Studies on Input Interface of Asian Characters Based on Common Syllabic Writing Systems

音節表記の共通性に基づいたアジア文字入力インタフェースに関する研究

October 2011

Graduate School of Global Information and Telecommunication Studies
Waseda University

Information Network System II

Ye Kyaw Thu

Keywords: Text entry, text input, Positional Mapping, Positional Gesture, Positional Prediction, Characters Per Minute (CPM), KSPC (Keystrokes Per Character), Myanmar, Bengali, Khmer, Nepali and Thai

ABSTRACT

This dissertation presents studies on input interface of Asian characters based on common syllabic writing systems. Keyboard mapping, gesture text input and predictive text input are considered for designing text input interfaces. Positional Mapping, Positional Gesture and Positional Prediction are the three concepts discussed in this dissertation.

Positional Mapping (PM) is a concept of keyboard or keypad mapping for mobile devices based on characters writing position of Asian syllabic languages. Its concept is dividing a mobile phone keypad into three levels, e.g. 1, 2 and 3 keys are upper, 4, 5 and 6 keys are normal or middle and 7, 8 and 9 keys are lower. In the normal level, 4 key is deemed as a front part (Left), 5 key as a central part and 6 key as a rear part (Right).

Positional Gesture (PG) is a concept of gesture text input for mobile devices based on characters writing position of Asian syllabic languages. Its concept is totally based on four simple gesture commands, which are “Left”, “Right”, “Up” and “Down”. “Left gesture command” is for left characters or symbols, “Right gesture command” is for right characters or symbols, “Up gesture command” is for upper characters or symbols and “Down gesture command” is for lower characters or symbols.

Positional Prediction (PP) is a concept of predicting possible combinations of vowels with a consonant or a syllable based on the positional information (i.e. Left, Right, Upper and Lower) of vowels, medials or consonant signs etc.

As part of the ongoing evaluation of PM, PG and PP, various text input prototypes were developed. An emphasis was placed on users’ typing speed, error rate and feedbacks in user study in order to recognize the feasibility and user-friendliness. PM and PG prototypes were developed for Myanmar (Burmese), Bengali and Khmer, and PP prototypes were developed for Myanmar, Khmer, Nepali and Thai respectively. These three approaches will contribute to find out common and user-friendly text input interfaces of Asian syllabic languages especially for mobile computing.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor, Prof. Yoshiyori URANO, for his kind support and understanding on my greedy working nature. It has been a real pleasure working with him since my master course that started in 2004.

My other professors, Prof. Wataru KAMEYAMA, Prof. Hidenori NAKAZATO and Prof. Yoshinori SAGISAKA, deserve special thanks for their valuable suggestions to improve this research.

A number of GITS students have worked together with me on Positional Mapping, Positional Gesture and Positional Prediction. In alphabetical order, they are: Huotely YIN, Jackchai Butsrikui, MD. Monzur Morshed, Ouk Phavy, Sai Aung Win Maung, Shrestha PRADYUMNA.

Also, I would like to thank GITS staff especially Yumiko KISHIMOTO and Ai MIYAGAWA for their kind and continuous support to complete my work as this dissertation.

My brother Ye Kyaw Thein has also assisted me in various ways such as typing contents of Myanmar dictionaries, helping user study implementation in Myanmar etc. My parents have ever supported me with their immense love.

Finally, to my wife Yoko, to whom this dissertation is dedicated: You have been my strength, my will and my informal Japanese and English teacher. Your delicious dishes have always provided me with moral support. Moreover, thank you for your taking care of my daughter Minori (Nyein Aye Thu). Without you, I could not accomplish this work. Thank you for your love.

To
Yoko NAKATSUKA

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1 Motivation

This dissertation is motivated by two primary concerns: the need for mobile phone text input in Myanmar language (Burmese) and common text input interfaces applicable for Asian syllabic languages based on character writing position. The primary goal of this research is to develop user-friendly and common text input interfaces extendable for multiple Asian syllabic languages.

1.1.1 The Need for Mobile Phone Text Input in Myanmar Language

Over the past few decades, wireless devices have revolutionized the communication industry, and text messaging services like Short Messaging Service (SMS) and Multimedia Messaging Service (MMS) have become necessary service in global wireless market today. A mobile phone is not only for voice communication but also for ubiquitous communication in developed countries. However, no text input method in Myanmar language was proposed for mobile phones in 2004 yet, and this situation led to my research.

In these years, information networking (e.g. twitter, launched in July 2006) [1][2] and social networking (e.g. facebook, launched in February, 2004) [3][4] services are getting more and more popular and daily accessible even via mobile browsers. User-friendly text input method in Myanmar language on mobile devices has never been more important than today for its native users.

1.1.2 The Common Text Input Interface for Asian Syllabic Languages

South and South-East Asian languages are syllabic languages and derived from Indic script or Brahmi around B.C. 3rd century. Due to this history, Nepali, Bhutan, Tibetan, Bengali, Myanmar, Lao, Thai, Khmer, etc. have common writing natures with Indian languages (e.g. Hindi, Marathi, Punjabi, Tamil). However, shapes of characters are different such as circular (e.g. Myanmar characters), distinctive horizontal line running along the top of characters (e.g. Devanagari characters) and Curlicues (Sinhala characters). Although shapes or glyphs of characters are different, structure of the consonant clusters or syllables are based on the same rule. Here, the common writing nature of Asian syllabic languages is explained with Myanmar, Khmer, Nepali and Bengali languages.

Myanmar is the official language in Myanmar that was formally known as Burma. It is a tonal and syllable-based language. Myanmar scripts are adopted from the Mon script (one of the main national races of Myanmar) that derived from Indian Brahmi script [5]. There are 34 basic consonants “က”, “ခ”, “ဂ”, “ဃ”, “င”, etc. (U+1000 to U+1021), 8 basic vowels “◌ါ”, “◌ာ”, “◌ိ”, “◌ီ”, “◌ု”, “◌ူ”, “ေ◌” and “◌ဲ” (U+102B to U+1032), 4 basic medial consonants or dependent consonant signs “◌ျ”, “ြ◌”, “◌ွ”, and “◌ှ” (U+103B to U+103E), dependent various signs “◌ံ”, “◌့”, “◌း”, “◌္” and “◌်” (U+1036 to U+103A) and independent various signs “၌”, “၍”, “၎”, “၏”, etc. (U+104C to U+104F), 2 punctuation marks “၊” and “။” (U+104A and U+104B) and 10 digits “၀”, “၁”, “၂”, “၃”, “၄”, “၅”, “၆”, “၇”, “၈” and “၉” (U+1040 to U+1049) [6]. Vowels can be divided into independent and dependent vowels. Independent vowels “ဣ”, “ဤ”, “ဥ”, “ဩ”, etc. (U+1023 to U+1027, U+1029 and U+102A) can stand alone, and dependent vowels (U+102B to U+1032) are written with a consonant. “U+” followed by hexadecimal numbers are Unicode code point of Myanmar characters [6].

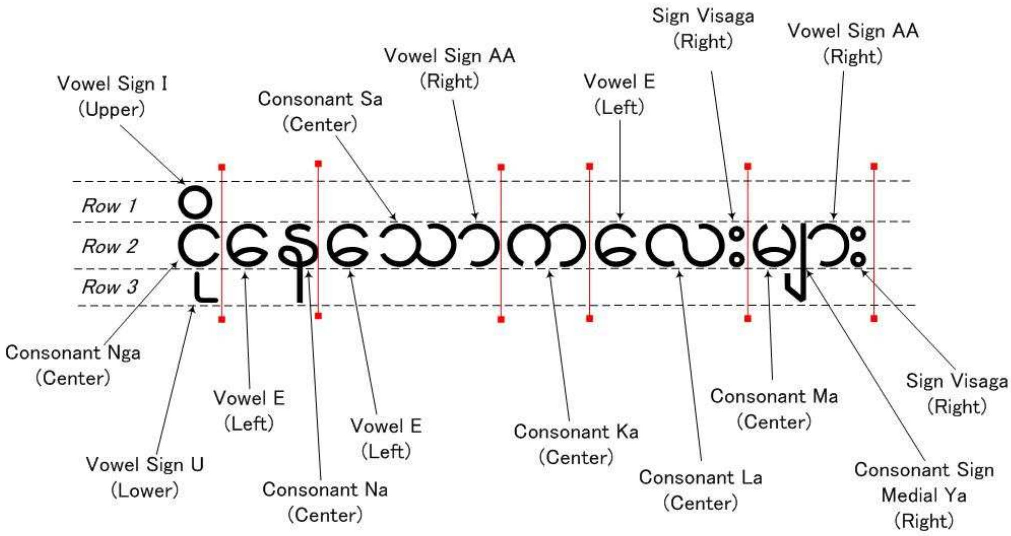


Figure 1-1. Character combination in Myanmar language

(Myanmar phrase "Ngo Nay Thaw Ka Lay Myaa" meaning “crying children”)

Image taken from my published paper [120]

Overall writing direction is from left to right, and the word order is SOV (Subject+Object+Verb). In Myanmar sentences, spaces are used to mark phrases, not to divide words. Making an analysis of Myanmar sentences, it can be basically considered that Myanmar characters are written in three levels (upper, middle and lower). Most characters have their defined positions compared to the base consonant that is usually written in the middle row and center (see Figure 1-1). For example, dependent vowel E or Tha Way Thou “ေ◌” (U+1031) should be written at the left side of the consonant, dependent various sign Dot Below or Auk Ka Myit “◌့” (U+1037) should be written at the lower side of the consonant, dependent vowel II or Lone Gyi Tin San Khat “◌ီ” (U+102E) should be written at the upper side of the consonant, stacked consonant Ka or Ka Gyi “္က” (U+1039 U+1000) should be written at the lower side of the consonant, medial consonant “◌ျ” (U+103B) should be written at the right side of the consonant, etc. In Figure 1-1, a Myanmar phrase "Ngo Nay Thaw Ka Lay Myaa" meaning “crying children” is formed by the combinations of left, right, upper and lower characters and it takes three rows. Six vertical lines indicate the pronunciation breaks of "Ngo", "Nay", "Thaw", "Ka", "Lay" and "Myaa".

The language of one of Southeast Asian countries, Cambodia, called Khmer belongs to the Mon-Khmer family of languages. Khmer alphabets closely resemble Thai and Lao alphabets, but Khmer is not a tonal language. Some vocabularies are borrowed from Sanskrit, Pali, French and Chinese in Khmer. Khmer uses a phonetic alphabet with 33 consonants (e.g. “ក”, “ខ”, “គ”, “ឃ”, “ង”, “ច”, “ឆ”, “ជ”, “ឈ”, “ញ”, etc.(U+1780 to U+17A2), 24 dependent vowels “ើ”, “ឿ”, “ៀ”, “េ”, “ែ”, “ៃ”, “ោ”, “ៅ”, etc. (U+17B6 to U+17C5), 12 independent vowels “ឥ”, “ឭ”, “ឧ”, “ឪ”, “ឩ”, “ឳ”, “ឫ”, etc.(U+ 17A5 to U+17B3), and some other diacritic signs “់”, “៉”, “៌”, “៍”, “័”, “៏”, etc.(U+17C6, U+17C7, U+17C8, etc.) [7]. Most consonants have reduced or modified forms, called subscripts (e.g. “្ក”, “្ខ”, “្គ”, “្ឃ”, “្ង”, etc.), when they occur as the second member of a consonant cluster. Vowels may follow or precede the consonants, or they may go above or below, or some combination of before, after, above, or below. 12 independent vowels can exist without a preceding consonant. Khmer writing begins on the top left of the page, and proceeds down and to the right. The word order follows SVO (Subject + Verb + Object) pattern as in English. Cambodian writing does not use spaces between individual words; instead, spaces are used to denote the end of phrases or sentences. Figure 1-2 shows an example of Khmer writing system.

Making an analysis of Khmer sentences, it can be basically considered that Khmer characters are written in three levels (upper, middle and lower) like in Myanmar language.

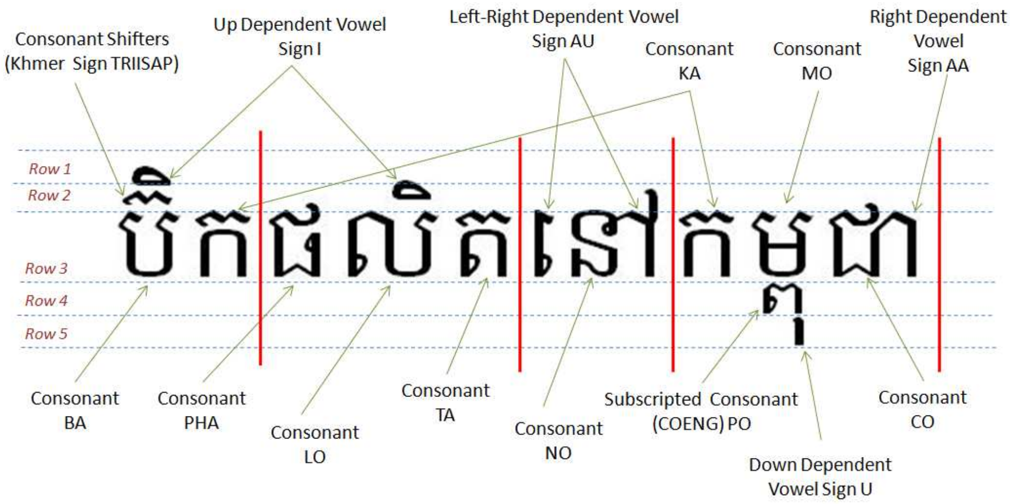


Figure 1-2. Character combination in Khmer language

(Khmer phrase “Bic Phalit Nau Campucha” meaning “a pen made in Cambodia”)

Image taken from my published paper [121]

Nepali is the official language in Nepal [8]. It is an Indo-Aryan language spoken in Nepal, Bhutan and some parts of India and Myanmar. It is closely related to Hindi and is commonly written in the Devanagari script [9] [10] [11] [12]. In Nepali consonants, last three cosonants क्ष (khsya), त्र (tra) and ज्ञ gyna are ligatures (i.e. क् + ष = क्ष, त् + र = त्र and ज् + ञ = ज्ञ). Nepali language is written from left to right, and there is no case distinction. In Nepali, word order is Subject + Object + Verb (SOV). Nepali uses independent vowels (e.g. "अ" (a), "आ" (aa) "इ" (i), "ई" (ii), etc.) and dependent vowels (e.g. "ा" (aa), "ि" (i), "ी" (ii) , "ु" (u), etc.), consonants (e.g. "क" (ka), "ख" (kha), "ग" (ga), "घ" (gha). "ङ" (nga), etc.), half consonants (e.g. "क्" (ka), "ख्" (kha), "ग्" (ga), "घ्" (gha), "ङ्" (nga), etc.), diacritics symbols ("ं" (anuswar), "ँ" (chandrabindu) and ":" (visarga)), punctuations ("्" (viraam or halanta), "।" (purnaviram) and "॥" (double danda)), special characters ("ॠ"(re), "ॐ" (om) and "रु" (rupee)), conjunct consonants or ligatures (e.g. "क्क" (क्क) (ka+ka), "ट्ट" (tta+tta), "त्त" (ta+ta), "न्न" (न्न) (na+na), etc.) and Nepali numbers (e.g. १ (1) २ (2) ३ (3) ४ (4), ५ (5), etc.).

Independent vowels are such as "अ", "आ", "इ" and "ई", which can be written alone. Dependent vowels are such as "ा", "ि","ी" and "ु", which have to be written together with a consonant (e.g. "का", "कि", "की" and "कु"). Nepali or Devanagari consonants have implied vowels like "क" (ka), which is a combination of "क" (k) and "अ" (a). Therefore, a pure consonant is written as "क्", which means that "अ" (a) vowel is not contained. Pure consonants such as "क्", "ख्", "ग्", "घ्" and "ङ्" are written contacting other consonants, and make dead consonants or half consonants. For example, a Nepali word "नमस्ते" ("Namaste" meaning "hello") contains half consonant of "स" (sa) (see Figure 1-3).

Consonants can be combined with one or more consonants and make a ligature. In most cases, ligatures formation can be easily recognized such as क् + र = क्र, ख् + र = ख्र and ग् + र = ग्र. But it is difficult to recognize some ligatures' combinations of consonants because of their new glyph or shape (e.g. क् + ष = क्ष, त् + र = त्र, ज् + ञ = ज्ञ, र + र = र्र, द + द = द्द and क् + त = क्त).

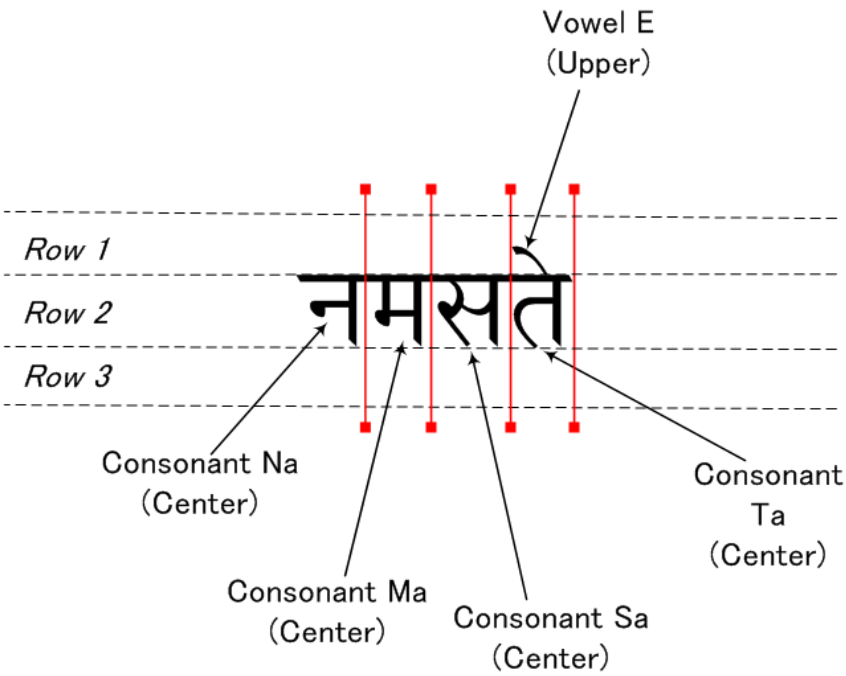


Figure 1-3. Character combination in Nepali language

(Nepali word "Namaste" meaning "hello")

Bengali is the official language of Bangladesh and one of 18 enlisted languages in the Indian constitution. It is the administrative language of the Indian states of Tripura and West Bengal and also one of the administrative languages of Kachar district, Assam. Bengali scripts consist of vowels, consonants and vowel-signs. There are total 60 symbols. Moreover many ligatures (conjuncts), few consonant symbols, and some special signs are also used in Bengali scripts. In the Bengali scripts, there are 11 vowels. All the vowels in Bengali are independent and in the initial position of a word retain their original shape. There are 39 consonants in

Bengali alphabets. The first vowel অ ‘a’ is implicit in all consonants. Each vowel except অ ‘a’ has another shape called vowel-sign to use with a consonant. There are 10 vowel signs.

Making an analysis of Bengali sentences, it can be basically considered that Bengali characters are written in three levels (upper, middle, lower) like in Myanmar language. And most of them have their defined positions (e.g. vowel signs "ি", "ে", "ৈ" have to be written in the left side of consonants, vowel signs "া", "ী", "ৌ" and consonant symbol "্য" (্য) should be written in the right side of the consonants, vowel signs "ু", "ূ", sign hashant "্" and consonant symbols "্র" (্র), "ৃ" should be written as a lower characters and consonant symbol "র্" (র্) and consonant "ঁ" should be written as an upper characters). And for the ligatures or (combine consonants) such as "ক্ক", "ক্র", "ক্স", "ক্ল", "ক্ষ", "ঙ্ক্ষ" and "প্ট" are written in the middle like consonants. In Figure 6-8, a Bengali word "Krissikaarza" (agriculture) is formed by the combinations of left, right, upper and lower characters and it takes three rows. Here, two vertical lines indicate the pronunciation breaks of "Kri" and "Ssikaarza".

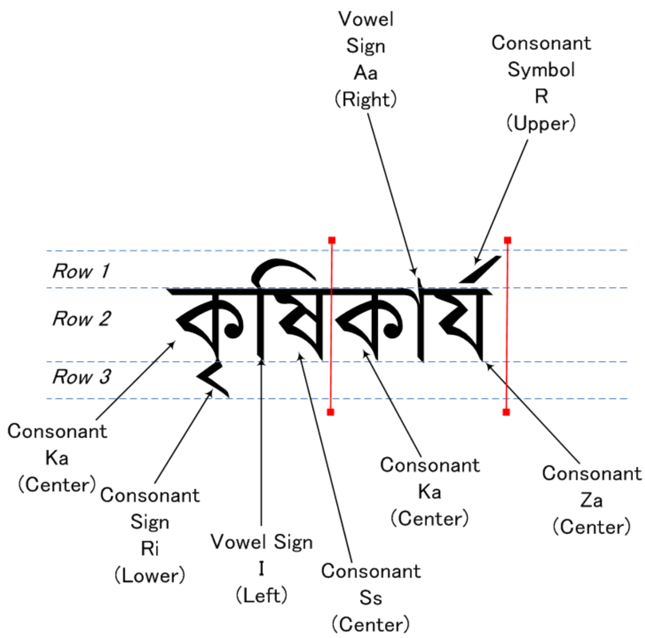


Figure 1-4. Character combination in Bengali Language
Image taken from my published papers [69, 127, 129]

As mentioned above, there are similarities in Asian syllabic languages for their writing, and thus, a common text input interface is applicable for Asian syllabic languages, and this finding motivated me to start my research.

1.1.3 The Challenge in Asian Syllabic Languages Text Entry

As we all know, QWERTY is the most common PC keyboard layout today. Although it is designed for English language, keyboard mappings of other languages including Asian syllabic languages are based on it. Asian syllabic languages differ from an alphabetic writing system not only in their shapes or glyphs but also in their higher number of characters than English, organizing and writing principles. Therefore, text entry with a PC keyboard in Asian syllabic languages is a challenge (see Chapter 2, Section 2.1).

Obviously, text entry on a mobile phone in Asian syllabic language is more challenging. Basically, keyboard mapping on 12-key keypad leads to an overloading even for English language (i.e. 26 characters). Text entry with limited number of keys and small screen involves a complex combination of mental and physical coordination.

All Asian syllabic languages derived from Brahmi and they are known as abugidas. Although Brahmi is no longer in use, it is a root of the scripts used today in South Asian and Southeast Asian syllabic languages such as Myanmar, Thai, Khmer and Hindi. A typical character in abugidas is formed by a combination of vowels with a consonant (see Chapter 1, Section 1.1.2). Moreover, differences in order of writing and typing, no support for some languages by mobile phone companies and luck of embedded fonts make a big challenge for text entry on mobile devices in Asian syllabic languages (See Chapter 2, Section 2.2 and 2.3).

1.1.4 Research Approach

This research is based on a highly iterative "design-prototype development-user study" cycle. All the developed text input interfaces in this research is based on the characters writing position or vowel combinations with a consonant.

To begin with, a possible text input interface was designed for Asian syllabic languages, and then, a prototype was developed rapidly for Myanmar language followed by user study. After making an analysis of the user study results, versions of that prototype was developed for other Asian syllabic languages. And thus, all prototypes were developed for Myanmar language and different versions were developed for some Asian syllabic languages. Bengali, Khmer, Nepali and Thai languages were studied to prove that text input interface based on characters writing position is user-friendly and applicable for all Asian syllabic languages.

Most of the text input interfaces in this research are designed for running on mobile devices such as a mobile phone, a Personal Digital Assistance (PDA), a MP3 player and a portable game player. The main design factor is that they are applicable to run with a limited numbers of button and small screen.

In the prototyping step, two methods were mainly used. One is “high-fidelity prototyping” and the other is a new approach called “PicMobi Prototyping”. High-fidelity prototyping was done by using Visual Basic, $C^{\#}$ (C Sharp) or J2ME (Java 2 Platform, Micro Edition) programming for testing with native users. PicMobi Prototyping is an intermediate solution of paper prototyping with mobile devices, which was proposed in 2009 (see Appendix G, Appendix H). It is an innovative prototyping technique with pictures and mobile devices that I learned from prototyping experiences of this research.

User study designed to evaluate typing speed, error rate and users’ feedbacks was conducted for all prototypes with real users in laboratories or in their native countries. Texts used for typing experiments consisted of various combinations of a consonant and vowels of Asian syllabic languages. Extensive log files were continuously written and video data of users’ typing behaviors were also taken for making a formal evaluation of each text input prototype. User study experiments were designed within-subjects test, where dependent variables were the completion time, number of errors and their feedbacks. Independent variables were the types of prototype (e.g. prototype for a mobile phone, prototype for a PDA, prototype for a PC keyboard). Each participant completed a pre-test and post-test questionnaire to collect demographic and subjective data about the user experience with prototypes.

1.2 Dissertation Organization

This chapter has illustrated the potential value of a common and user-friendly text input interfaces for Asian syllabic languages based on characters writing position. Here is a roadmap to subsequent chapters:

Chapter 2 describes background work that relates to this research in general, such as keyboard mappings and text input methods of several Asian syllabic languages for both a PC and mobile devices. A formal evaluation of text entry techniques is also mentioned briefly.

Chapter 3 introduces the concept of Positional Mapping (PM). This concept uses keyboard or keypad mapping based on the characters writing position. Developed PM prototypes for a mobile phone, a numeric keypad, a PDA, the Ergodex DX1, a game pad and an electronic whiteboard are described in detail. Results of user studies prove user-friendliness and applicability of PM for Asian syllabic languages.

Chapter 4 presents comparison results of the existing QWERTY based software keyboard and the software keyboard version of PM. This chapter clearly shows users' feedbacks on the developed Bengali and Khmer prototypes.

Chapter 5 presents comparison results of the PM software keyboard layout with two other possible keyboard layouts. Although there was no significant difference in text entry speed, Characters per Minute (CPM) of PM was the highest. Moreover, users rated the typing layout of PM the easiest.

Chapter 6 introduces the concept of Positional Gesture (PG). This concept uses left, right, up, down and center gestures or strokes commands for text input in Asian syllabic languages. Developed prototypes for Myanmar, Khmer and Bengali languages are described in detail. The user study results prove user-friendliness and applicability of PG for Asian syllabic languages.

Chapter 7 introduces the concept of Positional Prediction (PP). This concept uses four directional arrow keys, i.e., left, right, up and down arrow keys, for predicting possible combinations of consonant and vowels. Developed original PP prototypes for Myanmar and Khmer are described in detail. The user study results prove user-friendliness and applicability of PP for Asian syllabic languages.

Chapter 8 introduces how PP text input concept can be extended for small mobile devices. This chapter proposes new text input interface Positional Prediction Clickwheel (PP_Clickwheel) for mobile devices using clickwheel like the Apple iPod. Typing speed of novice users with PP_Clickwheel prototype was acceptable and positive feedbacks were received from them.

Chapter 9 begins with an introduction to the writing system and existing Romanization methods of Myanmar language, and then, a detailed explanation of Romanized handwriting text input interface is followed by a report of the user study results with my developed prototype. Finally, a discussion on my findings and suggestion for further text input interface research are made.

Chapter 10 presents font development, Direct Keyboard Mapping (DKM) and Romanized Positional Prediction (RomanPP_Fingerspelling) keyboard prototype for Myanmar language fingerspelling characters. Moreover, usability of my approach is discussed based on the user study results with the current fingerspelling software keyboard prototype.

Chapter 11 presents analysis results of Nepali and Thai syllables or vowels combinations with a consonant. There is also a discussion on why an analysis was made of possible combinations of vowels and a consonant, their patterns and maximum length in the current Nepali and Thai languages.

Chapter 12 introduces grapheme cluster segmentation tools developed for Myanmar language. Distributions of PP patterns of Myanmar and Khmer languages are described. Example usages of this tool for statistical analysis on distributions of Myanmar characters are presented in detail, and tool development for Khmer language is also introduced.

Chapter 13 concludes with prototypes design, major empirical results, a list of contributions, future work and closing remarks.

Figure 1-5 shows organization structure of main chapters of this dissertation.

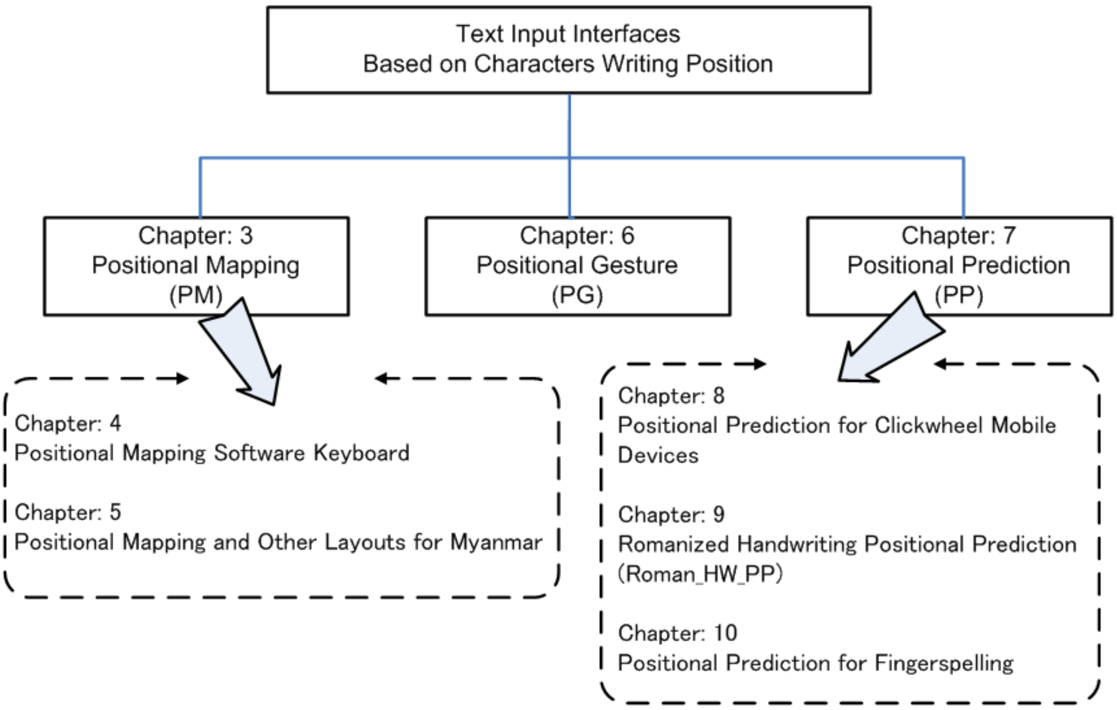


Figure 1-5. Organization structure of main chapters

CHAPTER 2
Related Works

2.1 Text Entry in Asian Syllabic Languages

In this chapter, existing keyboard layouts and text input methods of a PC and a mobile phone is presented. The focus is on text input methods of the abugidas such as Indian scripts, Khmer, Thai, Myanmar and Nepali languages [13].

As with alphabetic scripts, abugidas are written from left to right, but there is no explicit use of spaces or full stops to determine word or sentence boundaries. There are too many possible combinations of base consonants and vowels, marks in abugidas. In order to understand the common text entry issues of abugidas, five abugidas scripts are introduced briefly and explain their text input methods with standard or de facto standard keyboard layouts.

2.1.1 Indian Scripts PC Keyboard Layout

India is a highly multilingual country and around hundreds of languages are used. The principal official language is Hindi, the secondary official language is English and other languages such as Marathi, Bengali, Gujarati, Punjabi, Oriya, Sindhi, Nepali, Assamese, Urdu, Tamil, Telugu, Kannada, Malayalam, Sanskrit, Kashmiri, Manipuri, Konkani, Maithali and Santhali are also officially recognized. Indian languages belong to the Indo-European family of languages [14]. Languages of the north and western part of India belong to the Indo-Aryan family (spoken by about 74% of India’s speakers) while the languages of the south belong to the Dravidian family (about 24% of India’s speakers) [14]. Here, primary concentration is put on the most popular Indian scripts Devanagari. Devanagari consonants (occlusives), consonants (sonorants and fricatives), vowels (independent form), vowels (as diacritic with "प") can be seen in Figure 2-1, Figure 2-2, Figure 2-3 and Figure 2-4 respectively. Devanagari language is written from left to right, and there is no case distinction. Consonants can be combined with one or more consonants and make a ligature (e.g. ligature with "ह" ha: म्ह mha, न्ह nha, ण्ह ṇha, व्ह vha, ल्ह lha, ळ्ह ḷha, र्ह rha.).

क	ख	ग	घ	ङ
(ka)	(kha)	(ga)	(gha)	(ṅa)
च	छ	ज	झ	ञ
(ca)	(cha)	(ja)	(jha)	(ña)
ट	ठ	ड	ढ	ण
(ṭa)	(ṭha)	(ḍa)	(ḍha)	(ṇa)
त	थ	द	ध	न
(ta)	(tha)	(da)	(dha)	(na)
प	फ	ब	भ	म
(pa)	(pha)	(ba)	(bha)	(ma)

Figure 2-1. Devanagari consonants (occlusives)

य	र	ल	व
(ya)	(ra)	(la)	(va)
श	ष	स	
(śa)	(ṣa)	(sa)	
ह	ळ	्	
(ha)	(ḷa)	(halant)	

Figure 2-2. Devanagari consonants (sonorants and fricatives) and halant symbol

अ	आ	इ	ई
(a)	(ā)	(i)	(ī)
उ	ऊ	ऋ	ॠ
(u)	(ū)	(ṛ)	(ṝ)
ऌ	ॡ	ए	ऐ
(ḷ)	(ḹ)	(e)	(ai)
ओ	औ		
(o)	(au)		

Figure 2-3. Devanagari vowel (independent form)

प पा पि पी

(a) (ā) (i) (ī)

पु पू पृ पॄ

(u) (ū) (ṛ) (ṝ)

पॢ पॣ पे पै

(ḷ) (ḹ) (e) (ai)

पो पौ

(o) (au)

Figure 2-4. Devanagari vowel (as diacritic with “प”)

There are two possible way to type text in an Indic script, one is input the text using English (called transliteration or phonetic) or input it directly through an Indic script. Program such as Baraha and AAA+ (triple-A-plus) allow users to type in the Roman script and convert the output into an assigned Indian script [15][16]. For example, type ka for “क”, kra for “क्र”, kA for “का”, etc. in Baraha software, and type ka for “क”, kha for “ख”, ga for “ग”, etc. in AAA+ software. Transliteration or phonetic typing method is simple for users who already know English and foreigner who already have some basic knowledge of an Indic script. Another input method is a keyboard that maps characters for an Indic script on the keyboard (e.g. InScript (standing for Indian script) keyboard). InScript keyboard is the most commonly used for typing Indic scripts, and it was designed for many Indian abugidas [14]. It has been adapted by Microsoft and the keyboard layouts are shown in Figure 2-5 and Figure 2-6.

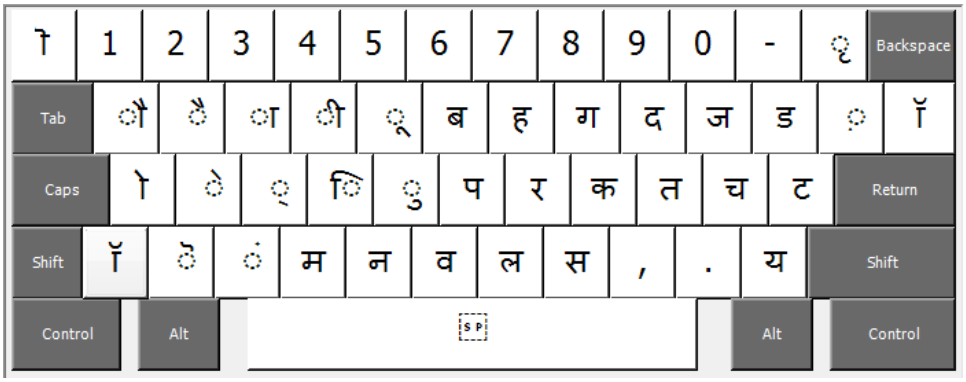


Figure 2-5. InScript keyboard for Devanagari (unshifted)

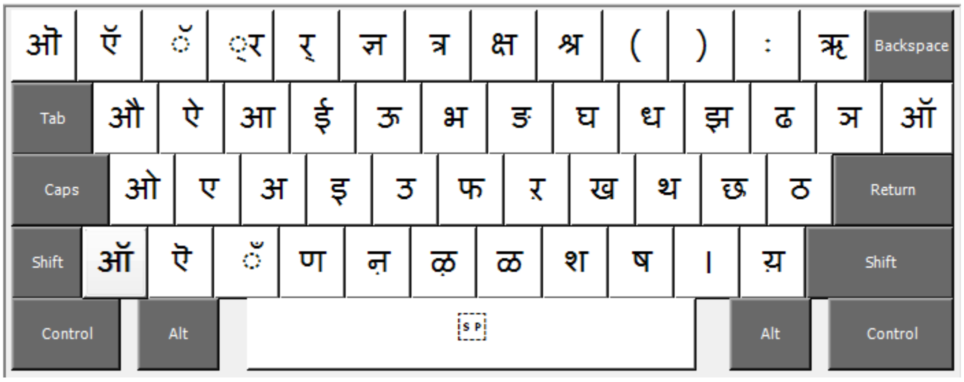


Figure 2-6. InScript keyboard for Devanagari (shifted)

In the InScript keyboard, vowels are placed on the left side of the keyboard while consonants are on the right. Similar sounds and related characters are mapped on the same key for easy to memorize the keyboard layout. For example, "क" (ka) is mapped on "k" key and "ख" (kha) is mapped on "K" (shift + k) key, "ज" (ja) is mapped on "p" key and "झ" (jha) is mapped on "P" (shift+p) key, etc. This keyboard is useful to type several Indic scripts because the key mappings are almost identical across Indic scripts.

2.1.2 Khmer PC Keyboard Layout

Khmer (also known as Cambodian) is the official language of the kingdom of Cambodia. It is belongs to the Mon-Khmer group of Austro-Asiatic languages and has been considerably influenced by Sanskrit and Pali [17]. It is not a tonal language and that is differs from neighboring languages such as Thai, Lao and Vietnam. Khmer language consists of 33 consonants, 32 subscript consonants with a pair of duplicates, 24 dependent vowels, 12 independent vowels, 2 consonant shifters and a dozen diacritic signs and other symbols including number signs [18] (see Figure 2-7, Figure 2-8, Figure 2-9). Almost all of the Khmer consonants have reduced or modified forms, called subscripts, when they combined as the second member of a consonant cluster (see Figure 2-7). Khmer is written from left to right, and there is no case distinction, using a space as a punctuation sign (i.e. used somewhat analogously to commas in Western languages) [19].

ក្ក ខ្ខ គ្គ ឃ្ឃ ង្ង

ច្ច ឆ្ឆ ជ្ជ ឈ្ឈ ញ្ញ

ដ្ដ ឋ្ឋ ឌ្ឌ ឍ្ឍ ណ្ណ

ត្ត ថ្ថ ទ្ទ ធ្ធ ន្ន

ប្ប ផ្ផ ព្ព ភ្ភ ម្ម

យ្យ រ្រ ល្ល វ្វ ស្ស

ហ្ហ ឡ្ឡ អ្អ

Figure 2-7. Khmer consonant and subscript consonants

Image adapted from [123]

អិ អី អឹ អឺ អុ អូ អួ

អើ អឿ អៀ អេ អែ អៃ អោ អៅ

អុំ អំ អាំ អះ អុះ អេះ អោះ

Figure 2-8. Khmer dependent vowels with consonant “អ”

Image adapted from [123]

Khmer independent vowels

ឥ ឦ ឧ ឩ ឪ ឲ ឫ

ឬ ឭ ឮ ឯ ឰ ឱ ឳ

Two Khmer consonant shifters (left side), Khmer diacritics (right side)

◌៉ ◌៊   ◌់ ◌៌ ◌៎ ◌៏ ◌័ ◌៍

Khmer numerals

០ ១ ២ ៣ ៤ ៥ ៦ ៧ ៨ ៩

Figure 2-9. Khmer independent vowels, two consonant shifters, diacritics (right side) and numerals
Image adapted from [123]

Typing Khmer word on Old Khmer typewriters and a PC keyboard is handwriting order (i.e. left to right). All subscripted consonants are also mapped on the keys. Today Khmer PC keyboards are designed based on Unicode encoding (see Figure 2-10, Figure 2-11 and Figure 2-12). And thus, the typing order is the same as spelling order (i.e. not handwriting order) and no need to consider keyboard mapping for subscripted consonants. For example, KhmerOS Unicode keyboard layout uses a subscript sign "◌្" for combining two consonants [20]. Although keyboard mapping for subscripted characters issue is solved by using Unicode subscript sign "◌្", keyboard mapping for other characters are also still challenge. This is because of the large character sets of Khmer and right-alter mode is used (e.g. right-Alter + t for typing "ឦ" independent vowel) (see Figure 2-12).

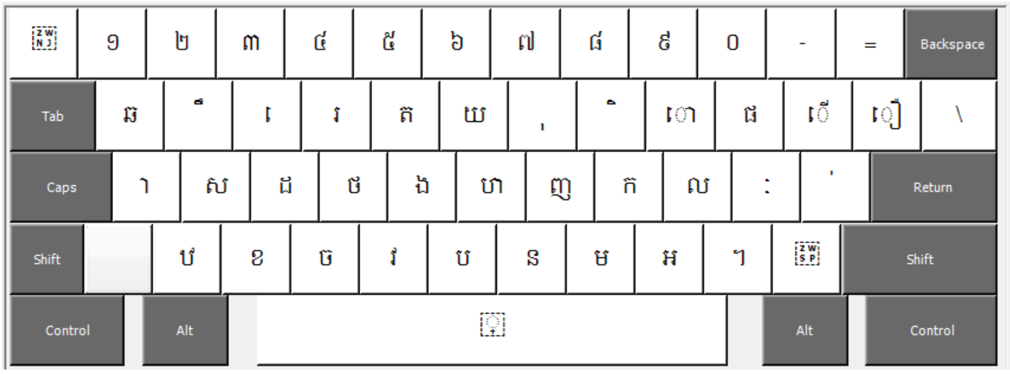


Figure 2-10. Khmer Unicode keyboard layout (unshifted)

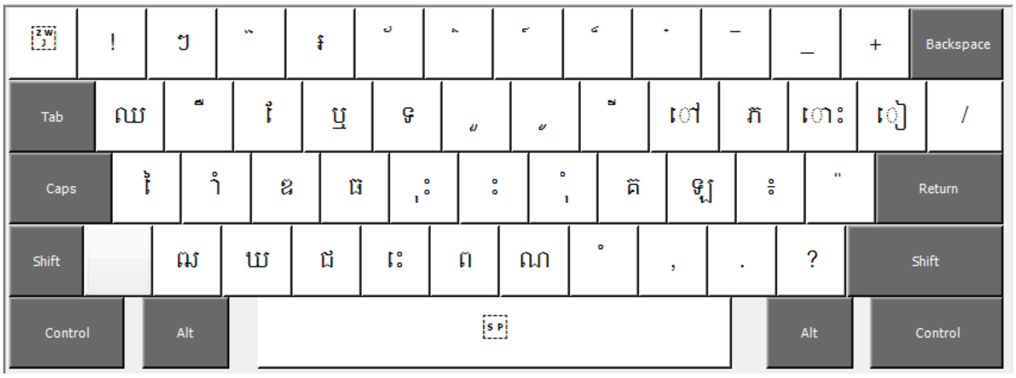


Figure 2-11. Khmer Unicode keyboard layout (shifted)

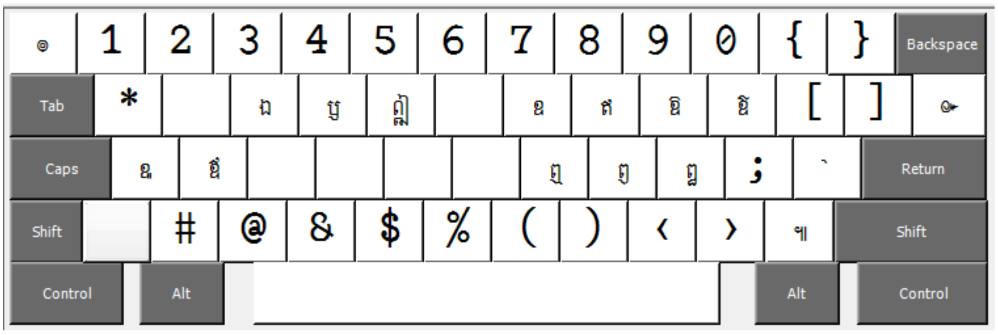


Figure 2-12. Khmer Unicode keyboard layout (Alt+Ctrl or Right Ctrl)

2.1.3 Thai PC Keyboard Layout

Thai is the official language of Thailand [21]. Thai script has been used for Thai, Pali, and Sanskrit languages in Buddhist text all over the country. Most of the dialects of Thai use the same script.

In 1283, King Ramkhamhaeng of Sukhothai created the first Thai alphabet by borrowing and adapting certain Mon and old Khmer words and called “Sukhothai Script”. In 1357, in the reign of King Li Thai, a new script called "King Li Thai script" came to be used. The shapes of the letters in the King Li Thai script are based on the ones of Sukhothai although some of them were modified. In 1680, during the reign of King Narai, a script called "King Narai script" was brought into use, and it has been developed and preserved as Thai national script up to now [22].

In Thai language, there are 44 consonants, 15 vowels, 4 consonant-vowels, 4 Tone marks, 5 Diacritic symbols (3 are not used presently), 10 Thai numbers and 2 Special symbols.

Consonant: 44 consonants that represent 21 consonant sounds are shown in Figure 2-1 and Figure 2-2. The extra letters represent sounds that are not distinguished anymore in modern pronunciation, but still carry etymological information and are used to distinguish tone [23].

ก	ข	ฃ	ค
(Ko Kai)	(Kho Khai)	(Kho Khuat)	(Kho Khwai)
ฅ	ฆ	ง	จ
(Kho Khon)	(Kho Rakhang)	(Ngo Ngu)	(Cho Chan)
ฉ	ช	ซ	ฌ
(Cho Ching)	(Cho Chang)	(So So)	(Cho Choe)
ญ	ฎ	ฏ	ฐ
(Yo Ying)	(Do Chada)	(To Patak)	(Tho Than)
ฑ	ฒ	ณ	ด
(Tho Nangmontho)	(Tho Phuthao)	(No Nen)	(Do Dek)

ต	ถ	ท	ธ
(To Tao)	(Tho Thung)	(Tho Thahan)	(Tho Thong)
น	บ	ป	ผ
(No Nu)	(Bo Baimai)	(Po Pla)	(Pho Phung)
ฝ	พ	ฟ	ภ
(Fo Fa)	(Pho Phan)	(Fo Fan)	(Pho Samphao)
ม	ย	ร	ล
(Mo Ma)	(Yo Yak)	(Ro Rua)	(Lo Ling)
ว	ศ	ษ	ส
(Wo Waen)	(So Sala)	(So Rusi)	(So Sua)
ห	ฬ	อ	ฮ
(Ho Hip)	(Lo Chula)	(O Ang)	(Ho Nokhuk)

Figure 2-13. Thai consonants (alphabetical order)

Image adapted from [122]

/ก/: ก	/บ/: บ
/ค/: ข ฃ ค ฅ ฆ	/ป/: ป
/ง/: ง	/พ/: ผ พ ภ
/จ/: จ	/ฟ/: ฝ ฟ
/ช/: ฉ ช ฌ	/ม/: ม
/ซ/: ซ ศ ษ ส	/ร/: ร
/ย/: ญ ย	/ล/: ฬ ล
/ด/: ฎ ด	/ว/: ว
/ต/: ฏ ต	/ฮ/: ห ฮ
/ท/: ฐ ฑ ฒ ถ ท ธ	/อ/: อ

/น/: ณ น

Figure 2-14. Thai consonants (ordered by pronunciation)

Image adapted from [122]

Vowel: There are 15 dependent vowels in Thai language and they can be categorized into 4 groups based on their position in combination with base consonant as follows:

Leading Vowels or Left Vowels: เ, แ, โ, ใ, ไ

Following Vowels or Right Vowels: ะ, า, ำ

Upper Vowels: ั, ิ, ี, ึ, ื

Lower Vowels: ุ, ู

Tone Mark: There are 4 marks " ่", " ้ ", " ๊ ", " ๋ " (5 tones) in Thai language. All of them are written above their related consonants or vowels.

Consonant-vowel: There are 4 Consonant-Vowel "ฤ" "ฤๅ" "ฦ" "ฦๅ" that are normally used in Pali, Sanskrit and old Thai words. But these 4 Consonant-vowels are rarely used today except in old style poetry. Especially in texting, these 4 consonant-vowels are never used.

Diacritic symbol: There are 4 upper diacritic symbols " ็ ", " ์ ", " ํ" and "๎". All of them are written above their associated consonants. " ฺ" is a lower diacritic symbol. It is written below its associated consonants. But the current Thai language uses " ็ " and " ์ " only.

Thai number: "๑" "๒" "๓" "๔" "๕" "๖" "๗" "๘" "๙" "๐". These 10 Thai numbers are used in some formal documents but not in daily life.

Special symbols: "ๆ" is a symbol used for word repetition and "ฯ" is a symbol used for word abbreviation.

Although Thai script derived from the Khmer alphabet, which is modeled after the Brahmi script from the Indic family, Thai doesn't have independent vowel signs or conjunct consonants like other Indic scripts [24].

