

# STANDARD KEYBOARD OVERLAY FOR THE USE OF KANNADA IN COMPUTERS

## Introduction

This document describes the standard keyboard overlay for using Kannada on computers. The layout has been notified as the standard by the Government of Karnataka and was arrived at by Kannada Ganaka Parishat.

For sometime now many attempts have been made to enable use of Indian languages in computers. Kannada is no exception. Though it is well understood that computers are not language specific, the use of Indian languages, nevertheless, poses problems since the various standards in use in computers today are dictated by the fact that the main language used in the computer industry is English. Naturally, there are problems when one embarks on enabling use of Indian languages on computers. In order to promote the use of Kannada on computers in general and address the problems associated with such use in particular, Kannada Ganaka Parishat was set up in 1997. Kannada Ganaka Parishat is a voluntary organization formed by professionals and experts in the field of computer software who are also knowledgeable in Kannada language and literature.

There are several Kannada software, which were developed and pressed into use by the Government and other users well before the Kannada Ganaka Parishat came into existence. The Kannada script software developers mainly worked in isolation, which led to non-standard implementation of Kannada on computers. The following are the main issues involved in such non-standard procedures adopted by the Kannada script developers.

1. Non-standard internal representation of Kannada:  
Each of the Kannada script software uses its own method of mapping of Kannada characters on to the ASCII codes in computers. This has resulted in non-portability of data among different font software.
2. Non-standard keyboard overlay:  
Each of the developer used his own mapping of Kannada characters on the otherwise standard QWERTY English keyboard for data entry. This has also created problems to the users.

Kannada Ganaka Parishat, since its inception, started addressing the above issues and organised many discussion meetings, seminars and conferences to find suitable solutions to enable the use of Kannada in computers effectively with ease. Two sub-committees were set up to achieve the desired goals. The present document contains information and the logic that has gone behind the evolution of a standard for mapping of Kannada characters on the existing keyboard.

The committee set up for the purpose of keyboard standards for Kannada embarked on an in-depth study of the same. This has resulted in a standard keyboard overlay for Kannada, which was arrived at after a detailed study, involving a wide range of users, writers and software developers.

The Government of Karnataka has not only accepted the keyboard layout suggested by Kannada Ganaka Parishat as the standard but also directed all Government buyers to demand that they supply this standard keyboard mapping with their software.

## **Computer, Keyboard and English:**

Though it is true that computers understand only the binary language, it is necessary, for the sake of humans, to provide an easy interface to the human language, such as English, Kannada etc. Since the primary users of computers interact with it in English, all computers are supplied with a keyboard with English letters and symbols painted or embossed on its keys. One will also notice that the keyboard uses a standard layout for English alphabets. For historical reasons, the computers make use of the same layout as the typewriter layout.

## **Kannada typing:**

Kannada typewriters came into being much before the computers came to be used for Kannada document preparation. Since typewriters use only fixed fonts, the formation of Kannada letters posed a challenge for typewriter designers. They used the idea of forming a particular letter by combining different pieces (called glyphs) together. For this reason, the keyboard for typewriters provides a keyboard layout, which has nothing common with the English layout. However, all Kannada typewriters provide the same layout even though it is not clear if the layout was ever documented as a standard. Also, there was no question of mixing languages in a typewriter media.

While it is perfectly fine for adopting the typewriter style of keyboard for computers as well, it does not make much of a sense to use a typewriter like layout, which uses as many as 46 keys for using Kannada in computers. The English and most European languages are much simpler in their written form adopting a very few alphabet set which are simply aligned on a horizontal line in a specific sequence to form the words.

## **Building blocks for Script display:**

In Indian languages and particularly in South Indian languages such as Kannada and Telugu, the letter shapes are quite large in numbers and as such the typewriter approach is to form the letters by using a set of building blocks. But to adopt such a layout for computers poses a different problem. It restricts the use of computers in the sense that the demand for many more keys other than the keys meant for English alphabets will come up. Computers are not merely word processing systems. They are much more than that and setting out to adopt typewriter like keyboard for computers reflect our mental restrictions. The software can take care of building a systematic structure of the letters of a language, through the elements of a set of building blocks.

