"AENASA": A NOVEL APPOARCH TOWARDS WORLD ICT PROSPERITY WITH MULTILINGUAL KEYBOARD LOCALIZED FOR GLOBAL LANGUAGES

Vishal Kaushik (Assistant Professor (Sel. Grade), CIT, UPES, Dehradun, India (vkaushik@ddn.upes.ac.in),
Rakesh Sharma, Professor, University of Maryland, USA
Neha Jain, M. Tech. AI, UPES, Dehradun, India (neha.capricorn19@gmail.com)

ABSTRACT

A keyboard design, once adopted, tends to have a long lasting and worldwide impact on daily user experience. There is a substantial body of research on touch screen keyboard optimization. Most of it has focused on English only. Applying rigorous mathematical optimization methods and addressing diacritic character design issues, this paper expands this body of work to other languages. More importantly and counter to the intuition that optimization by nature is necessarily specific to each language, this paper demonstrates that it is possible to find common layouts that are highly optimized across multiple languages in the world. The AENASA is named coded with the first letters of all the continents i.e. Asia, Europe, North America, Africa, South America and Antarctica. However, this present paper we have initially focused on Hindi or the languages based on the Devnagari script for the languages common in the Indian subcontinent.

Keywords: *Keyboard Layouts, World ICT, Global Keyboard Design, Localized Keyboards, Multilingual Layouts, Human Computer Interaction*

INTRODUCTION:

The web is playing a crucial role in bringing information and data to the populations around the world. Though a major amount of the content on the web is in a few languages, the web has initiated its content in multilingual form. Particular requirements come with linguistic and cultural diversity. A virtual keyboard is a software part that allows a user to enter data, characters. A virtual keyboard can work with multiple input devices, which may include a touch screen, an actual computer keyboard and a computer mouse. Virtual keyboards are usually used as an on-screen input method in devices with no physical keyboard, where there is no room for one, such as a pocket computer, personal digital assistant (PDA), tablet computer or touch screen enabled mobile phone. By tapping a virtual keyboard or finger tracing text is inputted. Virtual keyboards are also used as main item of imitating software for systems that have less buttons than a computer keyboard would have. Virtual touch screen keyboards use the same idea as projected keyboards but without the laser. Beneath the surface of the touch screen, your device detects changes in the voltage when your fingers touch the screen. Depending on the change in voltage and the location of the change, the device then determines which button you touched or which key you tapped. *(Huda Sarfaraz et.al, 2011; Grey et.al, 2015)*

The compatibility among different languages gives hope to the goal of simultaneously optimizing for multiple languages. The basic methodology used in the present study is to calculate the mean time of tapping a character and taking its average and minimizing this average by any algorithm thus enabling an optimized form of keyboard layout. *(Sornlertlamvanich, 2015; Grey et.al, 2015)*

There are several varieties of languages in the world. Primarily the languages were for the oral communication based on the graphical symbols/ signals used and sequenced in a phonological system to specify words or morphemes. Human's language in the world always has productivity, recursivity, and displacement properties in a complex structure to exhibit wider range of expressions. The biggest issue with the languages and is the great varieties in the language systems around the world. The differences are such severe that it is very difficult to understand the different languages with a single rule. *(Marc D. Hauser et.al, 2002)*

The written language can be put as the most revolutionary innovation in communication techniques. Writing opened the window for communication between individuals without meeting them real time and face to face. The document recording of information, ideas and knowledge of history was only possible with the help of written language

technology. All the modern day jobs of communication like email, chat, blog or micro blog or even the website etc. and the official reporting, document recording etc. are possible just because of one or the other written language.