In Thai writing system, there are no capital letters or spaces between words. In order to compose a syllable or a word, at least one base Consonant is required. Then adding other Consonant, Vowels, Diacritic Symbols and Tone marks to the base Consonant. The combinations can be Consonants without vowel and other marks, one Consonant or more with Vowels, one Consonant or more with Diacritic symbol, one Consonant or more with Tone mark, one Consonant or more with vowels and tone mark, one Consonant or more with vowels, diacritic symbol and tone mark. When a consonant has upper vowel and a tone mark together on it, the tone mark moves up and becomes smaller to make room for the vowel

Like in other Asian syllabic languages, there are too many possible combinations of base consonants and combining marks in Thai. And thus, base consonants and combining characters are encoded separately, rather than pre-combined. Thai combining marks can be divided into upper or lower vowels, tone marks, and diacritics. Moreover, a Thai base consonant can be combined with up to two combining marks, that is, zero or one upper or lower vowel and zero or one tone mark or diacritic [14]. The vowels “เ”, “แ”, “โ”, “ใ”, “ไ” that are written to the left of the preceding consonant are in visual order instead of phonetic order, unlike the Unicode representations of other Indic scripts [25]. This is because, since Thai localization had begun in 1968, by 1990 the Thai computer industry had equipped itself with libraries and tools to use visual order without a problem [14]. The visual order was formally announced as a national standard in TIS 1566-2541 (1998) [14]. Therefore, typing Thai words on a PC keyboard is visual order (i.e. not phonetic order). Figure 2-15 and Figure 2-16 show Thai Unicode keyboard layout.

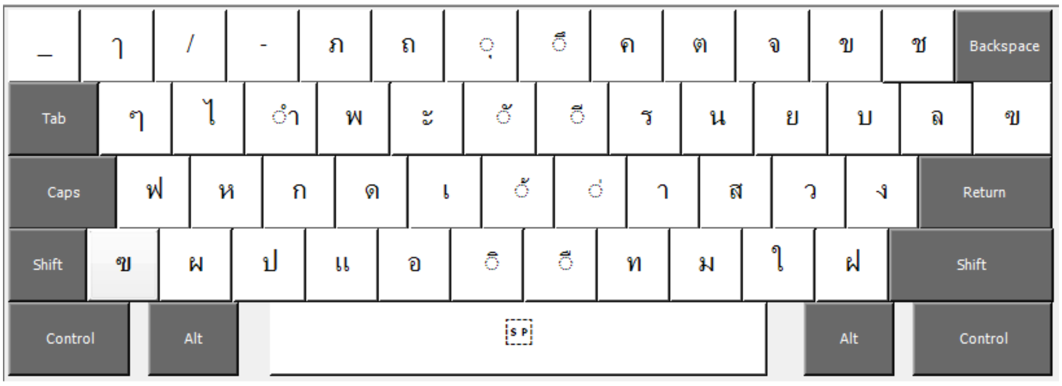


Figure 2-15. Thai Unicode keyboard layout (unshifted)

Figure 2-16. Thai Unicode keyboard layout (shifted)

2.1.4 Myanmar PC Keyboard Layout

Myanmar language is spoken by 32 million as a first language, and as a second language by minorities in Myanmar. It is a tonal and analytic language, which has five tones. There are three main (high, low, creaky) and two substandard (stopped and reduced) tones. The Myanmar writing system derives from the Brahmi-related scripts that flourished in India from about 500 B.C to over 300 A.D. Myanmar script is a system of writing constructed from consonants, consonant combination symbols (i.e. Medials), vowel symbols related to the relevant consonants, and diacritic marks indicating tone level (niggahita, visajjaniya). Despite great differences in appearance and detail, the Myanmar script follows the same basic principle as Devanagari.

There are 33 consonants from “က” (ka) to “အ” (a) (if “ဥ” and “ဿ” is added, 35 consonants) in Myanmar language ‘ThinBonGyi’, 12 basic vowels and some conjunction alphabets such as ၌, ၍, ၎, ၏ and ဤ. Nearly all syllables end with vowels. Unlike other Southeast Asian languages (e.g. Thai, Khmer), Myanmar language adopted words primarily from Pali rather than from Sanskrit. Verbs are placed at the end of the sentence in Myanmar language grammar. Overall writing direction is from left to right. The followings are Myanmar consonant, independent vowel, dependent vowel, various sign, consonant sign and numeral:

က	ခ	ဂ	ဃ	င	
ka	kha	ga	gha	nga	
စ	ဆ	ဇ	ဈ	ဉ	ည
ca	cha	ja	jha	nya	nnya
ဋ	ဌ	ဍ	ဎ	ဏ	
tta	ttha	dda	ddha	nna	
တ	ထ	ဒ	ဓ	န	
ta	ttha	da	dha	na	
ပ	ဖ	ဗ	ဘ	မ	
pa	pha	ba	bha	ma	
ယ	ရ	လ	ဝ	သ	ဿ
ya	ra	la	wa	sa	great sa
	ဟ	ဠ	အ		
	ha	lla	a		

Figure 2-17. Myanmar consonant

Dependent Vowel Sign

ာ	ါ	ိ	ီ	ု	ူ	ေ	ဲ
a	tall aa	I	ii	u	uu	e	ai

Various Sign

ံ	့	း	္	်
Anusvara	dot below	visarga	virama	asat

၍ ၎ ဪ ဪ

Dependent Consonant Sign

ျ	ြ	ွ	ှ
Ya	Ra	wa	ha

Punctuation

၊	။
Little section	section

Number

၀	၁	၂	၃	၄	၅	၆	၇	၈	၉	၀
zero	one	two	three	four	five	six	seven	eight	nine	ten

Figure 2-18. Dependent vowel, various sign, dependent consonant sign, punctuation and number

Basically, there are three types of typing order for Myanmar language, which are handwriting order, traditional typewriter order (Martin Tytell) and PC keyboard order [2]. General Handwriting order is 1) left vowel (i.e. “ေ”), 2) consonants (e.g. “က”, “ခ”, “ဂ”, “ဃ” , “င”, etc.), 3) medials (“ျ”, “ြ”, “ွ” and “ှ”), 4) upper vowels (e.g. “ိ”, “ီ”, “ဲ”, etc.) and lower vowels (e.g. “ု”, “ူ”, “့”, etc.), 5) killers (e.g. “က်”, “စ်”, “င်”, “န်”, “န့်”, etc.) and 6) right vowels (e.g. “ာ”, “ါ”, “း”, “ား”, etc.). Handwriting order for Myanmar word "မြို့"(city) is as follows:

မ + ြ + ိ + ု + ့ = မြို့

Traditional typewriter typing order depends on typists and there are various typing orders. According to the results of my survey with ten typists, there are four types of typing orders for Myanmar word " မြို့ "(city), but it is common that all of the typists type the consonant "မ" at the end as follows:

ြ + ိ + ု + ့ + မ = မြို့
ိ + ြ + ု + ့ + မ = မြို့
ိ + ု + ြ + ့ + မ = မြို့
ု + ိ + ့ + ြ + မ = မြို့

PC keyboard typing order can be divided into two groups; one is "typewriter based typing order" and the other is "Unicode encoding based typing order". The former is similar to typewriter order, which is 1) left vowel (i.e. “ေ”) and left medial (“ြ”), 2) consonants (e.g. “က”, “ခ”, “ဂ”, “ဃ”, “င”, etc.), 3) other medials (“ြ”, “ွ” and “ှ”), 4) upper vowels (e.g. “ိ”, “ီ”, “ဲ”, etc.) and lower vowels (e.g. “ု”, “ူ”, “း”, etc.), 5) killers (e.g. “က်”, “စ်”, “င်”, “န်”, “န့်”, etc.) and 6) right vowels(e.g. “ာ”, “ါ”, “း”, “ား”, etc.)

ြ + မ + ိ + ု + ့ = မြို့

The latter is almost the same as handwriting order except that consonants are typed at first for “ေ + consonant” combination. It is totally based on correct grammatical combination order of consonant and vowels or consonant and medials. The main difference is that typing "subscript or stack characters" needs a signal or Virama sign (U+1039) as follows:

ဗ + ္ + ဗ = ဗ္ဗ

According to my informal survey (i.e. visited around 25 desktop publishing shops in Yangon city and asked questions to the people of Mandalay city), there are three main PC keyboard layouts of Myanmar languages. They are “WinMyanmar”, “Zawgyi Myanmar” and “Myanmar3” keyboard layouts [27] [28] [29] (see Figure 2-19, Figure 2-20 and Figure 2-21).

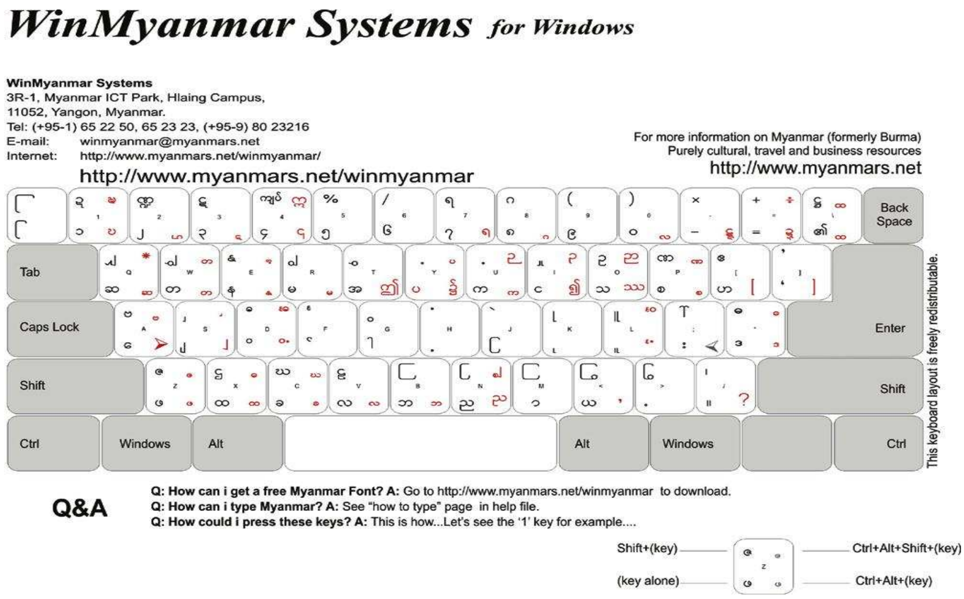


Figure 2-19. WinMyanmar keyboard layout (ASCII based)

Image from http://img458.imageshack.us/img458/9025/winmyanmarkeyboardbig9sx.jpg

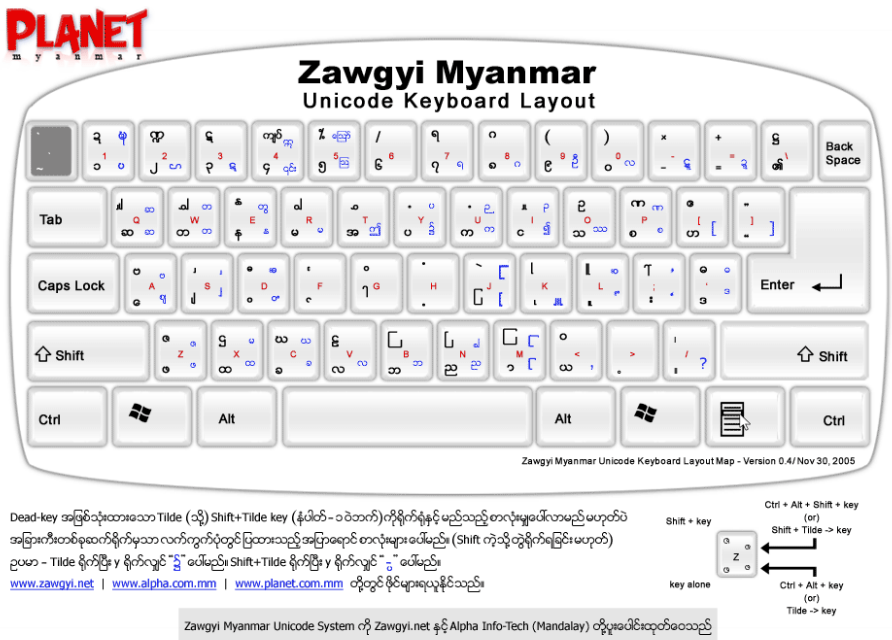


Figure 2-20. Zawgyi Myanmar keyboard layout (partial Unicode)

Image from http://sites.google.com/site/lynnseck/zawgyi_keyboard_map0.4.gif

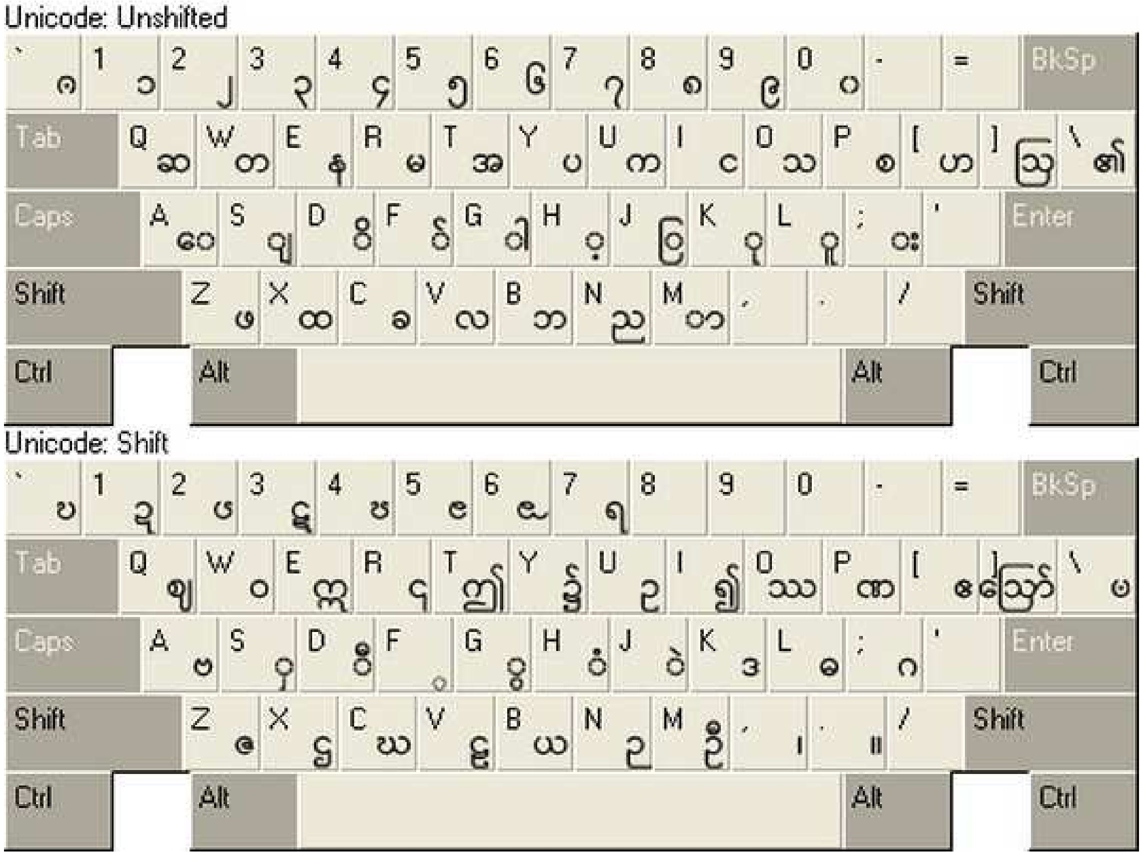


Figure 2-21. Myanmar3 keyboard layout (Unicode)

Image from http://www.myanmarnlp.net.mm/

The Myanmar3 keyboard layout and typing order is totally based on Unicode encoding. WinMyanmar keyboard has the older layout, but it is still used in desktop publishing shops (DTP shops) as well as most publishers in Myanmar. Zawgyi Myanmar is also widely used and especially for chatting, creating webpage, etc. Different glyphs of “ျ” and “ြ” are considered as different characters in WinMyanmar and Zawgyi keyboards and mapping them on several keys.

2.1.5 Nepali PC Keyboard Layout

Nepali language is the official language in Nepal [30]. It is an Indo-Aryan language spoken in Nepal, Bhutan, and some parts of India and Myanmar (Burma). It is closely related to Hindi and is commonly written in the Devanagari script [9] [31] [32] [33]. In Nepali consonants, last three cosonants क्ष (khsya), त्र (tra) and ज्ञ gyna are ligatures (i.e. क् + ष = क्ष, त् + र = त्र and ज् + ञ = ज्ञ). Nepali language is written from left to right, and there is no case distinction. In Nepali, word order is Subject + Object + Verb (SOV). Nepali uses independent vowels (e.g. “अ” (a), “आ”

(aa) “इ” (i), “ई” (ii), etc.) and dependent vowels (e.g. “ा” (aa), “ि” (i), “ी” (ii), “ु” (u), etc.), consonants (e.g. “क” (ka), “ख” (kha), “ग” (ga), “घ” (gha), “ङ” (nga), etc.), half consonants (e.g. “क्” (ka), “ख्” (kha), “ग्” (ga), “घ्” (gha), “ङ्” (nga), etc.), diacritics symbols (“ं” (anuswar), “ँ” (chandrabindu) and “:” (visarga)), punctuations (“्” (viraam or halanta), “।” (purnaviram) and “॥” (double danda)), special characters (“ॠ”(re), “ ॐ ” (om) and “रु” (rupee)), conjunct consonants or ligatures (e.g. “क्‌क” (क्क) (ka+ka), “ट्ट” (tta+tta), “त्त” (ta+ta), “न्‌न” (न्न) (na+na), etc.) and Nepali numbers (e.g. १ (1) २ (2) ३ (3) ४ (4), ५ (5), etc.).

Independent vowels are such as “अ”, “आ”, “इ” and “ई”, which can be written by themselves. Dependent vowels are such as “ा”, “ि”,”ी” and “ु”, which have to be written together with a consonant (e.g. “का”, “कि”, “की” and “कु”). Nepali or Devanagari consonants have implied vowels like “क” (ka), which is a combination of “क” (k) and “अ” (a). Therefore, a pure consonant is written as “क्”, which means that “अ” (a) vowel is not contained. Pure consonants such as “क्”, “ख्”, “ग्”, “घ्” and “ङ्” are written contacting other consonants, and make dead consonants or half consonants. For example, a Nepali word “नमस्ते” (namaste) contains half consonant of “स” (sa).

Consonants can be combined with one or more consonants and make a ligature. In most cases, ligatures formation can be easily recognized such as क् + र = क्र, ख् + र = ख्र and ग् + र = ग्र. But it is difficult to recognize some ligatures’ combinations of consonants because of their new glyph or shape (e.g. क् + ष = क्ष, त् + र = त्र, ज् + ञ = ज्ञ, र + र = र्र, द + द = द्द and क् + त = क्त).

Independent and dependent vowels, consonants, punctuations and numbers of Nepali are as follows:

क	ख	ग	घ	ङ
(ka)	(kha)	(ga)	(gha)	(nga)
च	छ	ज	झ	ञ
(cha)	(chha)	(ja)	(jha)	(nya)
ट	ठ	ड	ढ	ण
(tta)	(ttha)	(dda)	(ddha)	(nna)
त	थ	द	ध	न
(ta)	(tha)	(da)	(dha)	(na)
प	फ	ब	भ	म
(pa)	(pha)	(ba)	(bha)	(ma)
य	र	ल	व	
(ya)	(ra)	(la)	(va)	
श	ष	स	ह	
(sha)	(ssa)	(sa)	(ha)	
क्ष	त्र	ज्ञ		
(khsya)	(tra)	(gyna)		

Figure 2-22. Nepali consonant

Image adapted from my published paper [97]

Vowels:

अ	आ	इ	ई	उ	ऊ	ए	ऐ	ओ	औ	अं	अः
(a)	(aa)	(i)	(ii)	(u)	(uu)	(e)	(ai)	(o)	(au)	(a)	(a)
◌ा	◌ि	◌ी	◌ु	◌ू	◌े	◌ै	◌ो	◌ौ	◌ं	◌ँ	◌ः

Punctuations:

◌् । ॥

Numbers:

१ (1) २ (2) ३ (3) ४ (4) ५ (5) ६ (6) ७ (7) ८ (8) ९ (9) ० (0)

Figure 2-23. Nepali vowel, punctuation and number

Image adapted from my published paper [97]

Three major Unicode keyboard layouts exist for Nepali, which are “Traditional Layout”, “Romanized Layout” and “Devanagari Keyboard”. The Traditional Keyboard Layout derived from the layout of mechanical typewriters and Unicode version of this was developed by Madan Puraskar Pustakalaya (MPP) [34] [35] (see Figure 2-24). It might be easier for the user already familiar with the old traditional layout. However, it is not fully Unicode compatible and difficult to port to other OSs. The Romanized Layout has been developed by Madan Puraskar Pustakalaya (MPP) and has been designed to ease Nepali typing for those who are already familiar with English keyboard layout [34]. However, it is not proper arrangement of Nepali alphabets and not statistically optimized. As mentioned above, Nepali is closely related to written system of the Devanagari script and it is possible to type with Devanagari keyboard layout. The Devanagari Keyboard Layout has been developed by Dhak Prasad Upreti (Bishnu) [34]. Although it has been statistically modeled and corresponds to the frequency of single letters (monograph) in Nepali, it is not popular among native users.

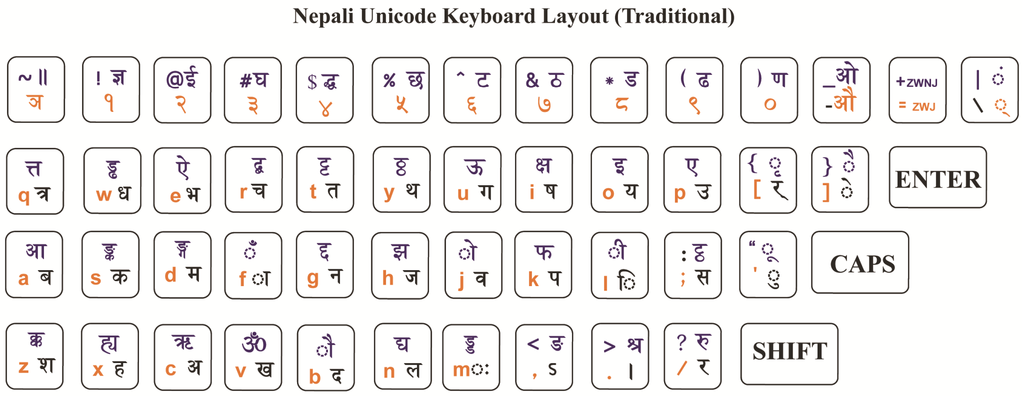


Figure 2-24. Nepali Unicode keyboard layout (traditional)

Image taken from

http://mpp.org.np/index.php?option=com_content&task=view&id=12&Itemid=81

2.2 Comparison on Thai Mobile Phone Keypad Layout

Unlike English and other Roman Character based languages; Thai language with 72 characters (Consonant, Vowel, Tone-mark and Special symbol) is challenge for designing the easy-to-use and user-friendly keypad mapping method. Currently, there are many methods of mobile keypad mapping which depend on brand and model of mobile handset. These keypad mapping methods also apply with Multi-tap and predictive method (T9, iTap). In the past there were other methods such as Thai SMS2 of TAOrange and Thai Matrix of Hutchison which are developed by local mobile operator, but those 2 methods are already obsolete from the market. This study focuses on 4 methods of well-known mobile phone manufacturers; Nokia, Sony Ericsson (SE), Samsung and Motorola which cover around 80% of mobile in Thai market. Though there are many methods of keypad mapping, Thai text input method is unable to gain good feedbacks from users because of its difficulty.

2.2.1 Keypad Mapping

Each leading mobile handset manufacturers (Nokia, SE, Samsung, Motorola, etc.) has their own keypad mapping methods while most of China-made and local mobile handset duplicated the keypad mapping and input method from leading global brand. Therefore, in this study, only 4 main brands covering around 80% of market share are considered.

Nokia, around 40% market share [36], the dominant brand in Thai mobile market, introduces their own keypad mapping by mapped consonants into “1”-“9”, upper and lower position vowel, special symbol and tone mark are assigned to “*”, “0” is for left and right position vowels, “#” is for changing input method (English multi-tap, English multi-tap capital letter, English T9, Thai multi-tap, Thai T9, number) as shown in Figure 2-25.

Figure 2-25. Nokia keypad layout

Image taken from my published paper [124]

The second ranking with 20% market share [36], Samsung has exactly the same keypad mapping method as Nokia but “#” is for space and using other buttons (not a 12-key keypad) for changing input method as shown in Figure 2-26.

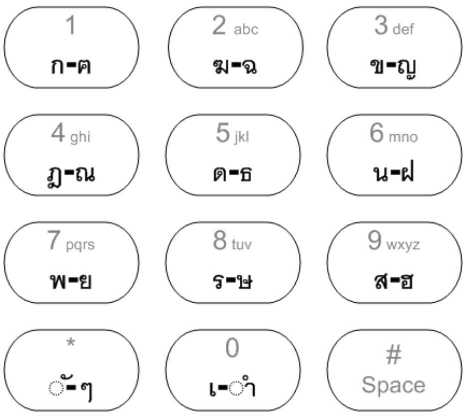


Figure 2-26. Samsung keypad layout

Image taken from my published paper [124]

SE, around 10% market share uses “1” for space and mapped consonants to “2” to “9”. “*” is for left and right position vowel, “0” is for upper and lower position vowel and “#” is for tone mark and special symbol and using other buttons for changing input method as shown in Figure 2-27.

Figure 2-27. Sony Ericsson keypad layout

Image taken from my published paper [124]

Motorola with more than 5% market share has a different mapping method; consonants are mapped into “1” to “6”. And “7” is for left position vowel, “8” is for upper and lower position vowel, “9” is for right vowel and right position special symbol, “0” is for tone mark and upper position special symbol , “*” is for space and “#” is for changing input method as shown in Figure 2-28.

Figure 2-28. Motorola keypad layout

Image taken from my published paper [124]

Nokia and SE show the list of characters (vowel, tone mark and special symbol) on screen when pressing “*” and “0” buttons. For Nokia, after users press “*” and “0”, a list of characters will be shown on the screen and users will be able to select the desired character from the list by using arrow keys. While Sony Ericsson just shows the list of characters on screen, users have to use normal multi-tap way to get the desired character, unable to select desired character by arrow keys. Moreover, Nokia shows only possible vowel and mark which help reducing the total keystroke while SE shows the entire list. For example, Nokia, when pressing “*” after main consonant all of the vowel, tone mark and special symbol will be shown but if there is no main consonant only special symbol will be shown. Samsung and Motorola do not show the list of characters on the screen.

2.2.2 Text Input Method

All of these keypad mapping methods are applied for both multi-tap and predictive input method. For multi-tap, users have to press the same key several times to get the desired character and it will be more difficult to get vowel and tone mark. For some characters, users have to press the button for more than 10 times and it will be much more annoying if they fail to press the correct keystroke which tends to occur with vowel and tone mark. For example, if users want to get the tone mark ◌๋, they need to press # button 10 times, but if they mistype, they have to delete the mistyped character and start all-over again.

Predictive input method seems to be easier to use than multi-tap, but practically users always face a difficulty at the very first time of using T9 or iTap and don't even try using it again. The first reason is that it's difficult to know exactly which button is for which character. It's not like English that there are only 3 characters in one button and all of those characters were printed on the button. Thai has only the first and last character printed on the button like ค-ร printed on 5-button while there are 5 characters in this button. Users have to memorize the position of characters (layout). Another reason is that users do not always get the desired word even if they press the right sequence and combination of buttons. It is because of the predictive algorithm that shows the frequently used word on the screen which is sometimes not the one that users want, and most of them do not know how to get the desired word. For these 2 reasons, users are not comfortable with predictive method and prefer multi-tap method. Comparing the text input method of Thai and English, Thai users who have English capability prefer English text input method because of its easiness.

2.2.3 Test Word/Phrase and Sentence

Nokia 5220, Samsung J700, SE K330 and Motorola RAZR V3 are used for a Keystroke per character (KSPC) test. KSPC is the number of keystrokes required, on average, to generate a character of text for a given text entry technique in a given language [37]. Frequently used words/phrases in Thai texting were chosen and 5 sentences were composed to be used in evaluating KSPC. For words/phrases, those words and phrases were grouped into 4 categories; "sentence ending and greeting", "normal place name", "daily verb" and "person and pronoun" as shown in Table 2-1. For sentences, 5 sentences covering consonant, vowel, tone mark, special symbol and number were composed as shown in Table 2-2.

Table 2-1. Thai words/phrases for KSPC evaluation

Table taken from my published paper [124]

Category	**Words/Phrases**
Ending and Greeting	ครับ คะ จ๊ะ มั้ย นะ ล่ะ ก็ อยู่ กับ ไม่ ป่าว อยาก อะไร ใคร ที่ไหน เมื่อไหร่ ยังไง สวัสดี ฝันดี คิดถึง
Place	สยาม สีลม ลาดพร้าว มหาลัย เซ็นทรัล พารากอน มาบุญครอง โลตัส บิ๊กซี โรงเรียน บ้าน ออฟฟิศ ที่ทำงาน โรงงาน ตลาด
Daily Verb	ไป มา กลับ กินข้าว นอน เที่ยว ทำงาน เดิน เล่น กำลัง โทร รับสาย ขึ้น ลง ดู หนัง
Person and Pronoun	แฟน พ่อ แม่ พี่ น้อง เพื่อน กิ๊ก ฉัน แก เธอ เรา มึง กู ที่รัก ไอ้

Table 2-2. Thai sentences for KSPC evaluation

Table taken from my published paper [124]

Sentence
สวัสดีครับ สบายดีรึป่าว (Hi, How are you doing?)
ทำอะไรอยู่ คิดถึงมากๆ (What are you doing? I miss you so much.)
เย็นนี้ 6 โมงครึ่งเจอกันที่พารากอนชั้น 4 นะ (See you at 6:30 p.m. at the 4th floor of Paragon.)
เราจะเปลี่ยนเบอร์ใหม่เป็น 0890304200 (I will change my phone number to 0890304200.)
หลับฝันดีนะ (Good night.)

Both Thai text input methods; multi-tap and predictive methods will be tested for KSPC. KSPC was calculated by counting the total keystroke required to get the specified word/phrase or sentence and divide by total character of the word/phrase or sentence. In case of sentence, the keystroke of spacing and text input mode changing was also counted. Contradictory, word/phrase case is regardless of spacing and text input mode changing.

2.2.4 Analysis on Results

For word/phrase test with 287 total characters, in multi-tap mode SE has the least KSPC of 3.39 then Motorola, Samsung and Nokia with 3.51, 3.58 and 3.63 KSPC respectively. In predictive mode, Nokia and Samsung have same KSPC of 1.51 the least among these 4 methods then SE and Motorola with 1.55 and 1.66 KSPC as shown in Figure 2-29.

For sentence test with 134 total characters including keystroke of spacing and changing text input mode, in multi-tap mode the least KSPC is SE of 2.86 then Motorola, Nokia and

Samsung with 3.22, 3.28 and 3.31 KSPC respectively. In predictive mode, Nokia is the least KSPC method of 1.38 then Samsung, SE and Motorola with 1.43, 1.51 and 1.60 KSPC as shown in Figure 2-30.

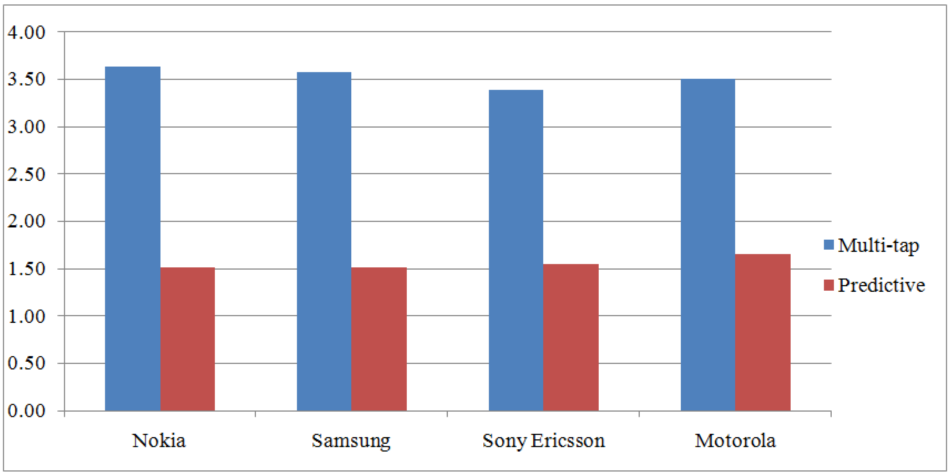


Figure 2-29. KSPC comparison of typing Thai words and phrases

Image taken from my published paper [124]

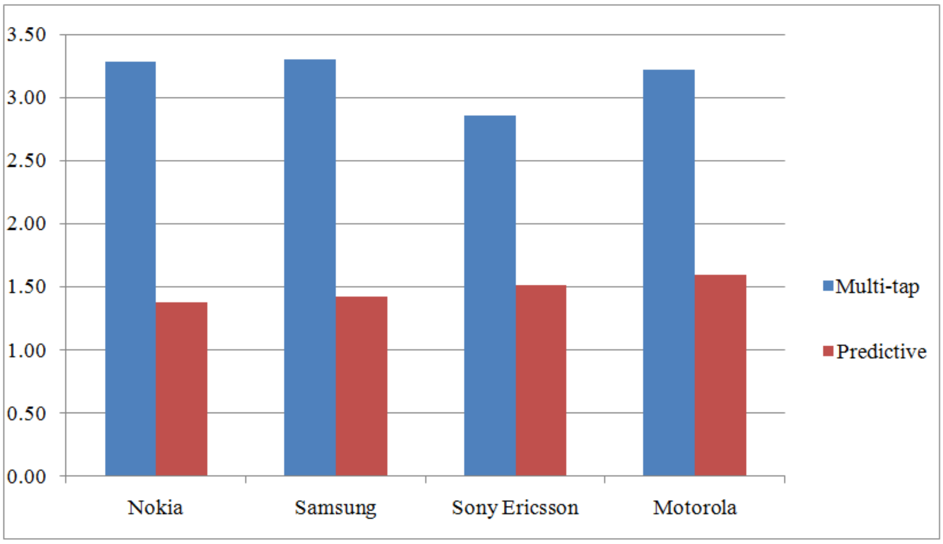


Figure 2-30. KSPC comparison of typing Thai sentences

Image taken from my published paper [124]

For multi-tap mode, in both word and sentence tests, SE and Motorola's keypad mapping which map vowel, tone mark and special symbol into 3 and 4 buttons have lower KSPC than Nokia and Samsung that use 2 buttons. Though Motorola uses 4 buttons for vowel and tone mark, Motorola's KSPC is higher than SE because Motorola maps all consonants into 6 buttons while SE uses 8 buttons. Therefore, Motorola needs more KSPC for consonant part.

Though Nokia and Samsung have the same keypad mapping, Samsung has lower KSPC than Nokia in multi-tap mode because Nokia shows the list and users have to select by pressing "select" button that requires one more extra keystroke.

But as mentioned in the previous section that Nokia has a feature to show the list and users are able to select by arrow key, this feature helps reducing the KSPC of Nokia from 3.63 to 2.99 in the case of word and from 3.28 to 2.74 in the case of sentence. With this feature, Nokia turned to be the least KSPC text input method in multi-tap. In order to show the list, however, it takes around 2-3 seconds after pressing the button that is not concerned in terms of KSPC. Considering the typing speed, however, it will be affected by this time lag.

For predictive mode, in both word and sentence tests, methods of Nokia and Samsung have lower KSPC than the ones of SE and Motorola. In word test, Nokia and Samsung have the same KSCP because Nokia and Samsung have the same keypad mapping and predictive algorithm. But in sentence test including numbers, Nokia's KSPC is lower than the one of Samsung because users can press and hold the button to type numbers for Nokia, but for Samsung if they press and hold the button, numbers are shown in Thai language that are rarely used in daily writing. Therefore, input mode needs to be changed to number mode before typing, which requires 5 keystrokes while Nokia requires only 1 keystroke.

2.2.5 Summary

In Thai text input method, all of the 4 methods divide characters into 2 groups; one is 44 consonants and the other is other characters including vowel, tone mark, consonant-vowel and special symbol. Normally, if the number of average characters mapped onto a key is small, KSPC is low, but this is not always true as seen from Motorola's method. Motorola tried to spare more buttons for the second group in order to reduce average number of characters per key, but this caused more average number of characters per key in consonant group. Since consonant is a major composition of Thai word, many words are the composition of consonant only. Therefore, average number of characters per key of consonant group should be kept as low as possible. I agree that 8 or 9 buttons should be assigned for consonant as Nokia, Samsung and SE did, but some obsolete consonants and rarely use consonants can be omitted to reduce the average number of characters per key.

As only 2 or 3 buttons are left for 23 characters of the second group, it is impossible to reduce the average number of characters per key for the second group, but there are still other keys available. I propose that arrow keys are used to reduce KSPC of Thai text input method by using Positional Prediction [38] [39].

2.3 Comparison on Myanmar Mobile Phone Keypad Layout

Today, a mobile phone is popular and SMS is also available in Myanmar, however, text input method for Myanmar language in a mobile phone is still a challenge research topic. There are only two keypad layouts or text input methods in the Myanmar mobile market, which are MyTap and M9. This section presents the results of my comparison between them.

2.3.1 Keypad Layout and Text Entry Method for Myanmar

Two keypad layouts of MyTap and M9 are presented as well as their text entry methods for Myanmar. Although those two keypad layouts are designed on the 12-key mobile phone keypad, their keyboard mappings differ largely from each other (see Figure 2-31 and Figure 2-32) but have a common approach for text entry process.

MyTap (Technomation Studios, Myanmar) is the Myanmar text input method run on MySM (Myanmar language Short Messaging System) software [40]. To my knowledge, it is the first text input method that came out as a product in 2006.

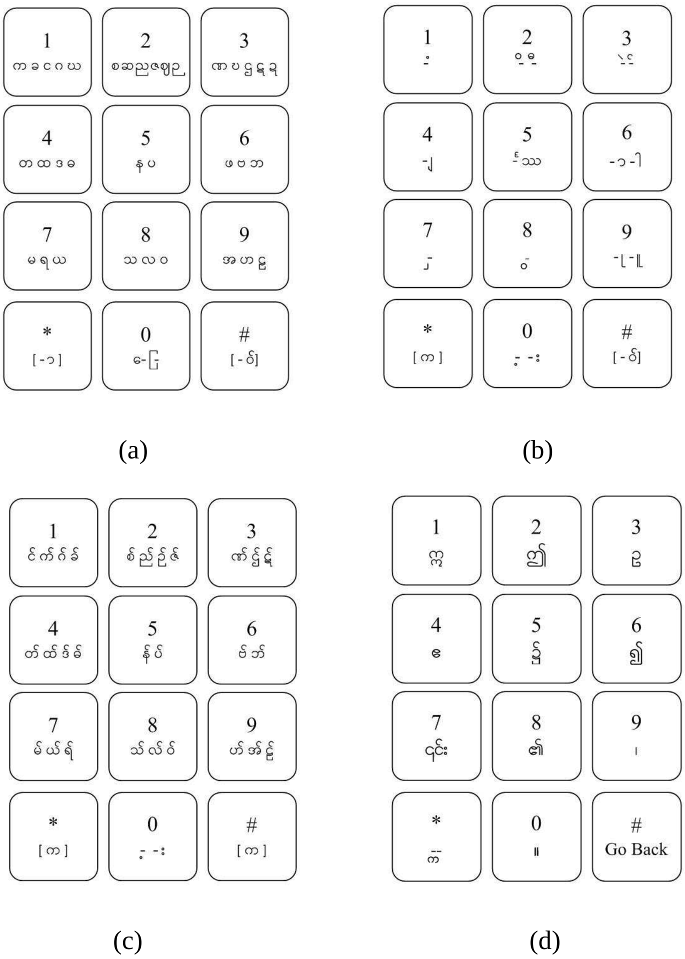


Figure 2-31. MyTap keypad layout

(a) consonant mode (b) vowel mode

(c) consonant character with Asat mode and (d) independent vowel mode

Image taken from my published paper [125]

MyTap keyboard mapping basically has three layers (1) consonant character mapping (က mode), (2) vowel character mapping (◌ာ mode) and (3) consonant character with Asat mapping (◌် mode) (see Figure 2-31). If keyboard mapping for independent vowels is counted as one, there exist four keypad layouts in total (see (d) of Figure 2-31). Consonant character mapping is mainly to map the Myanmar consonants "က" (Ka) to "အ" (A) to key 1 to key 9, and "ေ◌" (E) vowel and "◌ြ" (Ra) Medial to key 0. It does not adopt alphabetical order of consonant (e.g. က, ခ, ဂ, ဃ, င) but use frequency order (e.g. က, ခ, င, ဂ, ဃ) (see (a) of Figure 2-31). Vowel character mapping is based on the writing position or combination of vowels with a consonant, e.g. upper vowels " ံ" (Anusvara), "ိ" (I), "ီ" (Ii), "ဲ" (Ai) and "်" (Asat or killer) are mapped to key 1, key 2 and key 3 (upper row keys on a mobile phone keypad), and lower vowels or medials "ှ" (Ha), "ွ" (Wa), "ု" (U) and "ူ" (Uu) are mapped to key 7, key 8 and key 9 (lower row keys). But character mapping to key 5 is not based on the writing position of vowels because "င်္" (kinzi) is always written on the upper part of a consonant. Consonant character with Asat mapping or (် mode) is for typing the combination of Asat character with a consonant (e.g. န်= န + ်). The keypad layout for Myanmar independent vowels such as "ဣ" (I), "ဤ" (Ii) and "ဥ" (U) and Myanmar punctuation symbols "၊" (little section) and "။" (section) can be seen in (d) of Figure 2-31.

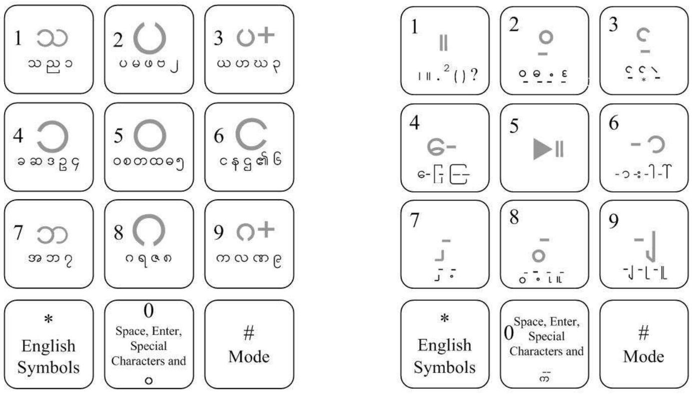
MyTap is a phase predictive text input method, which does not predict Myanmar syllable, word or phrase. It predicts a next typing step or a change from one mode to another especially from vowel mode to consonant mode, and the keyboard layout always pops up. For example, after a vowel "း" (Visarga) is typed, the mode will automatically change to consonant mode. Typing step of MyTap is the same as handwriting order (e.g. left vowel + consonant) and not Unicode typing order (e.g. consonant + left vowel). In MyTap keypad layout, key "*" and "#" are used for changing from one mode to another, but the function varies depending on the active layer. For example, key "*" is used to change exiting mode to consonant mode (က mode), vowel mode (ာ mode) and Paiksint or subscript mode (္က mode). Myanmar numbers can be typed with "Input Number" menu, and English letters can be typed with "Input English" menu. Space can be placed by pressing key 0 longer, and enter can be typed by pressing key "#". The following is the typing steps of "နေကောင်းလား။" (How are you?) with MyTap text entry method, which requires 18 keystrokes for 12 characters.

Table 2-3. Typing steps of Myanmar phrase "နေကောင်းလား။" (How are you?) with MyTap

Table taken from my published paper [125]

ေ	န	ေ	က	mode	ာ	mode	င်	း	လ	mode	ာ	း	mode	။
0	5	0	1	*	6	#	1	00	88	*	6	00	*	0

M9 (Myanmar9) is the Myanmar text input method (R & S Software, Myanmar), which is distributed with SM3 (Simple Myanmar Message) software. As far as I am concerned, it is the second text input method that came out as a product in 2007. Its keypad layout is based on the glyph or shape of Myanmar characters as follows:



(a) (b)

Figure 2-32. M9 keypad layout

(a) consonant mode (b) vowel mode

Image taken from my published paper [125]

Most Myanmar character glyphs or shapes consist of various combinations of circle or Landolt C structure such as consonants "က" (Ka), "ခ" (Kha), "ဂ" (Ga), "ဃ" (Gha) and "င" (Nga), vowels "ာ" (Aa), "း" (Visarga), "ိ" (I) and "ီ" (Ii), independent vowels "ဉ" (U) and "ဦ" (Uu), and numbers "၁" (1), "၃" (3), "၄" (4), "၈" (8) and "၀" (0) [6] [41]. And thus, they can be grouped as circle shaped characters (e.g. "ဝ", "စ" and "တ"), left gap shaped characters (e.g. "ခ", "ဆ" and "ဒ"), right gap shaped characters (e.g. "င", "န" and "ဋ"), bottom gap shaped characters (e.g. "ဂ", "ရ" and "ဇ") and up gap shaped characters (e.g. "ပ", "ဖ" and "ဗ"). It has been found that M9 keypad layout for Myanmar consonants are mapped based on the glyph or shape (see (a) of Figure 2-32), and its vowel mapping is based on the writing position or combination of vowels with a consonant (see (b) of Figure 2-32). English symbols such as "!", ":", ";", ",", "?" and "(" and ")" are assigned to key "*", and mode can be changed by pressing button "#".

M9 is also a phase predictive text input method like MyTap. The main difference is that users can turn on or off this feature (i.e. i-mode=on or i-mode=off option in the M9 setting). The system predicts a next possible phase with "i-mode=on" (i-mode is assumed as interactive or intelligent mode) setting, and it does not work with "i-mode=off" setting. Users can control the mode manually with "i-mode=off" setting (e.g. switching from consonant to vowel mode). There are five modes in total, which are "ကခဂ" (Myanmar consonant), "ေ, း, ္" (Myanmar vowel), "ABC" (English capital letter), "abc" (English small letter) and "123" (English number) mode. Key 5 works as "go back button" to Myanmar consonant mode while vowel mode is active. button Myanmar numbers 0 to 9 can be typed by pressing keys 0 to 9 longer (e.g.

Myanmar number 6 "၆" can be typed by pressing "key 6" longer). Paiksint or subscript characters can be typed by pressing key 0 while vowel mode is active.

"I-mode=on" setting works as Unicode typing order (e.g. consonant + left vowel) and timing on, and "i-mode=off" setting is handwriting order (e.g. left vowel + consonant) and timing off. The following is the typing steps of "နေကောင်းလား။" (How are you?) with M9 text entry method, which requires 18 keystrokes for 12 characters. One of my findings here is that it is difficult for novice users to type with 18 keystrokes due to time limitation.

Table 2-4. Typing steps of Myanmar phrase "နေကောင်းလား။ " (How are you?) with M9 (i-mode = on)

Table taken from my published paper [125]

န	◌ေ	က	◌ေ	◌ာ	mode	င်	◌း	လ	◌ာ	◌း	။
66	4	9	4	6	33	6	66	99	6	66	11

2.3.2 Methodology

Here, the information of participants, apparatus and procedures is presented for text input user study with MyTap and M9 text input methods.

Participants: 5 volunteer participants (3 males and 2 females) were recruited in Yangon city, Myanmar. Participants ranged from 24 to 34 years (mean = 29.8, sd = 3.4). One of them was familiar with a mobile phone and had prior experiences of Myanmar text typing with MyTap and M9, and for the rest of them, it was the first time to type Myanmar text with a mobile phone.

Apparatus: MySM Release 1.9.2 (MyTap version 1.5) and SM3 (free version) were installed on a Nokia mobile phone. A Nokia mobile phone (Model: 3110c) was used for user study with MyTap and M9 text input methods. A Nokia emulator of "Series 40 6th Edition SDK" was also used for counting required keystrokes to type a Myanmar character [42]. Myanmar text for user study consisted of 107 characters including 41 consonants, 52 vowels, 7 numbers, 6 symbols and 1 space, as shown below.

သူငယ်ချင်း၊	(Hi, friend!)
မတွေ့ရတာကြာပြီနော်။	(Long time no see.)
နေကောင်းရဲ့လား။	(How are you?)
ငါ့ဖုန်းနံပါတ်အသစ်က ၅၀၀၇၄၅၉၊	(My new phone number is 5007459.)
အားတဲ့အခါဖုန်းပြန်ဆက်ကွာ။	(Call me back when you have time.)
ဒါပဲနော်။	(Bye for now.)

Procedures: The procedures for user study are (1) explaining the keypad layout and text input method of MyTap and M9, (2) making demonstration of typing Myanmar text with a Nokia mobile phone, (3) allowing practice time to finish each model, (4) recording the participants' typing speed of short Myanmar message (6 sentences) for 10 times (including error correction time) (Note: M9 setting is imode=on and i-speed=6x) and (5) discussing with them and getting their evaluation.

2.3.3 Results and Discussion

Speed (Characters per Minute): Figure 3 shows the users' typing speed to finish short Myanmar message (see section 2.3.2) with MyTap and M9. The speed was calculated based on the Characters per Minute (CPM), which is generally calculated as [characters per second] x 60. Typing speed was evaluated with CPM instead of Words per Minute (WPM) [43]. This is because there is no standard definition for a word in Myanmar like in English (i.e. common definition of a word = 5 characters, including spaces) [44]. In order to know whether there exists significant difference between MyTap and M9, an analysis of variance (ANOVA) was used. The results are not statistically significant ($F_{1,4} = 0.144$, ns) (see Figure 2-33).