Hence a different approach is necessary indeed for the keyboard overlay for Kannada. The users should be allowed to mix languages and at the same time use all the symbols appearing on the keys other than the ones meant for English characters. Hence it makes perfect sense to go about providing a unique meaning to each one of the keys so as to enable a user to enter data in Kannada, with the use of minimum number of keys. It is not necessary, any more to think of forming letters by the user, using typewriter like building blocks.

**The software can take care of the letter building process. The user on his part has to think in terms of phonetic formation of the letters of his language. This concept is easily acceptable, as Indian languages are phonetically well structured and the scripts are also evolved to suit this structure. The script representation of Indian languages is unambiguous and follows almost all the properties of the language.**

## 26 keys for Kannada also:

With this as the background, it is now easy to show that the available keyboard is more than sufficient for use with Kannada. The keyboard provides 26 keys that represent all the English characters. Besides, additional 26 keys are available using the shift key, which are used for uppercase English characters. While this key structure plays no role for English, as far as the phonetics of the English language is concerned, it comes in very handy for using them to represent alphabets in Kannada.

Summarizing these thoughts, the following points can be made.

1. Make use of the existing standard keyboard for Kannada as well
2. Use only the keys representing the English alphabets set (a-z, A\_Z) for representing the basic character set in Kannada also
3. Do not make use of any other key representing various symbols such as \$,%,# etc., to represent Kannada characters.
4. The input speed and ease of use are two important factors to be borne in mind while specifying the layout.
5. The layout should be easy to learn even by a novice.

Since the input through the keyboard is now a software-related issue, it is possible for any software developer to associate any key to any character. Obviously this would lead to tremendous confusion among the users. Since the English keyboard layout is standard, users would have no difficulty to input data even blindly. The same practice should hold for Kannada as well. Hence there is the need for a standard keyboard layout.

The general philosophy followed in determining a standard layout for Kannada is to identify a unique key for each one of the Kannada alphabets. Of course, unlike in languages that use Roman script, this at first sight seems to be insufficient, but actually it is not so. This issue is discussed later.

The figure attached shows the keyboard layout. It depicts alongside the English alphabets, which simply indicate the location of the keys on the keyboard.

Let us now consider Kannada basic alphabets and symbols that are in use in modern Kannada writing. There are, in all 51 as shown below.

The alphabets consist of

1. 13 vowels (ಅ, ಆ, ಇ, ಈ, ಉ, ಊ, ಋ, ಎ, ಏ, ಐ, ಒ, ಓ, ಔ)
2. 25 varg consonants i.e., 5 groups of 5 consonants each  
(ಕ, ಖ, ಗ, ಘ, ಜ, ಚ, ಛ, ಙ, ರು, ಞ, ಟ, ಠ, ಡ, ಢ, ಣ, ತ, ಥ, ದ, ಧ, ನ, ಪ, ಫ, ಬ, ಭ, ಮ)
3. 9 non-varg consonants (ಌ, ಍, ಋ, ೠ, ೡ, ೢ, ೣ, ೤, ೥)
4. 3 post-fix symbols (anuswara ಂ, visarga ಃ & arkarvattu ಣ) and the vowel omission sign halant (ಱ), which is used to represent base consonants (traditionally called a half letter) i.e., a consonant that does not have the implicit vowel ಅ in it. This adds up to a total of 4 symbols.

First it is necessary to allocate a single key for each one of the vowels. It is customary to write consonants in their implicit form, i.e., together with the vowel A. It appears, therefore, logical as well to associate a unique key for each one of the 25 consonants, 9 non-varg consonants and the 4 special symbols.

Representation of consonants with the implicit ಅ (the first vowel) is advantageous since the basic consonants are seldom used. The software can account for the fact that when any vowel other than the first one (ಅ) is to be attached to the consonant, it can replace the implicit vowel ಅ. This saves a large amount of keystrokes of ಅ, as well. It is suggested that, if in case there is a need to represent a base consonant, then the corresponding consonant can be combined with the halant symbol. By imagining the half-letter symbol to be another vowel-like alphabet, this idea of using the halant or half-letter symbol to represent the base consonant is consistent with the logic that vowels other than the first vowel (ಅ) replaces the vowel (ಅ). Indeed, it is easy to incorporate such display formations in a software environment.

The rest of the composite characters that we get in modern Kannada writing can be obtained by suitable combination of these 51 characters. This is discussed next.