There many different types of writing techniques in the world based on one or the other different linguistic system logic. Each language has its own script, character set and style of writing. During industrial and computerisation era different types of keyboards were designed for the different languages in the different parts of the world. Most of it was in Europe and North America. Whereas some of the languages could not see the substantial development in terms of the automated writing instruments and tools such as the keyboard etc. Therefore certain countries/ societies still mark a very low penetration on modern ICT based technology. And because each language is based on a different alphabet set, script style and semantic system, a common keyboard layout for all the world languages is the most difficult task even today. Based on a small calculation we have tabulated the approx. percentage of the population and the language system. *(John C et.al., 2006)*

Table 1: Language Script and World Population percentage

S. No.	Linguistic System	Languages Associated	Population (Billion)	Percentage
1	Latin alphabet	English, Spanish, French, Portuguese, Malay-Indonesian, German, Turkish, Vietnamese, Italian, Polish, Dutch, Swedish, Latin, others	2.6 billion	36%
2	Chinese script	Chinese, Japanese, Korean Vietnamese, Zhuang	1.3 billion	18%
3	Devanagari script (India)	Hindi, Marathi, Konkani, Nepali, Sanskrit, Bangla, Oriya or Gurmukhi (Structural layout and grammar design) and others	1 billion	14%
4	Arabic alphabet	Arabic, Persian, Urdu, Punjabi, Pashto, Sindhi, Balochi, Malay, Uyghur, Kazakh (Chinese)	1 billion	14%
5	Cyrillic alphabet	Bulgarian, Russian, Serbian, Ukrainian, Macedonian	0.3 billion	4%
6	Dravidian script (South India)	Tamil, Kannada, Malayalam, Telgu	0.25 billion	3.5%

KEYBOARD DESIGN OBJECTIVES:

Objective function that is well stated can be used for quantitative optimization. Text entry is a complicated task. That involves cognitive, perceptual and motor components. Typing performance is limited mainly by hand movement on the keyboard and the layout of keys. Therefore keyboard optimization involves reducing movement time. *(Sornlertlamvanich, 2001)*

Both virtual projected keyboards and virtual touch screen keyboards can provide a great advantage to those who need a full-sized keyboard on the go. Not only are projected and touch screen keyboards lighter, the take up less, or no, additional space. However, neither type of virtual keyboard provides the same experience of typing and feeling the keys physically respond, which can be disorienting for some people. With virtual projected keyboards, you won't be able to type on anything other than a flat surface. You also can't rest your hands on either type of keyboard while you think of what to type; by doing so, you could type a string of nonsensical letters. *(Xiaojun Bi, 2012)*

POPULAR KEYOARD LAYOUTS COMPARISON:

QWERTY: This keyboard layout was designed by Sholes basically for the manual typing machine with improved performance of the machine, not the typist. Originally keys were laid out in alphabetical order which caused frequent jams of moving parts. It is patented US standard 101keys keyboard layout

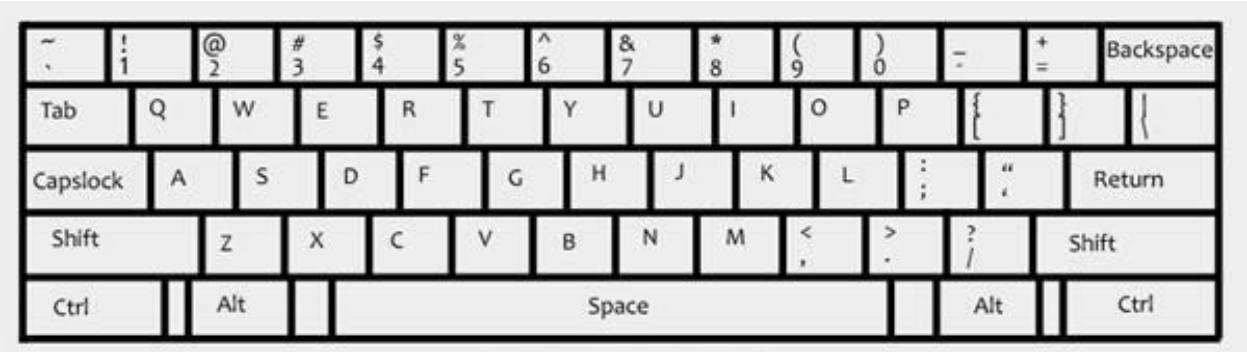

Figure 1: QWERTY Keyboard Layout (Image source: www.computerhope.com)

Dvorak: This keyboard layout was designed to enhance the efficiency of the typist. Thus by enhancing the typing speed it was more efficient and more comfortable. This layout has minimized movement.