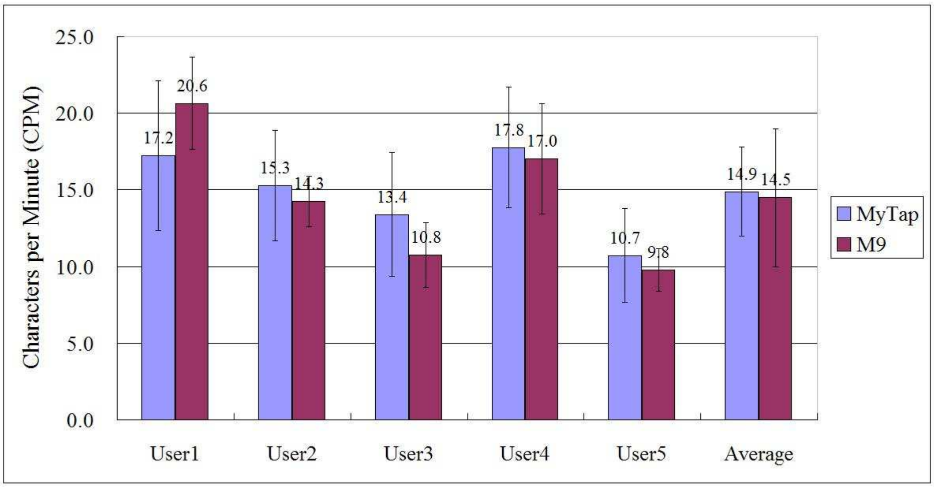


Figure 2-33. Characters per Minute (CPM) of five users with MyTap and M9

Image taken from my published paper [125]

Keystrokes: Keystrokes per Character (KSPC) is the number of keystrokes or stylus strokes required, on average, to generate each character of text using a given interaction technique in a given language [43] [37]. Here, however, KSPC formula was not used but keystrokes were counted manually to make comparison of keystroke requirements only (i.e. typing error and text editing process is excluded) to finish Myanmar text. Required keystrokes to type "frequently used Myanmar syllables" (see Table 2-5), "Myanmar words formed by two identical syllables" (see Table 2-6), "Pali and special words" (see Table 2-7) and "Myanmar text for user study" (see section 4.2) were counted.

To count keystrokes, MyTap needs to be counted twice and M9 needs to be counted four times. This is because the number of keystrokes depends on combination methods of consonant and Asat (e.g. "consonant + Asat" method or "(◌်) + consonant" method), "i-mode=on" (i.e. Unicode writing order and mode is changed automatically by timing) or "i-mode=off" (i.e. handwriting order and mode is not controlled by timing). And thus, keystrokes counting was implemented for "consonant + Asat" and "(◌်) + consonant" methods for MyTap and "i-mode=off", "i-mode=on" and "consonant + Asat", "i-mode=on" and "(*◌်) + consonant" methods and "in case of time out". The results clearly showed that M9 required higher number of keystrokes than MyTap in the typing process of Myanmar text (see Table 2-5, Table 2-6 and Table 2-7). User study has presented that all of the participants used "consonant + Asat" method mostly, and "(◌်) + consonant" and "(*◌်) + consonant" methods sometimes.

Table 2-5. Keystroke comparison of frequently used Myanmar syllables

Table taken from my published paper [125]

		MyTap		M9			
No.	Frequently Used Myanmar Syllables	C + Asat	(ှ) + C	i-mode (off)	i-mode (on) C + Asat	8 mode (on) (* ှ) + C	Time Out
1	အဲ	3	3	4	4	4	N/A
2	ကို	4	4	6	5	5	6
3	သည်	7	4	5	5	5	N/A
4	တယ်	7	4	6	6	6	N/A
5	နေ	2	2	5	3	3	N/A
6	ပါ	4	4	5	4	4	N/A
7	တွေ	4	4	8	5	5	6
8	ဖြစ်	7	5	11	8	9	N/A
9	ဆို	5	5	7	6	6	7
10	မှာ	4	4	5	4	4	5
11	တော့	5	5	10	7	7	8 or 9
12	တဲ့	4	4	8	8	8	9
13	တာ	3	3	5	4	4	N/A
14	လူ	5	5	7	6	6	N/A
15	များ	6	6	7	6	6	7 or 8
Total		**70**	**62**	**99**	**81**	**82**	

Many Myanmar words and names are formed by repeating the same character or syllable, e.g. common names "အောင်အောင်" (Aung Aung), "မြင့် မြင့် " (Myint Myint), "ကိုကို " (Ko Ko), etc., and words "ပါပါ" (dad), "မေမေ" (mom), etc. 10 Myanmar common words have been chosen for counting the number of keystrokes required to type with MyTap and M9 (see Table 2-6).

Table 2-6. Keystroke comparison of Myanmar words formed by two identical syllables

Table taken from my published paper [125]

		MyTap		M9			
No.	Myanmar Words Formed by Two identical Syllables	C + Asat	(ု) + C	i-mode (off)	i-mode (on) C + Asat	i-mode (on) (* ့) + C	Time Out
1	မင်းမင်း	18	10	15	14	14	15 or 16
2	ကိုကို	11	11	13	10	10	11 or 12
3	နိုင်နိုင်	22	12	23	16	18	18 or 19 or 20
4	ပါပါ	11	11	11	8	8	N/A
5	မေမေ	4	4	10	6	6	N/A
6	ဖေဖေ	4	4	12	8	8	N/A
7	ဘုန်းဘုန်း	24	18	27	20	22	23 or 24
8	အေးအေး	10	10	13	8	8	9 or 10
9	မြင့်မြင့်	20	12	21	16	18	17 or 18 or 21 or 22
10	အောင်အောင်	22	14	19	10	12	11 or 12 or 15 or 16
Total		**146**	**106**	**164**	**116**	**124**	

Myanmar has borrowed words from Pali and Sanskrit since the advent of contact with those languages. Some words have been borrowed in toto (e.g. "ဒါန", "သီလ", "အနိစ္စ", "ဒုက္ခ", "အနတ္တ"), and others have been adapted to suit Myanmar tongue and ear (e.g. "စိတ္တ" to "စိတ်", "ဓာတု" to "ဓာတ်", "စရိတ" to "စရိုက်", "မန္တြ" to "မန္တန်") [45]. Since many Pali words use subscript consonants, keystroke comparison was made for Pali words such as "တက္ကသိုလ်" (university), "သစ္စာ" (truth, veracity), "မေတ္တာ" (kindness, affection) and "နိဗ္ဗာန်" (Nirvana, heaven) with MyTap and M9 (see Table 2-7).

Table 2-7. Keystroke comparison of Pali or special words

Table taken from my published paper [125]

		MyTap		M9			
No.	Myanmar Words (Pali or Special Words)	C + Asat	(္) + C	i-mode (off)	i-mode (on) C + Asat	i-mode (on) (* ္) + C	Time Out
1	တက္ကသိုလ်	15	12	22	18	19	19 or 21
2	မန္တလေး	11	11	19	18	18	18
3	ယောက်ျား	15	14	14	9	10	11 or 12 or 13 or 14 or 15
4	ဣန္ဒြေ	11	11	19	16	16	17
5	မင်္ဂလာ	11	11	15	13	13	N/A
6	သစ္စာ	7	7	10	10	10	11
7	မေတ္တာ	8	8	17	15	15	15
8	ဥက္ကာ	9	9	15	14	14	N/A
9	ဧရာဝတီ	14	14	20	19	19	N/A
10	နိဗ္ဗာန်	17	14	22	17	18	18 and 20
Total		**118**	**111**	**173**	**149**	**152**	

Table 2-8. Keystroke comparison of Myanmar text for user study

Table taken from my published paper [125]

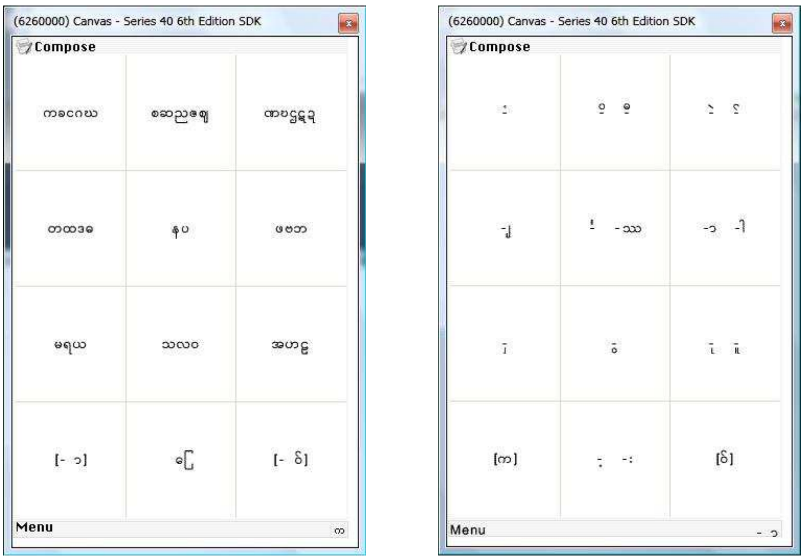
		MyTap		M9		
No.	Myanmar Text for User Study	C + Asat	(◌္) + C	i-mode (off)	i-mode (on) C + Asat	i-mode (on) (* ◌္) + C
1	သူငယ်ချင်း၊	29	21	23	16	17
2	မတွေ့ရတာကြာပြီနော်။	34	34	43	33	33
3	နေကောင်းရဲ့လား။	28	23	33	23	24
4	ငါ့ဖုန်းနံပါတ်အသစ်က ၅၀၀၇၄၅၉၊	51	43	61	48	51
5	အားတဲ့အခါဖုန်းပြန်ဆက်ကွာ။	45	39	59	45	47
6	ဒါပဲနော်။	18	18	23	17	17
Total		**205**	**178**	**242**	**182**	**189**
Total (including Enter key)		**210**	**183**	**252**	**192**	**199**

Questionnaire for Participants: A post-test questionnaire was completed by each participant. In it, MyTap was rated as easier to learn (2:3), quicker (2:3) and less typing mistake (2:3). As for the M9, it was rated as easier to learn (2:3), quicker (3:2) and less typing mistake (1:4).

Their comments or suggestions on MyTap and M9 were also requested and received as follows:

- Prefer MyTap because it is displaying keypad layout
- Both methods are annoying with timing
- M9 keypad layout is creative idea but difficult to memorize
- Prefer M9 with "i-mode=off" setting
- Why consonants are sorted by usage frequency order? Prefer alphabetical order and it is the best for the native people
- Both of the keypad layout are difficult to memorize
- Traditional handwriting order is better than Unicode typing order

Discussion: In user study, "i-mode=on" setting was used for M9 text input system. One reason for this is that MyTap changes mode automatically from vowel mode to consonant mode and it has no timing off function, and another reason is similar platform should be prepared for comparison. From the user study results as well as users' comments and suggestions, it was found that almost all of the users faced with timing problem. The number of required keystrokes can be increased because of time out when user is searching and typing two consonants or more than one vowel continuously. Typing process of M9 with "i-mode=off" is better than that of M9 with "i-mode=on" for the users. On the other hand, "imode=off" setting requires the highest number of keystrokes than other typing methods or options (see Table 2-5, Table 2-6, Table 2-7 and Table 2-8). MyTap keypad layout makes it easier to find desired consonants than M9, because M9 keypad layout is based on the glyph or shape of the consonant, which is difficult to memorize for first-time users. It is assumed that Myanmar consonants keypad layout by alphabetical order is suitable for native users. MyTap has an annoying keypad layout displaying feature for them, because it covers a text editing screen and therefore users cannot see the typing text (see Figure 2-34). Another user interface problem of MyTap is that users need to type number in English first and then convert it to Myanmar number. What MyTap and M9 have in common is that they try to predict next typing steps. For the most participants, typing speed or CPM of MyTap was slightly higher than that of M9 (i-mode =on), but the results of an analysis of variance (ANOVA) are not statistically significant ($F_{1,4} = 0.144$, ns).



(a) (b)

Figure 2-34. Nokia mobile phone emulator screen of MyTap

(a) consonant mode (b) vowel mode

Image taken from my published paper [125]

2.3.4 Summary

Two methods for Myanmar text entry on mobile devices namely MyTap and M9 were evaluated. This research is in progress, and the results of my current analysis have been just reported in terms of number of keystrokes, CPM and feedbacks from users. From this study, it has been found that both MyTap and M9 predict or prepare users' next typing step. Although M9 keypad mapping is interesting, it is difficult for first-time users to memorize. On the other hand, CPM of MyTap is slightly higher than that of M9 (i-mode=on) (14.9 CPM: 14.5 CPM), but MyTap has a complicated keypad layout displaying feature for typing and text editing process. It was also found that most users feel uncomfortable with time out problem. Follow-up analysis is to be made in detail and positional prediction text input interface is to be proposed for a mobile phone in the near future.

2.4 Evaluation of Text Entry Techniques

In chapters 3 to 10 of this dissertation, numerous text entry studies are conducted for Asian syllabic languages. Although different evaluation methods exist, the evaluation of Positional Mapping, Positional Gesture and Positional Prediction reported in this dissertation use the "unconstrained text entry user study". The unconstrained paradigm presents several words or sentences to native users (i.e. subjects) for transcription or copying. Depending on the experiments, physical document or text file is used for presenting words or sentences (see Figure 2-35).

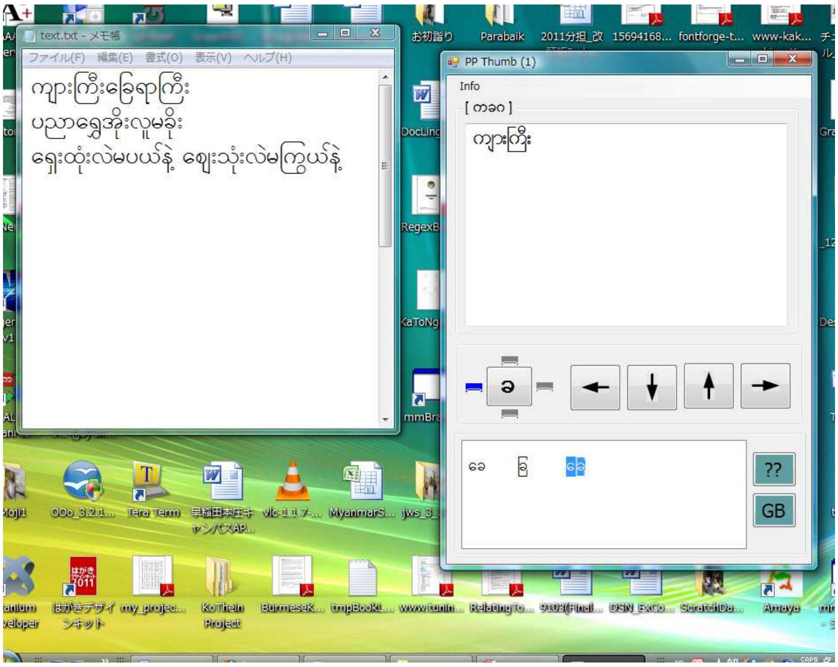


Figure 2-35. Example usages of text file for unconstrained text entry user study

(Here, three Myanmar sentences are presented for typing with text entry prototypes.)

During text typing or transcription, subjects are not allowed to use cursor control keys or a mouse for moving the cursor. In my prototypes, backspace software button is the only mean for error correction. Subjects are encouraged to transcript correctly as fast as they can. Speed and error measurements are done based on the keystrokes logs. Evaluation is made mainly by Characters per Minute, Keystrokes per Character, Likert scale and native users' feedbacks in this dissertation, and these are considered formal and reliable methods for evaluation of my proposed text input prototypes.

2.4.1 Speed

In the unconstrained paradigm, the calculation of typing speed for evaluation is straightforward. The length of the transcribed string *minus one* is used for calculating typing speed in this research [43]. This is because recording time for measuring typing speed begins with the entry of the first character and ends with the final character.

Borrowing MacKenzie's example [43, page no. 49]:

the quick brown fox jumps over the lazy dog (43 characters)

^ ^

t = 0 seconds t = 20 seconds

Words per Minute (WPM): is the unit for measuring typing speed and it is perhaps the most widely used. Common definition for an English word is 5 characters including spaces [44]. Some researchers prefer to report typing rates in Characters per Second (CPS). WPM calculation is as follows:

$$WPM = \frac{|T|-1}{S} \times 60 \times \frac{1}{5} \tag{2.1}$$

$$= \frac{43-1 \text{ characters}}{20 \text{ seconds}} \times 60 \times \frac{1 \text{ word}}{5 \text{ characters}} \tag{2.2}$$

The result is 25.2 WPM for this example. In this dissertation, typing speed was evaluated with *Characters per Minute (CPM)* instead of *Word per Minute (WPM)*. This is because there is no standard definition of word for Asian syllabic languages like in English. Thus, the formula for computing CPM is as follows:

$$CPM = \frac{|T|-1}{S} \times 60 \qquad (2.3)$$

Here, *T* is the typed transcribed string entered by a user, and |T| is the length of this string. *T* may contain characters, numbers, punctuation, spaces, etc. but not backspaces. S is seconds measured from the entry of the first character to the last.

2.4.2 Minimum String Distance (MSD)

Calculation of Minimum String Distance (MSD) is not a straightforward in the unconstrained testing paradigm. For *uncorrected errors* (i.e. errors like typing mistakes remain in the transcribed string) calculation, MSD statistic is used [46]. The MSD error rate formula is as follows:

$$\text{MSD error rate} = \frac{MSD\ (P,T)}{\max(|P|,|T|)} \times 100\% \qquad (2.4)$$

In Equation (2.4), *P* is the presented and *T* is the transcribed strings. The denominator represent the greater length of either *P* or *T*. Note MSD calculation cannot measure which characters are erroneous and just calculate the number of errors. A character-level analysis is needed to find out which characters are in error.

2.4.3 Keystrokes per Character (KSPC) for Error Metric Measure

For corrected errors, Keystrokes per character (KSPC) dependent measure is used [43]. The formula is a simple ratio of the number of entered characters to the number of characters in the final transcribed string.

$$\text{KSPC} = \frac{Keystrokes}{|T|} \qquad (2.5)$$

For example:

Keystrokes: thw<e quick brr<owm<n (21 keystrokes)

Transcribed: the quick brown (15 characters)

KSPC for error metric measure is 21/15 = 1.27

Note: KSPC for error metric measure is not the same as the KSPC characteristic measure [43]. Characteristic KSPC is a theoretical model of keystrokes per character.

2.4.4 Likert Scale for questionnaire

The Likert Scale is a popular format of questionnaire that is used in various kinds of surveys [47]. It is an ordered, one-dimensional scale from which respondents choose one option that best aligns with their view. There are typically between four and seven options. Five is very common and used in this dissertation. The following is an example Likert Scale table used for questionnaire after the typing experiment with a prototype.

Table 2-9. Five-point Likert scale questions on Positional Gesture prototype with various input devices

Likert Scales (range 1-5)	**Trackball**	**Mouse**	**Touchpad**	**Trackpoint**	**Tablet PC**
Difficult-Easy					
Painful-Enjoyable					
Slow-Fast					
Dislike-Like					

CHAPTER 3
Positional Mapping (PM)

3.1 Concept

Positional Mapping is a concept of keyboard or keypad mapping for mobile devices based on characters writing position of Asian syllabic languages. The following shows the application of the concept of Positional Mapping on a mobile phone keypad.

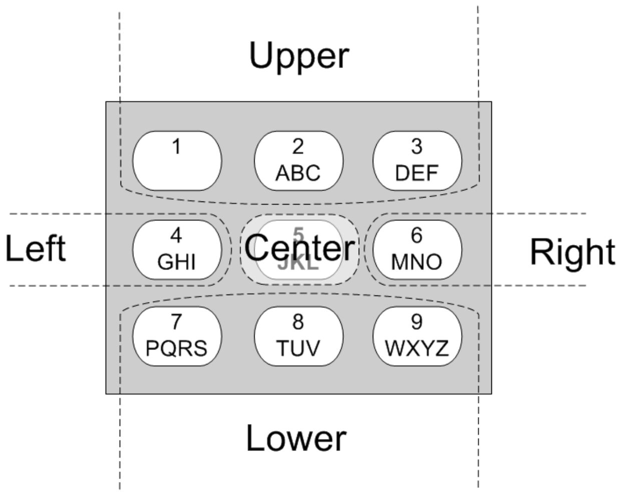


Figure 3-1. Positional Mapping for mobile phone keypad
Image taken from my published papers [126, 127, 128]

3.2 Existing Keypad Layouts for Mobile Phone

This section briefly discusses some of the existing keypad layouts or text input interfaces for a mobile phone.

3.2.1 Standard Keypad Layout

Standard 12-key keypad layout with Multi-tap input method is traditional and still popular because of its simplicity.

Figure 3-2. Standard ISO 12-key keypad layout
Image taken from my published paper [95]

In this input method, users press each key one or more times to get the desired letter. For example, the 3 key is pressed once for the letter 'd', twice for 'e' and three times for 'f'. In Multi-tap, users need to break (by waiting timeout or pressing timeout kill button) the key rotation sequence to type more than one letter from the same key consecutively. One distinct drawback of Multi-tap is the Keystrokes per Character (KSPC) [37], which increase based on the number of letters assignment on a key.

3.2.2 Less-Tap Keypad Layout

The arrangement of the letters on the Less-Tap [48] keypad depends entirely on letter frequencies in English. And thus, Less-Tap keypad layout is different from Multi-tap in which letters appear upon pressing a key. The idea is to get the most frequently used characters with one keystroke. For example, the letter assignment for number 2 key on Less-Tap keypad is a, c, b because the order of the most usage frequencies of a, b, c in English is "a, c, b". Here, it can be noticed that the Less-Tap keypad still keeps letters on the same key compared to Multi-tap keypad for compatibility with keypad standards. The keypad layout will change according to letter usage frequencies for other languages (e.g. German, Spanish and French). However, the big challenge is that there is no existing usage frequency table for phonetic based scripts (e.g. for Myanmar, Khmer, etc.), and digital format of proper lexicon and corpus are not created yet.

Figure 3-3. Less-Tap keypad layout

Image taken from my published paper [95]

3.2.3 MessagEase Keypad Layout

According to the table of usage frequency of letters in English, "e, t, n, r, o, i, a, s, h" are the most frequently used letters (71%). In the MessagEase, users can type these letters by one keystroke [49]. Letters such as "v, l, x, k, q, u, p, c, b, g, d, j, m, y, w, z and f" require 2 key strokes. For example, to type "c", first press "o" key once then "h" key. Another important consideration of MessagEase keypad design is to create a deterministic predictable entry system without need for any disambiguation (to avoid cognitive load). Two main disadvantages of MessagEase are that it is 1) difficult to memorize the keypad layout, and 2) not speedy because still 2 taps are required to input a single letter like "c, d, g, k, l, m, p, q" etc. And the keypad layout gets very complex by considering phonetic based scripts. This is because phonetic based scripts like Myanmar and Bengali languages have more alphabets than English.

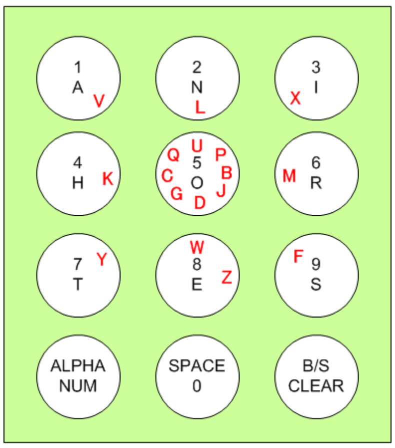


Figure 3-4. MessagEase keypad layout

Image taken from my published paper [95]

3.2.4 SIMKEYS Keypad Layout

SIMKEYS [50] follows an approach that is similar to how the shift key functions on a regular keyboard for assigning the rest of the characters onto the 12-key keypad. The most frequently used characters in English (i.e. A, E, I, L, N, O, R, S, T) called center plane can be typed with only one keystroke in SIMKEYS. And the remaining characters are separated into two groups, each containing the second most and least frequently used characters, respectively. To access the former group (i.e. B, C, D, F, G, H, M, P, U), users press the star key first and then the corresponding number key. Likewise, the latter group (i.e. J, K, Q, V, W, X, Y, Z) is accessed via the pound key. These two groups of letters are referred to as the star plane and pound plane. SIMKEYS concept cannot be applied directly for phonetic scripts because many letter combinations and key shifting are required in their written system.

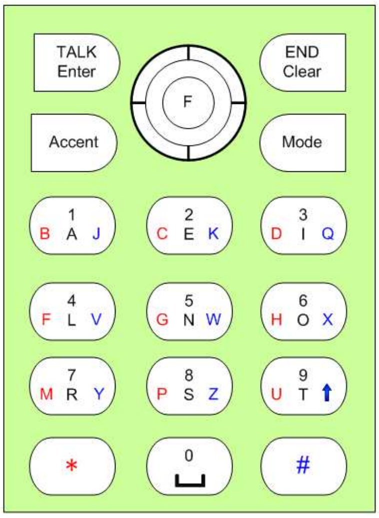


Figure 3-5. SIMKEYS keypad layout
Image taken from my published paper [95]

3.2.5 FasTap Keypad Layout

The idea of FasTap [51] is simple, which uses additional alphabetically ordered 26 buttons (i.e. English letters a to z). And thus, the KSPC = 1 for letters input. The texting is much easier in FasTap compared to traditional 12 buttons Multi-tap keypad. However, the size of the keys for the alphabets is small and inconvenient to use, and this inconveniency slows down the typing speed. Considering FasTap like keypad for phonetic based scripts, more keys are necessary for alphabets (e.g. 33 keys are required only for consonants characters in Myanmar language) and the size of the keypad increases, which is not suitable for mobile devices.

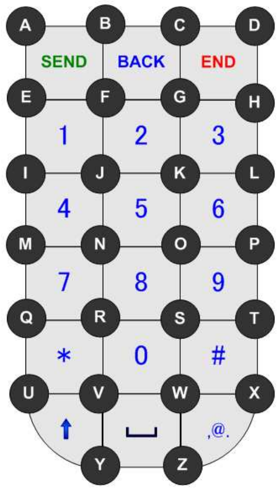


Figure 3-6. FasTap keypad layout

Image taken from my published paper [95]

3.2.6 DeltaII Keypad Layout

DeltaII Keypad [52] is one of the examples of rearranged QWERTY PC based keyboard layout used for a mobile phone keypad. Most PC users who are already familiar with QWERTY can get accustomed to DeltaII Keypad without difficulty.

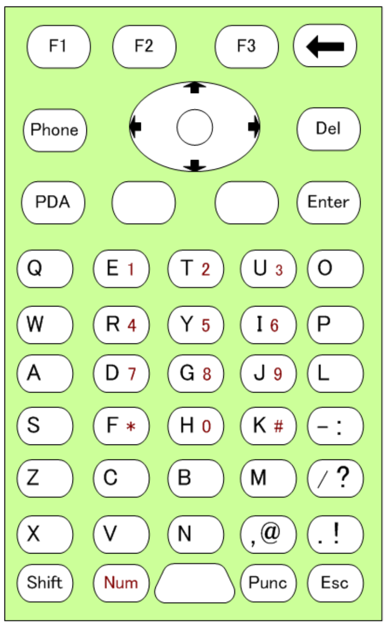


Figure 3-7. DeltaII keypad layout

Image taken from my published paper [95]

The developer of the DeltaII mentions that the average typing speed of DeltaII keypad is 37 WPM (Word per Minute). That is very much faster than the one of normal standard mobile phone keypad, which is 8 WPM. But the QWERTY keyboard was originally designed in two hand architecture, and it is said to slow down the typing speed to prevent type bars jamming in old typewriters. And a mobile phone should be one handed and usable with one thumb. In DeltaII keypad, at least 5x6 matrix keypad is required only for alphanumeric keys (i.e. excluding arrows keys and other keys), which will make users' thumb movements inconvenient.

3.3 Design

Taking common writing nature of Asian syllabic language into consideration, a new concept of characters mapping was developed for a mobile phone keypad. Positional Mapping is a concept of keyboard or keypad mapping for mobile devices based on common characteristics of phonetic scripts languages writing system.

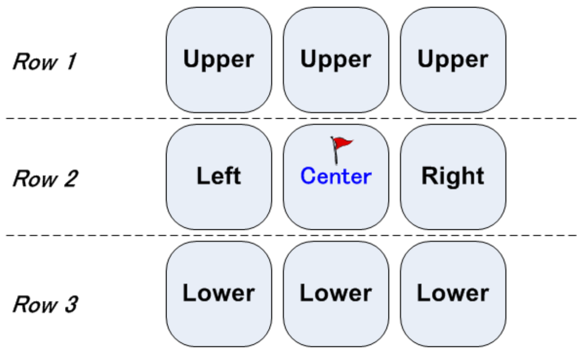


Figure 3-8. Positional Mapping for mobile phone keypad
Image taken from my published paper [95]

A mobile phone keypad is divided into three levels, e.g. 1, 2 and 3 keys are upper, 4, 5 and 6 keys are normal or middle and 7, 8 and 9 keys are lower. In the normal level, 4 key is seen as the front part (Left), 5 key as the center position and 6 key as the rear part (Right). According to the Myanmar language writing method, upper vowels and killer are assigned on 2 and 3 keys, front vowels on 4, main consonant on 5 key, rear vowels on 6 keys and lower vowels and medials on 7, 8 and 9 keys. Here, the consonants will be shown as a list. “Frequently used Myanmar characters and words” such as ‘ပါ’, ‘လား’, ‘၏’, ‘ထို’, ‘ဟို’, ‘ငါ’, ‘သူ’, ‘ကျမ’ and ‘ဒီ’ are assigned on number 1 key. The same concept of key mapping is applied for Bengali language. Figure 3-9 and Figure 3-10 show the detail key mapping for Myanmar language and Bengali language.

This mapping concept is simple, and its keypad layout is very easy to memorize for users who understand the Myanmar and Bengali language writing system. Although PM concept is based on key mapping of a mobile phone, it can be applied not only for a mobile phone but also for other mobile devices such as a PDA, an electronic dictionary and a tablet PC.

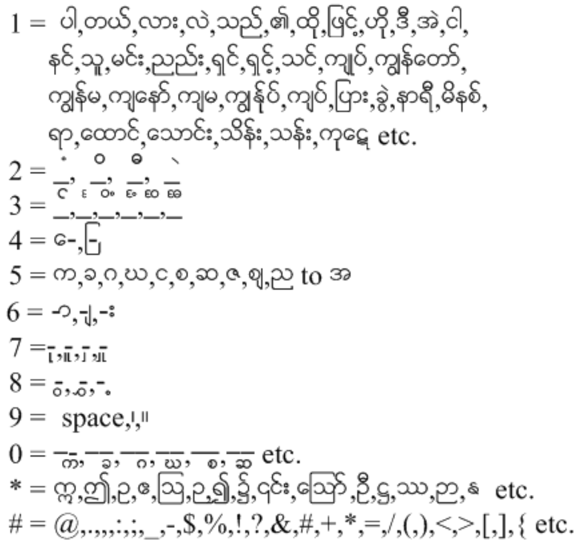


Figure 3-9. PM key mapping for Myanmar language
Image adapted from my published paper [94, 95]

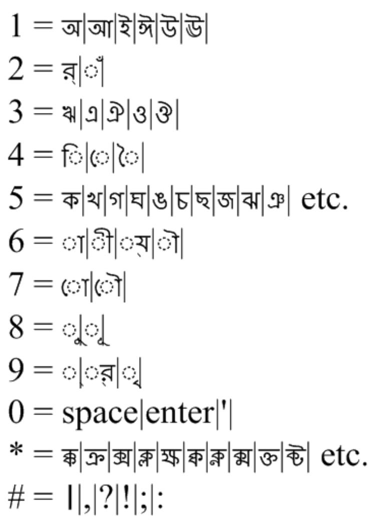


Figure 3-10. PM key mapping for Bengali language
Image taken from my published paper [95]

3.4 Prototyping

PM input prototypes have been implemented for a mobile phone, a PDA, a customizable keyboard Ergodex DX1 and a dual joystick game pad. Here, my 4 prototypes are presented, which are all developed with Microsoft Visual Studio .Net 2003.

3.4.1 Prototype for Mobile Phone

In a mobile phone prototype, key mapping can be changed through configuration file. Configuration file is the text file format that contains the key assignment for each button. By using different configuration file, it is possible to simulate the environment of other text input methods. For user study, four configuration files are used for Myanmar language; 1) Multi-tap with Myanmar consonants ordered by alphabetical, 2) Multi-tap with Myanmar consonants ordered by usage frequency, 3) PM with Myanmar consonants ordered by alphabetical and 4) PM with Myanmar consonants ordered by usage frequency. Note that there is no proper usage frequency table for Myanmar language yet and the usage frequency values used here are the results of my previous experiment data [53]. By using this simulation model with a PC mouse, users can experience how PM works.

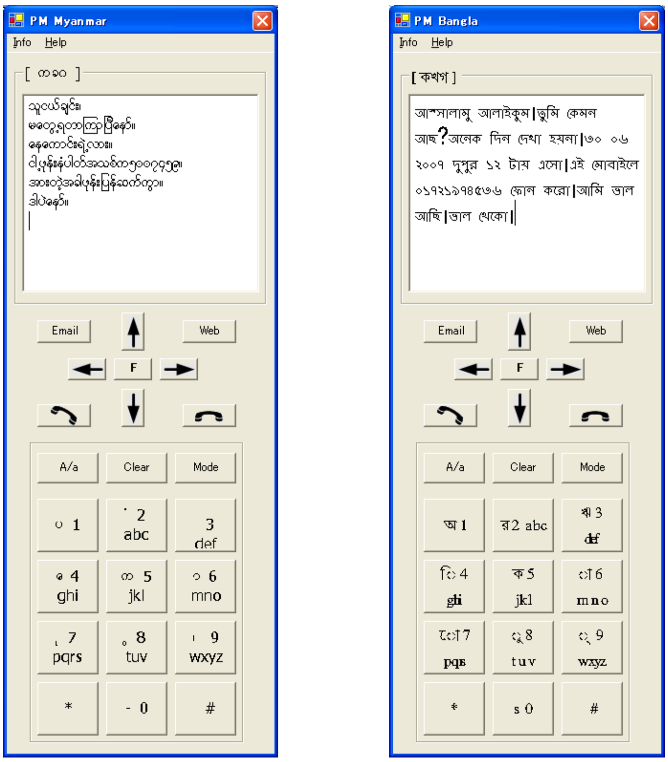


Figure 3-11. PM prototype for mobile phone

(Left: running with Myanmar language configuration file

Right: running with Bengali language configuration file)

Image adapted from my published paper [95]

3.4.2 Prototype for PDA

The second PM prototype is for a PDA, which user experiment was made for Myanmar language. This prototype was created to prove that my PM can be applied not only for key based text input methods but also for pen based text inputs. On a PDA, users can easily type Myanmar language with a stylus. Typing speed goes up compared to previous mobile phone prototypes, since users can directly reach the desired letters shown on the PDA screen. The logic of the letter assignment on a PDA is almost the same as a Myanmar language mobile phone prototype. An exception is that all of the lower medials and vowels are put all together on a button (i.e. ‘ု’ Button) (Fig.3-12). Two keystrokes or taps are required to type a Myanmar character on a PDA. The drawback of this prototype is that users have to go back to home menu screen after typing a character or some characters (i.e. in case of medial characters).

Figure 3-12. Myanmar language PM prototype for PDA
Image taken from my published paper [95]

3.4.3 Prototype for Ergodex DX1 Input System

My third PM prototype is the software interface for text inputting with the Ergodex DX1 input system (Figure 3-13). The Ergodex DX1 input system is a customizable keyboard that allows you to place removable keys anywhere you want and assign the buttons to anything you want. This input interface (Figure 3-14) is developed for experiment with my customized keyboard to figure out users' typing speed improvement.

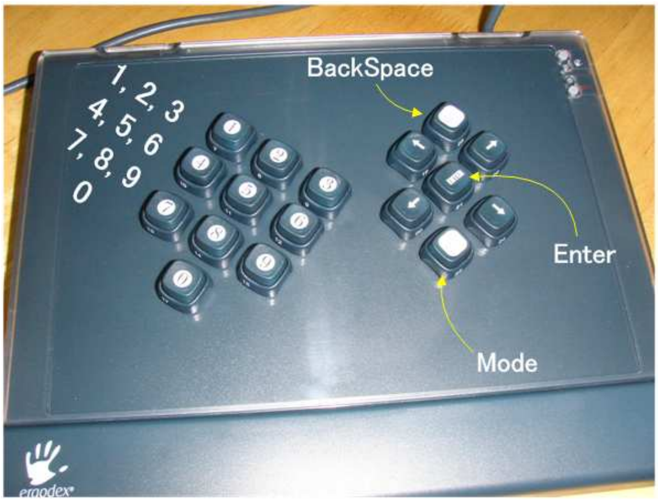


Figure 3-13. Key layout for Ergodex DX1 input system
Image taken from my published paper [95]

Here, 17 keys are used in total (0 to 9 keys, 4 directional arrow keys, 1 enter key, 1 backspace key and 1 mode key) for Myanmar text input. Figure 3-13 shows key arrangement. Here, 0 key is to type zero in Myanmar number mode. Key 1 to 9 are used for PM key mapping for Myanmar language, arrow keys are for moving the highlight of characters, enter key is for typing the current highlighted character and mode key is for changing character, number and symbol modes.

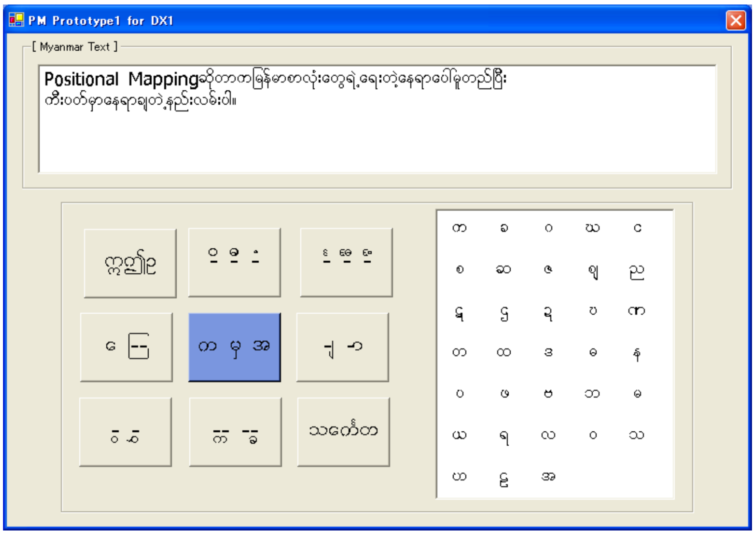


Figure 3-14. Myanmar language PM prototype for DX1
Image taken from my published paper [95]

3.4.4 Prototype for Dual Joystick Game Pad

This prototype is created in order to prove that my PM concept is applicable to various input devices. The software interface for this prototype is almost the same as user interface of DX1 input system (Figure 3-14), but this prototype takes the advantages of using 2 joysticks or thumb sticks (Figure 3-15). Users can move character group highlight with left joystick. Right joystick is used for character highlighting and typing. Some buttons are also used for text editing process like deleting characters.

Figure 3-15. Dual joystick game pad
Image taken from my published paper [95]

3.5 User Study and Evaluation

3.5.1 Subjects, Apparatus and Procedures

To evaluate the performance of PM, user study was conducted with five Myanmar native participants who are between 23 and 31 years old. All of the participants are familiar with a PC but have no experience of text input with a mobile phone and a numeric keypad. All are touch typists in English (with the QWERTY keyboard) but not in Myanmar language. My developed four simulation programs (i.e. a mobile phone, a PDA, the DX1 and a dual joystick game pad) were used for user experiments. No participants had any prior experience with my text input simulation prototypes. The experiment procedures are 1) explaining the concept of PM, 2) explaining the key assignment for all prototypes, 3) making demonstration of text input with each model, 4) allowing 20 minutes practice time for each user to learn text input methods with prototypes (4 multi-tap models of a mobile phone, a PDA, the DX1 and a dual joystick game pad) 5) recording user typing speeds including error correction time and 6) discussing with users and learning their responses. A record was made of the total time spent on typing Myanmar SMS message of 6 sentences (section 2.3.2) for 10 times with each model (Figure 3-19). And 7 days longitudinal study was performed for mobile phone models with 4 types of configuration files, with slowest typing user of 10 times experiments. Figure 3-16 shows the typing speed improvement for each configuration. For the evaluation process of Myanmar language PM prototypes, a soft keyboard "Win Myanmar Visual Keyboard" [27] was selected because this keyboard layout is old and widely used by desk top publishing (DTP) staff in Myanmar. Figure 3-18 shows 2 users' typing speed improvement with "Win Myanmar Visual Keyboard" for 7 days.

User study was also conducted for a Bengali mobile phone prototype with 5 Bengali native participants who are between 26 to 35 years old. All of them are familiar with a PC but only 3 of them are familiar with a mobile phone text input. The experiment procedures are the same as mentioned above, and the results can be seen in Figure 3-17. Figure 3-11 shows the short Bengali SMS message of 7 sentences used for experiment. Famous "AKTEL Bengali multi-tap" keypad layout is used for the evaluation process.

3.5.2 Evaluation of Myanmar language PM

From the graph of slowest user's typing speed improvement for 7days (Figure 3-16), it was found that average typing speeds of my two PM prototypes is higher than the ones of two multi-tap models for all 7 days (i.e. PM is 31% faster than Multi-tap and PM frequency is 29% faster than Multi-tap frequency). And it can be also proven that my PM keypad layout is user-friendly and easy to type even for novice users from the results of day1 (i.e. Multi-tap = 11min 09sec, PM = 6min 31sec, Multi-tap Frq = 10min 42sec, PM Frq=5min 09sec). Comparing the average typing speed among 4 prototypes (Figure 3-19), the fastest average typing speed is 3min 47sec, which is given by the DX1 prototype. Actually, it is difficult to make speed comparison among these 4 prototypes because different software prototypes and input devices (i.e. a mouse, a stylus, the DX1 keyboard, a dual joystick game pad, etc.) are used. For the rest of other 3 prototypes, the average typing speeds are: a PDA = 4min 26sec, a dual joystick game pad = 5min 30sec and a mobile phone = 4min 42 sec. Average typing speed of 2 users for "Win Myanmar Visual Keyboard" is 2min 30sec (Figure 3-18), and thus, it can be said that my PM keypad layouts can give appropriate typing speeds and applicable in the real environment. The fastest average speed given by the DX1 prototypes is of course 51% slower than "Win Myanmar Visual Keyboard", but only 17 keys, much less number compared to the whole keyboard are used.

3.5.3 Evaluation of Bengali language PM

For the Bengali language, an experiment was conducted with a mobile phone simulation prototype. Figure 3-10 shows my proposed PM key mapping for a Bengali mobile phone. The average typing speed of 5 users for PM prototype is 46% faster than the current AKTEL key mapping (Figure 3-17).

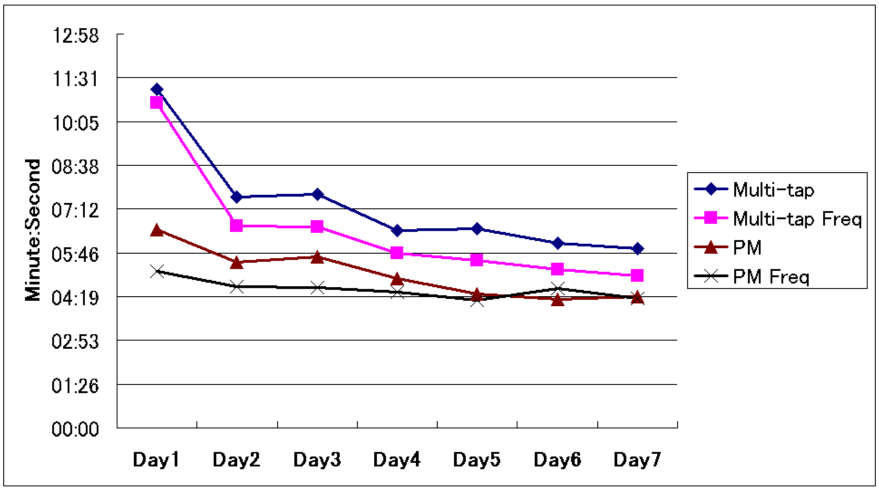


Figure 3-16. Typing speed improvement of slowest user for mobile phone prototype with 4 different configuration files (experiment results for 7 days)
Image taken from my published paper [95]

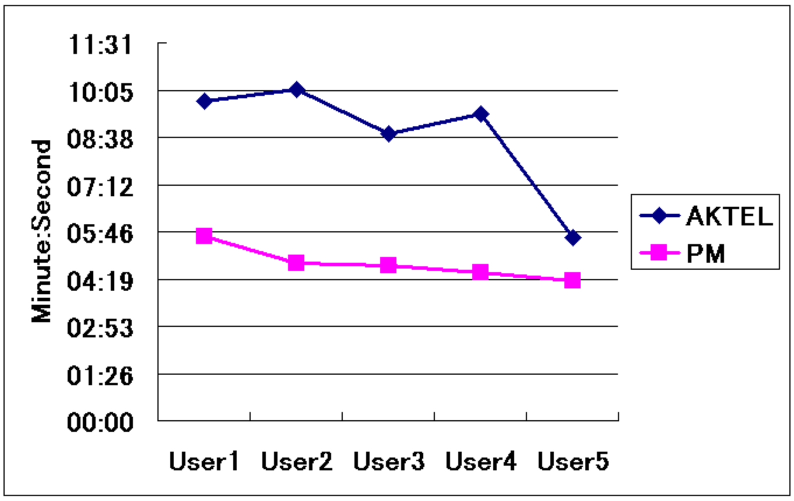


Figure 3-17. Average typing speed of 5 Bengali users for AKTEL and PM prototype
Image taken from my published paper [95]

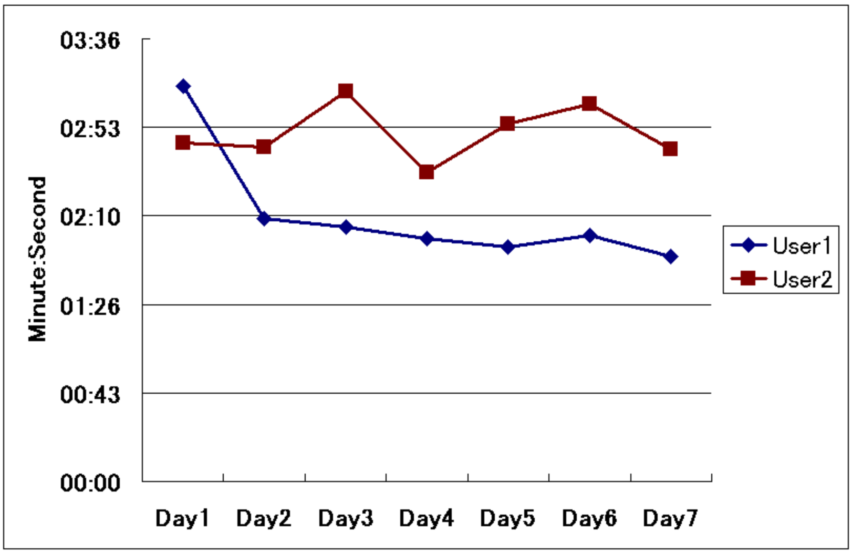


Figure 3-18. Average typing speed of 2 Myanmar users for "Win Myanmar Visual Keyboard" (experiment results for 7 days)

Image taken from my published paper [95]

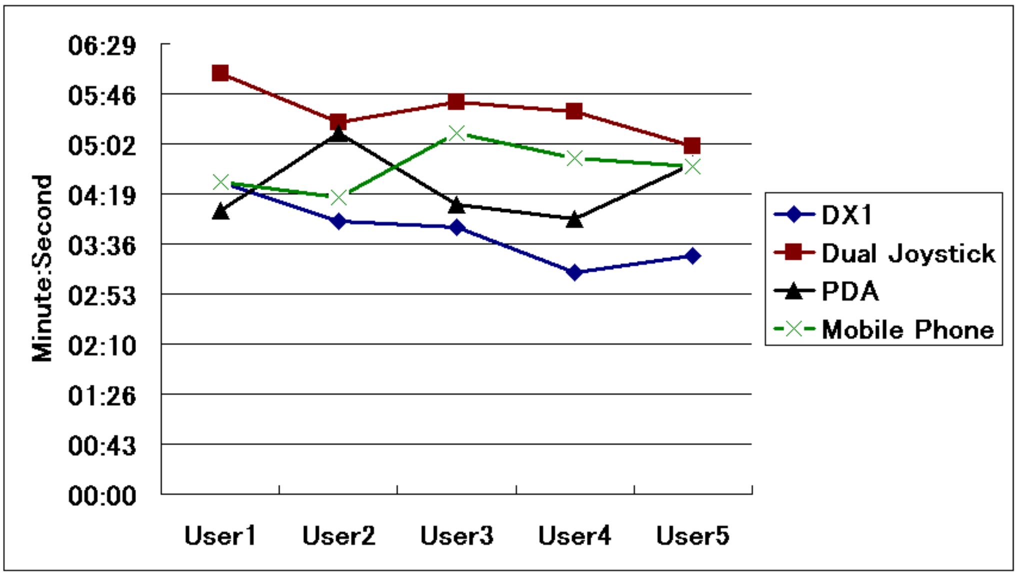


Figure 3-19. Average typing speed of 5 Myanmar users for 4 Positional Mapping prototypes

Image taken from my published paper [95]

3.6 Summary

A new key mapping method was proposed for phonetic scripts based on characters' writing positions. The advantages of this key mapping are 1) compact and easy to memorize the letter assignment on the keys (i.e. short learning curve), 2) applicable for many mobile devices such as a mobile phone, a PDA, a tablet PC and an electronic dictionary, 3) applicable for many similar languages such as Thai, Khmer and Hindi and 4) applicable not only for hard key but also for soft key. On the other hand, there are disadvantages like 1) this key mapping may be difficult to understand for non natives and 2) all of the consonants are assigned on the center key (usually number 5 key in a mobile phone keypad) and this may affect the typing speed (i.e. users have to use some arrow keys movements). However, the typing speed can increase by using PM key mapping with appropriate predictive text input techniques like T9, LetterWise and WordWise. From the experiment results, the PM enables appropriate typing speeds with various input devices, 31% faster than normal multi-tap in Myanmar language and 46% faster than AKTEL key mapping in Bengali language. There is also a plan to conduct user study with real a mobile phone in the near future and study other Asian syllabic languages continuously.