### **Formation of composite and conjunct letters**

#### **Composite letters:**

When the consonants are combined with vowels, we get the composite letters. Every combination of a consonant and each one of the 13 vowels gives rise to a unique unit of pronunciation called syllable and this also has a unique representation. Essentially, changing the basic shape of a consonant by adding a vowel sign (maatras) defines a composite letter.

Following the fact that the vowel signs have a unique form, the typewriter layout was arrived at. In particular, the typewriter associates a key to each of these maatras. We argue here that associating a key to each of the 12 maatras is unnecessary based on the following observations:

1. In Indian languages, if a vowel is preceded by a consonant, then the corresponding vowel sign attaches itself to the consonant to form a composite character. Generally, no vowel appears in its original or stand-alone form within a word.

This means, a vowel will retain its stand alone form only if it is preceded by a white space (common case) or by another vowel (rare case) (e.g. ಓಹೋ, ವಿಷುಅಲ್ ಬೇಸಿಕ್, ಉಸುಕು, ಮುಸುಕು).

Since these are the basic rules of the language, they can be easily adhered to in a software environment, thereby eliminating the need for 12 additional keys for vowel signs (maatras or swarna chinhas), for forming the composite letters.

#### **Conjuncts:**

Numerous conjuncts can be formed using consonants, which essentially are clusters of up to 4 consonants without the intervening implicit vowel (ಅ).

Thus, when two or more consonants are combined together and with a vowel, we get conjunct letters. If we go along with previous logic of obtaining the composite letters, then it

would be impossible to get conjunct letters. But, since we have already said that the half-letter key (link key ಠ) is used to omit the vowel (ಅ), it will serve the purpose here just fine to combine consonants in their basic form. Thus, to obtain ಶ್ರೀ it suffices to input ಶ followed by ಠ followed by ರ and then the vowel ಃ. Of course, the link key can be repeatedly used to combine more than two consonants to obtain the conjuncts.

Next question we ask is whether it is possible to use only the 26 keys which are painted with English characters on a Keyboard to represent the 51 alphabets in Kannada. The answer is evidently yes. Since each of the keys has a dual function of representing the small case (normal key) and Capital case (Shift key) letters in English. Thus these 26 keys represent 52 different symbols in English and hence are more than adequate for a complete representation of 13 swaras, 34 consonants and the four special symbols in Kannada.

However, while forming composite letters and the conjuncts, we encounter the following two exceptional cases, which can be resolved as under:

**Case 1:** Consider the case when a consonant is required to appear in its base form within a word (as in **ಆಗಿ**). How does one achieve this?

The logic used so far suggests that we must press a consonant key followed by the link key to reduce it to its base form. But, following it with any other consonant would cause it to form a conjunct letter instead. This can be overcome by pressing the link key twice. This double clicking is an indication to the software, which then takes care of our requirement. Thus, we press the 'ರ', link key twice and then the 'ಗ' to get ರಗ.

**Case 2:** Though it is not a general case to have a vowel appear in its stand-alone form preceded by a consonant, a software environment may be incomplete without allowing such a capability. Such situations do arise as exceptions or special cases. The problem is to avoid attachment of the vowel in its vowel sign (maatras) form with the preceding consonant. This problem is overcome by pressing the vowel ಅ after the consonant and before another vowel. The extra vowel ಅ between a consonant and a vowel is an indication for the software to avoid the formation of a compound letter. Here ಅ acts as the abstraction character.

**Now, we discuss the assignment or mapping of Kannada characters over the keys on which the 26 English keys are marked, of the computer keyboard.**

Though there is no need to bind English keys to Kannada keys in any particular order, it is often desirable and necessary to bind them based on the commonality of the sounds they represent, as closely as possible. This would enable keyboard usage more intuitive and enhance the input speed. It would also make the learning easier. This is more practical as computers are being used by one and all with many of them may not be having any formal training in Kannada typing.

This is done in 2 stages. The following table shows the associations based on this idea.

**Stage1.**

Vowels: ಅ, ಆ, ಉ, ಊ, ಋ, ೠ  
a, A, u, U, o, O

Consonants:

ಕ, ಖ, ಗ, ಘ, ಚ, ಛ, ಜ, ಝ  
k, K, g, G, c, C, j, J

ನ, ಣ, ಪ, ಫ, ಬ, ಭ, ಮ  
n, N, p, P, b, B, m

ಯ, ರ, ಲ, ವ, ಶ, ಸ, ಹ, ಳ  
y, r, l, v, S, s, h, L

Stage 2.