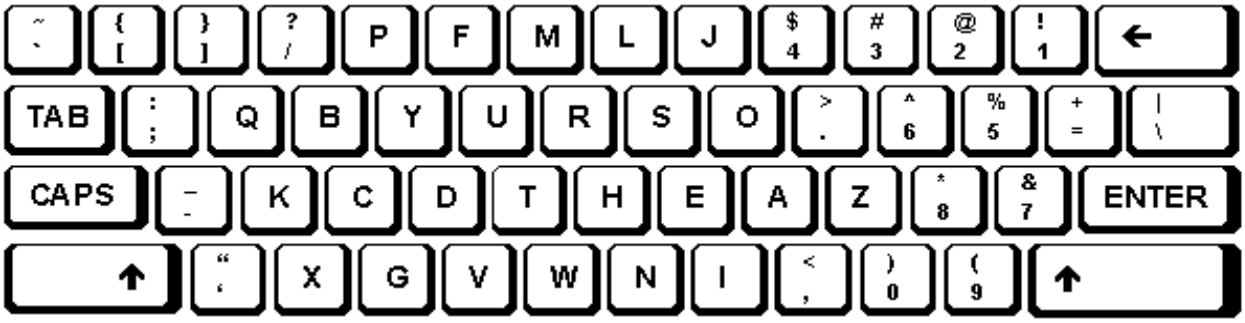

Figure 2: DVORK Keyboard Layout (Image source: www.computerhope.com)

Table 2: Comparison between the popular keyboard layouts

S.NO.	QWERTY	DVORAK
1.	Was designed for machine.	Was designed for typist
2.	Keys were laid out in alphabetical order	Keys were not in alphabetical order.
3.	Frequent jamming of letters enables fast speed.	It minimizes the movement to make typing easy and possible.

OPTIMIZING KEYBOARDS FOR MULTIPLE LANGUAGES:

Designing single keyboard hardware for multiple languages seems to be the task next to impossible. Although the electronic circuits can be designed but the outer cabinet casing displaying the key labels of different language systems will prove the actual impossibility. Also the manufacturing of such key stroke hardware cabinet infeasible as far as the business is concerned. Therefore we shall shift our focus from the usual key stroke keyboard to the pointer based keyboard. The pointer based keyboard can easily accommodate any number of language systems easily. Though this type of keyboard will be the virtual one but it will be as efficient as the normal keyboard. In the present paper proposes an optimized design for the common keyboard layout for the languages based on the Devanagari linguistic system.

As per our approach now with the point and click design keyboard, the efficiency will be ensured based on the distance of movement between the pointer and the place of click. The time calculated between a sample number of continuous point and clicks will be the total time spent on the keyboard. A comparison between the time calculated for a sample input through the proposed keyboard layout and the same sample input through the normal available keyboard layout would be the efficiency factor of the proposed layput.

For optimizing the keyboard layout we have used the Fitts Law. Fitts Law is basically an empirical model explaining speed accuracy trade-off characteristics of human muscle movement and it's similar to Shannon channel capacity theorem and it also predicts that the time required is a function of a ratio between the distance of the target and width of the target. Fitts Law is used to simulate the act of pointing, either by physically touching an object with a hand or finger, or virtually, by pointing to an object on a computer monitor using a pointing device. *(Katare, 2006; Sornlertlamvanich, 2015)*

$$MT = a + b \cdot ID = a + b \cdot \log_2\left(\frac{2D}{W}\right)$$

Equation 1: Fitts Law Equation

Here:

MT is the average time to complete the movement.
a and b are model parameters.
ID is the index of difficulty.
D is the distance from the starting point to the center of the target.
W is the width of the target measured along the axis of motion.
W can also be thought of as the allowed error tolerance in the final position, since the final point of the motion must fall within ±W/2 of the target's center. *(Fitts, Paul M., 1954)*

CONCLUSION:

In the present work we have proposed a common keyboard layout for the Devanagari linguistic system. We have optimized the keyboard layout using the Fitts law. We have tried to create a more useful keyboard layout for the wide range of users of the languages based on the Devanagari script such as Hindi, Marathi, Konkani, Nepali, Sanskrit, Sindhi, Maithili, Bodo, Bangla, Oriya or Gurmukhi and various other languages common in East Asia and many other places of the world. *(Cardona and Jain 2003)*

Devanagari script is based on the consonants which represent the oral sound produced through the combination of tongue and mouth. There are total thirty six 36 consonants (including Sanskrit symbols). Also the vocal sounds produced from the combination of mouth and vocal cord is represented as vowels. There are sixteen vowels (including Sanskrit symbols). Unlike Latin language system there is no capital or small letter. The character set of the consonant and vowel letters of the language is represented below –

क	ख	ग	घ	ङ	च	छ	ज	झ	ञ	ट	ठ
ड	ढ	ण	त	थ	द	ध	न	प	फ	ब	भ
म	य	र	ल	व	श	ष	स	ह	ळ	क्ष	ज्ञ

List 1: Consonant letters of Devanagari script

अ	आ	इ	ई	उ	ऊ	ए	ऐ	ओ	औ	अं	अः
ऋ	ॠ	ऌ	ॡ								

List 2: Vowel letters of Devanagari script

The each letter is the representation of a specific sound produced by a particular portion of mouth, lips, tongue, throat and nasals. The table below represents the classification of the letters based on vocal sound represented by each consonant letter –

Table 1: Vocal Sound Classification of Devanagari Consonant Letters

S.No.	Class of Vocal Sound	Vowel Letter Representation			
		Plosive (anunāsika)	Nasal (antastha)	Approximant antastha)	Fricative (ūṣma)
1.	Gutturals (Kanthaya)	क ख ग घ	ङ		ह
2.	Palatal (talavya)	च छ ज झ	ञ	य	श
3.	Labial (osthya)	प फ ब भ	म	व	श्र
4.	Retroflex (Murdhanya)	ट ठ ड ढ	ण	र	ष
5.	Dental (Dnatya)	त थ द ध	न	ल	स

The vowel and the short symbolic representation of vowel in the form of Matra are represented below based on the classification of vocal sound produced –

Table 2: Vocal Sound Classification of Devanagari Vowels Letters and Short Symbols (Matras)

S.No.	Class of Vocal Sound	Vowel Letter Representation	Short Symbol as Matra
1.	Gutturals (Kanthaya)	अ आ ॲ ऑ	ा ॅ ॉ
2.	Palatal (talavya)	इ ई	ि ी
3.	Labial (osthya)	उ ऊ	ु ू
4.	Palato Guttural (Kanthatalavya)	ए ऐ	े ै
5.	Labio Guttural (Kanthosthya)	ओ औ अं अः अँ ॐ	ो ौ ं ः ँ ॉ
6.	Retroflex (Murdhanya)	ऋ ॠ	ृ ॄ
7.	Dental (Dnatya)	ऌ ॡ	ॢ ॣ

The combination of vowels and consonants make the words. To represent the combination of the consonant letter and the vowel letter the words in Devanagari script are written with short symbolic representation of vowels associated with the consonants in the form of matra. Therefore the script is written in three segments vertically. The middle segment is the main segment where the main letter is written. The consonant vowel conjunction in the Matra's form is written either above or below the main letter in written in the middle segment.

The words also include the half letters to represent the sound of a letter not so prominent in the word. Along with some letters are used in combination with the other letters to represent the produce the mixed sound produced. These combining letters are combined various different forms all are represented below –

- र्व (Sound of r 'र ' comes before the main letter)
- ग्र (Sound of r 'र ' comes along with the main letter)

Devanagari scripts digits are also represented below -

० १ २ ३ ४ ५ ६ ७ ८ ९

List 3: Digits letters of Devanagari script

ALGORITHM:

```
# desibling the vowel and matra keys

String cons = " all the consonant";

String vol ="all the vowel"

String matra ="all matra"

String typech = Textbox1.text.lastchar;

If typetext= (Isfind(vols) ==1)

Then, desible the matra keys

//Is find check the whether the typed char is consonant

If typetext= (Isfind(matra) ==1)

Then, desible the vowel keys

If typetext= (Isfind(consonat) ==1)

Enable both matra and vowel keys.

End
```