CHAPTER 4
Positional Mapping Software Keyboard

This chapter compares the Positional Mapping software keyboard layout and other existing software keyboard layouts for Bengali and Khmer languages. User study was conducted with native participants.

4.1 Existing Software Keyboards

This section presents the existing software keyboards for Bengali and Khmer used for my user study.

4.1.1 Acharya Software Keyboard

Acharya multilingual editor for Indian languages is part of the open source IMLI (IIT Madras Language Initiative) project [54]. It supports not only Bengali but also Asamiya (Assamese), Hindi, Gujarati, Marathi, Oriya, Kannada, Tamil and Telugu. There are two keyboard layouts (InScript and Phonetic) for each language, and Bengali phonetic keyboard layout was used for user study (see Figure.4-1).

Figure 4-1. Acharya keyboard layout (version.1.0)
Image taken from my published paper [69]

4.1.2 Akkhor Bengali Software Keyboard

The Akkhor Bengali software keyboard is one of the built-in facilities of Akkhor word processor. It was developed by "Khan Md. Anwarus Salam" and released on April 14, 2004. Akkhor word processor supports both ASCII and Unicode. The keyboard layout exists on the left side of the program in alphabetical order (see Figure 4-2). Version 2 was used for making a comparison with the Positional Mapping software keyboard layout.

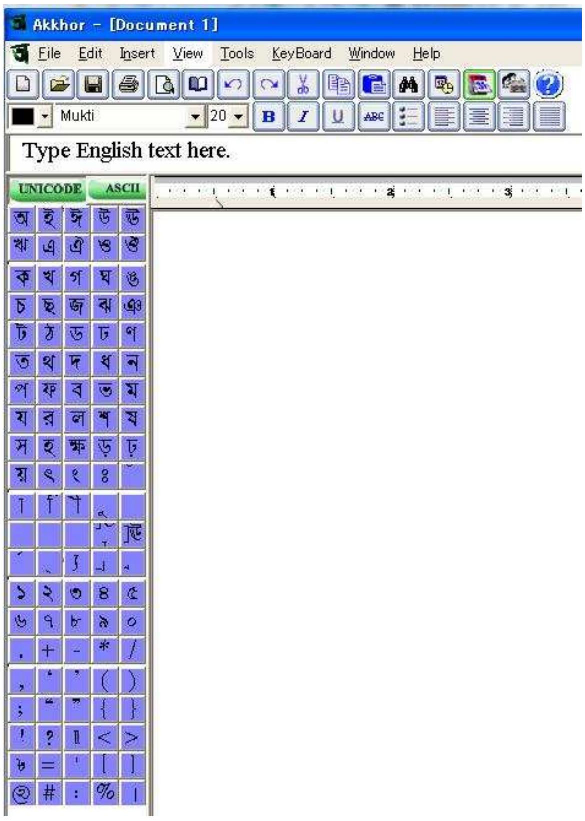


Figure 4-2. Akkhor word processor (version 2.0)
Image taken from my published paper [69]

4.1.3 NiDA Khmer Software Keyboard

A Khmer software keyboard is developed by modeling the standard NiDA keyboard (National Information Technology Development Authority) [55]. It was created just after Khmer Unicode 4.0 was released. Figure 4-3 shows the NiDA software keyboard layout.

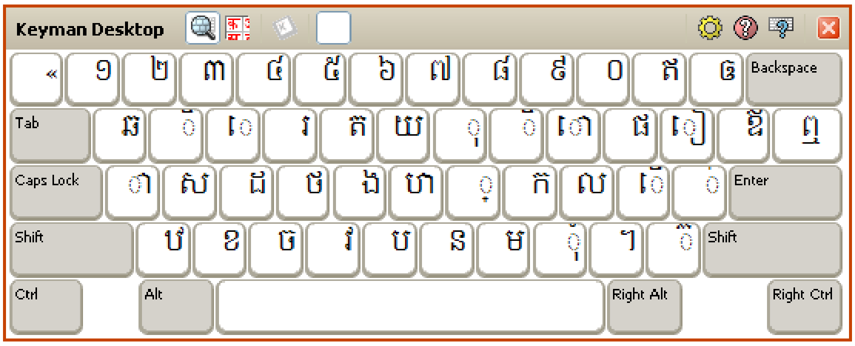


Figure 4-3 NiDA software keyboard layout (unshifted mode)
Image taken from my published paper [69]

The Khmer Unicode keyboard layout differs from the one of old version which is not in Unicode in which the subscripts of the consonants are not spread on the keyboard anymore. Instead, a subscript sign is used to indicate that the next consonant is subscript of the cluster, and that the typing order is not from left to right in the same order of hand writing, but in the order of word's spelling.

4.2 Prototyping

Positional Mapping software keyboard prototypes were implemented for user experiments. My 2 prototypes (for Bengali and Khmer languages) are presented here, which are both developed with Microsoft Visual Studio .Net 2003. These prototypes use Unicode, and users have to type following Unicode typing order. There is a plan to support hand writing order in the near future. “Mukti” font is used for Bengali, and “Khmer OS” font is used for Khmer.

4.2.1 Prototype for Bengali

In my Bengali Positional Mapping software keyboard, all Bengali characters are divided into nine groups according to their writing natures. They are consonant characters, left characters, right characters, upper characters, lower characters, left-right characters, independent vowels, ligatures and numbers. For easier searching, consonants are put in the center surrounded by left, right, upper, lower and left-right characters. 20 frequently-used ligatures are selected and put onto the keyboard for easier typing. Other ligatures can be typed by choosing “ligature menu” from the drop down list. Figure 4-4 shows prototype for Bengali.

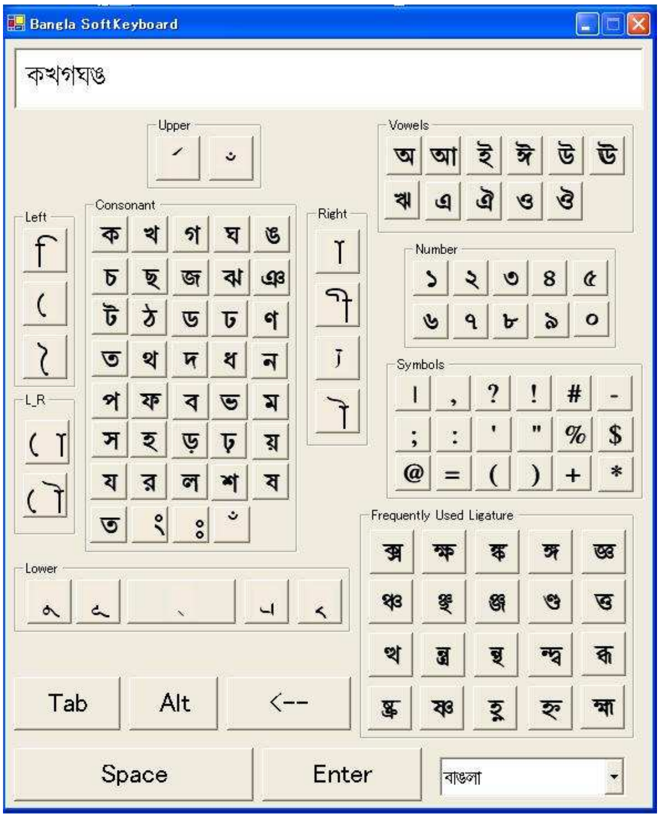


Figure 4-4. Positional Mapping software keyboard prototype for Bengali

Image taken from my published paper [69]

4.2.2 Prototype for Khmer

Key mapping concept for Khmer is the same as in Bengali. Consonants exist in the center surrounded by left, right, upper, lower and left-right characters.

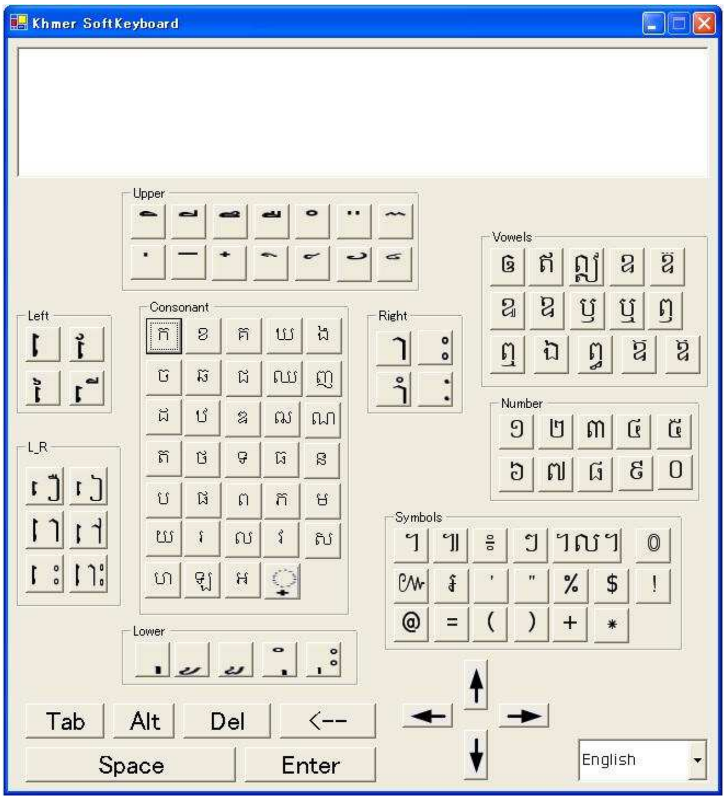


Figure 4-5. Positional Mapping software keyboard prototype for Khmer
Image taken from my published paper [69]

4 directional keys are put for easier editing process. For the subscript characters, users can type using Coeng character “្” button. Figure 4-5 shows prototype for Khmer.

4.3 User Study

To evaluate the performance of the Positional Mapping software keyboard, user study was conducted with two Bangladesh participants and a Cambodian participant. Their ages range from 23 to 32. They are all familiar with a PC, but none of them had any prior experience with my Positional Mapping software keyboard prototypes. Two of them are touch typists in English (with the QWERTY keyboard) but not in Bengali or Khmer language. One Bangladesh participant started to use a computer just recently.

User study was conducted in two of my university laboratories using a 1.06GHz Intel® Core™2 Duo CPU L7500 IBM Lenovo X61 tablet PC running Windows Vista™ Ultimate OS with 2006 MB RAM. Screen size is 12.1 inch and set to 1400X1050 resolution and 32 bit color. Users were required to use a stylus pen for text input experiments with a tablet PC. The

SMART electronic whiteboard SB660 64 inch was used for text entry experiments of the Positional Mapping keyboard with big screen environment [56]. The experiment procedures are as follows:

1) explaining the concept of Positional Mapping,
2) explaining the keyboard mapping of prototype,
3) making demonstration of text input with prototype,
4) allowing ten minutes practice time for each user to learn text input methods with prototypes (five minutes for a tablet PC and another five minutes for an electronic whiteboard),
5) recording user typing speeds on a tablet PC including error correction time,
6) recording user typing speeds on an electronic whiteboard including error correction time and
7) discussing with users and learning their responses.

A record was made of the total time spent on typing seven Bengali sentences (see Figure 4-6) and five Khmer sentences (see Figure 4-7) for 5 times.

আস্সালামু আলাইকুম।
তুমি কেমন আছ?
অনেক দিন দেখা হয়না।
৩০ ০৬ ২০০৭ দুপুর ১২ টায় এসো।
এই মোবাইলে ০১৭২১৯৭৪৫৩৬ ফোন করো।
আমি ভাল আছি।
ভাল থেকো।

Figure 4-6. Seven Bengali sentences for user study
Image taken from my published paper [69]

សួស្តី!

សម្លាញ់ឯងដឹងទេ គ្នាបានក្លាយជាបុគ្គលិកពេញសិទ្ធិហើយ។

គ្នាសប្បាយចិត្តខ្លាំងណាស់។

ថ្ងៃទី២០ ខែ៨នេះគ្នានឹងបានចូលធ្វើការហើយ។

ជួបគ្នាថ្ងៃក្រោយ

Figure 4-7 Five Khmer sentences for user study
Image taken from my published paper [69]

In the text entry experiment with Khmer prototype on a tablet PC, users were requested to type not only with a stylus pen but also with a mouse. This was for the purpose of clarifying the typing speed difference between with a stylus pen and with a mouse.

From the user study, it was found that all the three participants put a tablet PC on the table while they were typing (see Figure 4-8). And they faced difficulties with typing on an electronic whiteboard when their body shade covers the software keyboard. This is because front projection type of an electronic whiteboard was used (see Figure 4-9). This problem can be solved by using back side projection type of an electronic whiteboard instead.

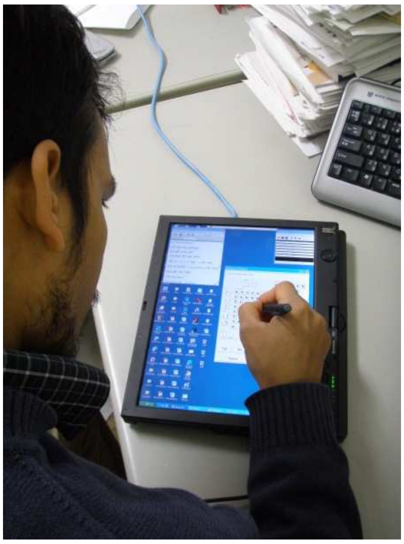


Figure 4-8 User study for Bengali with tablet PC
Image taken from my published paper [69]

Figure 4-9 User study for Khmer with electronic whiteboard
Image taken from my published paper [69]

4.4 Evaluation

Users' actual typing speed was recorded and CPM (Characters per Minute) was calculated for the evaluation process of my Positional Mapping keyboard layout. Small questionnaires were also conducted in order to figure out their feedbacks.

4.4.1 Evaluation by Users' Typing Speed

The average typing speed with the Positional Mapping software keyboard layout to finish seven Bengali sentences (see Figure 4-6) is 3 minutes 13 seconds on a tablet PC (see Figure 4-10) and 2 minutes 44 seconds on an electronic whiteboard (see Figure 4-11). The one with the Acharya software keyboard is 4 minutes 33 seconds on a tablet PC and 5 minutes 5 seconds on an electronic whiteboard. For the Akkhor software keyboard, the average typing speed on a tablet PC is 5 minutes 4 seconds and 4 minutes 43 seconds on an electronic whiteboard.

The average typing speed with the Positional Mapping software keyboard layout (running on a tablet PC) to finish five Khmer sentences (see Figure 4-7) is 2 minutes 34 seconds with a mouse and 3 minutes 54 seconds with a stylus pen (see Figure 4-13). And the one on the SMART electronic whiteboard is 3 minutes 16 seconds (see Figure 4-13). For the NiDA software keyboard layout, the average typing speed of five native users with a mouse is 4 minutes and 38 seconds. This value is taken from my previous work [57].

In general, all of the users' typing speeds increased during typing experiments with the Positional Mapping keyboard layout (see Figure 4-10, Figure 4-11 and Figure 4-13). It turned out faster with the Positional Mapping keyboard layout than with existing keyboards for both Bengali and Khmer. Detail comparison is shown in Table 4-1 and Table 4-2.

Table 4-1. Typing speed for Bengali

Table taken from my published paper [69]

Keyboard Layout	**With Tablet PC**	**with Whiteboard**
Positional Mapping	3 min 13 sec	2 min 44 sec
Acharya	4 min 33 sec	5 min 5 sec
Akkhor	5 min 4 sec	4 min 43 sec

Table 4-2. Typing speed for Khmer

Table taken from my published paper [69]

Keyboard Layout	**with Tablet PC**	**with Whiteboard**
Positional Mapping	3 min 54 sec (with stylus) 2 min 34 sec (with mouse)	3 min 16 sec
NiDA	4 min 38 sec	

4.4.2 Evaluation by Characters per Minute

Although Word per Minute (WPM) is the most widely-used measure of the text entry performance, my Positional Mapping keyboard layout is evaluated with Character per Minute (CPM), because it is not possible to find common definition of "word" in Bengali and Khmer language. Importantly, the WPM/CPM measure does not consider the number of keystrokes or gestures made during entry, but only the length of the resulting transcribed string and how long it takes to produce it [43]. The user study results indicate that the average CPM for Positional Mapping, Acharya, Akkhor and NiDA software keyboard is as follows respectively:

For Bengali:

Positional Mapping on a tablet PC: 42.59 CPM

Positional Mapping on an electronic whiteboard: 50.12 CPM

Acharya on a tablet PC: 30.11 CPM

Acharya on an electronic whiteboard: 26.95 CPM

Akkhor on a tablet PC: 27.04 CPM

Akkhor on an electronic whiteboard: 29.05 CPM

For Khmer:

Positional Mapping on A tablet PC with a stylus pen: 34.62 CPM

Positional Mapping on a tablet PC with a mouse: 52.60 CPM

NiDA: 29.14 CPM

CPM of the Positional Mapping keyboard layout is better than that of other keyboard layouts.

4.4.3 Evaluation by Users' Likert Scale

Questionnaires were conducted for users immediately after typing experiments. Four Likert Scales (1-5) were set on which to rate the Positional Mapping keyboard layout (with a tablet PC and an electronic whiteboard). Table 4-3 shows the average results of Likert scale questions. Labels for scale endpoints are in the most left columns and higher values are better.

From the questionnaire results with three participants, it can be noticed that most users preferred text entering with a tablet PC than with an electronic whiteboard. One of the reasons might be that they used an electronic whiteboard for the first time in the experiments. The interesting point is that although Likert scales for an electronic whiteboard responded by the three users are lower than those of a tablet PC (see Table 4.3), the average typing speed is nearly the same.

Table 4-3. Users' evaluation of Positional Mapping keyboard layout with tablet PC and whiteboard
Table taken from my published paper [69]

Likert Scales (range 1-5)	**Positional Mapping Tablet PC**	**Positional Mapping Whiteboard**
Difficult-Easy	3.5	3
Painful-Enjoyable	4.0	3.5
Slow-Fast	4.5	4
Dislike-like	4.5	4.5
Average	4.13	3.75

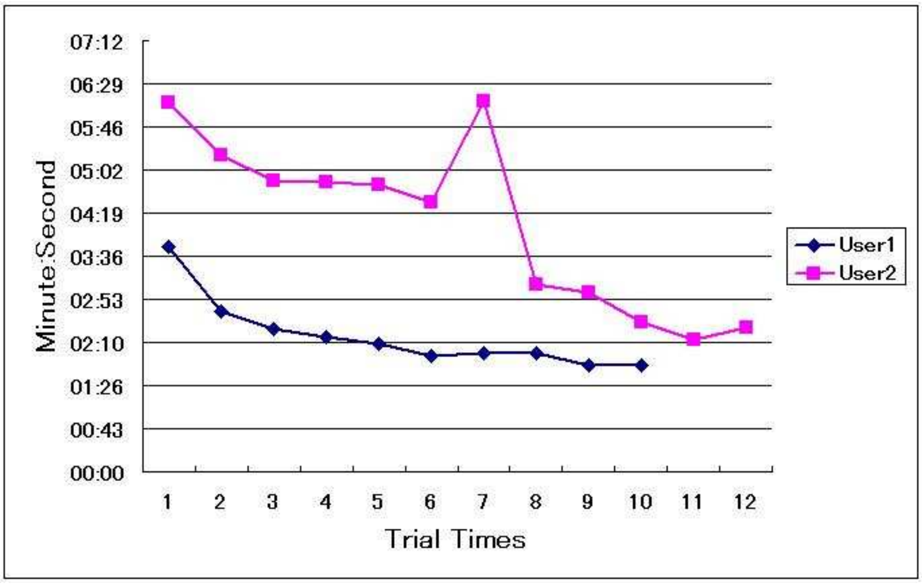


Figure 4-10. Typing speed of two Bengali users with Positional Mapping software keyboard running on tablet PC
Image taken from my published paper [69]

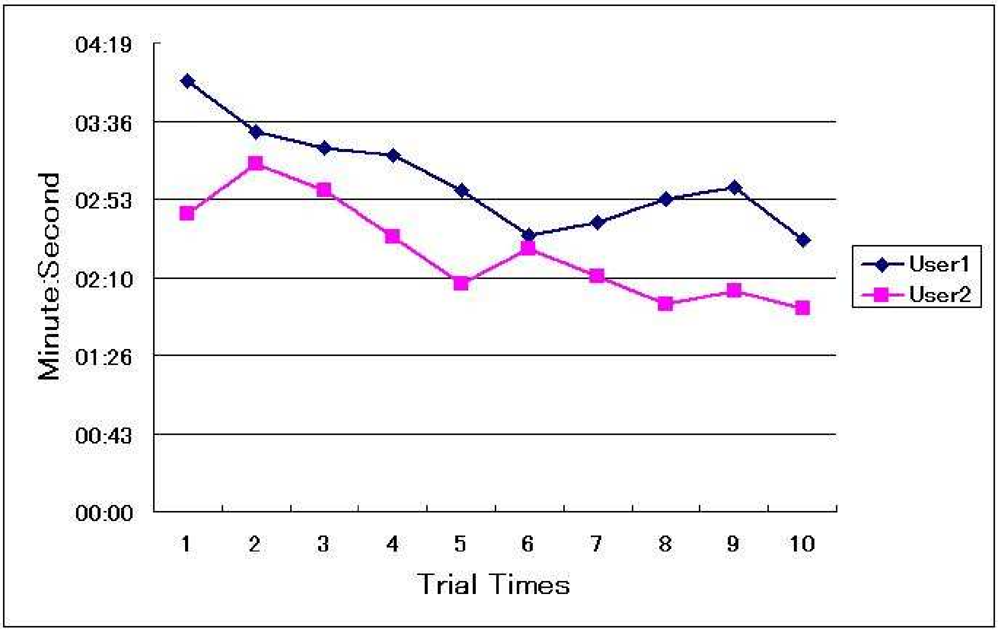


Figure 4-11. Typing speed of two Bengali users with Positional Mapping software keyboard running on SMART electronic whiteboard

Image taken from my published paper [69]

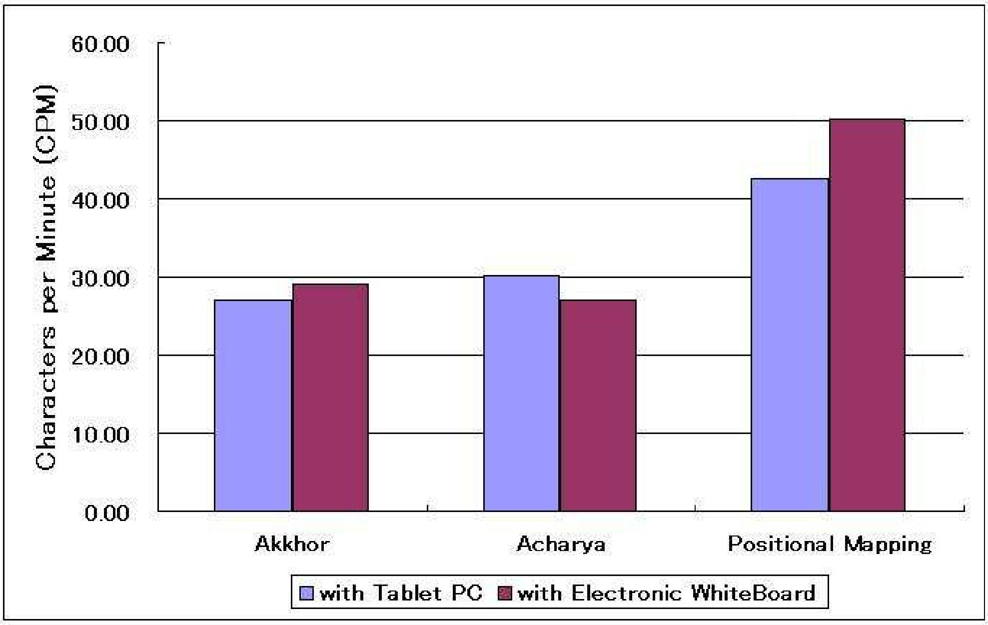


Figure 4-12. Characters per Minute (CPM) comparison for Akkhor, Acharya and Positional Mapping

Image taken from my published paper [69]

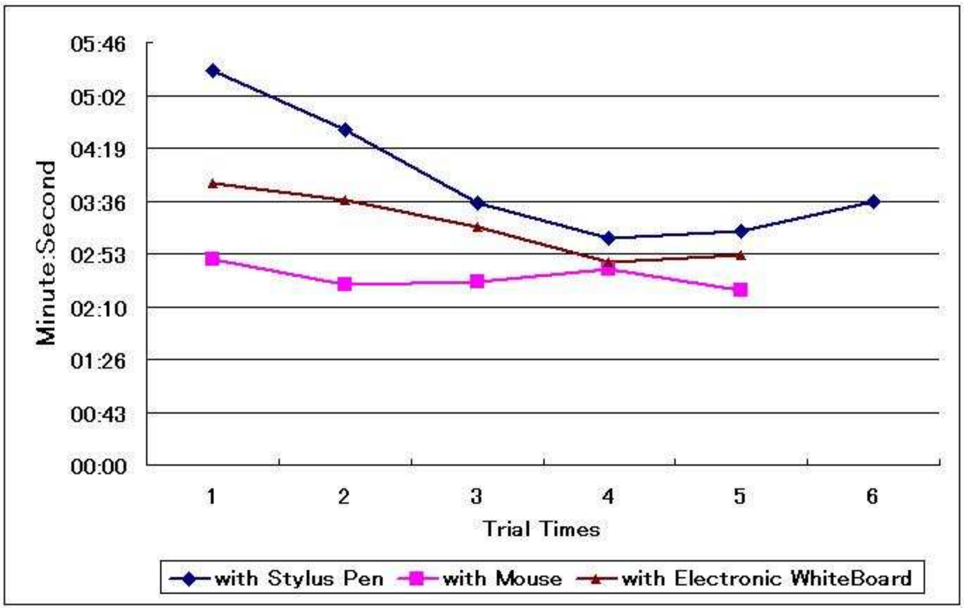


Figure 4-13. Typing speed of Cambodian users for Positional Mapping software keyboard
Image taken from my published paper [69]

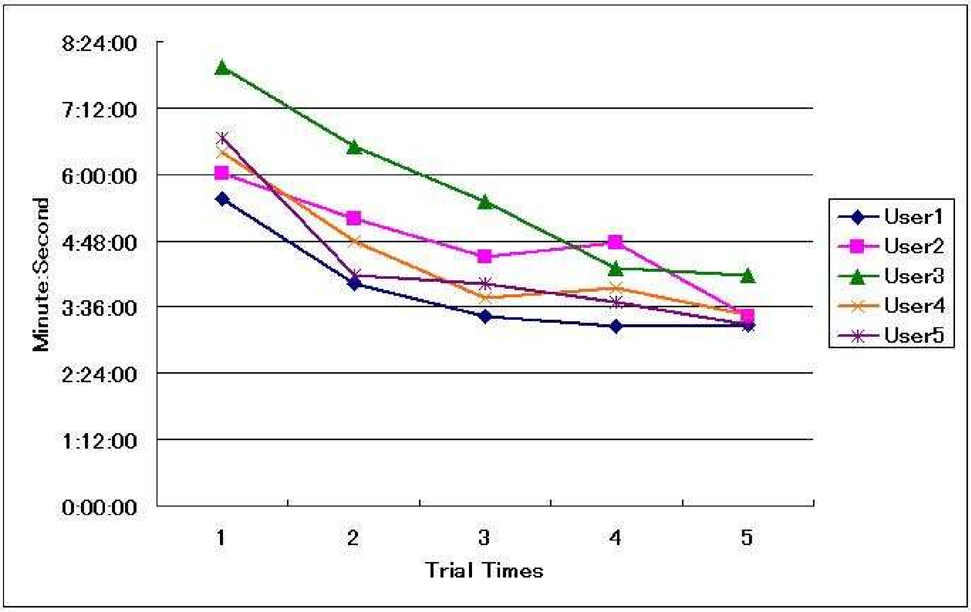


Figure 4-14. Typing speed of five Cambodian users for NiDA software keyboard layout
Image taken from my published paper [69]

4.5 Summary

The idea of new keyboard layout was proposed for Bengali and Khmer. From the users' typing speed and Likert scales questionnaires results, it is proven that the Positional Mapping keyboard layout is more user-friendly than the current keyboard layouts. Further refinements are to be made on the current prototypes, and an analysis is to be followed up on like error rate comparison. There is also a plan to apply this concept for similar syllabic scripts such as Myanmar and Hindi in the near future.

CHAPTER 5
Positional Mapping and Other Layouts for Myanmar

This chapter mainly compares the Positional Mapping software keyboard layout with two other possible keyboard layouts (Group by writing position and Group by vowel and medial) for Myanmar language. The purpose of this study is to see the comparison results of fist-time users' typing speed and their Likert scale evaluation on Positional Mapping with other layouts for Myanmar language.

5.1 Positional Mapping Software Keyboard Layout

Based on the concept of Positional Mapping, Myanmar characters are mapped depending on their writing positions [see Chapter 3]. In detail, consonants "က" (Ka) to "အ" (A) are mapped in the center, vowel sign "ေ◌" (E) and consonant sign "◌ြ" (medial Ra) are mapped on the left, consonant sign "◌ျ" (medial Ya), vowel sign "◌ာ" (Aa), vowel sign "◌ါ" (tall Aa) and sign "◌း" (visarga) are mapped on the right, sign "◌ံ" (Anusvara) , sign "◌ဲ" (Ai), sign "◌ိ" (I), sign "◌ီ" (ii) and sign "◌်" (Asat) are mapped on the top, and sign "◌ှ" (Medial Ha), sign "◌ု" (U), sign "◌ူ" (Uu), sign "◌့" (Aukmyit), sign "◌ွ" (Medial Wa) and Myanmar sign "◌္" (Virama) are mapped on the bottom. Figure 5-1 shows the Positional Mapping keyboard layout.

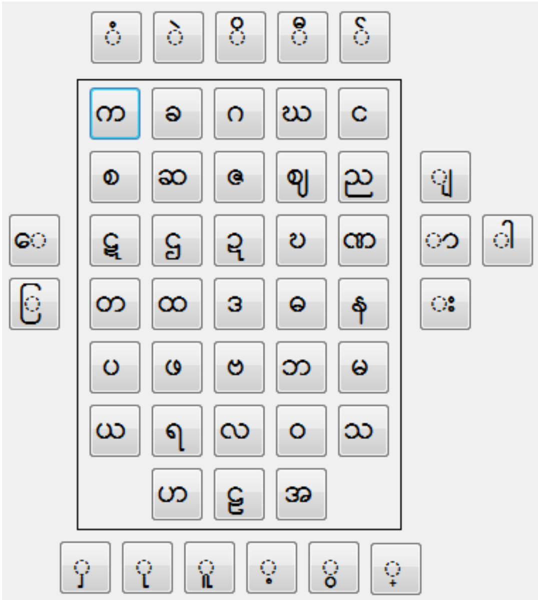


Figure 5-1. Positional Mapping software keyboard layout for Myanmar language

5.2 Group by Writing Position

One of the possible keyboard layouts that I consider is “Group by Writing Position”. Vowel and medial characters are grouped by their writing positions like in the Positional Mapping keyboard layout. However, the difference from Positional Mapping is to put all vowel and medial characters on the right side of the consonants. Left character group, right character group, upper character group and lower character group are mapped from top to bottom (see Figure 5-2). Here, Myanmar sign “◌္” (Virama) is considered as a lower character group.

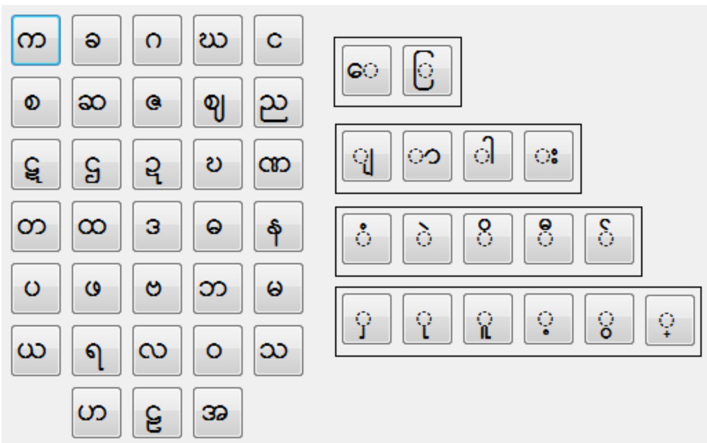


Figure 5-2. Group by writing position

5.3 Group by Vowel and Medial

Another possible keyboard layout that I consider is “Group by Vowel and Medial”. Myanmar language has four basic consonant combination symbols, which are consonant signs “ျ” (medial Ya), “ြ” (medial Ra), “ွ” (Medial Wa) and “ှ” (Medial Ha) (i.e. called Pint Yit Swe Htoo in Myanmar language). These four characters are divided into one group for easy finding, and other dependent vowels and various signs are divided into the other group. Myanmar sign “်” (Virama) is separated from consonant signs and vowels. Figure 5-3 shows a keyboard layout group by Vowel and Medial.

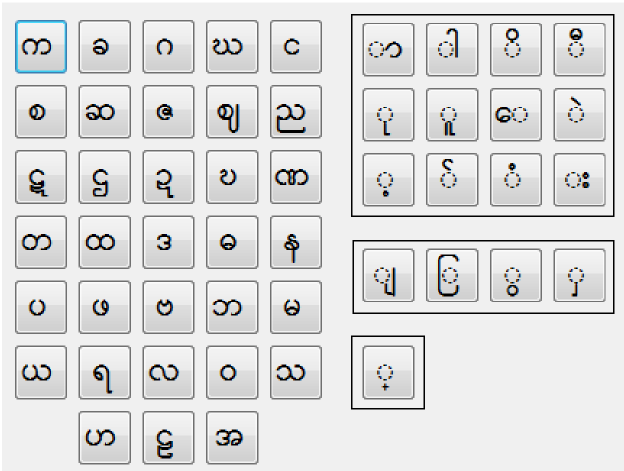


Figure 5-3. Group by Vowel and Medial

5.4 Prototyping

Text input prototype for three software keyboard layouts (i.e. Positional Mapping, Group by Writing Position and Group by Vowel and Medial) was developed with Microsoft Visual Basic. For the implementation, Visual Basic programming language was chosen, which is simple coding and suitable for rapid development. Figure 5-4 shows the prototype used for the comparison of three software keyboard layouts. Keyboard mapping of consonants, dependent vowels and consonant signs are mainly considered in this prototype, and independent vowels and punctuation signs are not considered.

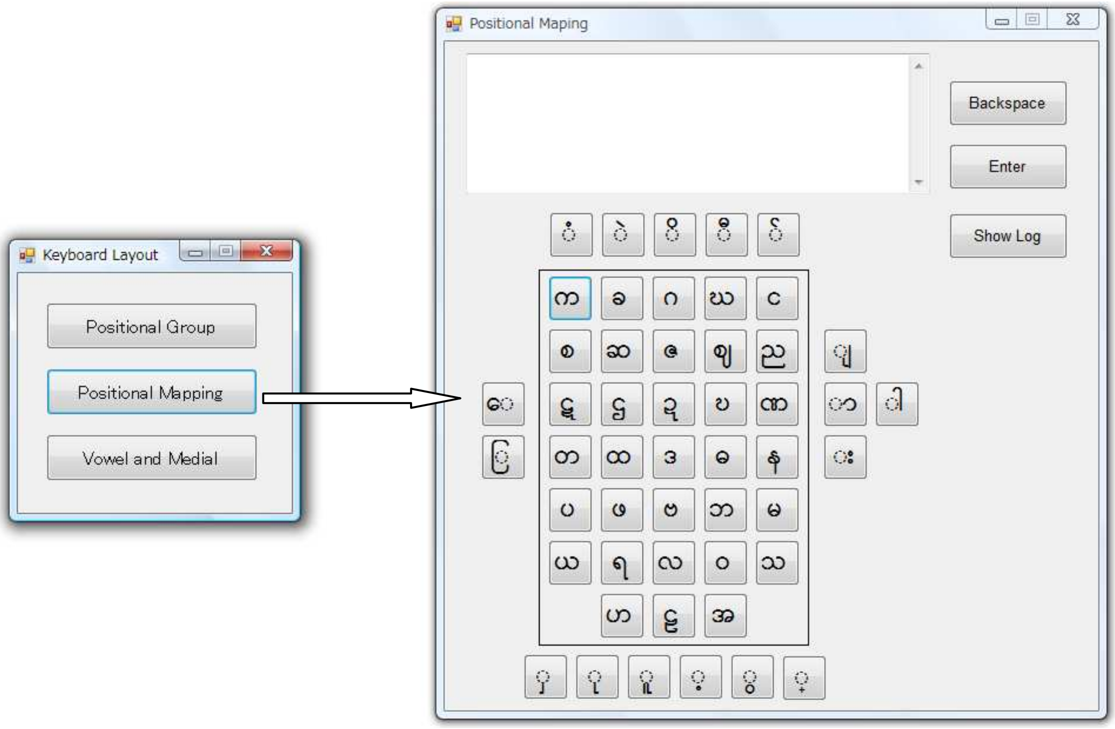


Figure 5-4. Prototype (left: main menu, right: by Positional Mapping keyboard layout)

5.5 User Study

5.5.1 *Participants*

Six volunteer native participants (4 males and 2 females) were recruited for user study. Participants ranged from 23 to 37 years (*mean* = 31.0, *sd* = 5.8). All of them were familiar with a PC but had no prior experience of Myanmar text typing with “Positional Mapping”, “Group by Writing Position” or “Group by Vowel and Medial” keyboard layouts.

5.5.2 *Apparatus*

The experiment was conducted in a quiet meeting room of Graduate School of Information and Telecommunication Studies, Waseda University. Two notebook computers equipped with normal mouses were used for typing three keyboard layouts (see Figure 5-5). The text consists of 70 Myanmar characters including 26 consonants, 35 vowels and 9 consonant signs (see Figure 5-6). If two enter keys are counted, there are 72 characters in total. I use 72 characters to calculate Characters per Minute (CPM) [43].

Figure 5-5. User study apparatus used for Positional Mapping, Group by Writing Position, Group by Vowel and Medial Keyboard layouts

ကျားကြီးခြေရာကြီး

ပညာရွှေအိုးလူမခိုး

ရှေးထုံးလဲမပယ်နဲ့ ဈေးသုံးလဲမကြွယ်နဲ့

Figure 5-6 Three Myanmar proverbs for user study

5.5.3 Procedures

The experiment procedures are as follows:

1) explaining the keyboard layouts of Positional Mapping, Group by Writing Position and Group by Vowel and Medial,
2) making demonstration of text input with three keyboard layouts text input prototype,
3) allowing 5 to 10 minutes practice time to each user to learn text input with my prototype,
4) recording the users' typing speeds of three short Myanmar proverbs for ten times (including error correction time) and
5) discussing with users and getting their responses.

5.5.4 Design

Each participant performed ten trials of typing for each of three keyboard layouts. Two participants started with Positional Mapping layout, another two participants with Group by Writing Position layout and the rest participants with Group by Vowel and Medial layout.

5.6 Evaluation by Typing Speed

Typing speed was evaluated with *Characters per Minute (CPM)* instead of *Word per Minute (WPM)* [43][44]. This is because there is no standard definition for a word in Myanmar like in English (i.e. common definition of a word = 5 characters, including spaces) (Yamada, 1980) [44]. No significant difference was found in entry speed measured in CPM among Positional Mapping, Group by Writing Position and Group by Vowel and Medial ($F_{1,9} = 0.204$, ns). Figure 5-7 shows that CPM for the layout of Positional Mapping averaged 54.50, the one of Group by Writing Position averaged 50.50 and the one of Group by Vowel and Medial averaged 52.00. From this user study, it can be said that the current design of Positional Mapping software keyboard layout does not affect the typing speed and the result is almost the same comparing with Group by Writing Position and Group by Vowel and Medial. However, this result is the outcome of user study with six users, and the result might be different in case user study is held with more users and other factors are measured such as key distance and preparation time spent between key pressings. Here, Keystrokes Per Character (KSPC) is equal (i.e. 72 KSPC) for all three keyboard layouts, and there is no need to make comparison [43].

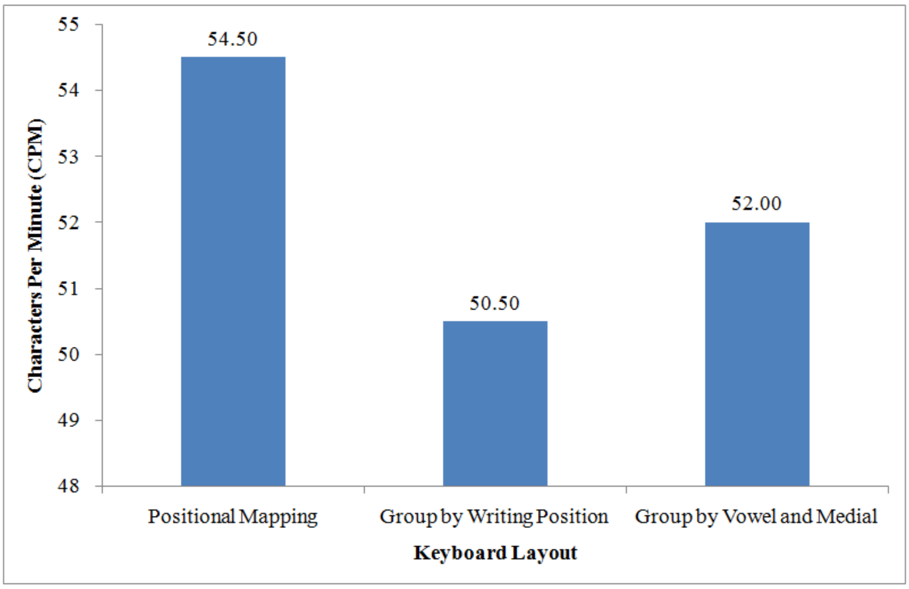


Figure 5-7 Typing speed (CPM) by keyboard layouts

Figure 5-8 shows the results of typing speed improvement of six users for ten times with the Positional Mapping keyboard layout. For the first time, users entered text at 44 CPM (*SD* = 7.70) and for the tenth time, users entered text at 65 CPM (*SD* = 7.39) on average.

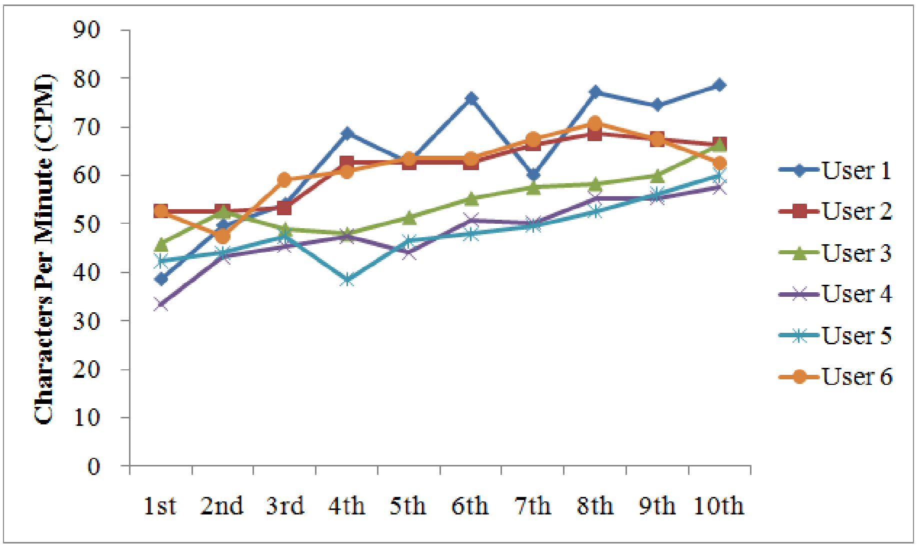


Figure 5-8 Typing speed improvement of six users for ten times

with Positional Mapping keyboard layout

5.7 Evaluation by Users' Likert Scale

Four Likert scales questions (1 to 5) were set to rate the user-friendliness of Positional Mapping text input method (see Chapter 2, Section 2.4.4). These four scales are (1) Difficult-Easy (2) Painful-Enjoyable (3) Slow-Fast and (4) Dislike-Like. Table 5-1 shows the average or arithmetic mean results of Likert scale questions. From the results, it can be generally said that all of the users preferred text entering with the Positional Mapping keyboard layout. They were also requested to rank the three keyboard layouts in order of easy typing. Positional Mapping ranked top, Group by Writing Position ranked second and Group by Vowel and Medial ranked third. In particular, Positional Mapping was the first choice by five users.

Table 5-1. Users' evaluation of Positional Mapping keyboard layout

Likert Scales (range 1-5)	**User1**	**User2**	**User3**	**User4**	**User5**	**User6**
Difficult-Easy	5	5	5	5	5	5
Painful-Enjoyable	5	1	5	5	4	3
Slow-Fast	5	5	4	5	4	4
Dislike-like	5	5	4	5	5	5
Average	5.00	4.00	4.50	5.00	4.50	4.25

5.8 Summary

Over ten trials, there was no significant difference in entry speed measured in CPM among Positional Mapping, Group by Writing Position and Group by Vowel and Medial ($F_{1,9} = 0.204$, ns). However, Figure 5-7 shows that CPM for the layout of Positional Mapping averaged 54.50 is the highest among the three keyboard layouts. Users' evaluation of Positional Mapping by Likert scale (see Table 5-1) and by ranking show its user-friendliness.

CHAPTER 6
Positional Gesture (PG)

6.1 Related Works

This section discusses some of the handwritten or gesture based text input methods for touch screen interfaces.

Since traditional handwritten recognition had many limitations like slow typing speed, Goldberg and Richardson introduced Unistroke alphabet in 1993 [58]. In Unistroke, each character is represented by a single stroke and can be used even by blind people (see Figure 6-1). And thus, there is no segmentation problem in the recognition process. However, users have to spend time on learning Unistroke characters, which are difficult to learn and recall [59]. If this concept is applied to Asian syllabic scripts, it will become more complex and require more learning time of users.

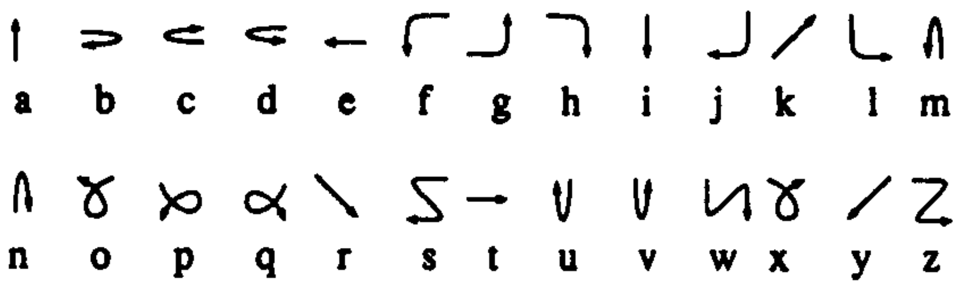


Figure 6-1. Unistroke handwriting method for English alphabets
Image from http://www.pitecan.com/presentations/PenInput/page26.html

Graffiti is another handwriting alphabet developed by Palm Computing for the Palm Pilot PDA product series (see Figure 6-2). Graffiti was developed by Jeff Hawkins who had previously created "PalmPrint" to recognize natural handwriting [60] [61]. It requires minimal time for learning Graffiti alphabet because it is very similar to normal English alphabet. For the recognition process, however, Graffiti strokes are more complex compared to Unistroke strokes.

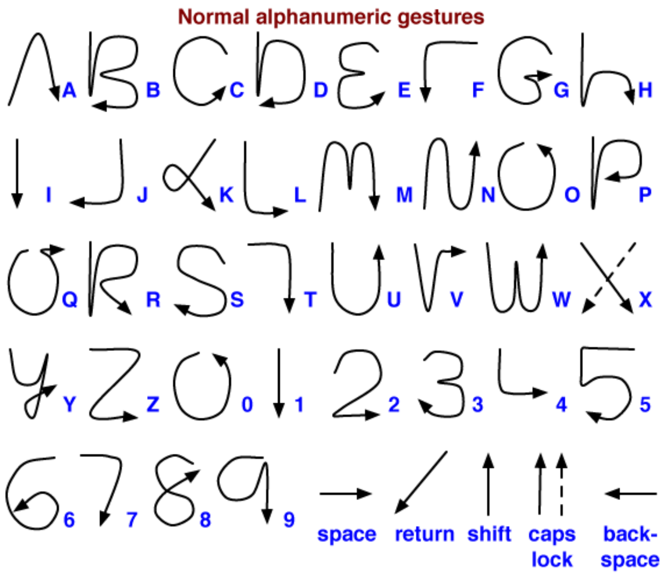


Figure 6-2 Graffiti handwriting method for English alphabets
Image adapted from http://en.wikipedia.org/wiki/Graffiti_(Palm_OS)

EdgeWrite is also based on unistroke text entry idea for handheld devices like PDAs, and is designed for people with motor impairments [62]. Text can be entered by traversing the edges and diagonals of a square hole imposed over the usual text input area of a PDA (see Figure 6-3). In EdgeWrite, recognition algorithm is checking not only pattern recognition but also the sequence of corners that are hit. The authors of EdgeWrite mentioned that users can type 18% more accurate than Graffiti (p<.05), with no significant difference in speed.