1. Of the two post-fix characters, we assign **M** for '೦' and **H** for 'ಃ', which in a way close to the sounds represented by these symbols (note that Upper case letters are used).
2. Similarly, the vowels ಋ, ಌ and ೠ are assigned to the keys **R, Y** and **V** respectively.
3. The vowels ಎ and ಏ are represented using the keys **e** and **E** respectively.
4. The vowels ಇ and ಈ are denoted by the keys **i** and **I** respectively.
5. Since, the consonants ತ, ಥ, ದ, ಧ find more use than the consonants ಟ, ಠ, ಡ, ಢ and also keeping in mind the accessibility of the keys in the layout, the consonants ತ, ಥ, ದ, ಧ, ಟ, ಠ, ಡ, ಢ are denoted by the keys **t, T, d, D, q, Q, w, W** respectively.
6. The non-varg ಷ is represented by the key **x** (lower case).
7. The consonants ಞ and ಙ are very infrequent and are denoted by **z** and **Z** keys respectively.
8. The post-fix character ಼ and link character ಽ are represented by the keys **F** and **f** respectively. Notice that we have used **f** for link key to form a conjunct of consonants, since it finds more frequent use than the character ಼

With this, we have represented all the vowels, consonants and generally used symbols in Kannada with the 51 options available with the keyboard.

### Special cases:

There is a need to modify some of the following characters:

1. Vowel ಋ and its maatra ಠ (which are rarely used) as modified forms of ಋ and ಠ
2. The nukta characters ಙ and ಞ as modified forms of ಜ and ಘ.
3. The old Kannada consonants ಞ and ಙ (which are rarely used and sound very close to the consonants ರ and ಳ) as the modified forms of ರ and ಳ.

Since the 51 key positions have already been used while the keyboard provides 52 possibilities, we enjoy an additional freedom of using the key **X** (upper case) that has not

been assigned any character. This can be very effectively used to resolve the above special cases and any other that may come up in future.

**Case 1:** The vowel ಋ is rendered by pressing the keys **R** (which gives ಋ) and **X** consecutively. If this vowel is preceded by a consonant then the maatra ೠ is rendered.

**Case 2:** The nukta character ಞ or ಞ̣ is rendered by pressing the key **X** after we press the key **j** (for ಞ) or **P** (for ಞ̣) respectively.

**Case 3:** As above, the key **X** can be used after **r** (ಠ) and **L** (ಱ) to get ಠ and ಠ̣ respectively. Their conjunct forms are rendered as is done in the case of ಠ and ಱ.

The method of obtaining almost all the characters in Kannada from the basic 51 key positions has been used earlier by the Prof. Kasturi Rangachar and Prof. K P Rao for Kannada word processors. The present keyboard layout is fairly close to the layout evolved by Prof. K P Rao who still used some additional keys to represent some compound and other characters. They have been dispensed with and some modifications have been made while arriving at this present scheme.

The present keyboard overlay has been arrived at after engaging in the study of several layouts in practice. Many discussions have been carried out with software developers, Kannada typewriter users and other concerned persons. The unique feature of this layout is the use of **q, Q, w** and **W** for ಳ, ಠ, ಡ, ಢ and **z** and **Z** for ಞ and ಞ̣. Such a representation does not in any way place burden on a veteran typewriter user or, for that matter, a novice who is accustomed to the English Keyboard. It is true that the key association from English to Kannada does use closeness in the sound as a criterion. This has been done only to help the users to switch between languages as easily as possible. It is wrong to think that the Kannada input is achieved by using English as a via media. Also by using keys to represent consonants in their natural form (together with the sound u as in cut, but etc.), input becomes faster since the keystrokes are minimized.

The depiction of Vedic hymns, music scores, grammar texts and dictionaries use many special symbols that guide pronunciation and timing. Again the key X comes in as a handy tool for such renderings.

Kannada Ganaka Parishat acknowledges the support extended by the Government departments, users, Kannada Script software developers, writers and all others.

Written by:  
C. V. SRINATHA SASTRY  
General Secretary, Kannada Ganaka Parishat  
Gokhale Institute of Public Affairs  
Narasimharaja Colony  
BANGALORE 560 019, INDIA  
Ph:080-6615972