above rules. The problems could, however, be solved by having an override option to suppress their implementation or, more sensibly, by removing all linguistic considerations from the software.

Another problem concerning vowel sequences within a word will remain, however. For instance, "மஅரத்து", keyed in as "மஅரத தஉ" which will be rendered as "மரத்து". This could be cured by making the consonant symbol represent the pure consonant rather than the vowel consonant of the *a* series.

The main problem of the Tamil-99 keyboard, nevertheless, concerns its breach of ergonomic principles concerning operational sequences in the formation of the letters. A positive feature should, however, be noted, namely that the layout enables alternation of strokes between the two hands, although an imbalance in the load on the fingers is evident, with much overloading of the right hand index finger.

The solution

The problem of the Tamil keyboard has no easy solution, especially when the cost of remedial measures is a consideration. There is, nevertheless, a pressing need to work towards a unified keyboard for typewriting, true font, phonemic and hybrid word processing, especially in view of problems of training and/or retraining.

One solution is to use the typewriter keyboard as standard. It poses difficulties, but not insurmountable, for the designer of the electronic data systems. Another is to examine how typing and true font word processing could be accommodated in the Tamil-99 keyboard. A possibility is proposed in Figure 2, where the Tamil-99 keyboard layout is adapted for the typewriter (and on a similar basis for true font word processing). The Tamil letters that will be moved to other locations are shown in grey, next to those to be placed on the keyboard.

Better compromises are possible, but the present purpose is to demonstrate that a fair compromise is possible, that could be accommodated without much difficulty in the electronic

data storage, transmission and retrieval systems for Tamil-99. Ideally, the Tamil-99 keyboard should be critically reviewed and redesigned to retain its strong features, especially the allocation of all vowels to the left hand and the frequent consonants to the right, and in a way that the keyboard could be as close as possible to a final version of true font and typewriter keyboards.

Arguments relating to the cost of the changeover are weak, as most of the changes could be accommodated by rewriting software. The benefits to the user by way of enhanced performance will justify the investment.

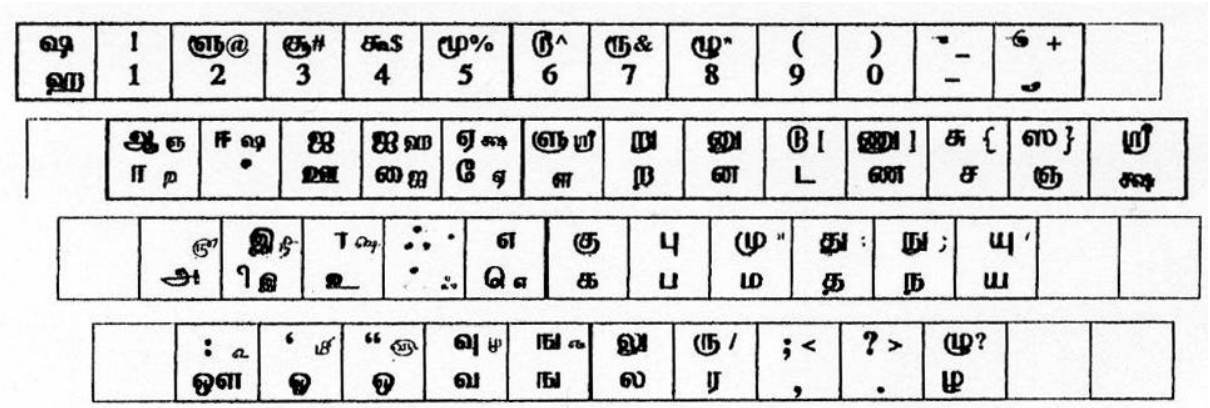


Figure 2: Adaptation of Tamil-99 for Typewriter Keyboard

References

Barnes, RM (1964) Motion and Time study: Design and Measurement of Work, John Wiley, 1964.

Cheran SK, Vaseeharan, T & Cheran, E,

http://tamillinux.sourceforge.net/vasee/cfreq/

Jayatillake, CLV & Sivasegaram, S (1975) Sinhala Typewriting Sysytems: Towards Greater Economy and Higher Efficiency, Proceedings, Section C, Annual Sessions of the Sri Lanka Association for the Advancement of Science, also published in slightly abridged form in Sinhala in Economic Review, Sinhala edition in 1976.

Kovinthapillai, V (1971) Tamil Typewriting Systems: a Critique of Existing Keyboards and a Proposal, Department of Tamil, University of Peradeniya.

Siromoney, G (1963) Efficient Methods of Telegraphy and Teleprinting in Tamil, *Tamil Culture*, Vol.1, no. 4.

Sivasegaram, S (1976a) Application of Work Study to the Design of the Tamil Typewriter Keyboard, *Management Studies*, Vol. 1, No. 2, pp. 9-14.

Sivasegaram, S (1976b) தமிழ்த் தட்டச்சுப்பொறியின் சாவிப்பலகை /மேலுங் கையடக்கமானதும் திறமையானதுமான ஒரு விதானம், பொருளியல் நோக்கு, 1979, pg. 25-30.

Tamil Inayam 2000. Preparatory Conference, August 12th 1999, Colombo http://www.thamizh.com