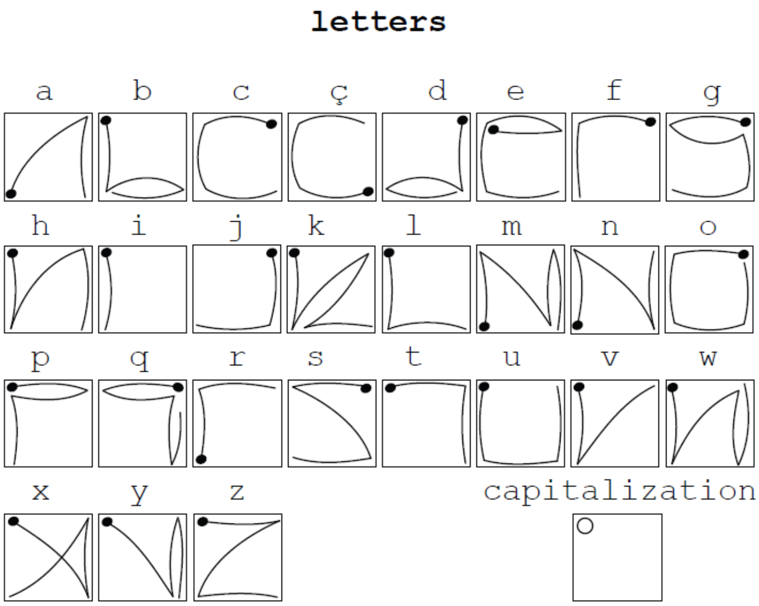


Figure 6-3 EdgeWrite strokes for writing English letters
Image adapted from http://depts.washington.edu/ewrite/downloads/EwChart.pdf

All of the input methods mentioned above are based on English alphabet, and though it is possible to create Graffiti like characters or Unistroke based characters for syllabic scripts e.g. in Myanmar language, the writing positions may need to be defined for combination characters (e.g. vowel signs, tones, subscript characters, etc.) because these characters have to be written according to their defined writing positions. (See example word for Myanmar language in Figure 6-4). And most of the Myanmar characters are similar in shape or glyph such as (က, ဃ, တ, ထ, ဘ, ယ, လ, သ, ဟ), (ခ, စ, ဗ) and (ဂ, င, ပ, ဝ). In handwriting text input, it will be difficult for these kinds of characters to be distinguished by the recognition engine, and this will affect the accuracy.

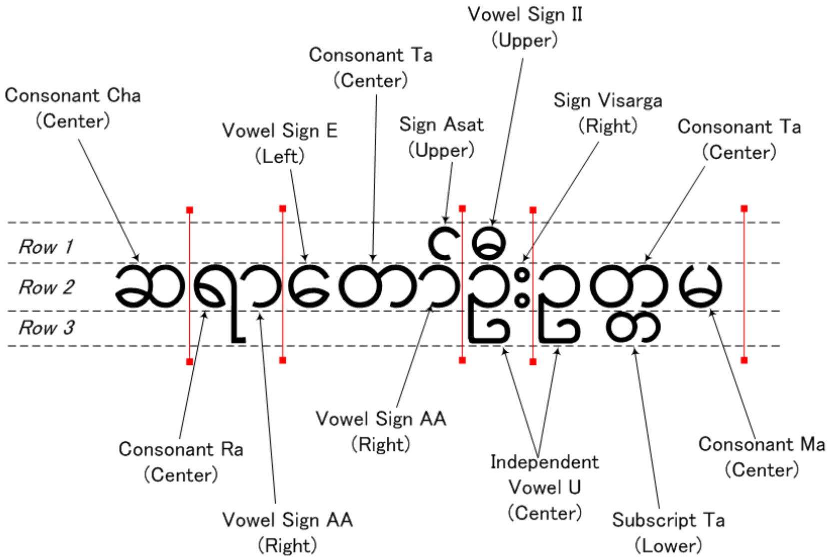


Figure 6-4 Myanmar phrase "Sayadaw U Ottama"
Image taken from my presented poster [126]

Gesture Keyboard (GKB) for Devanagari (one of the Indic scripts) is based on handwritten gesture recognition technology, which was proposed by R. Balaji, V. Deepu, Sriganesh Madhvanath and Jayasree Prabhakaran [63] [64]. GKB input method is very appropriate for the nature of syllabic scripts writing system. In this input method, users do not need to write down Devanagari consonants and they are already shown on a tablet keyboard. Users can type a consonant by giving a special gesture command (i.e. strike through over a consonant). Other matras (i.e. vowel signs) can be typed by writing at a specific position relative to the glyph of the base consonant (see Figure 6-5). This input method is smart and applicable to other similar syllabic scripts. Recognition accuracy and typing speed can increase compared to normal handwritten techniques. However, users have to write down vowel signs or other combination symbols correctly, which may lead to reduce typing speed. And then, it is still necessary to create recognition engine for vowel signs or other characters.

Figure 6-5. Gesture Keyboard (GKB) layout for Devanagari
Image taken from HP Labs India GKB Demo Flash Program
(Download link: http://www.hpl.hp.com/india/demos/demo.html)

A Khmer software keyboard is developed by modeling the standard NiDA (National Information Technology Development Authority) keyboard [55] (see Figure 4-3). It was created just after Khmer Unicode 4.0 was released. The Khmer Unicode keyboard layout differs from the one of old version which is not in Unicode in which the subscript of the consonants is not spread on the keyboard anymore. Instead, a subscript sign is used to indicate that the next consonant is subscript of the cluster, and that the typing order is not from left to right in the same order of hand writing, but in the order of word's spelling.

Although soft keyboards or visual keyboards are possible solutions for syllabic languages, typing syllabic languages is still difficult for novice users. In addition, they are not suitable for small mobile devices, because it is difficult for most syllabic languages characters (e.g. Myanmar, Bengali and Khmer characters) to be distinguished from other similar characters in small soft keyboards.

6.2 Logical Combination Structure of Myanmar, Khmer and Bengali

As mentioned in Chapter 1, section 1.1, Myanmar, Khmer and Bengali languages writing systems have many common characteristics. And it has been clearly found that the writing systems are based on adding of left, right, upper and lower characters to a consonant basically. Figure 6-6, Figure 6-7 and Figure 6-8 show the logical combination structure found for syllabic scripts writing systems such as Myanmar, Khmer and Bengali. Taking this into consideration, a new gesture method has been developed for syllabic scripts.

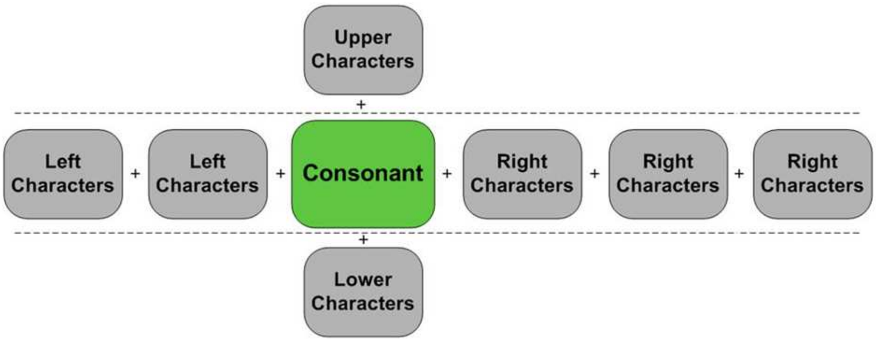


Figure 6-6. Combination structure of other characters with a consonant in Myanmar language
Image taken from my published paper [128]

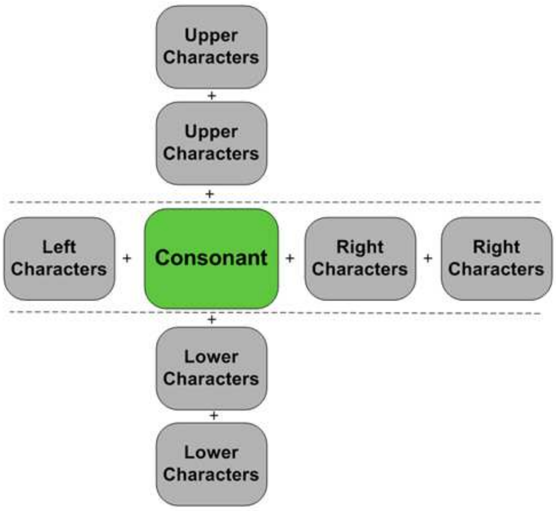


Figure 6-7. Combination structure of other characters with a consonant in Khmer language
Image taken from my published papers [39, 57, 127]

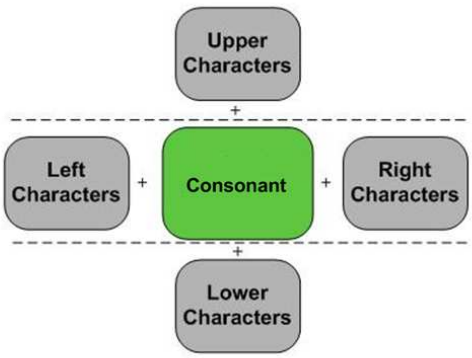


Figure 6-8. Combination structure of other characters with a consonant in Bengali language
Image taken from my published paper [128]

6.3 Concept

Positional Gesture (PG) is a simple gesture text input method for computing devices based on “common characteristics” or “characters writing position” of Asian syllabic scripts writing system. The concept is totally based on four simple gesture commands, which are "Left", "Right", "Up" and "Down". "Left gesture command" is for left characters or symbols, "Right gesture command" is for right characters or symbols, "Up gesture command" is for upper characters or symbols and "Down gesture command" is for lower characters or symbols. Here, as a concept, "Left gesture command" can be "dragging mouse pointer to left" or "moving data glove to left" or "pressing left arrow key" or "moving eye ball to left" or anything. For the consonant characters, additional gesture is usable like "drawing dot" or "writing circle" or anything. In my prototypes, "Left gesture command with short distance" is used for Myanmar language and “Double Click” is used for Khmer language to make it simple. Figure 6-9 shows Positional Gesture text input concept.

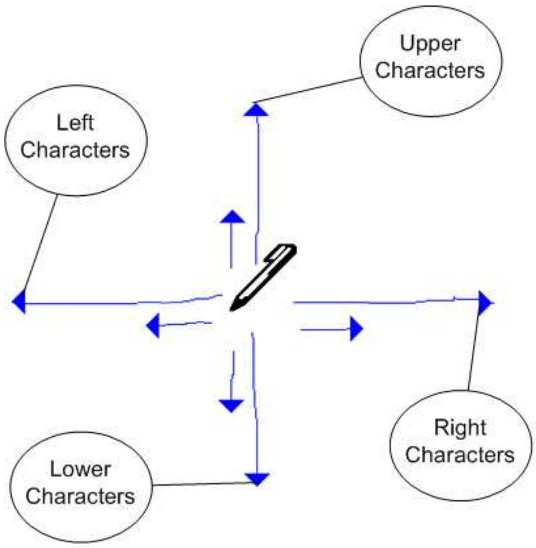


Figure 6-9. Positional Gesture with stylus

Image taken from my published papers [57, 126, 127, 128, 129]

6.4 Prototyping for Myanmar and Khmer Languages

Positional Gesture text input interface prototypes in Myanmar and Khmer languages (see Figure 6-10 and Figure 6-11) are presented here, which were developed with Microsoft Visual Studio .Net 2003. These prototypes can be used with a stylus pen, a trackball or a mouse.

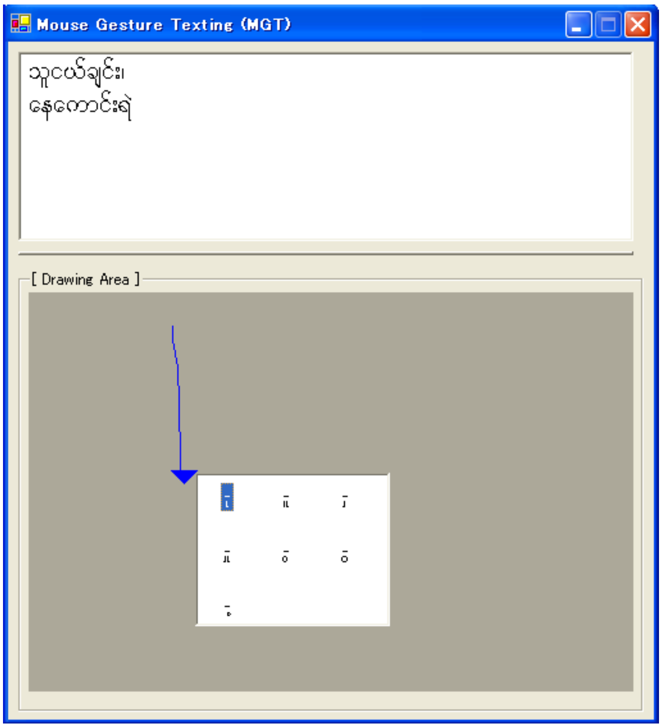


Figure 6-10. Positional Gesture prototypes for Myanmar
Image taken from my published papers [127, 129]

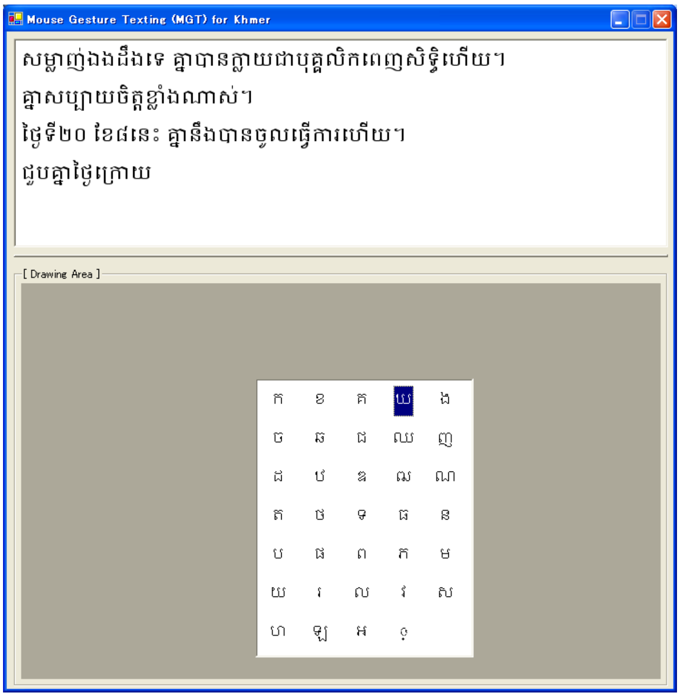


Figure 6-11. Positional Gesture prototypes for Khmer
Image taken from my published papers [57, 129]

For the basic text editing, "mouse dragging to the left direction by pressing right click button" is used for "Back Space" function and "mouse dragging to the down direction by right click button" is used for "Enter" function. Recognition algorithm for my prototype is very simple because it is only necessary to check drawing direction, left click or right click and length of the path. And there is no need to be straight line, and drawing within angle of 60 degree range is allowed for each direction (i.e. Left, Right, Up and Down). Table 6-1 and Table 6-2 show gesture commands for Myanmar and Khmer prototype.

Table 6-1. Gesture commands for Myanmar

Table taken from my published papers [127, 129]

Gesture Commands	**Character Assignments**
Left (long)	Left characters (" ေ" and "ြ")
Right (long)	Numbers ("၁", "၂", "၃", "၄", "၅", etc.)
Up (long)	Symbols ("@", "!", "&", "$", "%", etc.)
Down (long)	Subscript consonants ("္က", "္ခ", "္ဂ", "္စ", "္ဆ", etc.)
Right (short)	Right characters ("ျ", "ာ", "း", "၊", "။", etc.)
Up (short)	Upper characters ("ိ", "ီ", "ဲ", etc.)
Down (short)	Lower characters ("ု", "ူ", "ျ", "ွ","ှ", etc.)
Double Click	Consonant characters ("က", "ခ", "ဂ", "ဃ", "င", etc.)

Table 6-2. Gesture commands for Khmer

Table taken from my published papers [57, 127, 129]

Gesture Commands	**Character Assignments**
Left (long)	Left, Right characters (“ឿ”, “ៀ”, “ោ”, etc.)
Right (long)	Numbers (“១”, “២”, “៣”, “៤”, etc.)
Up (long)	Symbols (“។”, “,"”, “?”, “!”, “#”, etc.)
Down (long)	Independent vowels, Symbols and frequently used characters (“ឲ”, “ឮ”, “។ល។”, “៖”, etc.)
Left (short)	Left characters (“េ”, “ែ”, “ៃ” and “ើ”)
Right (short)	Right characters (“ា”, “ះ”, “ាំ” and “ៈ”)
Up (short)	Upper characters (“ិ”, “ី”, “ឹ”, “៊”, etc.)
Down (short)	Lower characters (“ុ”, “ូ”, “ួ”, etc.)
Double Click	Consonant characters (“ក”, “ខ”, “គ”, “ឃ”, “ង”, etc.)

In my prototype, all of the gesture lines are shown with four directions of arrow heads. And the color of the gesture lines will change from blue to red for text editing commands. One limitation of the current prototype is that stylus pen should provide left and right click feature.

6.5 User Study

User experiments were conducted for Myanmar and Khmer languages with my developed prototypes in order to figure out users' typing speed for Positional Gesture text input (Figure 6-14). A PC mouse and a stylus pen with a tablet were used for user study with five Myanmar native participants who are between 21 and 33 years old. And for Khmer language, a PC mouse and a trackball were used for user study with five Cambodian native participants who are between 24 and 26 years old. All of the participants are familiar with a PC but don't have an experience of using a pen and a trackball. The experiments procedures are as follows:

1) explaining the concept of Positional Gesture text input,
2) making demonstration of text input with prototype,
3) allowing 10 minutes practice time for each user to learn text input with prototype,
4) recording users' typing speed of short message (6 Myanmar sentences/5 Khmer sentences) for 5 trial times (including error correction time) and
5) getting users' feedbacks for each prototype with small questionnaires.

Figure 6-12 and Figure 6-13 show the Myanmar text (containing 107 characters including spaces) and Khmer text (containing 135 characters including spaces) used for user study respectively.

သူငယ်ချင်း၊	(Hi, friend!)
မတွေ့ရတာကြာပြီနော်။	(Long time no see.)
နေကောင်းရဲ့လား။	(How are you?)
ငါ့ဖုန်းနံပါတ်အသစ်က ၅၀၀၇၄၅၉၊	(My new phone number is 5007459.)
အားတဲ့အခါဖုန်းပြန်ဆက်ကွာ။	(Call me back when you have time.)
ဒါပဲနော်။	(Bye for now.)

Figure 6-12. Six Myanmar sentences for user study
Image taken from my published papers [38, 127, 129]

សួស្តី!

(Hi!)

សម្លាញ់ឯងដឹងទេ គ្នាបានក្លាយជាបុគ្គលិកពេញសិទ្ធិហើយ។

(You know, friend, I am now accepted to be a contract employee.)

គ្នាសប្បាយចិត្តខ្លាំងណាស់។

(I'm extremely happy.)

ថ្ងៃទី២០ ខែ៨នេះគ្នានឹងបានចូលធ្វើការហើយ។

(I will start my work on 20 of August.)

ជួបគ្នាថ្ងៃក្រោយ

(See you next time.)

Figure 6-13. Five Khmer sentences for user study
Image taken from my published papers [39, 127, 129]

Figure 6-14. Photos from user study with mouse, trackball, tablet with stylus, touch pad and tablet PC

6.6 Evaluation

6.6.1 Evaluation by Users' Typing Speed

Users' actual typing speed was recorded and CPM (Characters per Minute) was calculated for the evaluation process of my Positional Gesture prototype. From the user study results, the average typing speed of Positional Gesture text input to finish six Myanmar sentences (see Figure 6-12) with a stylus pen is 7 minutes 29 seconds (see Figure 6-15) and 5 minutes 53 seconds with a mouse (see Figure 6-16). And the average typing speed of Positional Gesture text input to finish five Khmer sentences (see Figure 6-13) with a mouse is 7 minutes 38 seconds (see Figure 6-20) and 7 minutes 37 seconds with a trackball (see Figure 6-21). In general, all of the users' typing speed with my prototypes increased during five trial experiments (see Figure 6-19 and Figure 6-24). The results of the previous user study "Win Myanmar Visual Keyboard" typing speed by 2 users were referred for comparison (see Figure 6-18) [65]. Average typing speed of 2 users for "Win Myanmar Visual Keyboard" is 2 minutes 30 seconds. It was found that Positional Gesture text input method is 199% (with a pen) and 135% (with a mouse) slower than "Win Myanmar Visual Keyboard". For Khmer, the Tavultesoft software keyboard that adopted the NiDA keyboard layout was used, and average typing speed of 5 users is 4 min 38 sec (see Figure 6-22). The Tavultesoft software keyboard is referred to as Khmer software keyboard hereafter. It was found that Positional Gesture text input method for Khmer language is 64% (64.39% with a trackball and 64.75% with a mouse) slower than "Khmer software keyboard".

6.6.2 Evaluation by Characters per Minute (CPM)

Although Word per Minute (WPM) is the most widely used measure of the text entry performance, my Positional Gesture text input is evaluated rather with Character per Minute (CPM). It is because the common definition of "word" cannot be found in syllabic scripts such as Myanmar and Khmer. Importantly, the WPM/CPM measure does not consider the number of keystrokes or gestures made during entry, but only the length of the resulting transcribed string and how long it takes to produce it [43]. The Myanmar text used for user study contains 107 characters, and the Khmer text contains 135 characters. Based on the user study results, average CPM for Positional Gesture and software keyboards is as follows:

For Myanmar language,

Positional Gesture with a mouse: 19.27 CPM

Positional Gesture with a pen: 14.29 CPM

Win Myanmar Software Keyboard: 42.8 CPM

For Khmer language,

Positional Gesture with a mouse: 17.69 CPM

Positional Gesture with a trackball: 17.72 CPM

Khmer software keyboard: 29.13 CPM

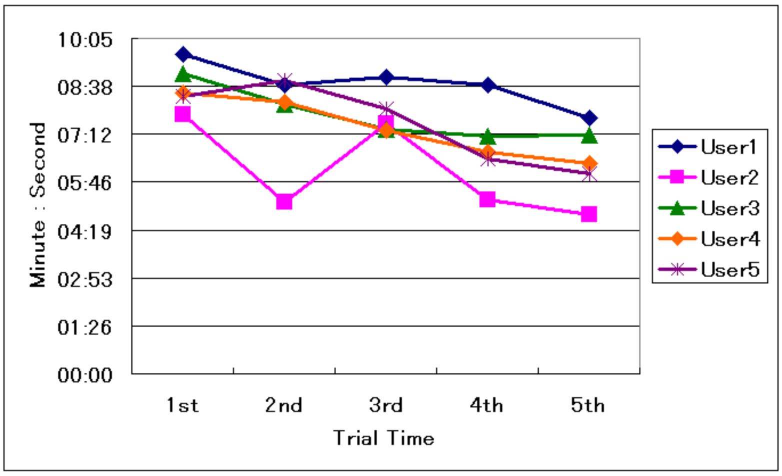


Figure 6-15. Typing speed of five Myanmar users with pen

Image taken from my published paper [129]

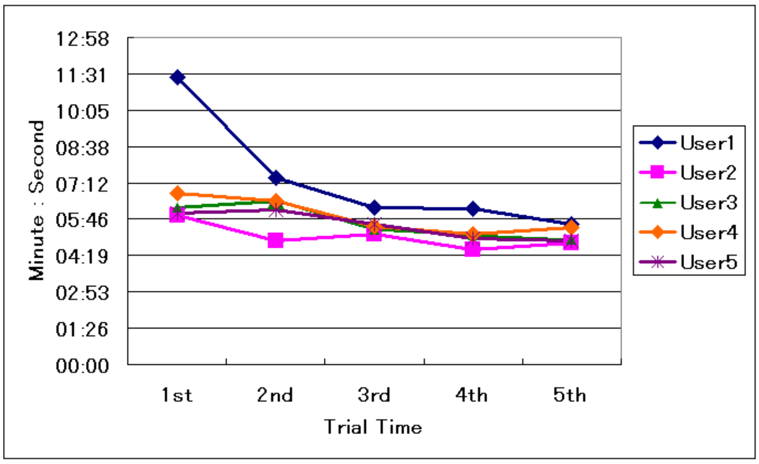


Figure 6-16. Typing speed of five Myanmar users with mouse

Image taken from my published paper [129]

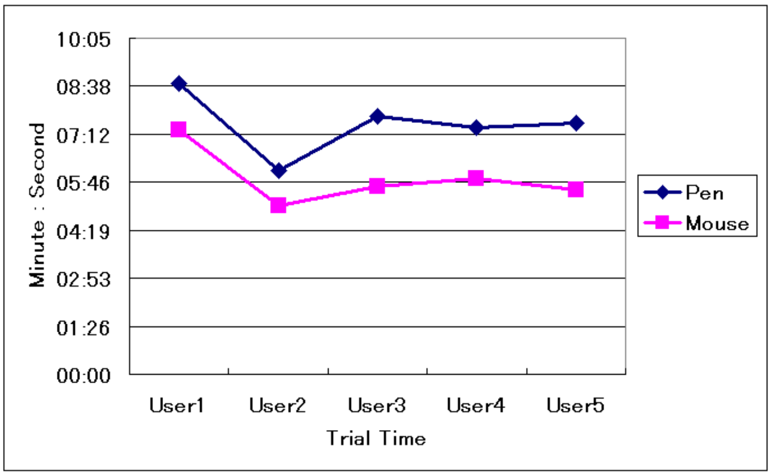


Figure 6-17. Typing speed comparison of five Myanmar users with pen and mouse
Image taken from my published paper [129]

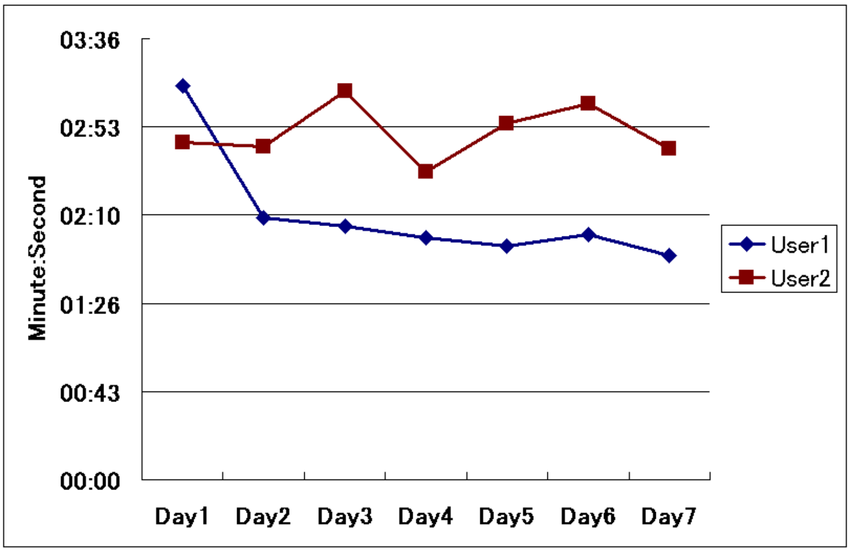


Figure 6-18. Average typing speed of two users for "Win Myanmar Visual Keyboard"
Image taken from my published paper [129]

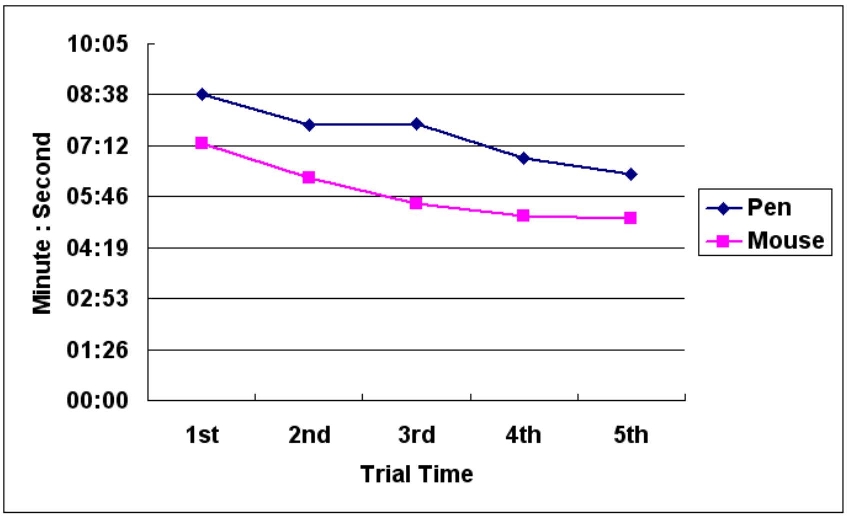


Figure 6-19. Typing speed improvement of five Myanmar users for Positional Gesture with mouse and pen

Image taken from my published paper [129]

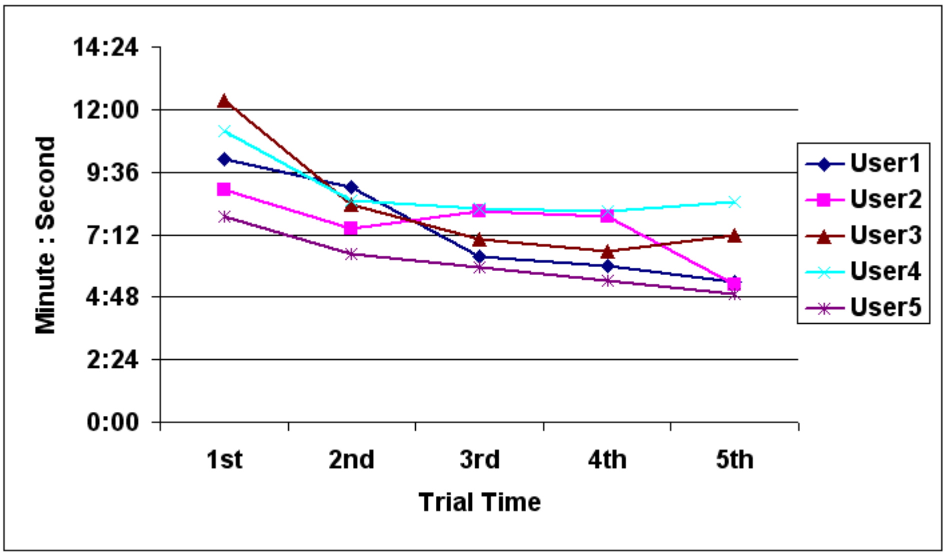


Figure 6-20. Typing speed of five Cambodian users with mouse

Image taken from my published paper [129]

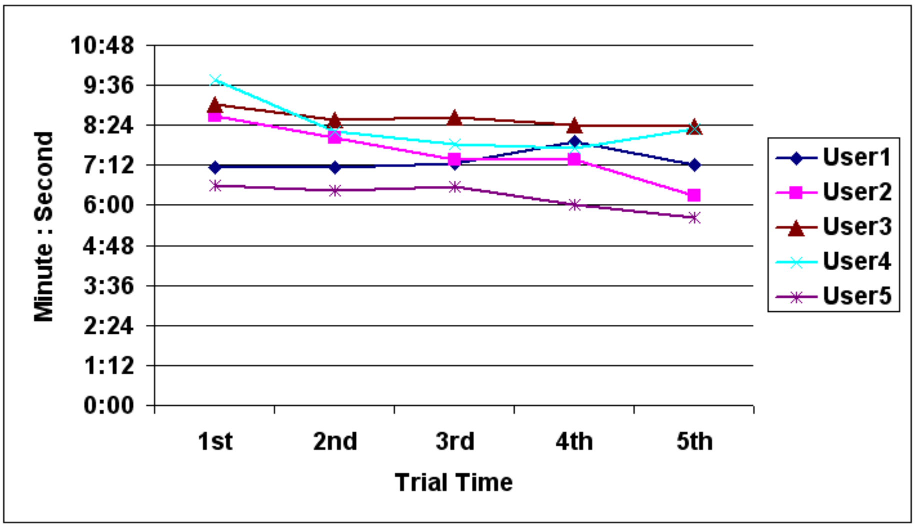


Figure 6-21 Typing speed of five Cambodian users with trackball

Image taken from my published paper [129]

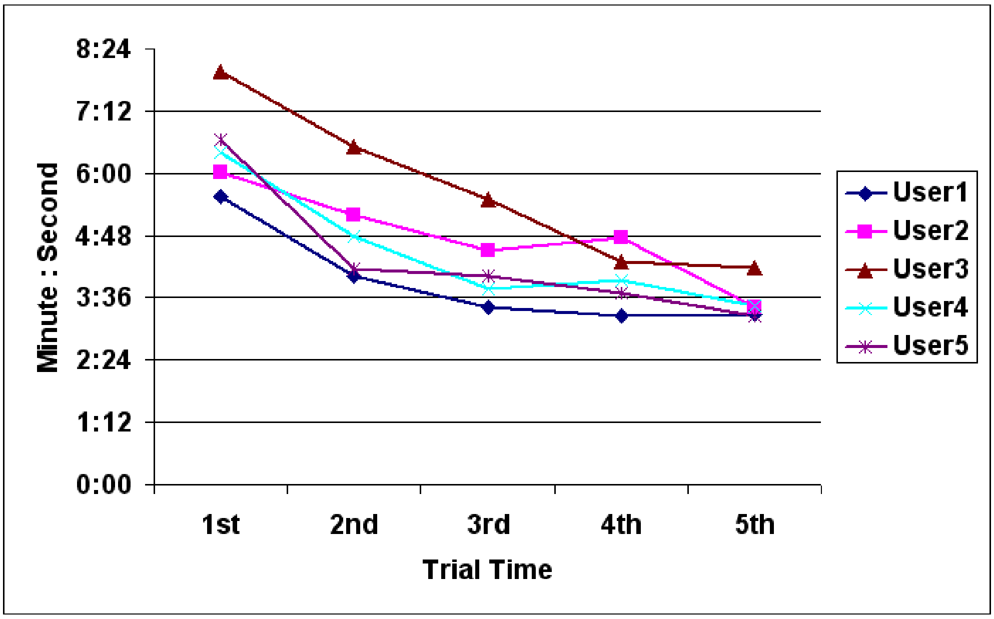


Figure 6-22 Typing speed of five Cambodian users with Khmer software keyboard

Image taken from my published paper [129]

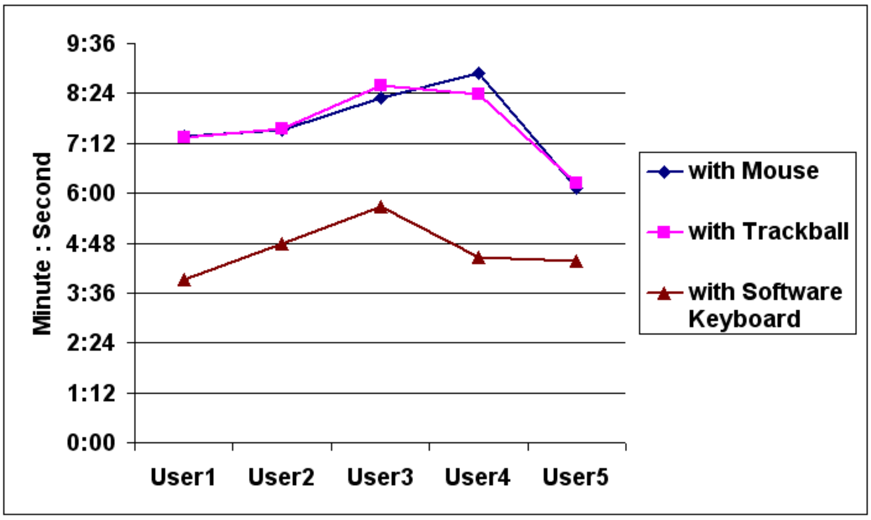


Figure 6-23 Typing speed comparison of five Cambodian users with mouse, trackball and software keyboard

Image taken from my published paper [129]

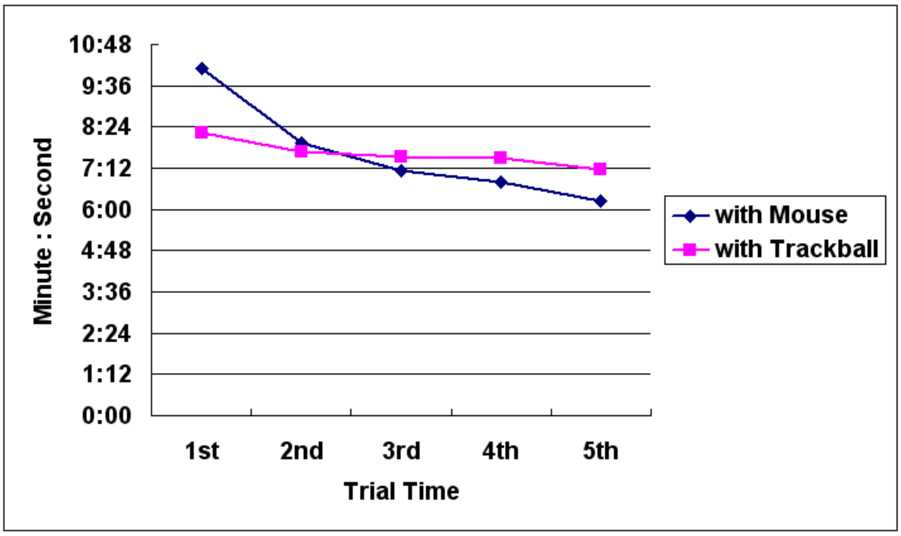


Figure 6-24 Typing speed improvement of five Cambodian users for Positional Gesture with mouse and trackball

Image taken from my published paper [129]

6.6.3 Evaluation by Users' Likert Scale

Small questionnaires were conducted to participants in order to figure out their feedbacks on Positional Gesture.

Table 6-3. Mean (standard deviation) responses by five Myanmar users for five-point Likert scale questions

Table taken from my published paper [129]

Likert Scales (range 1-5)	**PG with Pen**	**PG with Mouse**
Difficult-Easy	3.4 (0.89)	4.6 (0.55)
Painful-Enjoyable	3.6 (0.55)	3.4 (1.14)
Slow-Fast	3.0 (0)	4.2 (0.84)
Dislike-Like	4.4 (0.89)	4.6 (0.55)

Table 6-4. Mean (standard deviation) responses by five Khmer users for five-point Likert scale questions

Table taken from my published paper [129]

Likert Scales (range 1-5)	**PG with Trackball**	**PG with Mouse**	**Software Keyboard**
Difficult-Easy	2.0 (1.22)	3.2 (0.84)	4.2 (1.30)
Painful-Enjoyable	2.6 (1.14)	3.6 (0.89)	3.8 (1.10)
Slow-Fast	2.0 (0.71)	3.4 (1.14)	3.8 (1.10)
Dislike-Like	2.8 (1.79)	4.0 (0.71)	4.0 (1.22)

The questionnaires were conducted immediately after typing experiments. Participants were given four Likert Scales (1-5) on which to rate the Positional Gesture text input with 3 different input devices (a mouse, a pen and a trackball). Table 6-3 and Table 6-4 show mean and standard deviation of users for Likert scale questions. Labels for scale endpoints are in the most left columns and higher values are better. From the questionnaire results, it was noticed that most users preferred text entering with a mouse than a stylus pen or a trackball: Myanmar (Mouse>Pen, 4.6>4.4) and Khmer (Mouse>Trackball, 4.0>2.8). One of the reasons might be that the participants used a pen with a tablet and a trackball for the first time in the experiments. The interesting point is that although Likert scales responds by the Cambodian users for a trackball are lower than a mouse (see Table 6-4), the average typing speed is nearly the same (7min 38sec with a mouse and 7min 37sec with a trackball). Overall results of evaluation by users are satisfactory.

6.7 Summary

It is desirable to compare the user study results of my prototype with handwritten; however, there is no handwritten text input system for Myanmar and Khmer language yet. Although some thesis papers regarding off line handwritten recognition for Myanmar characters were available, no paper was found for online handwritten recognition. And thus, "Win Myanmar Visual Keyboard" and "Khmer software keyboard" were used for the evaluation process. Average CPM of Myanmar Positional Gesture prototype is 14.29 with a pen and 19.27 with a mouse respectively. And average CPM for Khmer Positional Gesture prototype is 17.72 with a trackball and 17.69 with a mouse. Comparing Positional Gesture typing speed with software keyboards, it is slower for both for Myanmar and Khmer language. The reason is that very few gesture commands are used in my prototypes to enable easier typing but software keyboards use one to one key mapping (e.g. 58 keys + shifted mode + right-alt-mode for Khmer software keyboard). Another important reason is that though typing with the QWERTY based software keyboard is familiar text input method, Positional Gesture is a new one for all of the participants. The typing speed or CPM can increase or decrease according to the type of input device used. And it also depends on how characters are mapped (i.e. character assignments). From the mean and standard deviation values of Likert Scale questions (see Table 6-3 for Myanmar language and Table 6-4 for Khmer language) and the user study results (see Figure 6-15 to Figure 6-24), it is proven that Positional Gesture is one of the possible text input interfaces for small mobile devices.

CHAPTER 7
Positional Prediction (PP)

7.1 Related Works

The research work for predictive text entry system began in China and Japan around 1960s and gradually got popular together with word processing software. Some predictive methods are based on phonetics (e.g. Romaji (or) Kana codes in Japanese and Pinyin codes in Chinese). Other methods are based on the shape of Chinese characters (called Kanji in Japanese) or stroke orders (e.g. Four Corner Method, Wubi, Q9, etc.) [66].

The dictionary based predictive methods like Tegic's T9 and Zi Corp's eZiText predict the word or phrases according to the key sequences type by users. For example, when users press "6" (mno) and "3" (def) on a standard ISO mobile phone keypad, the system will predict the words like "of" and "me", and then order by usage frequency. Ambiguity arises when more than one word matches a given key sequence (e.g. "5477" = "kiss", "lips") and "next" key is used for choosing other alternative words. Or users can stop predicting system and spell the word which involves rewriting the whole word using multi-tap. To improve the performance of disambiguation algorithms, some research work has been already proposed like dictionary based predicting with used context of the phrase or document being typed [67] and LetterWise [68] that used prefix-based disambiguation instead of using dictionary of stored word by MacKenzie, Kober, Smith, Jones and Skepner (2001).

On the other hand, traditional multi-tap text input method is still popular because of its simple nature. As far as I know, there is no predictive text input that supports Myanmar language yet.

7.2 Concept

In making an analysis of Myanmar word formation, it was found that writing system largely depends on adding left, right, upper and lower characters to consonant (i.e. consonant clusters). Here, left, right, upper and lower characters mean Myanmar dependent vowels, directives and subscript consonants that are always written with a consonant (e.g. left: "ေ◌" (E) and "◌ြ" (RA), Right: "◌ာ" (AA) and "◌း" (VISARGA), etc., Up: "◌ိ" (I), "◌ီ" (II), "◌ံ" (ANUSVARA), etc. and Down: "◌ု" (U), "◌ူ" (UU), "◌ှ" (HA THO), "◌ှု" (HA THO TA CHAUNG NGIN), "◌ွ" (WA SWE), "◌ွှ" (WA SWE HA THO), "◌့" (AUKMYIT), "◌္က" (Subscripted KA), etc.). They have to be written always with consonants (i.e. dependent), and their positions are defined when they are combined with consonants. Taking this into consideration, a new consonant cluster prediction method is proposed based on given positional information of dependent characters. For example, "Ka + Right" for "ကာ", "ကား", "ကျ", "ကျာ"and "ကျား, "Ka + Down" for "ကု", "ကူ" and "ကွ", "Ka + Down + Up" for "ကို", etc. The vowel positional information adding order can be "Consonant + Left + Down + Up + Right" or Consonant + Left + Up + Down + Right. This text entry order is similar to Myanmar Unicode input order. Logically, it can also support hand writing order (i.e. Left + Consonant + Down + Up + Right or Left + Consonant + Up + Down + Right). Users only need to mention positional vowel information once even for the words that need to combine more than one positional vowel. Based on the given positional vowel information, the system calculates all possible consonant cluster combinations, then, removes impossible vowel combinations for given consonant (i.e. unpronounceable or no meaningful combinations), and after that, sorts possible consonant clusters with average usage frequency. Here, note that there is no proper or standard usage frequency table for Myanmar characters and words yet and my own usage frequency table is used, which is the result of my previous work. Finally, the system sorts the consonant cluster according to users' typing history. Figure 7-1 shows the example process of positional prediction for "Kha + Left + Right + Up".

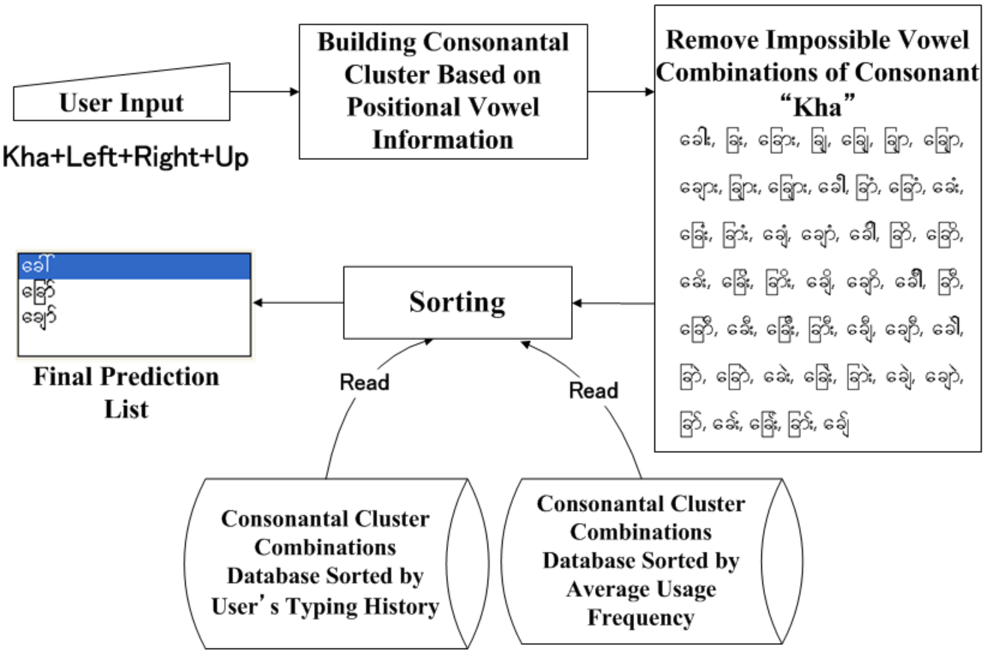


Figure 7-1. Process flow of Positional Prediction for Myanmar consonant "Kha"
with vowel information (Left + Right + Up)
Image taken from my published papers [38, 128]

7.3 Prototyping for Myanmar and Khmer Language

Here, Positional Prediction text entry prototype for Myanmar and Khmer languages is presented, which were developed with Microsoft Visual Studio .Net 2003. The prototypes are shown in Figure 7-2 and Figure 7-3. To reduce the candidate selection time, just a self subscript character of given consonant in predictive list is mentioned. Consonants subscripting with different characters and Myanmar short form writing can be typed with subscript button (i.e. the button labeled with "◌္" character).

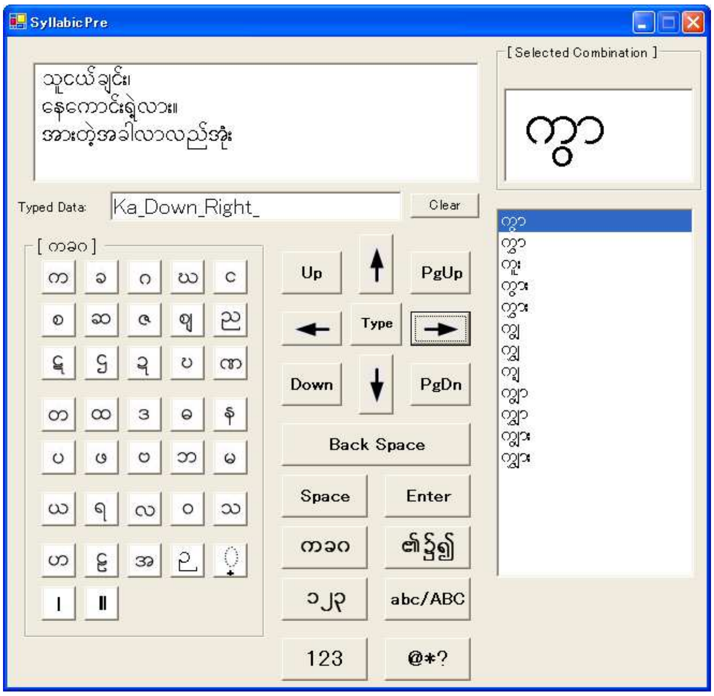


Figure 7-2. Positional Prediction text entry prototype for Myanmar language

(Up, Down, PgUp and PgDn buttons are used for moving the highlight selection of consonant clusters of list box.)

Image taken from my published papers [38, 128]

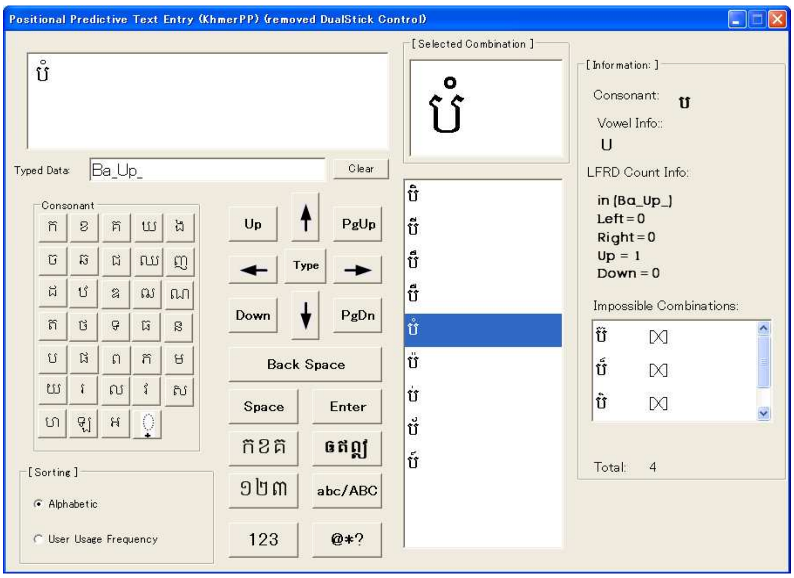


Figure 7-3. Positional Prediction text entry prototype for Khmer language

(Up, Down, PgUp and PgDn buttons are used for moving the highlight selection of consonant clusters of list box.)