CONCLUSION AND FUTURE SCOPE:

In this paper, we have focused on an optimized Hindi keyboard layout. We have tried to create a more useful keyboard layout for Hindi speaking, users. Using the Fitts law we have optimized the keyboard layout. In the future work, we will create a multilingual keyboard that will be able to incorporate different languages of India. Users will have a virtual keyboard that will be able to accept any Indian language input. Creating such a keyboard would help ease the communication gap in India and would allow more languages to be used in computer and mobile communication.

In the future work we will create a multilingual keyboard that will be able to incorporate different languages of India. Users will have a virtual keyboard that will be able to accept any Indian language input. Aiming for Indian users it is going to ease the communication gap between people and English is not going to be the only language in which you can communicate via computer or mobiles.

REFERENCES:

Bi, X., Smith, B., and Zhai, S. (2010). Quasi-qwerty soft keyboard optimization. ACM CHI. 283-286.

Bi, X., Smith, A. B., and Zhai, S. (2012). Multilingual Touchscreen Keyboard Design and Optimization. Human-Computer Interaction, 27(4), 352 - 382.

Katre, D. S. (2006). Position Paper On Cross-cultural Usability Issues of Bilingual (Hindi & English) Mobile Phones, In the Proceedings of Indo-Danish HCI Research Symposium, Hosted by Department of Design, Indian Institute of Technology, Guwahati, India.

Potipiti, T., Sornlertlamvanich, V., & Thanadkran, K., (2001). "Towards an Intelligent Multilingual Keyboard System", National Electronics and Computer Center, National Science and Technology Development

Agency, Ministry of Science and Technology Environment, Bangkok, Thailand, Proceedings of HLT2001, San Diego, 346-349.

Katre, D. S. (2006). Position Paper On Cross-cultural Usability Issues of Bilingual (Hindi & English) Mobile Phones, In the Proceedings of Indo-Danish HCI Research Symposium, Hosted by Department of Design, Indian Institute of Technology, Guwahati, India.

Sarfaraz, H., Dilawari, A., Hussain, S. (2011). "Assessing Urdu Language Support on the Multilingual Web", Center for Language Engineering, Al-Khawarizmi Institute of Computer Science, University of Engineering and Technology, Pakistan, Conference on Human Language Technology for Development, Alexandria, Egypt, 2-5.

MacKenzie, I.S. (1992). Fitts' law as a research and design tool in human-computer interaction. Human Computer Interaction, 7 (1), 91-139.

Sornlertlamvanich, V. (2001). "Bridging the Information Gap between Urban and Rural Lifestyle,Collect, Analyze, Visualize", Global Expert Meeting on Multilingualism in Cyberspace, Khanty-Mansiysk, Russia Federation, July 4-9.

Sornlertlamvanich, V., Pacharawongsakda, E., & Charoenporn, T. (2015). "Understanding Social Movement by Tracking the Keyword in Social Media", Proceedings of Multiple Approaches Lexicon (MAPLEX2015), Yamagata, Japan, February 9-10.

Paolillo, John C.; Das, Anupam (2006). "Evaluating language statistics: the Ethnologue and beyond" (PDF). UNESCO Institute of Statistics. pp. 3–5.

Hauser M. D., Chomsky N., Fitch W. T. (2002), "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?", Science: Vol. 298 no. 5598 pp. 1569-1579)

Fitts, Paul M., (1954) The information capacity of the human motor system in controlling the amplitude of movement, Journal of Experimental Psychology, Vol 47(6), Jun, 381-391

Cardona G. and Jain D.(2003). The Indo-Aryan Languages, Routledge, pages 68-69, ISBN 978-0415772945,

Image sources : www.computerhope.com

Image Sorces: www.webopedia.com

http://www.powertyping.com/dvorak/keyboard.html