Image taken from my published papers [39, 128]

These prototypes are usable with a mouse, a stylus pen, a Logitech dual thumbstick game pad and a Nintendo Wii remote controller. With a dual thumbstick game pad, “left thumbstick” works for selecting consonants and other editing buttons and “right thumbstick” works for giving Left, Right, Up and Down information. With a Wii remote controller, the mouse pointer can be controlled by moving the remote controller and the four directional keys pad is used for giving Left, Right, Up and Down information.

Although this current prototype is designed for the use with a PC, a tablet PC or an electronic whiteboard, by changing the user interface, it is also applicable for mobile devices such as a mobile phone, a PDA and a portable game player. Because of that, most current mobile devices have four directional arrow keys (See Figure 7-4 and 7-5).

Figure 7-4. Four directional arrow keys (left) and Nokia N76 mobile phone (right) Dell X51 PDA

Image adapted from my published papers [38, 128]

Figure 7-5. Four directional arrow keys (left) and Sony PSP portable game player (right) XO laptop

Image adapted from my published papers [38, 128]

7.4 User Study

Initial user study was conducted for Myanmar language and Khmer language with the current Positional Prediction text entry prototypes in order to know users' typing speed and to get their comments.

7.4.1 User Study and Evaluation of Myanmar Language

A PC mouse, a stylus pen with a tablet PC, a Logitech dual thumbstick game pad and a Nintendo Wii remote controller were used with three Myanmar native users. All of the participants are familiar with a PC but don't have an experience of using a stylus pen, a dual thumbstick and a Wii remote controller. The experiments procedures are as follows:

1) explaining the concept of Positional Prediction text entry,
2) making demonstration of text input with prototype,
3) allowing 15 minutes practice time for each user to learn text input with prototype,
4) recording users' typing speeds of short Myanmar message (6 sentences) for 5 times (including error correction time) and
5) discussing with users and learning their responses.

Figure 6-15 shows the Myanmar text used for user study. Recorded users' typing speeds are calculated to CPM (Characters per Minute) for the evaluation process. The user study results show that the average typing speed of Positional Prediction text entry method to finish six Myanmar sentences with a mouse is 3 minutes 15 seconds (32.62 CPM), by a stylus pen is 3 minutes 21 seconds (31.64 CPM), with a dual thumbstick game pad is 7 minutes 5 seconds (14.96 CPM) and with a Wii remote controller is 5 minutes 30 seconds (19.27 CPM) respectively. With the current prototype, the fastest user's typing speed for each device is: 2 minutes 33 seconds (41.57 CPM) with a mouse, 2 minutes 32 seconds (41.84 CPM) by a stylus pen, 7 minutes 11 seconds (14.76 CPM) by dual thumbstick and 5 minutes 10 seconds (20.52 CPM) with a Wii remote controller. From these results, it is found that the faster typing speed is available with a mouse or a stylus pen. The first reason is that all of the users are unfamiliar with a dual thumbstick game pad and a Wii remote controller. And the second reason is that users' cannot directly point or highlight the consonants like in a mouse and a stylus pen. The third reason is because of the program bugs (e.g. in my current prototype, the mouse pointer movement time with a Wii remote controller is slow and sometimes difficult to control).

To make comparison with the existing Myanmar text input method, "Win Myanmar Visual Keyboard" was used [27] [65]. Actually, "Win Myanmar Visual Keyboard" uses more keys or buttons than my prototype because it uses the QWERTY keyboard interface. The results of the previous user study indicate that the average typing speed of two users for "Win

Myanmar Visual Keyboard" is 2 minutes 30 seconds (42.4 CPM) (see Figure 7-6). The average typing speed of Positional Prediction text entry method with the current proposal (see Figure 7-7) is 30 % slower than "Win Myanmar Visual Keyboard". But fastest typing speed 41.57 CPM is almost the same as average typing speed of "Win Myanmar Visual Keyboard" (i.e. 42.4 CPM).

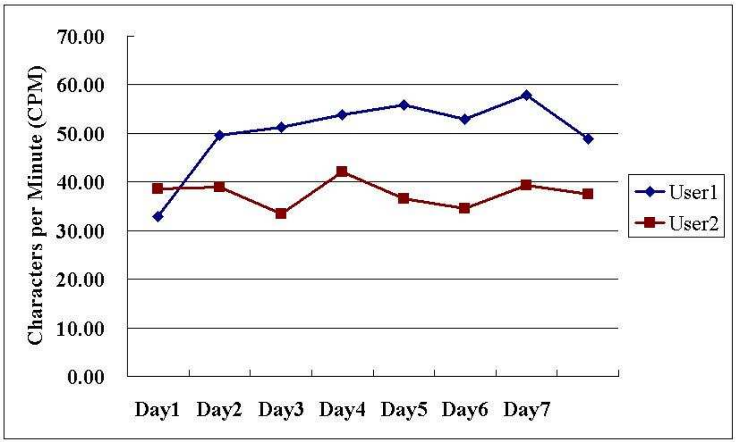


Figure 7-6. Average typing speed of two users to finish six Myanmar sentences with Win Myanmar Visual Keyboard (experiment results for seven continuous days)
Image taken from my published paper [38]

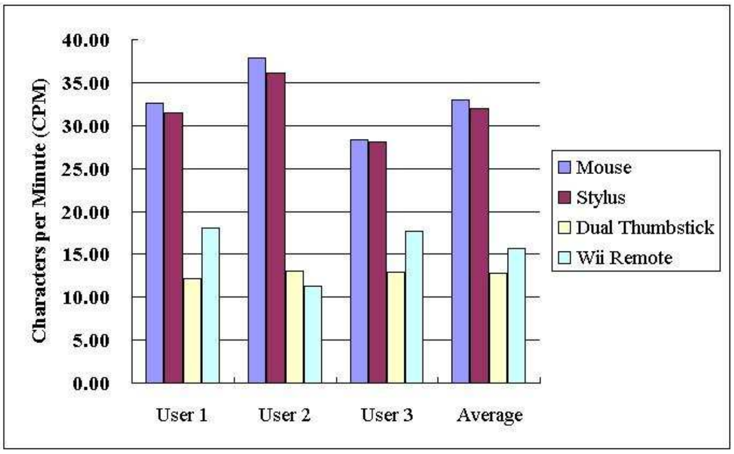


Figure 7-7. Average typing speed of three Myanmar users with Positional Prediction text entry prototype
Image taken from my published paper [38]

7.4.2 User Study and Evaluation of Khmer Language

A mouse and a PC were used for initial user study with three native users. All of the participants were familiar with a PC but didn't have an experience of Khmer text typing with my Positional Prediction text input method. The experiments procedures are as follows:

1) explaining the concept of Positional Prediction text entry,
2) making demonstration of text input with Positional Prediction text input prototype,
3) allowing 5 minutes practice time for each user to learn text input with prototype,
4) recording users' typing speeds of short Khmer message (5 sentences) for 10 times (including error correction time) and
5) discussing with users and learning their responses.

Figure 6-16 shows the Khmer text used for user study. The users' actual typing speed was recorded and CPM (Characters per Minute) was calculated for the evaluation process of my Positional Prediction text input method. Small questionnaires were also conducted to them in order to figure out their feedbacks.

Evaluation by Users' Typing Speed:

The user study results show that average typing speed of Positional Prediction text entry method to finish five Khmer sentences with a mouse is 5 minutes 44 seconds (24.62 CPM). With the current prototype the fastest typing speed is 4 minutes 11 seconds (32.27 CPM) and the slowest typing speed is 9 minutes 48 seconds (13.78 CPM) to finish 5 Khmer sentences that contain 135 characters in total. Figure 7-8 shows CPM comparison for three native users for ten trial times to finish 5 Khmer sentences.

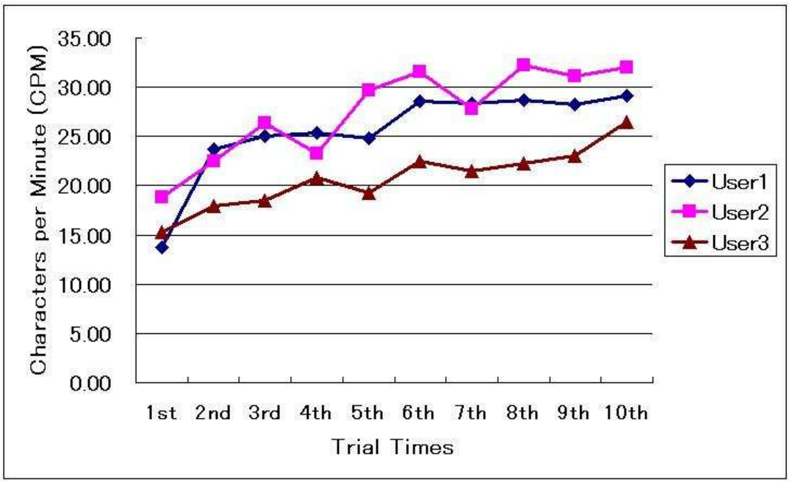


Figure 7-8 CPM comparison for three native users with Positional Prediction prototype
Image taken from my published paper [39]

From my previous studies [69], the average CPM for the NiDA software keyboard layout to finish the same 5 Khmer sentences (Figure 6-16) is 29.14. Figure 7-9 shows the CPM comparison of five native users for the NiDA software keyboard layout. Figure 7-10 shows CPM comparison between Positional Prediction text input method and the NiDA software keyboard input.

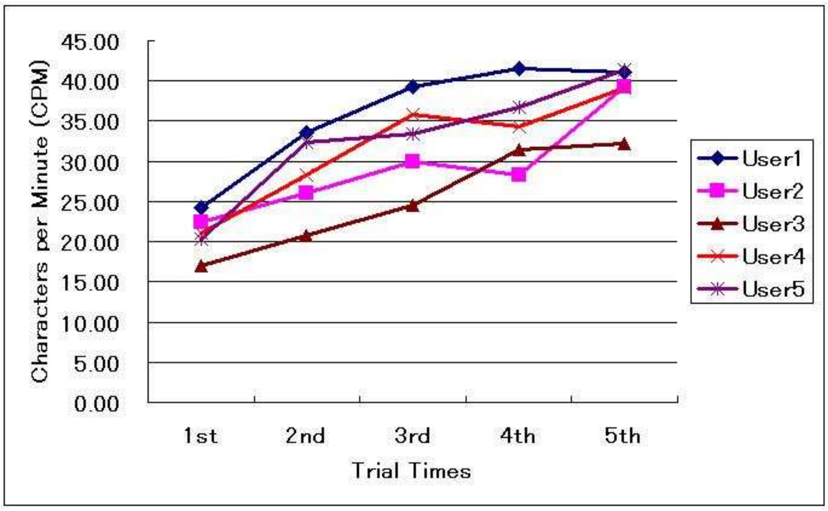


Figure 7-9 CPM comparison for five native users with NiDA software keyboard
Image taken from my published paper [39]

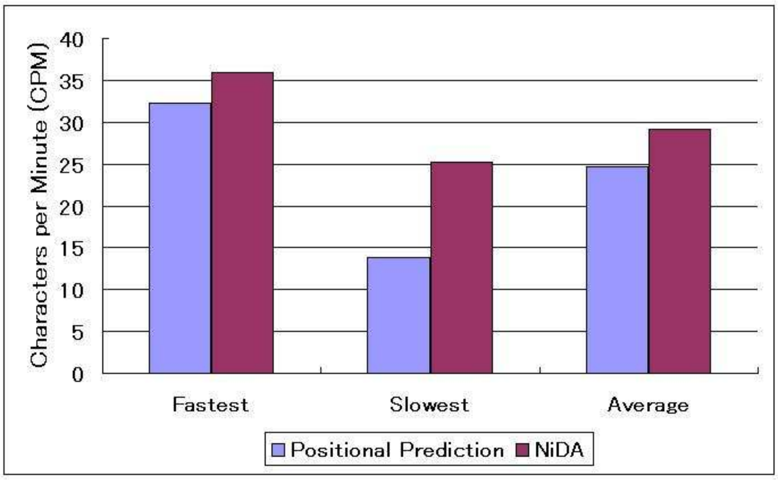


Figure 7-10. Fastest CPM, slowest CPM and average CPM between Positional Prediction and NiDA software keyboard

Image taken from my published paper [39]

Evaluation by Likert Scale Questions

Questionnaires were conducted for users immediately after typing experiments. Four Likert Scales (1-5) were set on which to rate the Positional Prediction. Table 7-1 shows the average results of Likert scale questions. Labels for scale endpoints are in the most left columns and higher values are better. From the questionnaire results, it can be generally said that all three users preferred text entering with Positional Prediction.

Table 7-1. Users' evaluation of Positional Prediction text input (U1, U2 and U3 mean User1, User2 and User3, and higher value is better.)

Table taken from my published paper [39]

Likert Scales (range 1-5)	U1	U2	U3	Average
Difficult-Easy	4	4	4	4
Painful-Enjoyable	4	5	3	4
Slow-Fast	5	4	2	3.67
Dislike-Like	5	4	3	4

7.5 Summary

This chapter has presented a new predicting text entry method that is applicable for many kinds of mobile devices. Initial user study was held for Myanmar and Khmer languages and the results are presented. There is no predictive text entry method for Khmer language yet. And thus, the NiDA software keyboard was chosen for the evaluation process. In average CPM, NiDA is greater than Positional Prediction but not so much different (i.e. 29.14>24.62). Although Positional Prediction is new text input concept, even first-time users can understand and type with appropriate typing speed. But it was noticed that users sometimes faced difficulties in dividing one consonant cluster from a word. However, I believe that this difficulty can be gradually overcome with some typing practice.

CHAPTER 8
Positional Prediction for Clickwheel Mobile Devices (PP_Clickwheel)

8.1 Introduction

Mobile devices such as a mobile phone, a PDA and a music player are now popular in Southeast Asian developing countries such as Cambodia, Laos, Myanmar, Thailand and Bangladesh. However, there are many difficulties to operate with their mother languages because all menus are displayed in English or Chinese, which does not support efficient typing or tapping method for their native languages. I believe that mobile devices will become more popular and necessary for daily work in Southeast Asian developing countries in the future. It is clear that text typing on mobile devices such as sending email, writing notes, searching information and saving the music title is becoming daily work in developed countries. And thus, many researcher of Human Computer Interaction or Human Computer Interface (HCI) area focus on text input interfaces on small mobile devices even for English. To the best of my knowledge, most mobile phone users in Southeast Asian developing countries rarely do SMS messaging with their native languages but use a mobile phone only for speech communication. One of the main reasons is that text typing is not easy on their mobile phones. Users have to memorize the keypad mapping and need much practice to get accustomed to the text typing method. Another reason is that typing speed is slow with annoying process on a mobile phone keypad. In the days to come, mobile devices will become smaller and very important for communication. The investigation of efficient text input interfaces on small mobile devices should not be neglected for mother languages in Southeast Asian developing countries.

This chapter proposes the new text input interface PP_Clickwheel (Positional Prediction with Clickwheel) for Khmer (language of Cambodia). Khmer is a syllabic language, and thus, Positional Prediction (PP) text input concept is used for predicting possible combinations of vowels with a consonant or a syllable [38]. PP_Clickwheel text input interface will be useful for typing similar syllabic languages such as Lao, Myanmar (Burmese), Thai, Nepali, Hindi and Bengali with small mobile devices.

8.2 Related Works

This section presents some text input methods with clickwheel on small mobile devices like the iPod [113, 117].

8.2.1 Keypad on Clickwheel

This input method logically divides clickwheel into eight sections like a mobile phone keypad and detect users' tapping [114, 115]. Keyboard mapping is the same as a mobile phone keypad mapping, e.g. 2 for (a, b, c), 3 for (d, e, f), 4 for (g, h, i) and so on. This method uses eight sections of the click wheel (tapping only and no clicking) for text typing.

Figure 8-1. Keypad layout of iPod clickwheel

Image taken from my published paper [121]

8.2.2 Cursive

Cursive text input method is based on the imagination of writing letters onto the scrollwheel [114, 116]. For example, to write a lowercase letter “a”, from the top right (between the menu and next buttons on 1 and 2G iPods, below the play/pause on 3Gs), start rotating your thumb leftwards around the scrollwheel one complete turn. Then, for the tail of “a”, scroll right one 1/4 turn. Example codes used by the current Cursive text input method are as follows:

Table 8-1 Example codes used by Cursive text input

Table taken from my published paper [121]

Unshifted	Shifted	Sequence
a	A	llllr
b	B	lllro
c	C	lll
d	D	rrrr
e	E	olll

Here, “l” or “r” represents a quarter turn in each direction (“l” represents anti-clockwise, “r” represents clockwise and “o” represents the center button).

8.2.3 Thumbscript

Thumbscript text input method is based on a simple 9 pixel drawing tablet and drawing or tapping pictograph of English alphabets on it [117]. For example, drawing orders for “a”, “b”, “c”, “u”, “w” and “y” can be seen in Figure 8-2. All letters are formed from top to bottom, but when “start” and “stop” are on the same row, the order is from left to right. Reverse drawing or taping order will be used for punctuation, and symbols like reverse drawing order of “c” will be typed “(“ and reverse drawing order of “u” will be typed “¥”.

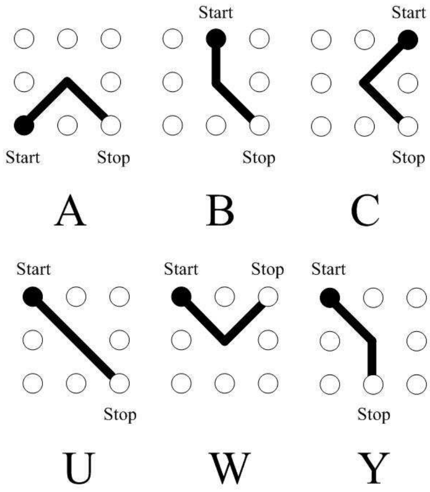


Figure 8-2 Thumbscript text input method

Image taken from my published paper [121]

This method will work with fourth and fifth generation, second generation mini, photo, and color iPods. It uses eight sections of the clickwheel and the action button as a telephone keypad, and rewind and forward buttons to move the cursor [114].

8.2.4 Discussion on Current Clickwheel Text Input Methods

A brief discussion is made here on the current clickwheel text input methods.

"Keypad on Clickwheel" method is usable for alphabetic languages such as English, Spanish and French, but is not directly applicable to Khmer. This is because Khmer is a syllabic language, which has triple numbers of characters (i.e. 74 characters excluding subscript characters) of English. And thus, keyboard or keypad mapping on 9 buttons is a challenge and KSPC (Keystrokes per Character) will be higher [43].

Text input interface of "Cursive" and "Thumbscript" methods are interesting but difficult even for English. First-time users have to memorize the coding for "a" to "z", and several practice times are required for getting used to it. Some coding and tapping pictographs do not look like actual glyph or shape of the English alphabets. "lll" is for "c" or "C", "llll" is for "o" or "O", and "ll" is for "l" or "L", which is easy to understand, however, "lllro" is for "b" or "B", "ror" is for "i" or "I", "lro" is for "k" or "K", "rrrro" is for "m" or "M", and "irrr" is for "n" or "N", which is very difficult for first-time users in Cursive method [116]. Another example is pictographs of "U", "W" and "Y" (see Figure 8-2). Although it is possible to create Cursive coding or Thumbscript tapping pictograph for syllabic scripts e.g. in Khmer language, it may be necessary to design many codes or pictographs (e.g. 33 consonants, 24 dependent vowels, 12 independent vowels and some other directive signs) And another important factor is that the shape or glyph of Khmer is difficult to create pictograph or shorthand writing (see an example of Khmer words in Figure 8-7, Figure 8-16).

Other text input interfaces for mobile devices with clickwheel (e.g. the iPod) have been proposed, e.g. “On-Screen Keyboard”, “Four-Button Telephone Keypad”, “Unicode Hex Input”, “Kana Palette”, “Wheelboard and Multilingual Wheelboard” and “Morse Code” [114]. Most of the input methods mentioned above are based on English alphabet, and not all interfaces take into consideration writing natures of syllabic languages. Among them, “On-Screen Keyboard” and “Wheelboard and Multilingual Wheelboard” input interfaces are applicable to Khmer language, but it may be difficult to distinguish Khmer characters due to their similarity in shape on a limited-sized screen (e.g. “ក” (ka), “គ” (ko), “ត” (ta), “ភ” (pho), etc.)

8.3 Positional Prediction Text Input Interface with PP_Clickwheel

Clickwheel provides only 7 operations or commands, which are left click, right click, up click, down click, center click, clockwise scrolling and anti-clockwise scrolling as shown in Figure 8-3. Therefore, text input method or text input interface was designed for Khmer (i.e. consonants, independent vowels, dependent vowels, punctuations and numbers) with the 7 commands.

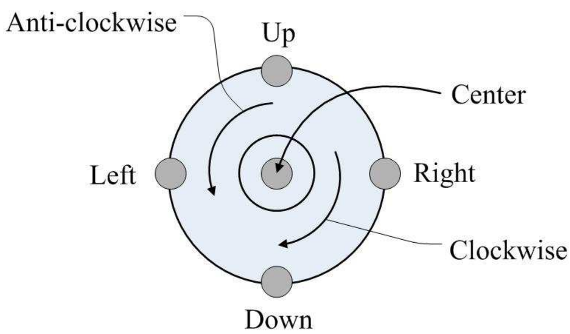


Figure 8-3 Operations provided by clickwheel

Image taken from my published papers [97, 121]

7 commands were assigned for Khmer text typing with clickwheel as follows:

Clockwise and Anti-clockwise Scrolling:
Highlighting a group of characters in the main menu or a character in a candidate list

Left, Right, Up and Down Click:
Giving parameter for vowel combinations (e.g. “ក” (ka) + Right for “កា”)

Center Click:
Selecting a character group or typing a character

The main menu or characters groups are 1) កខគ (consonants), 2) ឲឥឮ (independent vowels), 3) ១២៣ (numbers), 4) ។៕៖ (punctuations, ligatures and symbols), 5) Freq (frequently used Khmer words such as “ខ្ញុំ” (I, me, my), “វា” (he, she, they, it), “គាត់” (he, him, his, she, her), “អញ” (I, me, my) and “នាង” (Miss, young lady, girl, you)) and 6) Mode (for other languages or text editing process) (see Figure 8-4). The typing steps for Khmer syllable are 1) select a group, 2) select a consonant, 3) give parameters for vowel combinations and 4) select a syllable from a suggested candidate list.

8.4 PP_Clickwheel Prototype Implementation

Here, Positional Prediction text entry prototype with clickwheel (PP_Clickwheel) for Khmer language is presented, which was developed with Microsoft Visual Basic. For the implementation, Visual Basic programming language was chosen, which is simple coding and suitable for rapid development. The prototype can be seen in Figure 8-4.

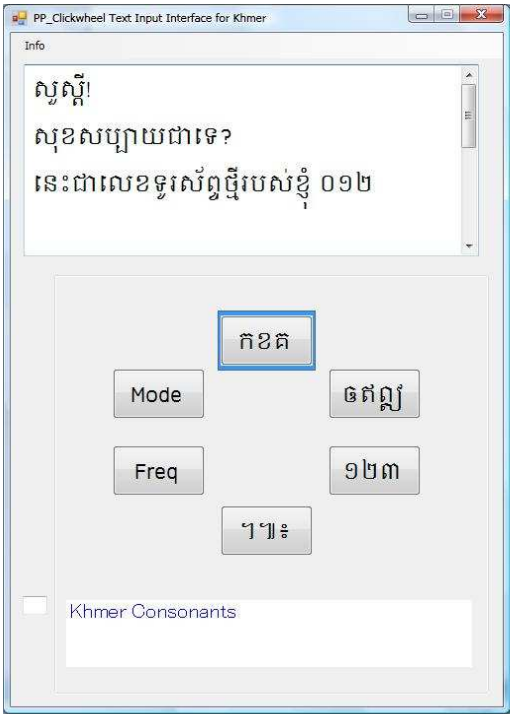


Figure 8-4 Main menu of PP_Clickwheel prototype

Image taken from my published paper [121]

8.4.1 PP_Clickwheel interface

My PP_Clickwheel text input interface for Khmer syllables, words and conjunct consonants is based on the following 4 simple steps:

1) Choose “កខគ” (Consonant) menu from the main menu

2) Choose a consonant that you want to type from a candidate consonant list

3) Give parameter or vowel combination structure. (e.g. [Ka + Left] for “កេ”, “កែ”, “កៃ” and “កើ”, [Ka + Right] for “កា”, “កាំ”, “កេះ”, “កោ”, “កៅ”, “កឿ” and “កៀ”, and [Ka + Down + Up] for “កុំ”, “ក៊ូ”, etc.) This concept is based on Positional Prediction (PP) [38].

4) Select a syllable or a word from a candidate list of suggested possible vowel combinations with the consonant that you have chosen.

Typing steps of “តា” (“Taa”, grandfather) with PP_Clickwheel is as follows:

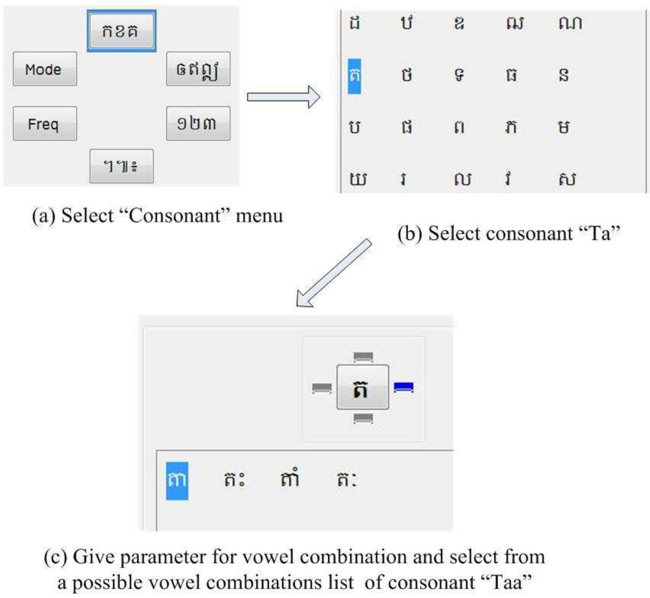


Figure 8-5 Typing steps for "តា" ("Taa" meaning grandfather) with PP_Clickwheel

Image taken from my published paper [121]

8.4.2 Keystroke per character (KSPC)

Keystrokes Per Character (KSPC) is the number of keystrokes or stylus strokes required, on average, to generate each character of text using a given interaction technique in a given language [37] [43] [70]. It is difficult to calculate exact KSPC value for PP_Clickwheel because my prototype uses clickwheel for selecting menu, character, syllable and word. Time required to press a key is different from the one to scroll a clickwheel. And thus, the typing speed was evaluated with Characters Per Minute (CPM); how many characters are typed in a minute, for PP_Clickwheel prototype [43].

8.5 Method

8.5.1 Participants

Ten volunteer participants (6 males and 4 females) were recruited in the Phnom Penh city, Cambodia. Participants ranged from 22 to 35 years (*mean* = 26.4, *sd* = 3.6). All of them were familiar with a PC but had no prior experience of Khmer text typing with Positional Prediction text input method, and it was the first time for them to use clickwheel.

8.5.2 Apparatus

The experiment was conducted with native users in Phnom Penh. A notebook computer equipped with an optical clickwheel mouse (BOMU-RHW01/SWH, Buffalo Inc.) was used for simulation of using mobile device with clickwheel 12) (see Figure. 8-6). Figure 8-16 shows a Khmer text used for user study. The text consists of 109 characters including 44 consonants, 26 vowels, 10 subscript consonants, 3 diacritic signs, 9 numbers and 5 symbols.

Figure 8-6 User study with PP_Clickwheel prototype

Image taken from my published paper [121]

8.5.3 Procedures

The experiments procedures are as follows:

1) explaining the concept of Positional Prediction text entry,
2) making demonstration of text input with PP_Clickwheel text input prototype,
3) allowing 5 to 10 minutes practice time to each user to learn text input with my prototype,
4) recording the users' typing speeds of short Khmer message (5 sentences) for 10 times (including error correction time) and
5) discussing with users and getting their responses.

8.6 Results and Discussion

Typing speed was evaluated with *Characters per Minute (CPM)* instead of *Word per Minute (WPM)* [43]. This is because there is no standard definition for a word in Khmer like in English (i.e. common definition of a word = 5 characters, including spaces) (Yamada, 1980) [43].

8.6.1 Characters Per Minute (CPM)

The user study results show that the average typing speed of first-time users with PP_Clickwheel prototype to finish five Khmer sentences by a clickwheel mouse is 18.9 CPM. With the current PP_Clickwheel prototype, the fastest typing speed is 28.9 CPM and the slowest typing speed is 8.2 CPM. Figure 8-7 shows CPM comparison for ten native users for ten trial times to finish five Khmer sentences.

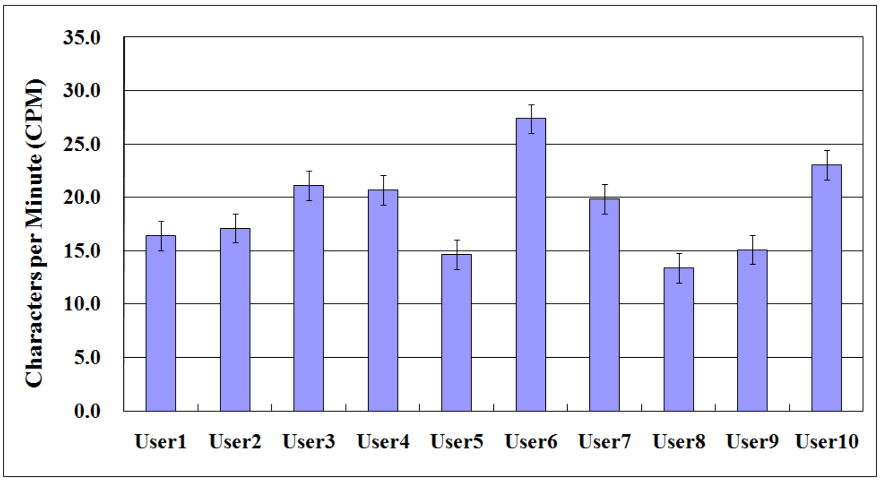


Figure 8-7. Characters per Minute of 10 users with PP_Clickwheel

Image taken from my published paper [121]

8.6.2 Participants Questionnaire

Questionnaires were conducted to the users immediately after the typing experiments. Four Likert scales questions (1 to 5) were set to rate the user-friendliness of PP_Clickwheel text input method. These four scales are (1) Difficult-Easy (2) Painful-Enjoyable (3) Slow-Fast and (4) Dislike-Like. Table 8-2 shows the average or arithmetic mean results of Likert scale questions. From the results, it can be generally said that all of the users preferred text entering with PP_Clickwheel.

Table 8-2 Ten users' evaluation of PP_Clickwheel prototype

Table taken from my published paper [121]

User	**Evaluation with Likert Scales (1 to 5)**			
	Difficult-Easy	**Painful-Enjoyable**	**Slow-Fast**	**Dislike-Like**
1	4	4	4	5
2	4	5	4	4
3	3	2	4	3
4	5	5	5	5
5	3	4	4	4
6	4	5	5	5
7	4	4	4	5
8	3	4	3	4
9	4	5	4	5
10	4	4	5	4
Mean	3.8	4.2	4.2	4.4
Note: Likert Scales (1-5): 1 = worst, 3 = neutral and 5 = best				

8.6.3 Discussion

There is no predictive text entry method for Khmer language yet. Although PP_Clickwheel is a new predictive text input interface, even first-time users can understand and type with appropriate typing speed. However, it has been noticed that users sometimes face difficulties in dividing one consonant syllable from a word. This is because there are various syllable structures such as a consonant, a consonant with a vowel and a consonant with some vowels, and some vowels have a combined structure. As far as I know, there are around 45 syllable structures in total. The followings are 10 Khmer words formed by various syllables:

1. [C+C]: អ + រ = អរ (happy)
2. [C+DV]: ទ + ◌ៅ = ទៅ (go)
3. [C+DV+C]: យ + ◌ា + យ = យាយ (grandmother)
4. [C+DV+C+DV]: ស + េ◌ + រ + ◌ី = សេរី (free)
5. [C+DV+C+DV+C]: ច + ◌ំ + ណ + េ◌ + ញ = ចំណេញ (profit)
6. [C+DV+C+C]: ស + ◌ា + គ + រ = សាគរ (sea)
7. [IDV+C]: ឯ + ង = ឯង (you)
8. [IDV+C+DV]: ឯ + ណ + ◌ា = ឯណា (where)
9. [IDV+C+DV+C]: ឪ + ព + ◌ុ + ក = ឪពុក (father)
10. [C+CS+C]: ប + ◌៉ + ង = ប៉ង (wish)

Here, C = consonant, IDV = independent vowel, DV = dependent vowel and CS = consonant shifter.

It was found that most of the Khmer syllables contained (DV), and thus, positional prediction text input method is suitable. Based on the PP text input concept, users have to type consonants (C) and independent vowels (IDV) one by one. But possible vowel combinations with a consonant (C+DV) are predictable with left, right, up and down clicks. And thus, users have to divide syllables formed by one consonant from a word. For example, users have to divide (C+DV), (C+DV) and C from [C+DV+C+DV+C] syllable. An explanation of this grouping concept is necessary for first-time users but native users can understand within one or two practice times.

Although it can be said that the PP text input concept is applicable to Khmer, it is difficult to define “left vowel”, “right vowel”, “up vowel” and “down vowel” clearly. This is because native people deem some of the vowel combinations as a vowel or a character such as “◌ៅ” (au), “◌ោ” (oo), “◌ើ” (oe), “◌ឿ” (ya) and “◌ៀ” (ie). These vowels are known as “two-part dependent vowel signs” in the Khmer Unicode table. These two-part dependent vowel signs have glyph pieces which stand on both sides of the consonant [7]. In my PP_Clickwheel prototype, users can type these vowels with [Left+Right] or [Right] parameter.

8.7 Summary

This research is in progress, and PP_Clickwheel text input interface has been introduced here. The user study results proved that PP_Clickwheel is a possible text input interface for Khmer language, and Positional Prediction text input method is applicable for clickwheel mobile devices. Positive feedbacks were received from the native users such as “This is more user-friendly than a software keyboard”, “There is no need to memorize the keyboard mapping” and “If I get used to this clickwheel mouse, the speed will be faster”. Further refinements on the current prototype and follow-up analysis are to be made. There is also a plan to extend

PP_Clickwheel prototype for other similar syllabic languages such as Myanmar, Nepali and Thai in the near future [97].

CHAPTER 9
Romanized Handwriting Positional Prediction (Roman_HW_PP)

9.1 Introduction

Graphical user interfaces, web meeting, speech recognition, etc. are reliable, however, writing or typing is still the most common way in human communication with computing devices. Several software keyboards and text input methods have been developed for Myanmar language, yet none of them is designed for non-native users, and therefore, a new Myanmar language text input interface is presented for non-natives. My approach combines Romanized handwriting of Myanmar consonants and Positional Prediction (PP) text input method [38] [39].

This chapter begins with an introduction to the writing system and existing Romanization methods of Myanmar language. Next is an explanation of the concept of Positional Prediction text input method. Then, a detailed explanation of Romanized handwriting text input interface is followed by a report of the user study results with my developed prototype. Finally, a discussion is made on my findings and suggestions for further text input interface research.

9.2 Romanization of Myanmar Language

Currently, there is no Romanization standard for Myanmar language. In Myanmar language, many words are spelled differently from the way they are pronounced. For example, the word for "snack" is pronounced as *tha-ye-sar* (သရေစာ) but spelled as *thwa-ye-sar* (သွားရည်စာ), and therefore, replicating Myanmar sounds in Roman script is difficult [71, 72]. There is a Pali-based Romanization system, but it fails to replicate consonants in contemporary Myanmar language. There are various Romanization methods such as Duroiselle’s System, Latter’s System, Grant Brown’s System, Stewart’s System (IPA), Cornyn’s System (typewritten), Minn Latt’s 1966 System and Myanmar Language Commission’s Pronunciation System (see Table 8-1) [45, 71]. The following table indicates a comparison among existing Romanization methods for Myanmar word "သုတေသန" (research).

Table 9-1. Comparison among Romanization methods for Myanmar word "သုတေသန" (research)

Table adapted from my published papers [72, 120, 130, 131]

Romanized Word	**Romanization Method**
sutesana	Duroiselle’s System
thŏŏtéthană	Latter’s System
thu.te-thăna.	Grant Brown’s System (Conventional)
θúteθəná	Stewart’s System (IPA)
thu.tei-thana.	Cornyn’s System (typewritten)
thutéithăna	Minn Latt’s 1966 System
thu. tei thana.	Myanmar Language Commission's Pronunciation System

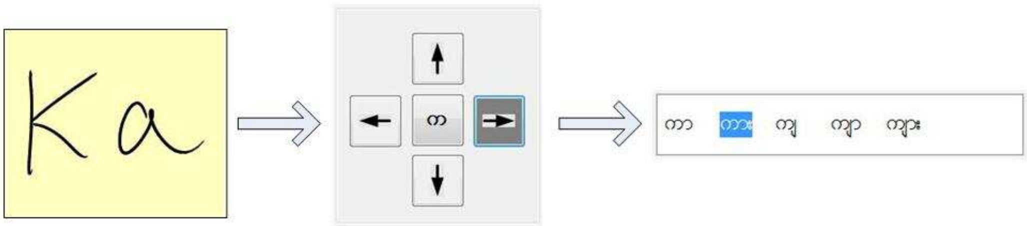
Commonly used Romanization for Myanmar language is sometimes confusing even for native users, and it is hard to apply to text input process on small mobile devices. For example, Burglish (Burmese English) Romanized text input system has no concrete definition for Romanization rule, because it tries to cover all possible or similar pronunciation (“ye” presents “ရီ”, “ရဲ”, “ရည်”, “ရယ်”, “ရေ”, “ရေး”, “ရည့်”, “ရဲ့”, “ယီ”, “ယဲ”, “ယည်”, “ယယ်”, “ယေ”, “ယေး”, “ယည့်”, “ယဲ့”, “လျီ”, “လျဲ”, “လျည်”, “လျယ်” and “လျေ”, “thape” and “thi pf” for typing “သိပ်” (very) and “ta ka Ka thol” for typing “တက္ကသိုလ်” (university)) [73]. And thus, Burglish

Romanization text input method is difficult to apply directly for mobile devices because a lot of candidates are listed and capital letters are used for subscript characters etc.

9.3 Proposal of Romanized Handwriting Text Input Interface

Romanization is supposed to be the easiest text input method for non-native users. In this proposal, Romanization is used only for consonants due to the absence of standard Romanization for Myanmar, which may cause ambiguities in Romanizing Myanmar words. Another reason is that if Romanization is used for the whole Myanmar language like Japanese Romaji text input method, non-native users need to know the pronunciation of all Myanmar words. My Romanized handwriting text input interface for Myanmar is based on the following 3 simple steps (see Figure 9-1):

1) Write a Roman character for a consonant.
   (e.g. ka for "က", kh for "ခ", ga for "ဂ", gh for "ဃ", **ng** for "င", etc.)
2) Give parameter for vowel combination structure.
   (e.g. [kh + Left] for "ခေ", "ခြ", and "ခြေ", [ma + Right] for "မာ", "မား", "မျ", "မျာ", and "များ" and [pa + Down + Up] for "ပုံ", "ပွံ", "ပို", "ပွီ" and "ပွိ")
3) Select a syllable or consonant cluster from a list of possible vowel combinations with the consonant that you wrote in Roman alphabets.



(1) Write a Roman character for a consonant "ka"

(2) Give parameter for vowel combination

(3) Select from a possible vowel combinations list of a consonant "ka"

Figure 9-1. Typing Myanmar word "ကား" (Car) with Romanized handwriting text input interface

Image taken from my published papers [120, 130]

9.4 Prototype Implementation

Here, Myanmar language Romanized handwriting prototype for non-native users is presented, which was developed with Microsoft Visual Basic. For the implementation, Visual Basic programming language was chosen, which is simple coding and suitable for rapid development.

In my prototype, new features are introduced for easier and faster Myanmar language typing, i.e. "SD", "=R" and "LRUD" buttons. Here "SD" button is for deleting typed syllable, "=R" button is for repeat typing and "LRUD" is for editing vowels. Although exiting Myanmar software keyboards have "Delete" and "Backspace" buttons, these buttons allow users to delete only one character. A Myanmar word is formed by the combination of syllables such as "ကျောက်စိမ်း" (jade) ("”ကျောက်"+"စိမ်း"), "ကိုယ်လေးလက်ဝန်" (advanced stage of pregnancy) ("ကိုယ်"+"လေး"+"လက်"+"ဝန်") and ဘုန်းကြီးကျောင်း" (monastery) ("ဘုန်း”+"ကြီး"+"ကျောင်း"). And most syllables contain more than one character (e.g. seven characters in "ကျောင်း" syllable). And thus, it is believed that "deleting typed syllable" or "SD" button is important for Myanmar language text editing. "=R" button or "repeat typing of lastly typed character or syllable" is also important for Myanmar language because many Myanmar words and names are formed by repeating the same character or syllable such as "မမ" (sister), "ဟောဟောဒိုင်းဒိုင်း" (openly, frankly), "လောက်လောက်လားလား" (dependable), "ရရစားစား" (just make both ends meet) and "စုစု" (common Myanmar female name). "LRUD" or "Left, Right, Up and Down" button is for adding or editing dependent vowels with 4 directional arrow keys. Users can add missed vowels with "LRUD" button, even if they selected similar syllable in the suggested candidate list and typed by mistake. "Frq" or "Frequently used Myanmar words" is for typing frequently used Myanmar words such as "ကျွန်တော်" ("I" for male), "ကျပ်" ("Kyat" Myanmar currency unit) and "သည်" (word indicating the verb ending of a Myanmar sentence). Figure 9-2 shows the Romanized handwriting text input prototype.

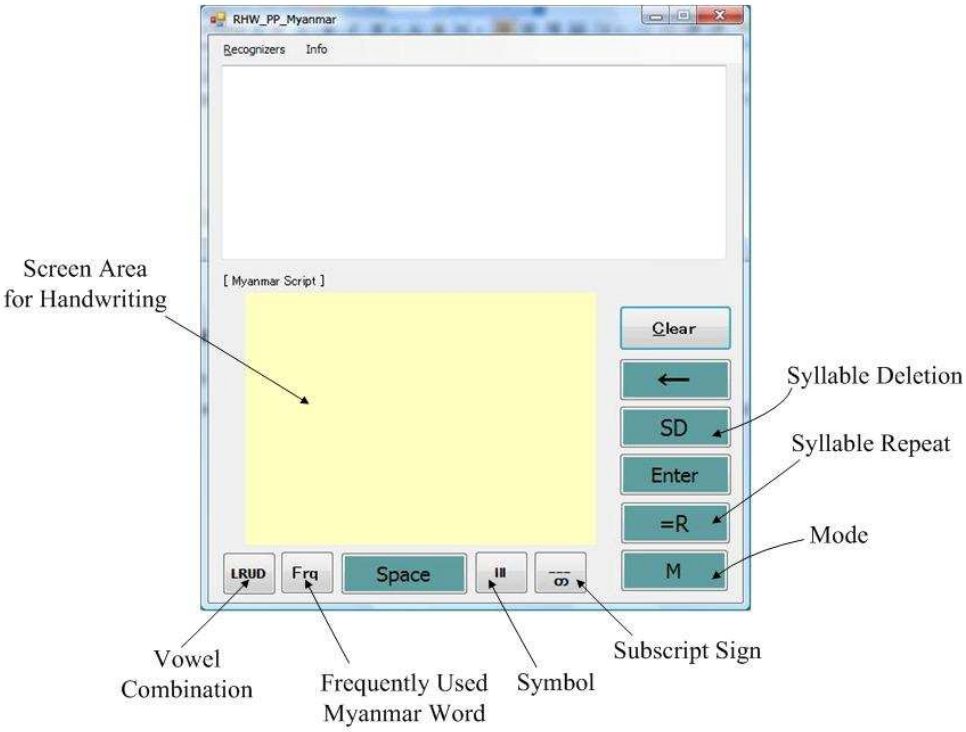


Figure 9-2. Romanized handwriting text input prototype for Myanmar language

Image taken from my published papers [120, 130]

9.5 User study

9.5.1 Participants

User study was conducted with 5 natives (3 males and 2 females) and 5 non-natives (3 males and 2 females). Native participants were recruited in Yangon, Myanmar and non-native participants were recruited from the local university campus. Non-native participants consisted of a Mongolian, a Chinese, a Japanese, a Cambodian and a Thai. Ages ranged from 24 to 36 with an average of 29.00 (SD = 4.6). All of them were familiar with a PC but had no prior experience of Myanmar text typing with Romanized Handwriting text input method, and seven of them (4 natives and 3 non-natives) had no experience of using a stylus and a touch screen.

9.5.2 Apparatus

The experiment with native users was conducted in Yangon, Myanmar and the one with non-native users was conducted in my university laboratory located in Tokyo, Japan (see Figure 9-3). A 1.06GHz Intel Core 2 Duo CPU L7500 IBM Lenovo X61 tablet PC running Windows Vista Ultimate OS with 2006 MB RAM was used. Users were required to use a stylus pen for text input experiment with a tablet PC. A Myanmar text used for user study consisted of 107 characters including 41 consonants, 52 vowels, 7 numbers, 6 symbols and 1 space. Figure 6-15 shows the Myanmar text used for user study.

Figure 9-3. User study with non-native users
(from left to right: Japanese, Thai, Mongolian, Chinese and Cambodian)

9.5.3 Procedures

The experiment procedures are (1) explaining the concept of Positional Prediction text entry, (2) explaining the consonant Romanization, (3) making demonstration of Romanized handwriting text input with prototype, (4) allowing 5 to 10 minutes practice time to users to learn text input with prototype, (5) recording their typing speed of short Myanmar message (6 sentences) for 10 times (including error correction time) and (6) discussing with them and getting their evaluation.

9.6 Results and Discussion

Typing speed was evaluated with *Characters per Minute (CPM)* instead of *Words per Minute (WPM)* [43]. This is because there is no standard definition for a word in Myanmar like in English (i.e. common definition of a word = 5 characters, including spaces) (Yamada, 1980) [44].

9.6.1 Characters Per Minute (CPM)

The user study results show the average typing speed of first-time users with Romanized handwriting text input prototype to finish 6 Myanmar sentences, which is 27.2 CPM by natives and 26.9 CPM by non-natives. With the current prototype, the fastest typing speed of natives is 39.6 CPM and that of non-natives is 35.5 CPM, and the slowest typing speed of natives is 16.8 CPM and that of non-natives is 12.4 CPM. Figure 9-4 shows CPM comparison among 5 non-natives for 10 trial times to finish 6 Myanmar sentences. It was found that there is no big difference in typing speed improvement between natives and non-natives (see Figure 9-5).

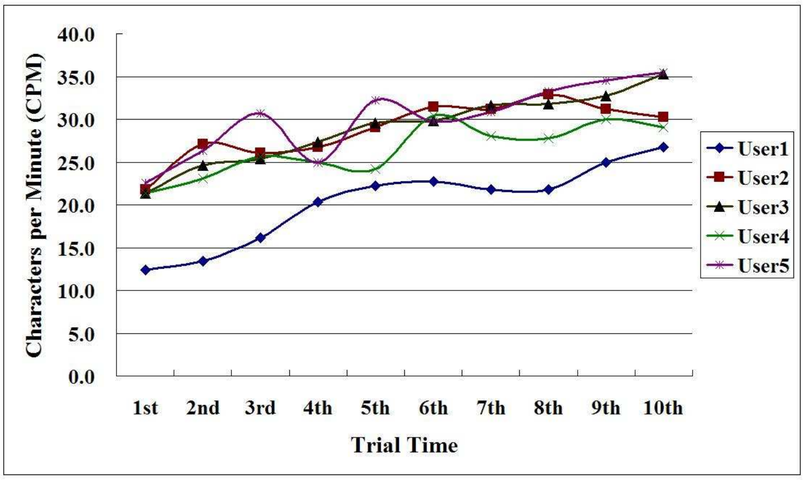


Figure 9-4. Characters Per Minute (CPM) of five non-natives

Image taken from my published paper [130]

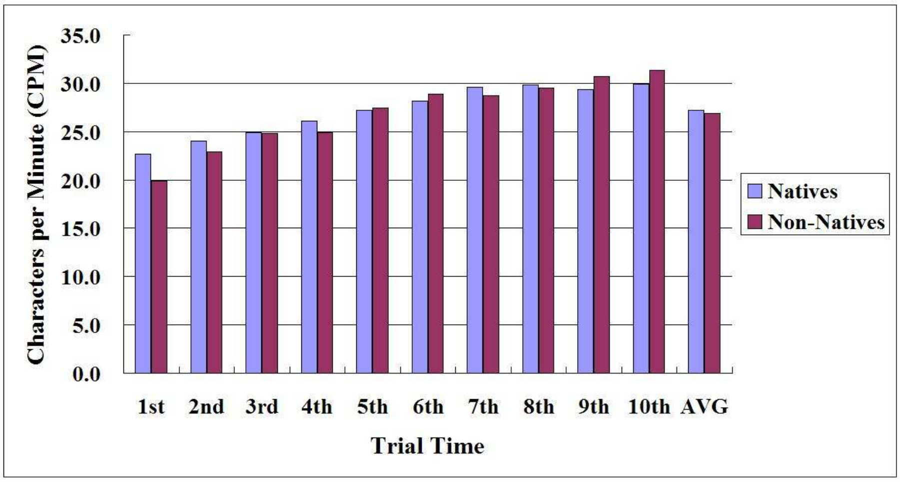


Figure 9-5. Characters Per Minute (CPM) of natives and non-natives

Image taken from my published paper [130]

9.6.2 Participants Questionnaire

Questionnaires were conducted to the participants immediately after the typing experiments. 4 Likert scales (1 to 5) questions were set to rate the user-friendliness of Romanized handwriting text input method. The scales are (1) difficult-easy (2) painful-enjoyable (3) slow-fast and (4) dislike-like. Table 9-2 shows the average or arithmetic mean results of Likert scale questions. From the results, it can be generally said that all the participants were fond of text entering with Romanized handwriting text input prototype.

Table 9-2. Evaluation by natives and non-natives

Table taken from my published paper [130]

Likert Scale (Range 1 to 5)	**Natives**	**Non-Natives**
Difficult-Easy	4.4	4.2
Painful-Enjoyable	3.6	4.4
Slow-Fast	3.4	3.8
Dislike-Like	4.2	5.0
Average	**3.9**	**4.3**

9.7 Discussion

The Evaluation of my Romanized handwriting text input prototype was made with CPM not with *Keystrokes Per Character (KSPC)*. KSPC is the number of keystrokes or stylus strokes required, on average, to generate each character of text using a given interaction technique in a given language [37] [43]. It is difficult to calculate exact KSPC value for Romanized handwriting text input interface because my prototype uses Romanized handwriting for Myanmar consonants, pressing software button (i.e. LRUD button) for predicting vowel combinations and selecting a syllable from a candidate list. Therefore, the typing speed was evaluated with CPM; how many characters are typed in a minute, for my prototype [43].

As shown in Figure 9-5, the difference of average typing speed between non-natives and natives is only 1.11%. However, natives performed better in the fastest and slowest typing speed than non-natives. From user study, it was found that Romanization for some Myanmar consonants is still difficult for natives. This is because "Romanization of Myanmar Unicode version 5.1" was used for Myanmar consonants to reduce ambiguities such as "sa" for consonant "သ", "ca" for consonant "စ" and "ra" for consonant "ရ". [6] In commonly-used Romanization, "tha" is for "သ", "sa" is for "စ" and "ya" is for consonants "ယ" and "ရ". And thus, natives need to memorize the Unicode Romanization, and they can sometimes make mistakes in Romanization. This fact can influence typing speed of natives. For non-natives, highlight was made on each consonant character to divide consonant and vowels clearly, because it was the first time for them to see Myanmar characters. From the results of Likert scale evaluation by non-natives, it can be said that my approach is a possible solution to find out user-friendly Myanmar language text input interface for non-natives (see Table 9-2). Positive feedbacks were also received from the natives such as "This is more user-friendly than a software keyboard", "There is no need to memorize the keyboard mapping" and "I don't know how to type with the current Myanmar keyboard layout but this method is very easy for me".

Although Romanization or Latinization was used in this prototype, this kind of text input interface is applicable to languages of non-natives (i.e. transliteration). Transliteration or mapping of 33 Myanmar consonants into another language is possible. For example, if a mother tongue of non-natives is Japanese, Myanmar consonants can be written with Japanese hiragana or katakana such as "か" or "カ" for "က" (ka), "かは" or "カハ"for "ခ" (kh), "が" or "ガ" for "ဂ" (ga), "がは" or "ガハ" for "ဃ" (gh) and "んが" or "ンガ" for "င" (ng).

9.8 Summary

It was proven that Positional Prediction text input concept is applicable to non-natives of Myanmar language. It is believed that Romanization handwriting text input interface is useful for small mobile devices such as a PDA, a mobile phone and an electronic dictionary, and that my approach is applicable to other Asian syllabic languages. The limitation of this text input approach is that non-natives are required to memorize 33 Myanmar consonants together with equivalent Romanization and have knowledge of how to write English alphabets. In the near future, there is a plan of user study with non-natives who are learning Myanmar language.

CHAPTER 10
Positional Prediction for Fingerspelling

Information and communication technologies (ICT) should be available not only for normal users but also for users with disabilities. Gap control is still needed especially for users with disabilities in developing countries. This chapter presents font development, the Direct Keyboard Mapping (DKM) and Romanized Positional Prediction (RomanPP) keyboard prototype for Myanmar language fingerspelling characters. Moreover, usability of my approach is discussed based on the user study results with the current fingerspelling software keyboard prototype. The evaluation was made in terms of typing speed CPM (characters per minute) and feedbacks from three types of users, i.e., hearing-impaired users, general users and desktop publishing staff. I believe that the outcome of this research is useful for design of Myanmar fingerspelling keyboard layout, creation of fingerspelling educational contents and communication between hearing-impaired and general users.

10.1 Introduction

Myanmar fingerspelling is the representation of Myanmar characters and numbers with hands in Myanmar sign language. It is a necessary communication method for hearing-impaired or deaf people, and is used especially for signing names, city names and words, which do not exist in sign language. The motivation for this research is the present state that hand drawn pictures are used in sign language textbooks because there is no fingerspelling font or keyboard layout for Myanmar fingerspelling. Here, two text input interfaces are proposed for Myanmar fingerspelling using my developed TrueType font for Myanmar fingerspelling text input. One is a keyboard mapping of fingerspelling characters based on the current Myanmar PC keyboard layouts called "Direct Keyboard Mapping (DKM)", and the other is Romanized text input method based on the Positional Prediction called "Romanized Positional Prediction for Myanmar Fingerspelling (RomanPP_Fingerspelling)". User study was conducted with a software keyboard prototype to evaluate the DKM and RomanPP_Fingerspelling keyboard layouts. The main focus is on user-friendliness of DKM and RomanPP_Fingerspelling as well as user interface of the DKM fingerspelling software keyboard, i.e., which is better labeled by Myanmar character or labeled by fingerspelling. The evaluation was made based on typing speed CPM and feedbacks from hearing-impaired users, general users and desktop publishing staff. The results show that the mentioned layouts provide appropriate typing speed and are highly accepted not only by hearing-impaired users but also by general users and desktop publishing staff.

10.2 Myanmar Fingerspelling

For all I know, there are two different fingerspelling character sets for Myanmar language; one is used in northern Myanmar (e.g. used at "Mandalay School for the Deaf" in Mandalay city) and the other is used in southern Myanmar (e.g. used at "Mary Chapman School for the Deaf" in Yangon city). The latter was invented by Dr. Maliwan Tammasaeng in collaboration with Myanmar sign language teachers and students in 1987 [74]. The mentioned two are similar in consonant but mainly different in vowel, medial and symbols. Myanmar fingerspelling consonants are based on the American Manual Alphabet such as "a" for "" (အ), "b" for "" (ဘ), "d" for "" (ဒ), "z" for "" (ဇ) [74], [75]. In this chapter, fingerspelling characters set for font development is used from a Myanmar sign language dictionary book published by Department of Social Welfare, Ministry of Social Welfare, Relief and Resettlement [76], because the fingerspelling characters set mentioned in the book was recognized as a standard in 2007. Myanmar fingerspelling characters used in this research are as follows:

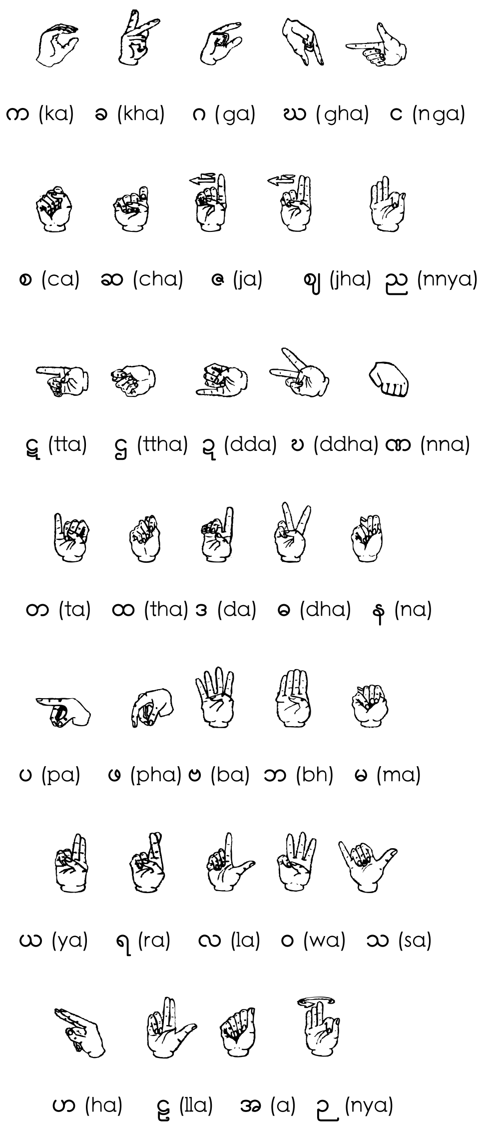


Figure 10-1. Consonant

Image adapted from my published papers [132, 133]

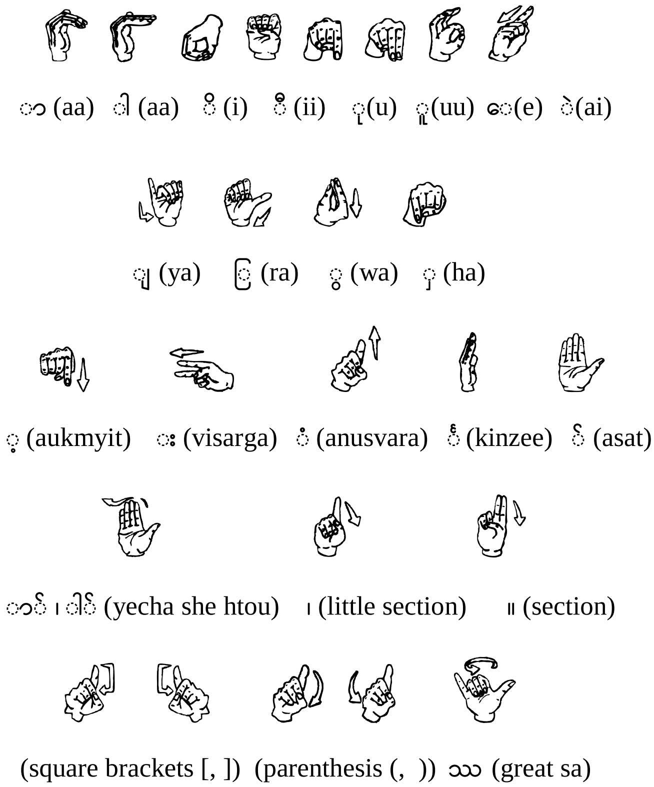


Figure 10-2. Vowel, consonant sign, various sign and symbol

Image adapted from my published papers [132, 133]

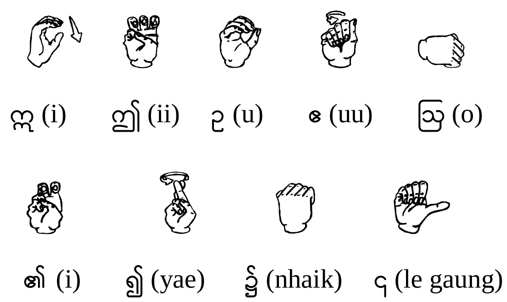


Figure 10-3. Independent vowel and symbol

Image adapted from my published papers [132, 133]

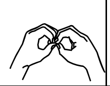
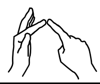
The presentation order of fingerspelling is the same as Myanmar language handwriting order. The main difference is that glyph and writing position of fingerspelling characters never change according to what it surrounds unlike, for example, the width of Myanmar character medial ြ (ra), which changes to ကြောက် (fear) and ခြောက်(frighten), using only the glyph  in fingerspelling.

10.3 Related Works

There exist British fingerspelling font namely "BDA (British Deaf Association) Fingerspelling" and American Sign Language (ASL) fingerspelling font namely "Gallaudet TrueType Fingerspelling", and they enable typing with the QWERTY and Dvorak keyboard layouts [77][78]. Although the predominant spoken language of the United Kingdom and the United States is English, BDA fingerspelling and ASL fingerspelling are quite different from each other. BDA fingerspelling expresses with two handed except alphabet "C (c)", and ASL fingerspelling uses one handed manual alphabet (see Table 10-1). As far as I know, there is no such font development for Myanmar fingerspelling yet.

Table 10-1. BDA and ASL fingerspelling alphabets "A (a)" to "E (e)"

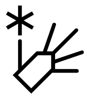
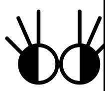
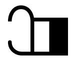
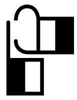
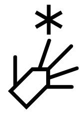
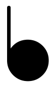
Table adapted from my published paper [133]

Alphabet / **Fingerspelling**	A (a)	B (b)	C (c)	D (d)	E (e)
British Deaf Association					
American Sign Language					

SignWriting is a system of writing sign languages using a combination of iconic symbols or abstract pictures for handshapes, facial expressions, body locations, contacts, movements, etc. [79]. It was developed in 1974 by dancer Valerie Sutton and already applied to BDA and ASL fingerspelling [80] (see Table 10-2). "Sutton UK" and "Sutton US" fonts were used for typing test of SignWriting alphabets, and it was learned that SignWriting alphabets of BDA and ASL can be typed with the QWERTY or Dvorak keyboard like BDA and ASL fingerspelling.

Table 10-2. SignWriting for BDA and ASL fingerspelling alphabets “A (a)” to “E (e)”

Table adapted from my published paper [133]

English Alphabet	A (a)	B (b)	C (c)	D (d)	E (e)
Sutton UK **(BDA)**					
Sutton US **(ASL)**					

Currently, SignWriting is not used in Myanmar deaf education, and a special word processor needs to be developed for SignWriting typing. In this research, the focus is only on text input method of fingerspelling characters and not on whole Myanmar sign language, and thus, the design of BDA and ASL fingerspelling fonts as well as their keyboard mappings are taken into consideration.

Finger-Chat system is an instant messaging system designed to support teaching of fingerspelling alphabets, which covers BDA fingerspelling, SignWriting fingerspelling and English typing [81], and this system was also studied to design my software keyboard prototype.

10.4 Direct Keyboard Mapping (DKM)

Direct Keyboard Mapping (DKM) is a method to map Myanmar fingerspelling characters based on the current Myanmar PC keyboard layouts. It is assumed that DKM is easier and short learning curve for users who are already familiar with Myanmar keyboard layouts. Basically, it is one to one mapping and simple process compared to designing a new keyboard layout, but the main challenge is selecting one of the Myanmar PC keyboard layouts for DKM. This is because there are many keyboard layouts or key mappings and typing methods for Myanmar language today. Old Myanmar PC keyboard layouts and typing methods are totally based on ASCII encoding, and existing ones are moving to Unicode encoding (although some of which are not fully Unicode standard encoding yet). Old Myanmar keyboard layouts consider various glyphs of the same medial or possible combinations of the medial with other medials (e.g. “ြ”, “ြ”, “ြ –”, “ြွ”, “ြှ”, etc.) and subscript consonants (e.g. “–္တ”, “–္ခ”, “–္ဂ”, “–္ဓ”, “–္ဏ”, “ၬ”, etc.) However, Unicode encoding automatically renders the appropriate glyphs depending on the character combination, and therefore, keyboard mapping for different glyphs of a character is not necessary in Unicode keyboard layout. Moreover, each fingerspelling character has only one glyph, and thus, there is no need to consider glyph changing process and keyboard mapping for different glyphs. For those reasons, it is possible to use one of the existing Unicode keyboard layouts for fingerspelling characters.

As a first step, existing Myanmar PC and typewriter keyboard layouts were studied. From this study, it can be said that almost all Myanmar PC keyboards mappings are based on traditional typewriter and similar keyboard mapping especially for consonants, but some different mappings exist for vowels, medials and subscript consonants (see Table 10-3) [82]. Following this result, it was decided to use "Myanmar3 keyboard layout", which is one of the existing Myanmar Unicode keyboard layouts developed by Myanmar Unicode and NLP (Natural Language Processing) Research Center [29].

Table 10-3. Keyboard mapping of WinMyanmar, Zawgyi Myanmar and Myanmar3 keyboard for 'q', 'w', 'e', 'r', 't' and 'y' keys

Table adapted from my published paper [82]

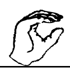
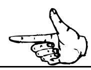
Key (Shift) (Unshift)	Q q	W w	E e	R r	T t	Y y
WinMyanmar **ASCII Ver. 2.6**	ျှ ဆ	ျွှ တ	[illegible] န	ျွ မ	ွှ အ	[illegible] ပ
Zawgyi Myanmar **(Unicode)**	ျှ ဆ	ျွှ တ	[illegible] န	ျွ မ	ွှ အ	[illegible] ပ
Myanmar3 **(Unicode)**	ဈ ဆ	ဝ တ	ဣ န	၎ မ	ဤ အ	၌ ပ

In this chapter, consideration is made mainly on keyboard mapping for consonants, vowels and medials because these characters are very important in Myanmar word formation. All of the fingerspelling consonants are assigned the same as the Myanmar3 keyboard layout.

Table 10-4 shows keyboard mapping for Myanmar fingerspelling consonants “က” (ka) to “င” (nga). For the keyboard mapping of Myanmar fingerspelling vowels, consonant signs and various signs, there is no need to consider different glyphs of those characters (e.g. different glyph of Medial “Ra” such as“ြ”, “ြ”, “ြ” and “ြ”). This is because all of the fingerspelling characters have only one glyph except “ါ” (tall aa). Other combined vowels, e.g., “ာ်” (aa + asat or yecha shae htoo), “ါ်” (tall aa + asat or yecha shae htoo)” are defined as a fingerspelling character (see Figure 10-2). For other Myanmar fingerspelling characters such as independent vowels and symbols, it is possible to follow one of the existing Myanmar keyboard layouts or consider a new keyboard mapping.

Table 10-4. Direct keyboard mapping for Myanmar fingerspelling consonant “က” (ka) to “င” (nga)

Table adapted from my published papers [132, 133]

Myanmar Consonant	က	ခ	ဂ	ဃ	င
Equivalent Fingerspelling					
Key	u	c	Shift 8	Shift c	i

10.5 Romanized Positional Prediction for Fingerspelling (RomanPP_Fingerspelling)

RomanPP is a Romanized text input interface based on the concept of the Positional Prediction [83]. Here, brief explanation is made on the challenge of Romanization for Myanmar language and RomanPP text input interface for Fingerspelling.

10.5.1 Romanization of Myanmar Language

In Myanmar language, many words are spelled differently from the way they are pronounced. For example, the word for "snack" is pronounced tha-ye-sar (သရေစာ) but spelled thwa-ye-sar (သွားရည်စာ), and therefore, replicating Myanmar sounds in Roman script is difficult [71], [72]. There is a Pali-based Romanization system, but it fails to replicate consonants in contemporary Myanmar language. There are various Romanization methods such as Duroiselle’s System, Latter’s System, Grant Brown’s System, Stewart’s System (IPA), Cornyn’s System (typewritten), Minn Latt's 1966 System and Myanmar Language Commission's Pronunciation System [45, 71]. The current Romanization rules for Myanmar language are difficult to understand for first-time users, and it will be more difficult to apply for text input process on small mobile devices. For example, Burglish Romanized text input system has no definite or concrete definition for Romanization rule for their system because they try to cover all possible or similar pronunciation ("ye" represents "ရီ", "ရဲ", "ရည်", "ရယ်", "ရေ", "ရေး", "ရည့်", "ရဲ့", "ယီ", "ယဲ", "ယည့်", "ယယ် ", "ယေ", "ယေး", "ယည့်", "ယဲ့", "လျီ", "လျဲ", "လျည်", "လျယ်" and "လျေ", "thape" and "thi pf" for typing "သိပ်" and "ta ka Ka thol" for typing "တက္ကသိုလ်") [73].

10.5.2 RomanPP

To reduce the ambiguities of Romanization, RomanPP text input method uses Romanization only for Myanmar consonants, independent vowels, special characters and punctuation. In other words, RomanPP is not Romanization for Myanmar words. The Romanization for Myanmar consonants are based on Unicode naming system, but only 2 or 3 English characters were used for a Myanmar consonant to reduce number of keystrokes to type a character. The followings are the Romanization used in this chapter:

က (ka) ခ (kh) ဂ (ga) ဃ (gh) င (ng)

စ (ca) ဆ (ch) ဇ (ja) ဈ (jh) ည (nn)

ဋ (tt) ဌ (tth) ဍ (dd) ဎ (ddh) ဏ (nna)

တ (ta) ထ (th) ဒ (da) ဓ (dh) န (na)

ပ (pa) ဖ (ph) ဗ (ba) ဘ (bh) မ (ma)

ယ (ya) ရ (ra) လ (la) ဝ (wa) သ (sa)

ဟ (ha) ဠ (ll) အ (a) ဉ (ny)

The text input process of RomanPP is almost the same as PP but the different is typing Romanization of a consonant (e.g. "la" + → for typing "လာ", "လား", "လျ", "လျာ", "လျား", etc.)

10.5.3 RomanPP for Myanmar Fingerspelling

Thinking of the fact that few Myanmar computer users are familiar with Myanmar PC keyboard layouts, RomanPP is another possible text input interface for typing Myanmar fingerspelling.

This is because Myanmar computer users are used to the QWERTY keyboard for English typing.

“s”, “f”, “e”, “c” and “j”, “l”, “i”, “m” keys were used for typing parameter of vowel combinations instead of four directional arrow keys. If four directional arrow keys are used for predicting possible vowel combinations, users need to move their right hand from the home keys (a, s, d, f, j, k, l and ;) to the positions of the four directional arrow keys. Taking into consideration that this moving time will affect users’ typing speed, it was decided to use alphabet keys near home keys. Here, “s”, “f”, “e” and “c” keys are for left-handed users, and “j”, “l”, “i” and “m” keys are for right-handed users (see Figure 10- 4).

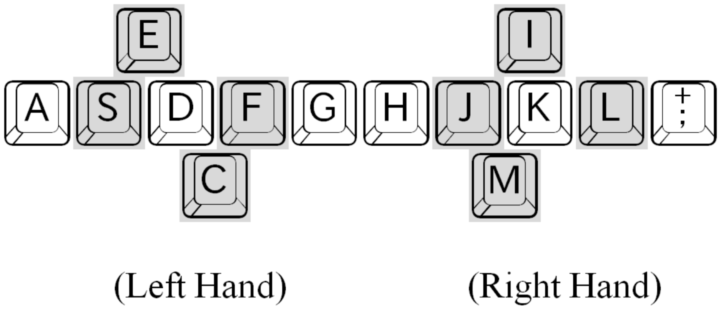


Figure 10-4. Four directional arrow keys assignment for RomanPP_Fingerspelling

10.6 Prototype Development

Font developing process for Myanmar fingerspelling and software keyboard prototype for DKM and RomanPP_Fingerspelling are presented here. Prototype was developed with Microsoft Visual Studio 2005, and Visual Basic programming language was chosen for the prototype implementation, which is simple coding and suitable for rapid development.

10.6.1 Font Development for Myanmar Fingerspelling

As far as I know, no font had been developed for Myanmar fingerspelling characters. As a first step, TrueType fonts were developed for Myanmar fingerspelling before software keyboard implementation. The reason why TrueType fonts were selected rather than other font types such as PostScript and OpenType [84] is because fonts that are easy to read and print well on the monitor screen are enough. Moreover, TrueType fonts worked fine on Windows as well as Mac. A scanner (Sharp UX-MFSOCL), Adobe Illustrator CS3 version 13.0.0 and FontLab Studio version 5 were used, and a fingerspelling font was created by the following steps:

1) Scanning the Myanmar fingerspelling character set with 600 dpi from a Myanmar sign language dictionary book,
2) Tracing each fingerspelling character and then converting to vector format using Adobe Illustrator CS3,
3) Copying the vector image from Adobe Illustrator CS3 and pasting it to the FontLab Studio (Here, name and/or Unicode code point is assigned to each fingerspelling glyph), and then, adjusting each fingerspelling character's path and
4) Generating TrueType fonts (mmFingerspelling.ttf) and testing all fingerspelling characters.

Note: There are many possible ways to develop a font, and among them, the font creation steps above were used for development of Myanmar fingerspelling font (see Figure 10-5).

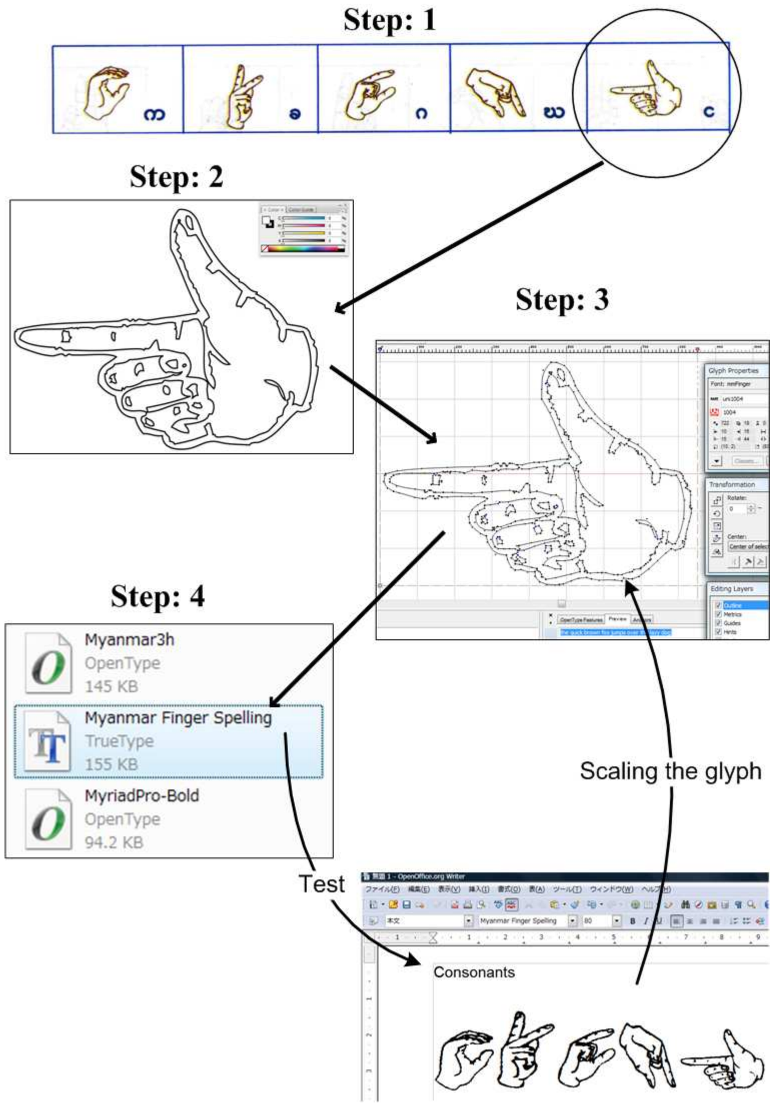


Figure 10-5. Font developing steps for fingerspelling “Nga”

Image adapted from my published papers [132, 133]

(Step:1 Scan fingerspelling “Nga” picture, Step:2 Trace and convert to vector format, Step:3 Copy the vector image and paste it to FontLab, and then, adjust “Nga” character path with other fingerspelling characters and Step:4 Generate true type font and test with OpenOffice.org Writer)

To my knowledge, there are several font rendering engines such as Pango, Microsoft’s Uniscribe engine, IBM’s ICU (International Components for Unicode), QT and FreeType font engine [84, 85, 86, 87, 88, 89], and among them, IBM’s ICU (test with OpenOffice.org Writer) and Microsoft’s Uniscribe engine (test with notepad) were used for measuring and positioning all fingerspelling characters.

In the designing process of fingerspelling font, writing position of each character does not need to be considered. In Myanmar language, characters are written in three levels (upper, middle and lower) and most characters have their defined positions (e.g. “ေ◌–” and “ြ◌” should be written as a front vowel, “◌ု”, “◌ူ”, “◌ှ”, “◌ှု”, “◌ွ”, “◌့”, etc. should be written as a lower character and “◌ိ”, “◌ီ”, “◌ံ”, etc. should be written as an upper vowel and “◌ာ”, “◌ျ”, “◌း”, etc. should be written on the right side of consonants). In the case of fingerspelling, however, all of the characters including vowels and medials are represented from left to right in one row. Therefore, the step of adjusting each character’s paths of fingerspelling is not so complex compared to that of Myanmar characters. My developed fingerspelling font (mmFingerspelling.ttf) is used to display Myanmar fingerspelling characters in this chapter (see Figure 10-1, Figure 10-2, Figure 10-3, Figure 10-4 and Table 10-4).

10.6.2 DKM Software Keyboard Development

In the DKM prototype, the focus is on Myanmar consonants, vowels, medials and symbols that are defined as a standard fingerspelling character set (see Figure 1, Figure 2 and Figure 3) [76]. Based on the Myanmar3 keyboard layout, 33 fingerspelling characters (20 consonants, 9 independent and dependent vowels, 2 Medials and 2 Symbols) were assigned to the unshift mode, and 23 fingerspelling characters (14 consonants, 3 independent and dependent vowels, 2 medials and 4 symbols) were assigned to the shift mode (see Figure 10-6 and Figure 10-7).

Keyboard layout can be switched to display fingerspelling or Myanmar characters. The selected keyboard layout options only affect the appearance of the software keyboard keys and do not affect the output text (i.e. the output text always shows fingerspelling characters). Basically, only 1 keystroke is required to type a fingerspelling character with the DKM software keyboard, but the number of keystrokes required to type a Myanmar character is not the same as the one to type an equivalent fingerspelling character. For example, 2 keystrokes usually need to be pressed for “◌ါ်” (“aa” vowel + “asat” vowel called “yecha shae htoo”) in Myanmar PC keyboards, however, only 1 keystroke is required to type with the DKM software keyboard. This is because “aa” vowel plus “asat” character is defined as one fingerspelling character (i.e.) in standard fingerspelling character set. Another important factor for counting number of keystrokes per character is “shift mode”, and another keystroke is required for pressing or clicking shift key. In this prototype, typing fingerspelling characters such as "ဍ", "ဋ", "ရ", "ဂ" and "ဝ" requires 2 keystrokes. Figure 10-6 shows the DKM software keyboard layout of Myanmar fingerspelling unshift mode, and Figure 10-7 shows that of shift mode.

Figure 10-6. DKM software keyboard layout (unshift mode)

Image adapted from my published papers [132, 133]

Figure 10-7. DKM software keyboard layout (shift mode)

Image adapted from my published papers [132, 133]

10.6.3 RomanPP_Fingerspelling Software Keyboard Development

In the RomanPP_Fingerspelling prototype, the QWERTY layout is used and English alphabets are shown. As mentioned in section 10.5.3, “s”, “f”, “e”, “c” and “j”, “l”, “i”, “m” keys are assigned for typing parameter of vowel combinations. Those keys are highlighted and work as “vowel combinations parameter key” or “four directional arrow keys” when users type Romanization of Myanmar consonant such as ka for “”, kh for “” and ga for “”. Vowel combinations parameter such as “left only”, “left and right” and “up only” is shown in LRUD Info Panel (see Figure 10-8). Possible vowel combinations with a consonant typed by users in Roman alphabets are shown in a candidate list (see Figure 10-8). When users type a consonant or syllable, highlight of “s”, “f”, “e”, “c”, “j”, “l”, “i” and “m” are removed and work as other software keys.

Figure 10-8 is a picture of the RomanPP_Fingerspelling prototype when “kal” or “kaf” is typed by users. Here, “ka” represents fingerspelling ka “” and “l” or “f” represents right vowels. Possible combinations of fingerspelling ka “” with right vowels such as “” (ကာ), “” (ကား), “” (ကျ) and “” (ကျာ) are shown in the candidate list.

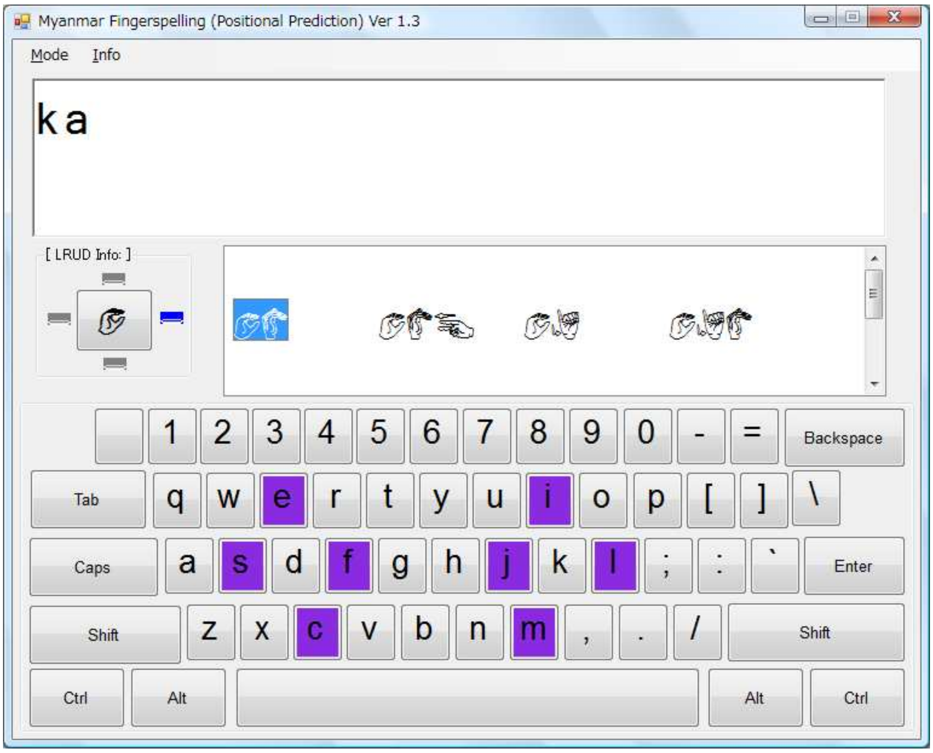


Figure 10-8. RomanPP_Fingerspelling software keyboard layout (version 1.3)

10.7 Methodology

10.7.1 Participants

User study was held in Yangon city, Myanmar with three types of volunteer users (i.e. hearing-impaired users, general users and desktop publishing staff), i.e., 48 native participants in total.

18 hearing-impaired users (9 female and 9 male) ranging in age from 13 to 22 years (mean = 17.1, SD = 2.9) were recruited who are students at Mary Chapman School for the Deaf and unfamiliar with personal computer. General users consist of 16 users (10 female and 6 male) ranging in age from 18 to 33 years (mean = 26.8, SD = 4.0), most of whom are students at KMD Computer Training Center, Yangon city, Myanmar and already familiar with Myanmar keyboard layouts (WinMyanmar or Zawgyi). 14 desktop publishing (DTP) staff (5 female and 9 male) ranging in age from 16 to 33 years (mean = 25.7, SD = 4.8) were recruited who routinely use one of the exiting Myanmar keyboards at work. None of them had prior experience with the DKM or RomanPP_Fingerspelling software keyboard.

10.7.2 Apparatus

The DKM and RomanPP_Fingerspelling software keyboard prototype ran on standard Microsoft Windows 7 notebook with a 2.0 GHz Intel Core2 duo CPU and 4 GB of memory. Another standard Microsoft Windows XP notebook with 1.5 GHz Intel Core duo CPU and 1 GB of memory was also used for user study with general users. Figure 10-9 shows the user study environment with hearing-impaired users at Mary Chapman School for the Deaf.

Figure 10-9. User study with hearing-impaired students

10.7.3 Procedures

The experiment procedures are as follows:

1) explaining the current Myanmar keyboard “Myanmar3”, the DKM and RomanPP_Fingerspelling keyboard layouts (as well as introducing to general users and DTP-staff what fingerspelling is and how it is used by hearing-impaired people),
2) making typing demonstration of the DKM and RomanPP_Fingerspelling software keyboard prototype with both a physical keyboard and a computer mouse,
3) allowing users to type a few trial fingerspelling words (maximum 10 minutes of practice time for each keyboard layout) to get familiar with the DKM and RomanPP_Fingerspelling software keyboard layout,
4) recording their typing speed of five fingerspelling words (see Figure 10-10) for 10 times (including error correction time) and
5) discussing with them and getting their feedbacks.

10.7.4 Design

A user study method was designed for the DKM and RomanPP_Fingerspelling keyboard layout with three types of users (hearing-impaired users, general users and DTP staff) and three independent factors (typing with a software keyboard labeled by fingerspelling characters, a software keyboard labeled by Myanmar characters and the RomanPP_Fingerspelling software keyboard).

Fingerspelling text for user study contains two actors’ names “ရန်အောင်” and “ကျော်သူ”, an actress’ name “ဆွေဇင်ထိုက်”, one of common female names “ဖြူဖြူဝင်း” and a city name “စစ်တွေ”. All those names are already familiar to native users and do not contain subscript consonants. The names without subscript consonants were selected because “subscript symbol” or “subscript consonants” are not defined in the current standardization of Myanmar fingerspelling character set. In practical, hearing-impaired people show subscript characters by moving their fingers downwards. Fingerspelling text contains 18 consonants, 11 vowels, 4 medials and 8 various signs. The mentioned names cover most combination patterns of vowel(s) and medial with a consonant (see Figure 8).

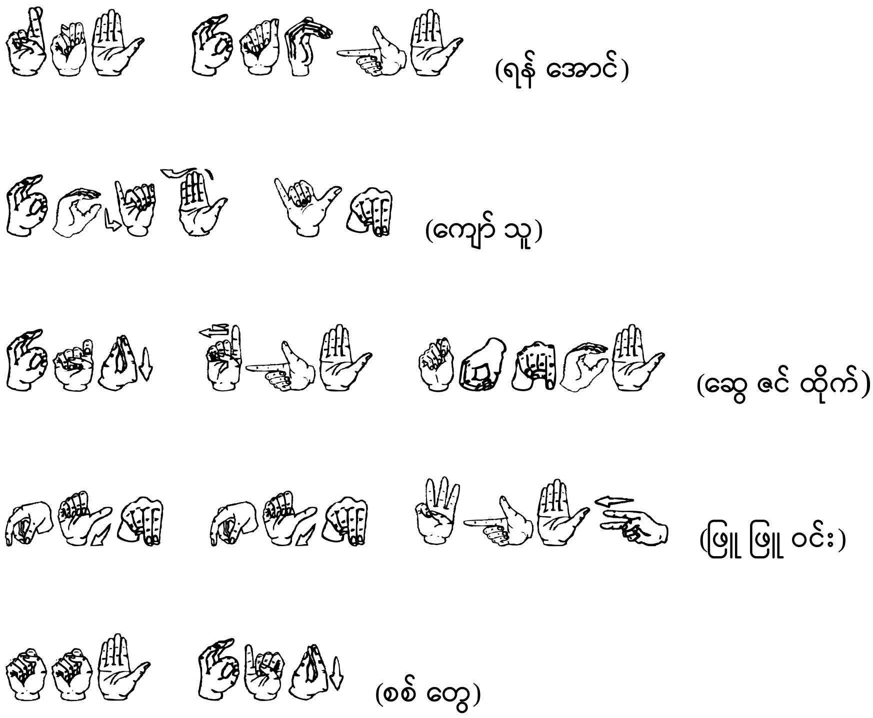


Figure 10-10. Five fingerspelling names for user study

(Equivalent Myanmar characters are presented with brackets.)

Image adapted from my published paper [133]

Formal and informal discussion was held with each hearing-impaired user to get his/her comments and suggestions relating to the DKM and RomanPP_Fingerspelling software keyboard prototype. 16 general users and 14 DTP staff were also recruited in order to figure out typing speed and feedbacks of the users who are already familiar with existing Myanmar PC keyboard layouts. Moreover, an analysis was made of “a software keyboard labeled by fingerspelling characters”, “a software keyboard labeled by Myanmar characters” and RomanPP_Fingerspelling based on Likert scale evaluation by users.

10.8 Results and Discussion

10.8.1 Evaluation by Characters Per Minute (CPM)

Typing speed was evaluated with Characters per Minute (CPM) instead of Words per Minute (WPM) [43]. Although space is put between fingerspelling syllables in Figure 10-10 for easier reading, users did not have an instruction to type a space between syllables, and thus, there is no need to consider a space for calculating CPM. My “T” value for CPM calculation is 42 (not 41) because fingerspelling character “” (“aa” vowel and “asat”) is counted as two.

Figure 10-11 shows typing speed improvement of hearing-impaired users for 10 times with DKM (labeled by Fingerspelling), DKM (labeled by Myanmar characters) and RomanPP_Fingerspelling. Average typing speed of hearing-impaired users is 15.6 CPM for DKM (labeled by Fingerspelling), 18.0 CPM for DKM (labeled by Myanmar characters) and 7.5 CPM for RomanPP_Fingerspelling, respectively. It was found that CPM of typing with DKM (labeled by Fingerspelling) is 209% faster and that of typing with DKM (labeled by Myanmar characters) is 241% faster than RomanPP_Fingerspelling.

Figure 10-12 shows average CPM of hearing-impaired users, general users and DTP staff for RomanPP_Fingerspelling user study. Average CPM of hearing-impaired users is 7.5, that of general users is 15.1 and that of DTP staff is 13.4, respectively. It was found that average CPM of general users 15.1 is the highest among three types of users, and 112% faster than DTP staff and 202% faster than hearing-impaired users.

Figure 10-13 shows average CPM comparison between DKM (labeled by Fingerspelling) and DKM (labeled by Myanmar characters). Average CPM of hearing-impaired users typing with DKM (labeled by Fingerspelling) is 15.6 and that of typing with DKM (labeled by Myanmar characters) is 18.0. Average CPM of general users typing with DKM (labeled by Fingerspelling) is 38.5 and that of typing with DKM (labeled by Myanmar characters) is 44.7. Average CPM of DTP staff typing with DKM (labeled by Fingerspelling) is 18.9 and that of typing with DKM (labeled by Myanmar characters) is 54.4. From the mentioned results, it can be said that CPM of typing with DKM (labeled by Myanmar characters) is higher than that of typing with DKM (labeled by Fingerspelling). In this study, there was no significant difference between DKM (labeled by Fingerspelling) and DKM (labeled by Myanmar characters) on CPM ($F_{1, 2} = 1.9$, $p > .05$).

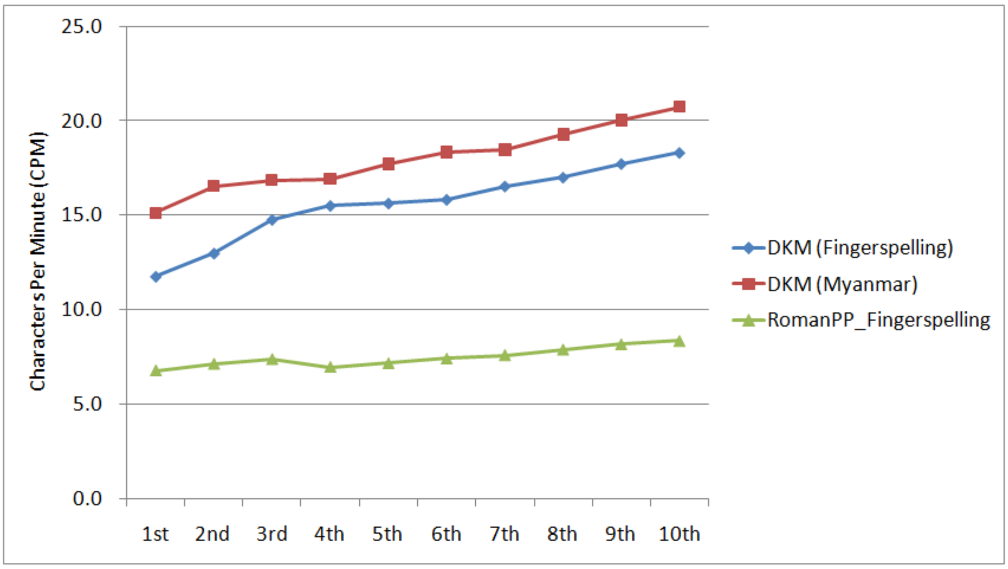


Figure 10-11. Typing speed improvement of hearing-impaired users for 10 times with DKM (labeled by Fingerspelling), DKM (labeled by Myanmar characters) and RomanPP_Fingerspelling

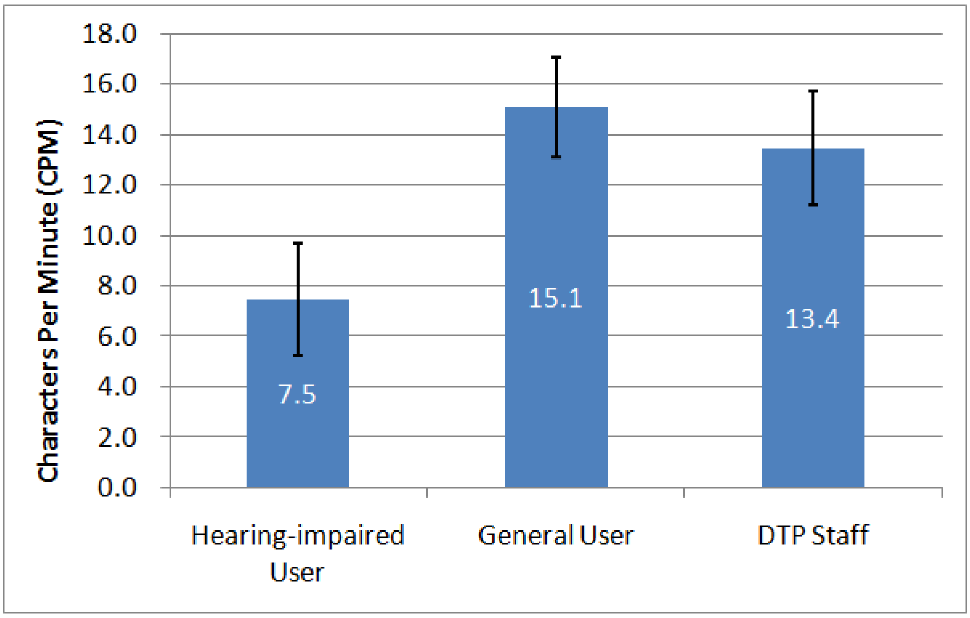


Figure 10-12. Average typing speed of hearing-impaired users, general users and DTP staff for 10 times with RomanPP_Fingerspelling

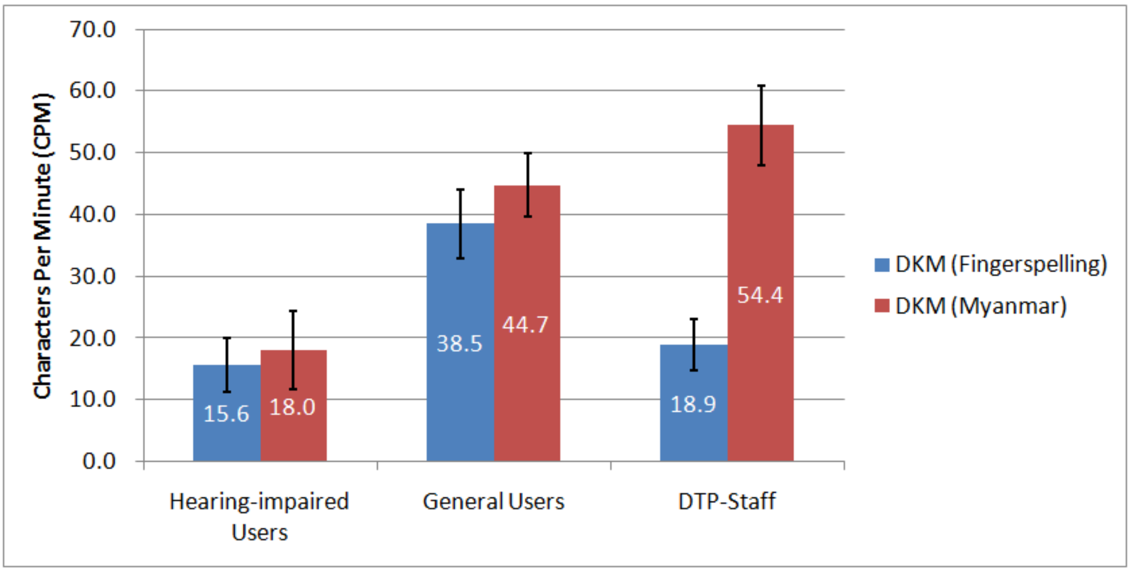


Figure 10-13. Average typing speed of hearing-impaired users, general users and DTP staff for 10 times with DKM (labeled by Fingerspelling) and DKM (labeled by Myanmar characters)

10.8.2 Evaluation by Questionnaires and Likert Scale

Questionnaires were conducted to the participants immediately after the typing experiments. Informal questionnaires were held to hearing-impaired users in order to get their comments and suggestions on my proposed DKM and RomanPP software keyboard. Communication with hearing-impaired users was made through sign language teachers' translation as well as writing messages. The followings are four questions that were asked:

1) Is a fingerspelling software keyboard necessary for hearing-impaired people?
2) Are you familiar with one of the existing Myanmar PC keyboard layouts?
3) Do you think Direct Keyboard Mapping (DKM) and RomanPP_Fingerspelling are usable for fingerspelling typing?
4) Do you have any comments or suggestions?

The answers to questions 1, 2 and 3 from 18 hearing-impaired users are summarized as "Yes", "No" and "Yes", respectively. As for the question 4, very positive answers were received such as "Although I am a first-time user of the DKM and RomanPP_Fingerspelling software keyboard prototype, I can type fingerspelling characters", "I am sure this will be very useful for fingerspelling characters typing", "Text input method for fingerspelling is necessary and it will be useful for creation of fingerspelling educational contents" and "Please keep making research on fingerspelling text input”. At the same time, there was an unexpected comment like “Fingerspelling used at Mary Chapman is different from standardized fingerspelling characters (11 differences in consonant characters and complete difference in vowel) and it caused ambiguity to us”. Moreover, there was also a comment like “Some fingerspelling characters are similar in DKM (labeled by Fingerspelling), and it is difficult to distinguish them, e.g., “” (na) and “” (tha), “” (ma) and “” (ca), “” (da) and “” (ya), and thus, bigger software keys might be better for a software keyboard with fingerspelling.”

4 Likert scales (1 to 5) questions were set to rate the user-friendliness of DKM (labeled by Fingerspelling), DKM (labeled by Myanmar characters) and RomanPP_Fingerspelling software keyboard layouts. The scales are (1) difficult-easy (2) painful-enjoyable (3) slow-fast and (4) dislike-like. Table 10-5, Table 10-6 and Table 10-7 show the average or arithmetic mean results of Likert scale questions to hearing-impaired users, general users and DTP staff, respectively. Here, Likert scales value 1 is the most negative, value 3 is neutral and value 5 is the most positive.

Table 10-5. Evaluation by hearing-impaired users

Likert Scale (Range 1 to 5)	**DKM (Fingerspelling)**	**DKM (Myanmar)**	**RomanPP**
Difficult-Easy	4.3	4.2	3.7
Painful-Enjoyable	3.8	4.1	4.7
Slow-Fast	3.8	4.4	3.5
Dislike-Like	4.5	4.1	4.5

Table 10-6. Evaluation by general users

Likert Scale (Range 1 to 5)	**DKM (Fingerspelling)**	**DKM (Myanmar)**	**RomanPP**
Difficult-Easy	4.8	4.4	4.9
Painful-Enjoyable	4.5	4.2	4.4
Slow-Fast	4.7	4.3	4.7
Dislike-Like	4.8	4.5	4.2

Table 10-7. Evaluation by DTP staff

Likert Scale (Range 1 to 5)	DKM (Fingerspelling)	DKM (Myanmar)	RomanPP
Difficult-Easy	4.5	4.5	3.1
Painful-Enjoyable	4.1	4.1	2.7
Slow-Fast	4.6	4.5	3.5
Dislike-Like	4.8	4.8	3.0

According to the mentioned results, it can be generally said that hearing-impaired users, general users and DTP staff are fond of text entering with DKM (labeled by Fingerspelling), DKM (labeled by Myanmar characters) and RomanPP_Fingerspelling software keyboard layouts. To see clearly their evaluation on "DKM software keyboard labeled by fingerspelling", "DKM software keyboard labeled by Myanmar character" and "RomanPP_Fingerspelling software keyboard", average value was calculated and a comparison graph was drawn (see Figure 10-14).

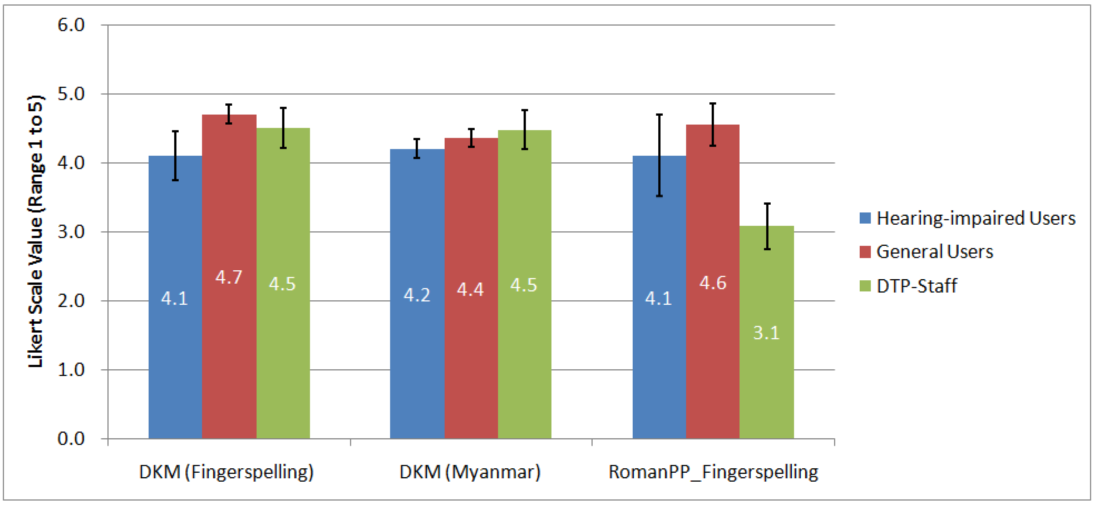
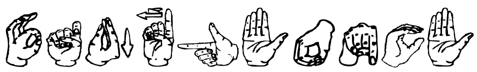
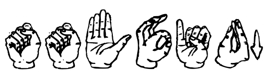


Figure 10-14. Comparison of Likert scale evaluation results of DKM (labeled by fingerspelling), DKM (labeled by Myanmar) and RomanPP_Fingerspelling

If only Likert scale evaluation results are considered, there is no significant user acceptance level difference on three keyboard layouts such as 4.1 : 4.2 : 4.1 (from hearing-impaired users), 4.7 : 4.4 : 4.6 (from general users) and 4.5 : 4.5 : 31 (from DTP staff).

However, it can be said that most participants attained good results with a software keyboard labeled by Myanmar character, judging from their typing speed (in terms of CPM), suggestions and comments. Taking their comments on DKM, only one of them prefer a software keyboard labeled by fingerspelling and the rest of them are comfortable with a software keyboard labeled by Myanmar character. CPM of RomanPP_Fingerspelling was affected because of its higher KSPC (Keystrokes per character) value (see Table 10-8). Although 50 KSPC was required for typing five fingerspelling names with the DKM keyboard layout, 107 KSPC was required with the RomanPP_Fingerspelling keyboard layout, which means that KSPC of RomanPP_Fingerspelling is 214% higher than DKM. As expected, most general users and DTP staff did not dare to give strong comments or suggestions relating to fingerspelling text input, and they gave just general positive comments such as “both DKM and RomanPP_Fingerspelling are possible to use” and “this might be useful for hearing-impaired users”. To cover this problem, additional questionnaires were held in terms of Likert scale evaluation after the typing experiment.

Table 10-8. Keystrokes Per Character (KSPC) comparison between DKM and RomanPP_Fingerspelling

Fingerspelling Text	**DKM**	**RomanPP_ Fingerspelling**
	10	23
	8	13
	13	32
	12	21
	7	18

10.9 Summary

This chapter has proposed DKM (labeled by fingerspelling), DKM (labeled by Myanmar character) and RomanPP_Fingerspelling keyboard layouts. An experiment compared three text input methods, i.e., typing with the DKM software keyboard labeled by fingerspelling, typing with the DKM software keyboard labeled by Myanmar character and typing with the RomanPP_Fingerspelling software keyboard with 18 hearing-impaired users, 16 general users and 14 DTP staff. For each text input method, user study was held to 10 users who entered a total of 5 fingerspelling names for 10 times (including error correction process). From user study with the current software keyboard prototype (version 1.3), average CPM values of 48 users are acceptable especially for DKM and it is proven that my proposals are possible solutions for fingerspelling typing. Moreover, based on the users' comments, suggestions and Likert scale value, it is proven that the fingerspelling software keyboard layout labeled by Myanmar character has better user interface for first-time users than the one labeled by fingerspelling. For the RomanPP_Fingerspelling implementation, work is planned to reduce KSPC based on Romanization of consonants such as "k" for "", "K" for "", "g" for "" and "G" for "". In the near future, user study is to be conducted with hearing-impaired users who are already familiar with standardized fingerspelling characters.

CHAPTER 11
Analysis on Possible Combinations of Vowels with a Consonant

This analysis is based on pronounceable, meaningful and syllable formation structure of vowels with a consonant on Nepali and Thai languages. My proposed Positional Prediction text input prototype is used as a tool for producing candidate lists of possible vowels combination patterns with a consonant. The results of this analysis will be useful for designing predictive text input methods for Nepali and Thai languages.

11.1 In Nepali

This section presents my analysis results on Nepali syllables or vowels combinations with a consonant. The motivation is to find out a user-friendly predictive text input interface based on Nepali syllable formation. Independent vowels were divided into four groups based on their writing positions. These groups are “Left” (ि◌), “Right” (◌ा, ◌ी, ◌ौ, ◌ौं, ◌ो, :), “Up” (◌े, ◌ै, ◌ँ, ◌ं, ◌ैं, ◌ैँ, र्) and “Down” (◌ु, ◌ू, ◌ृ, ◌्, ., ◌्र). The possible vowels combinations lists (e.g. Left+Consonant, Consonant+Right, Consonant+Up, Consonant+Down, Left+Consonant+Right, etc.) for Nepali consonants are produced with Positional Prediction (PP) text input prototype [38, 39]. Then, an analysis was made of vowels combinations, conjunct consonants or ligatures and combination patterns manually. Impossible combinations (i.e. unpronounceable and not found in modern Nepali language) were removed and the number of possible combinations was counted. The differences of vowels combination patterns of each consonant are shown clearly and the maximum length of a consonant syllable can be found.

11.1.1 Analysis on Vowels Combinations with a Consonant

As a first step, all vowels combinations with a Nepali consonant were collected by using PP text input prototype. My PP prototype uses Nepali Unicode font (Madan2) [35]. Although a range of vowels combination patterns (i.e. not allow to combine impossible vowels combination patterns for Nepali such as Left + Left + Consonant, Consonant + Right + Right + Right and Consonant + Down + Right + Right + Right) are controlled in my PP prototype, there are still some mixed vowel combinations contained in the candidate list. For example, in pressing “क (ka) + right arrow key”, the prototype will list “का”, “की”, “कौ”, “कौं”, “को”, “क:”, “र्का”, “र्की”, “र्कौ” and “र्को”. In this case, the last four syllables (i.e. “र्का”, “र्की”, “र्कौ” and “र्को”) were taken and put in the “Up + Right” combination list (refer to Table 11-1). Although Positional Prediction concept groups vowels with a consonant based on their glyph or shape, some mixed vowel combinations are supported for faster typing. Another example is that in pressing “त” (ta) + down arrow key + right arrow key”, the prototype will list “त्रा”, “त्री”, “त्रौ”, “त्रूौ”, “त्रौं”, “त्रो”, “र्ता”, “र्ती”, “र्तौ” and “र्तो”. Here, the consonant “त्र” (tra) is contained in the output list. This is because the consonant “त्र” (tra) is formed by the combination of consonant “त” (ta) and “र” (ra), as mentioned in the section 2. From my studies, there are 9 possible vowels combination patterns for Nepali language, which are as follows:

1) Left (Left + Consonant)
2) Right (Consonant + Right)
3) Up (Consonant + Up)
4) Down (Consonant + Down)
5) Down, Up (Consonant + Down + Up)
6) Left, Up (Consonant + Left + Up)

7) Up, Right (Consonant + Up + Right)
8) Down, Right (Consonant + Down + Right)
9) Down, Left (Consonant + Down + Left)

Character combination order is important in the Unicode fonts, but typing order depends on Input Method Editor (IME) like Tavultesoft keyman [90]. This is because Unicode mainly focuses on encoding standards for text processing such as sorting, searching and word breaking. The output character or syllable changes according to the combination order. The followings are examples:

[Example 1: र् (r) + क (ka)]

Combination Order 1:

र् + क = र्क

Combination Order 2:

क + र् = कर्

[Example 2: र् (r) + ि (i) + क (ka)]

Combination Order 1:

1) क + ि= कि

2) र् + कि = र्कि

Combination Order 2:

1) र् + ि= र्ि

2) र्ि + क = र्ि क

The correct writing order of Nepali in handwriting is 1) left vowel, 2) consonant, 3) down vowel, 4) upper vowel and 5) right vowel.

Table 11-1. Possible combinations of Nepali consonant “क” (ka)

(Highlighted cells indicate the longest “ka” consonant combined with vowels.)

Table taken from my published paper [134]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
L	कि																		
R	का	की	कौ	कौं	को	क:													
U	के	कै	कँ	कं	कैं	कैँ	र्के	र्कै											
D	कु	कू	कृ	क्	क.	क्र	क्रू	क्रु											
DU	क्रे	क्रै	कुँ	कूँ	कृँ	क्रँ	क्रूँ	क्रुँ	कुं	कूं	क्रं								
LU	किँ	र्किँ	किं	र्किं	र्कि														
UR	काँ	कीँ	कौँ	र्काँ	र्कीँ	र्कौँ	र्कोँ	कां	कीं	कौं	कों	र्कां	र्कीं	र्कौं	र्कों	र्का	र्की	र्कौ	र्को
DR	क्रा	क्री	क्रौ	क्रौं	क्रो														
DL	क्रि																		

11.1.2 Analysis on Conjunct Consonants

Other interesting and complex combinations are the formation of conjunct consonants or ligatures. A conjunct consonant looks like a single character, however, it is actually a combination of more than one characters, e.g., क्क, क्ख, र्क (consonants क्ष, त्र and ज्ञ are also conjunct consonants). According to my knowledge, there are four types of combination or writing styles for ligatures, which are as follows:

1) Vertical Combination Form

   (Later consonant append to the bottom of the earlier consonant e.g. क्क, ख्ख, etc.)

2) Horizontal Combination Form

   (Later consonant append to the right side of the earlier consonant e.g. क्ख, ग्ग, च्क, ग्छ, ञ्ज, etc.)

3) Alternative Combinations Form

   (e.g. र्क, प्र, त्र, ट्र, ड्र)

   Especially, “र” (ra) conjunct consonants or conjunct consonants containing “र” (ra) are complex. Combinations form changes depending on the combinations order such as (र् + क = र्क and क् + र = क्र). It also depends on the shape of combined consonants such as त्र, ह्र and व्र.

4) Special Combinations Form
   Vertical combinations usually occur when the same consonants are combined, however, combinations of different consonants can sometimes present a vertical form like द + ध = द्ध. Some of the conjunct consonants have special presentation such as द + म = द्म and द + ग = द्ग.

Some of the complex or vertical combinations are not used in today's Nepali such as ङ्क, ङ्ख, ङ्ग and ङ्घ (see Table 11-2). The total number of possible conjunct consonants is 954 excluding unpronounceable and impossible conjunct consonants. Here, only the conjunct consonants that contain 2 consonants (36 x 36 = 1,296 conjunct consonants) are considered. Although conjunct consonants or ligatures that contain more than 2 consonants exist in Devanagari scripts, they are not used in the current Nepali.

Table 11-2. Conjunct consonants or ligatures for consonants "क" (ka) to "ञ" (nya)

(Cells with two diagonal lines indicate impossible conjunct consonants or ligatures.)

Table taken from my published paper [134]

◌्	क	ख	ग	घ	ङ	च	छ	ज	झ	ञ
क	क्क	क्ख	क्ग	क्घ	क्ङ	क्च	क्छ	क्ज	क्झ	~~क्ञ~~
ख	ख्क	ख्ख	ख्ग	ख्घ	ख्ङ	ख्च	ख्छ	ख्ज	ख्झ	~~ख्ञ~~
ग	ग्क	ग्ख	ग्ग	ग्घ	ग्ङ	ग्च	ग्छ	ग्ज	ग्झ	~~ग्ञ~~
घ	घ्क	घ्ख	घ्ग	घ्घ	घ्ङ	घ्च	घ्छ	घ्ज	घ्झ	~~घ्ञ~~
ङ	~~ङ्क~~	~~ङ्ख~~	~~ङ्ग~~	~~ङ्घ~~	~~ङ्ङ~~	~~ङ्च~~	~~ङ्छ~~	~~ङ्ज~~	~~ङ्झ~~	~~ङ्ञ~~
च	च्क	च्ख	च्ग	च्घ	च्ङ	च्च	च्छ	च्ज	च्झ	~~च्ञ~~
छ	~~छ्क~~	~~छ्ख~~	~~छ्ग~~	~~छ्घ~~	~~छ्ङ~~	~~छ्च~~	~~छ्छ~~	~~छ्ज~~	~~छ्झ~~	~~छ्ञ~~
ज	ज्क	ज्ख	ज्ग	ज्घ	ज्ङ	ज्च	ज्छ	ज्ज	~~ज्झ~~	~~ज्ञ~~
झ	झ्क	झ्ख	झ्ग	झ्घ	झ्ङ	झ्च	झ्छ	झ्ज	झ्झ	~~झ्ञ~~
ञ	~~ञ्क~~	~~ञ्ख~~	~~ञ्ग~~	~~ञ्घ~~	~~ञ्ङ~~	~~ञ्च~~	~~ञ्छ~~	~~ञ्ज~~	~~ञ्झ~~	~~ञ्ञ~~

11.1.3 Results of Analysis

In my analysis, all possible combinations of dependent vowels with a Nepali consonant are considered. There are 9 possible combination patterns for Nepali as mentioned above, and it is assumed that “Down + Consonant + Up” and “Up + Consonant + Down” are the same here. Handwriting order may differ from person to person, but there is no need to consider this factor to count possible combinations. PP prototype was used to list possible combinations for each consonant (i.e. Ka to Gnya). The variable declaration for Left, Right, Up and Down vowel parameters to develop PP prototype are as follows:

LeftList = New String() {"ि"}

RightList = New String() {"ा", "ी", "ौ", "ौं", "ो", ":"}

UpList = New String() {"े", "ै", "ँ", "ं", "ैं", "ैँ", "र्"}

DownList = New String() {"ु", "ू", "ृ", "्", ".", "्र"}

Here, it is difficult even for native people to understand combination patterns of "र्" and "्र". "र्" is for upper vowel combinations such as (र + ् + क = र्क), (र् + के = र्के) and (र + ् + क + ौ = र्कौ) , and "्र" is for lower vowel combinations such as (क + ्र) = क्र, (ख + ्र) = ख्र and (ग + ् र) = ग्र.

The total number of possible combinations of dependent vowels counts 2,156 for all Nepali consonants. Figure 11-1 shows a graph of possible combinations of dependent vowels with a consonant. The highest possible combination value is 72 with “ठ” (ttha) consonant, the lowest value is 42 with “ण” (nna) consonant and the average value is 59.89.

A comparison was made of 9 possible combination patterns of dependent vowels with a consonant as shown in Figure 11-2 As expected, “Up + Consonant + Right” combination pattern is the highest value, which is 694. The lowest combination pattern is “Left + Consonant +

Down”, which counts 25. What’s interesting is that there are as many as 6 down vowels and 6 right vowels in the array variable list but the number of “Consonant + Down + Right” combination patterns is only 123 (see Figure 11-2).

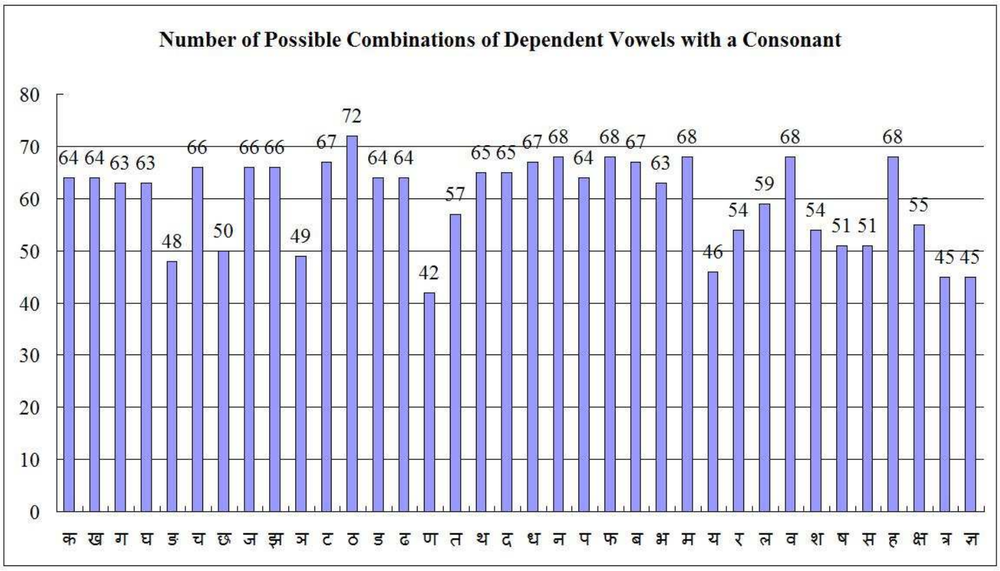


Figure 11-1. Number of possible combinations of dependent vowels with a consonant (Ka to Gnya)

Image taken from my published paper [134]

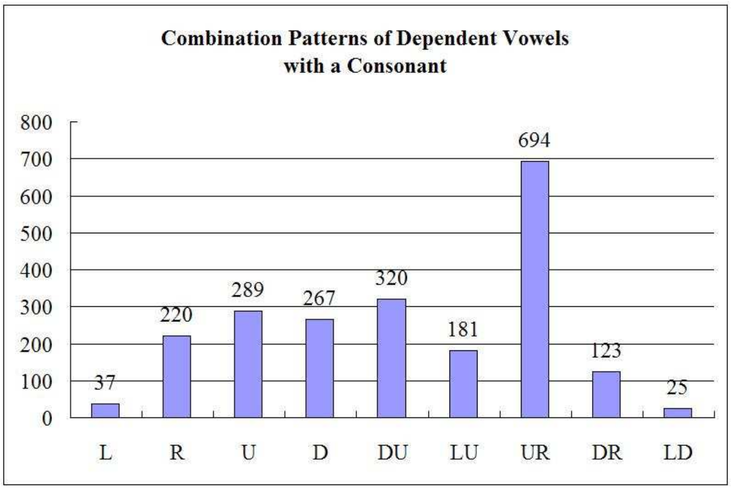


Figure 11-2. Combination patterns of dependent vowels with a consonant

(Here, L = Left, R = Right, U = Up, D = Down, DU = Down + Up, LU = Left + Up, UR = Up + Right, DR = Down + Right and LD = Left + Down)

Image taken from my published paper [134]

The maximum length of a syllable (combination of vowels and a consonant) in the current Nepali language is 4. Here a consonant is counted as one character (see Table 11-1). The followings are examples of maximum combinations length for “ख” (kha).

“ख” (kha) Consonant:

in [DU] combination: ख्रूँ, ख्रुँ

in [LU] combination: खिँ, खिं

in [RU] combination: खाँ, खीँ, खौँ, खोँ, खों, खां, खीं, खौं, खों

in [DR] combination: ख्रौं

11.1.4 Discussion

This section discusses why an analysis was made of possible combinations of vowels and a consonant, their patterns and maximum length in the current Nepali language.

It is believed that text typing on small mobile devices will become more popular and necessary communication in Asian developing countries such as Myanmar (Burma), Bangladesh, Nepal, Bhutan, Laos and Cambodia. In these countries, however, there is no efficient and user-friendly text input method for mobile devices yet. Asian languages are syllabic languages that derived from Indic script or Brahmi around BC third century. And thus, Myanmar language or Burmese, Bengali, Nepali, Dzongkha (language of Bhutan), Lao and Khmer have common writing natures with Indian languages such as Hindi, Marathi, Punjabi and Tamil. But current mobile devices key-mapping or text input methods such as multi-tap or T9 [91] are based on English and not directly applicable to syllabic languages, because writing natures of Asian syllabic languages are different and have larger numbers of characters than English alphabets (e.g. Khmer has triple numbers of characters (i.e. 74) of English, if counting conjunct consonants the numbers of characters will be higher). My research looks for common and user-friendly keyboard mappings and text input methods for Asian syllabic languages based on their word formation or writing natures for mobile devices. Positional Mapping (PM) [53] [69], Positional Gesture (PG) [57] and Positional Prediction (PP) [38] [39] are the proposed text input interfaces based on Asian syllabic languages’ writing nature and their combination structure of vowels with a consonant. And thus, analysis on possible combinations of vowels with a consonant, possible vowels combination patterns with a consonant and the maximum length of vowels with a consonant for Asian syllabic languages are very important for PG and PP text input interfaces.

This section presents the analysis process based on left, right, up and down combinations of vowels with a consonant. My study is on several Asian syllabic languages, but the results relating to only Nepali are stated here. It cannot be said that these results indicate the exact number of possible combinations because the removing process of impossible combinations is a complicated task. However, it is certain that there are 9 possible combination patterns of vowels with a consonant. Although some vowel combinations and conjunct consonants are not in use in the current Nepali, they were used in old Nepali, and might be used again in the future (e.g. for writing foreign names, technical terms, slang words, names of actors and actresses, etc.) It is also possible to use them in other Devanagari languages or other ethnic groups' languages in Nepal such as Maithili, Bhojpuri, Newari, Tamang and Nepalbhasa.

The decision-making of possible or impossible combinations is difficult because Devanagari script is shared by many Indic languages. In Asian syllabic languages, words are formed by one or more syllables. There are various syllable structures such as a consonant, a consonant with a vowel, and a consonant with some vowels. The followings are Nepali words formed by syllables:

दि (give) (VC), लि (take) (VC), खा (eat) (CV), बिर्सी (forget) (VCCLV), को (who) (CV), के (what) (CV), कहाँ (where) (CCVV), किन (why) (VCC), कसरी (how) (CCCV), कहिले (when) (CVCCV), अनुसन्धान (research) (VCVCLVC), असल (good, pure) (VCC), आज (today) (VC), ईज्जत (prestige, honor, self-respect, status) (VLC) and गाह्रो (difficult) (CVLV)

Here, C = Consonant, V = Independent Vowel or Dependent Vowel and L = Ligature

From this analysis, the number of combination patterns of dependent vowels with a consonant is clearly seen as well as their combination ratios (see Figure 11-2). Their average values for 36 consonants are as follows:

Table 11-3. Average values of possible combinations for Nepali

Table taken from my published paper [134]

Combination Pattern	Average
Left + Consonant	1.03
Consonant + Right	6.11
Consonant + Up	8.03
Consonant + Down	7.42
Consonant + Down + Up	8.89
Left + Consonant + Up	5.03
Consonant + Up + Right	19.3
Consonant + Down + Right	3.42
Consonant + Left + Down	0.69

FontLab Studio was used to check the glyphs of all possible conjunct consonants or ligatures [92]. Although glyphs and code number of all conjunct consonants or ligatures can be seen, it is not possible to find possible combinations of vowels with a consonant. The possible combinations list of conjunct consonants which was produced by PP prototype and the conjunct consonants list of Madan2 font table were confirmed.

11.1.5 Summary

This dissertation has reported the results of my analysis on possible combinations of vowels with a consonant in the current Nepali language. From this analysis, it was found that there are 9 possible vowels combination patterns with a consonant and that the maximum length of vowels with a consonant is 4. The total number of possible combinations of vowels with a consonant is 2,156, and that of conjunct consonants is 954. It is difficult to prove that these numbers are correct and exact values because of the nature of syllabic languages. These numbers are the results based on left, right, up and down combinations of vowels with a consonant, which will be useful for Positional Prediction text input interfaces. Further refinements on the current prototype and follow-up analysis on possible combinations of syllables or word formation structure are to be made. An analysis is now being made of possible combinations of vowels with a consonant for Khmer (language of Cambodia) and Thai (language of Thailand), and the results are to be presented hopefully in the near future.

11.2 In Thai

This dissertation presents my analysis results on Thai Vowels, Diacritic Symbols and Tone marks combination with a Consonant. The motivation is to find out a user-friendly predictive text input interface based on Thai syllable formation. Vowels were divided into four groups based on their writing position in combining to a base consonant. These groups are “Left” (เ , แ , โ , ใ , ไ), “Right” (ะ , า , ำ) , “Up” (ั, ิ, ี, ึ, ื) and “Down” (ุ, ู). With the concept of Position Prediction text input prototype [38] [39], only one consonant with Vowels, Diacritic symbols and Tone marks combination case will be taken into this analysis. The possible Vowels, Diacritic symbols and Tone marks combinations lists (e.g. Left + Consonant + Up, Consonant + Right, Consonant + Down, etc.) were produced for Thai consonants manually based on grammatical rule. Consonant was grouped into 3 groups by their level of tone; high-tone, medium-tone and low-tone, and each tone will have the same characteristic of possible combination of tone marks. Then, an analysis was made of Vowels, Diacritic Symbols and Tone marks combinations and combination patterns. Impossible combinations (i.e. unpronounceable, meaningless and never used in Thai language) were removed and the number of possible combinations was counted. The differences of Vowels, Diacritic symbols and Tone marks combination patterns of each consonant can be seen clearly and it is possible to find out the maximum length of a consonant syllable.

11.2.1 Analysis on Vowels Combination with a Consonant

The current Thai keypad mapping for mobile device, 44 alphabets are mapped into 8-9 buttons and only 2 or 3 buttons are left for Vowels, Diacritic symbols, Tone marks and Special symbols. Therefore, input of Vowels, Diacritic symbols and Tone marks is very difficult. Some characters require almost 10 times pressing of the same button. My proposed PP model for Thai will help predicting Vowels, Diacritic symbols and Tone mark by pressing arrow keys to address their position in combination with base consonant. But users still have to input consonant by multi-tap. Therefore, PP will concern only possible combination of vowels, Diacritic symbols and Tone marks with one consonant.

Table 11-4. Three-tone consonant group and tone rules

Table taken from my published paper [122]

Group	Tone Rule
Low-tone ค ฆ ง ช ซ ฌ ญ ฑ ฒ ณ ต ท ธ น พ ฟ ภ ม ย ร ล ว ฬ ฮ	Only 3 tone, 2 tone marks " ◌่ " and " ◌้ " can be applied
Medium-tone ก จ ด ฎ ต ฏ บ ป อ	All 5 tone, 4 tone marks can be applied
High-tone ข ฃ ฉ ฐ ถ ผ ฝ ศ ษ ส ห	Only 3 tone, 2 tone marks " ◌่ " and " ◌้ " can be applied

As a first step, all of the combinations of each consonant with vowels and diacritic marks were listed, and then, grammatical rule of tone marks was applied. In Thai, there are 3 categories of consonant called “Triyang”, Low-tone consonant, Medium-tone consonant and High-tone consonant. Table 11-4 shows each group’s rule and members.

From the steps above, 14 possible combination patterns were found out for Thai language in case one consonant with vowel, tone marks and diacritic marks are as follows:

1) Left (Left + Consonant)
2) Right (Consonant + Right)
3) Up (Consonant + Up)
4) Down (Consonant + Down)
5) Left, Right (Left + Consonant + Right)
6) Left, Up (Left + Consonant + Up)
7) Right, Right (Consonant + Right + Right)
8) Up, Right (Consonant + Up + Right)
9) Up, Up (Consonant + Up + Up)
10) Down, Up (Consonant + Down + Up)
11) Left, Up, Right (Left + Consonant + Up + Right)
12) Left, Up, Up (Left + Consonant + Up + Up)

13) Left, Right, Right (Left + Consonant + Right + Right)
14) Left, Up, Right, Right (Left + Consonant + Up + Right + Right)

Character combination order is important in the Unicode fonts, and it should be typed from left to right and down to up, which is to say, the writing order of Thai in handwriting is 1) Left Vowel 2) Consonant 3) Down Vowel 4) Up Vowel or Up diacritic mark 5) Tone mark and 6) Right Vowel. For example, the typing order of the word เหล้า (Lao) is "เ" + "ห" + "ล" + "้" + "า". But in this analysis, there is no need to consider this factor for counting the possible combinations because there is no effect of different order on the same combination as follows:

1) Right + Left = Left + Right
2) Right + Up = Up + Right
3) Up + Left = Left + Up
4) Up + Down = Down + Up
5) Up + Up + Left = Left + Up + Up
6) Left + Right + Up = Up + Right + Left = Left + Up + Right
7) Left + Right + Right + Up = Left + Up + Right + Right

After listing all of the possible combinations of each consonant with vowels, diacritic mark and tone marks by grammatical rule, some combinations that are unpronounceable or meaningless word or never used in Thai are included in this list. Neatly, an unpronounceable or meaningless word or a never used word in Thai was removed, and only a pronounceable or meaningful syllable or part of syllable was counted as shown in Table 11-5.

11.2.2 Results of Analysis

The total number of possible combinations of vowels, diacritic symbols and tone marks with one consonant counts is 1,770 for all Thai consonants. Figure 11-3 shows a graph of possible combinations of vowels, diacritic symbols and tone marks with one consonant. The highest possible combination value is 88 with "ต" (To tao) consonant, the lowest value is 4 with "ฒ" (tho phu-thao) (excluding 2 obsolete consonants) and the average value is 40.23.

Looking at the 10 highest possible combination values, it was found that 7 of them are in Medium-tone group consonant. The reason is that the consonants in this group are able to combine 4 tone marks. From this analysis, it was also found that the 2 obsolete consonants are not used and among the 10 lowest possible combination value 8 of them are the consonants used only in Pali and Sanskrit word [93]. This can be a good proof for Thai keypad mapping design to neglect those 2 obsolete consonants and group the rarely use consonants in the same button not in the alphabet order as the current key mapping.

Furthermore, the number of appearance of each Vowel and Diacritic mark in Possible Combinations was also counted in order to figure out which character is frequently used as shown in Figure 11-4. These marks are shown in alphabetical order. These marks were analyzed group by group. Left position group, the most frequently used mark is “เ” (Sara e) at 452 times or 58.17% of this group while the second frequently used in this group is “แ” (Sara ae) at 132 times. “เ” is also the most frequently used among all of Vowels and Diacritic marks. For Right group, the most frequently used mark is “า” (Sara ah) at 207 times or 49.64% of this group. “ ี” (Sara ee) is the most frequently used in Up group at 176 times and follow by “ ิ” (Sara i) at 159 times. For Down group, there is no significant difference between two of them, “ ุ ” (Sara u) at 100 times and “ ู” (Sara uh) at 84 times. This result is usable to design the order of these characters in each button to create more user-friendly keypad mapping.

A comparison of 14 possible combination patterns of dependent vowels with a consonant was also made as shown in Figure 11-5. The combination pattern that has the highest possible combination is “Left + Consonant + Up” combination pattern which is 356. It is no surprise that the lowest combination pattern is the most complicated combination “Left + Consonant + Up + Right + Right”, which counts 3.

Table 11-5. Possible combinations of Thai consonant "ป" (po pla)

Table taken from my published paper [122]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
L	เป	แป	โป	ไป														
R	ปะ	ปา	ปำ															
U	ป่	ป้	ป๊	ป๋	ปิ	ปี	ปึ	ปื	ปั	ป็								
D	ปุ	ปู																
LR	เปะ	เปา	แปะ	โปะ														
LU	เป่	เป้	เป๊	เป๋	เปิ	เปี	เปื	เป็	แป่	แป้	แป๊	แป๋	โป่	โป้	โป๊	ไป๋	ไป่	
RR	-																	
UR	ป่า	ป้า	ป๊า	ป๋า	ป่ะ	ป้ำ												
UU	ปิ่	ปิ้	ปิ๊	ปิ๋	ปี่	ปี้	ปี๊	ปี๋	ปึ่	ปึ้	ปึ๊	ปึ๋	ปื้	ปื๋	ปั่	ปั้	ปั๊	ปั๋
DU	ปุ่	ปู่	ปุ้	ปู้	ปุ๊	ปุ๋												
LUR	เป่า	เป้า	เป๊า	เป๋า	เป๊ะ	โป๊ะ												
LUU	เปิ่	เปิ้	เปิ๊	เปี่	เปี้	เปี๊	เปื่	เปื้										
LRR	เปา ะ																	
LURR	เป๊า ะ																	

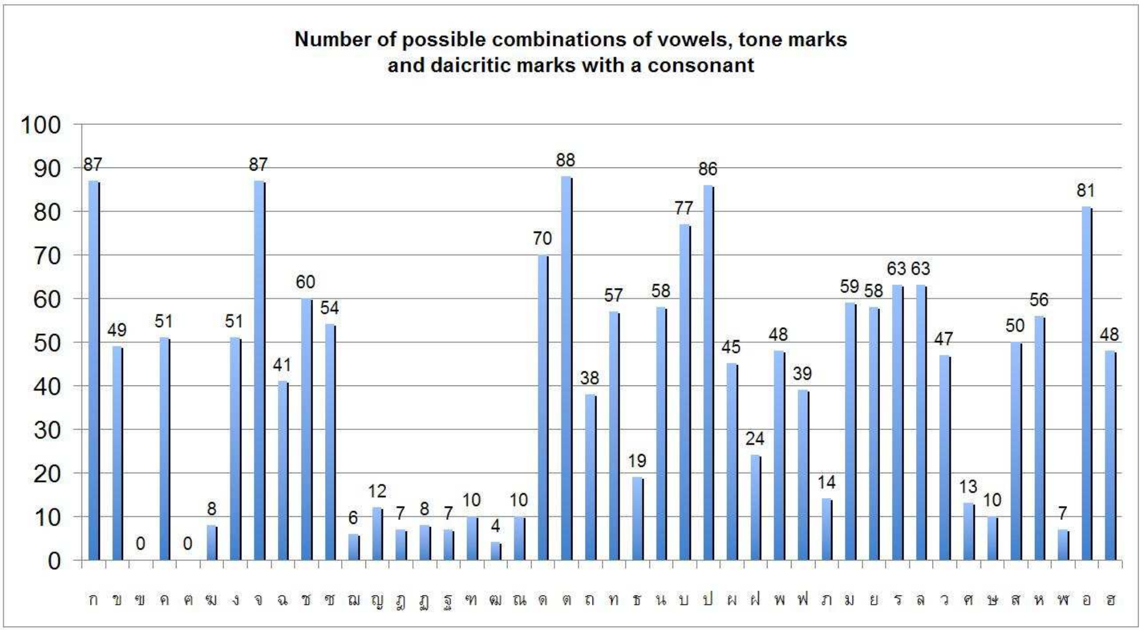


Figure 11-3. Number of possible combinations of each consonant

Image taken from my published paper [122]

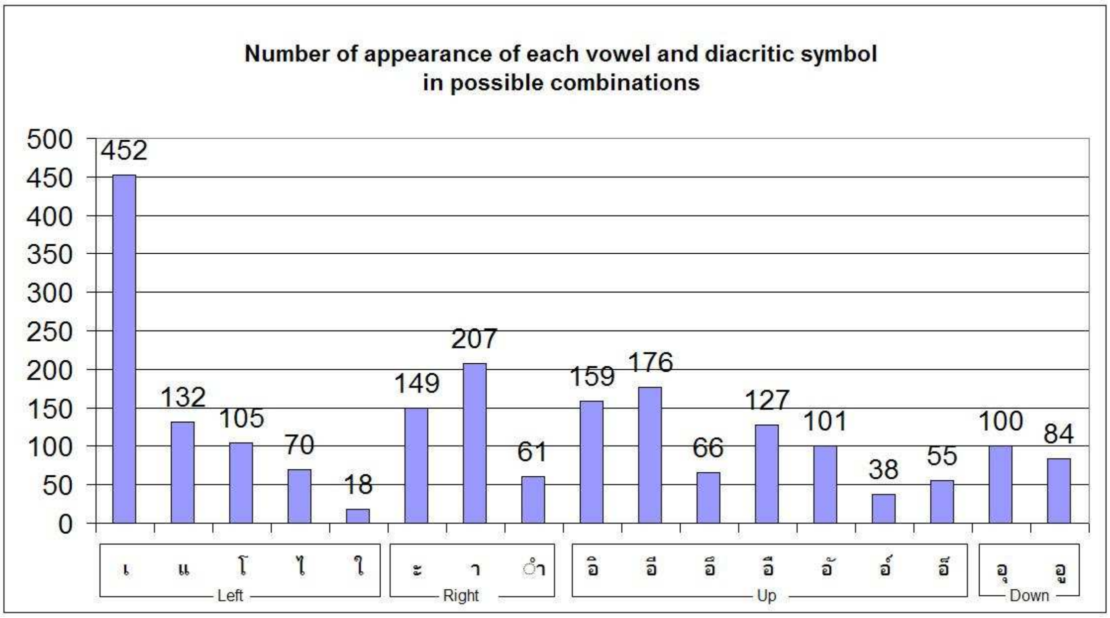


Figure 11-4 Number of appearance of each vowel and diacritic symbol in possible combinations

Image taken from my published paper [122]

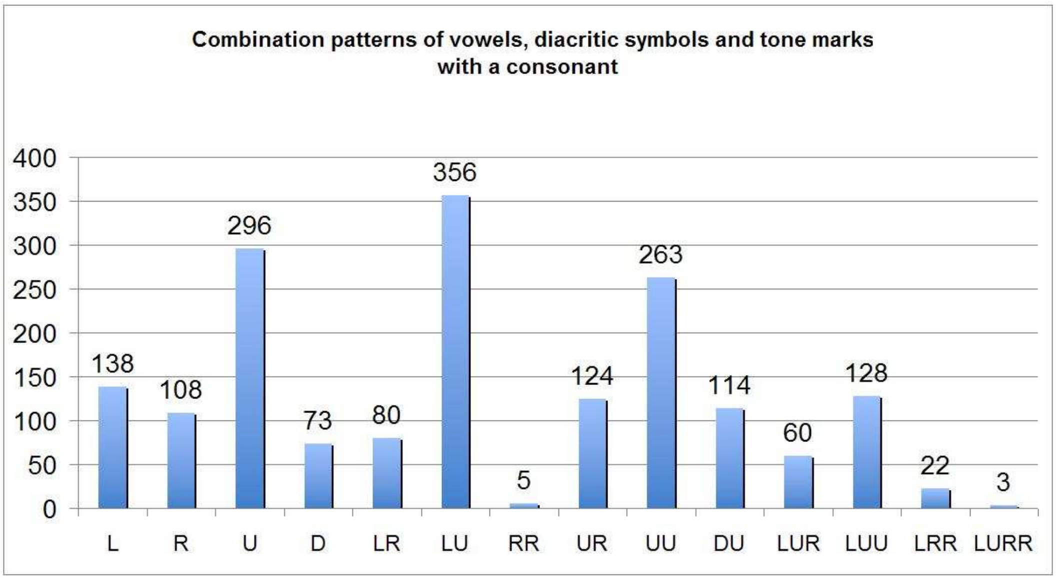


Figure 11-5. Combination patterns of vowels, diacritic symbols and tone marks with a consonant

(Here, L = Left, R = Right, U = Up, D = Down, LR = Left + Right, LU = Left + Up, RR = Right + Right, UR = Up + Right, UU = Up + Up, DU = Down + Up, LUR = Left + Up + Right, LUU = Left + Up + Up, LRR = Left + Right +Right and LURR = Left + Up + Right + Right)

Image taken from my published paper [122]

11.2.3 Discussion

This section discusses why an analysis was made of possible combinations of vowels and a consonant, their patterns and maximum length in the current Thai language.

It is believed that text typing on small mobile devices will become more popular and necessary communication in Asian developing countries such as Thailand, Myanmar (Burma), Bangladesh, Nepal, Bhutan, Laos and Cambodia. In these countries, however, there is no efficient and user-friendly text input method for mobile devices yet. Asian languages are syllabic languages that derived from Indic script or Brahmi around BC third century. And thus, Thai, Myanmar language (Burmese), Bengali, Nepali, Dzongkha (language of Bhutan), Lao and Khmer have common writing natures with Indian languages such as Hindi, Marathi, Punjabi and Tamil. But current mobile devices key-mapping or text input methods like multi-tap or T9 are based on English and not directly applicable to syllabic languages, because writing natures of Asian syllabic languages are different and have larger numbers of characters than English alphabets (e.g. Thai has triple numbers of characters (i.e. 74, excluding 10 of Thai number) of English). My research looks for common and user-friendly keyboard mappings and text input methods for Asian syllabic languages based on their word formation or writing natures for mobile devices. Positional Mapping (PM) [53] [94] [95], Positional Gesture (PG) [57] and Positional Prediction (PP) [38] [39] are the proposed text input interfaces based on Asian syllabic languages' writing nature and their combination structure of vowels with a consonant. And thus, analysis of possible combinations of vowels with a consonant, possible vowels combination patterns with a consonant and the maximum length of vowels with a consonant for Asian syllabic languages are very important for PG and PP text input interfaces.

This section presents the analysis process based on left, right, up and down combinations of vowels with a consonant. My study is on several Asian syllabic languages, but the results relating to only Thai is stated here. It cannot be said that these results indicate the exact number of possible combinations because the removing process of impossible combinations is a complicated task. However, it is certain that there are 14 possible combination patterns of vowels, diacritic symbols and tone marks with a consonant. Although some of the combinations are not in use in the current Thai, they were used in old Thai, and might be used again in the future (e.g. for writing foreign names, technical terms, slang words, names of actors and actresses, etc.)

The decision-making of possible or impossible combinations is difficult because words are formed by one or more syllables. There are various syllable structures such as a consonant, a consonant with a vowel, and a consonant with some vowels. The followings are Thai words formed by syllables:

ณ (at) (C), ก็ (and, also) (CD), ขม (bitter) (CC), รวม (include) (CCC), กรรม (sin, action) (CCCC), ภรรยา (wife) (CCCCV), ใจ (heart) (VC), ขา (leg) (CV), พระ (priest, monk) (CCV), งาม (beautiful) (CVC), เรา (we) (VCV), เก่า (old) (VCVT), เจาะ (drill) (VCVV), พระสงฆ์ (Buddhist monk) (CCVCCCD), หมู่ (group) (CCVD), เหล้า (alcohol) (VCCVT), เหมาะ (suitable, fit) (VCCVV), อาหาร (food) (CVCVC), สวัสดี (Hello) (CCVCCV), หนังสือ (book) (CCVCCVC), นิตยสาร (magazine) (CVCCCVC), มหาวิทยาลัย (university) (CCVCVCCVCVC)

Here, C = Consonant, V = Vowel, T = Tone mark, D = Diacritic mark

From this analysis, I can clearly see the number of combination patterns of dependent vowels, diacritic symbols and tone marks with a consonant as well as their combination ratios (see Figure 11-5). Their average values for 44 consonants are as follows:

Table 11-6. Average values of possible combinations for Thai

Table taken from my published paper [122]

Combination Pattern	Average
Left + Consonant	3.14
Consonant + Right	2.45
Consonant + Up	6.73
Consonant + Down	1.66
Left + Consonant + Right	1.82
Left + Consonant + Up	8.09
Consonant + Right + Right	0.11
Consonant + Up + Right	2.82
Consonant + Up + Up	5.95
Consonant + Down + Up	2.59
Left + Consonant + Up + Right	1.36
Left + Consonant + Up + Up	2.91
Left + Consonant + Right + Right	0.50
Left + Consonant + Up + Right + Right	0.07

11.2.4 Summary

A report has been made on the results of my analysis on possible combinations of vowels, diacritic symbols and tone marks with a consonant in the current Thai language. From this analysis, it was found that there are 14 possible vowels combination patterns with a consonant and that the maximum length of vowels with a consonant is 5. The total number of possible combinations of vowels with a consonant is 1770. It is difficult to prove that these numbers are correct and exact values because of the nature of syllabic languages. These numbers are the results based on left, right, up and down combinations of vowels, diacritic symbols and tone marks with a consonant, which will be useful for Positional Prediction text input interfaces. Further refinements on the current prototype and follow-up analysis on possible combinations of syllables or word formation structure are to be made. An analysis is now being made of possible combinations of vowels with a consonant for Khmer (language of Cambodia) and Sinhala (language of Sri Lanka) and the results are to be presented hopefully in the near future.

CHAPTER 12
Grapheme Cluster Segmentation Tool

12.1 Development for Myanmar Language

This chapter introduces my developing grapheme cluster segmentation tool and its segmentation methods, and presents output results from experiments with a Myanmar-English dictionary. Motivation of this research is to produce Positional Prediction text input candidate list sorted by frequently used grapheme clusters. This candidate list is necessary for Positional Prediction text input method. Section 12.1.2 refers to my proposed Positional Prediction text input concept. Rule-based segmentation method is used, which is totally based on combination of vowels with a base consonant. Four segmentation rules are proposed for determination of grapheme cluster boundaries in Myanmar text. Briefly introducing, Rule1 is for the break point in front of a character, Rule 2 is for the break point at the end of a character, Rule 3 is for the non-Myanmar character and Rule 4 is for checking errors such as wrong break points, mixed syllables and exceptional cases. Discussions are also held on difficulties of preparing a plain text data from existing Myanmar electronic documents for initial study on my four segmentation rules, most of which are related to typing mistakes, wrong typing order of Myanmar syllables and encoding methods. The program has been designed for providing two additional output formats based on the results of grapheme cluster segmentation; one is in terms of “Positional Prediction combination pattern” and the other is in terms of “character”. This software tool is useful not only for analysis on grapheme clusters distribution but also for various Natural Language Processing (NLP) of Myanmar language.

12.1.1 Existing Syllable and Word Segmentation Methods

A syllable is a unit of speech sound and is typically constructed by combinations of vowel(s) and consonant(s). Syllable segmentation or syllabification is an important process in natural language processing and can be applied for word segmentation or word breaking by merging segmented syllables. Here is a brief introduction to existing syllable and word segmentation methods used for Myanmar language.

Naïve string matching algorithm or longest character matching method is applicable for segmentation of Myanmar text. The basic idea of this algorithm is simple; it matches pattern string against input string character by character. If a mismatch occurs, shift the whole input string down by one character in relation to the pattern string, and start matching again at the beginning [98]. The pseudo code of Naïve string matching algorithm is as follows:

```
Input: Strings Pattern and Text
length_Text ← |Text|;
length_Pattern← |Pattern|;

for i ←0 to length_Text- length_Pattern do
    if Pattern[1..m] = Text[i+1, .. , i+ length_Pattern] then
        PRINT(Pattern matches at index i+1);
    end
end
```

Myanmar Syllables segmentation using character level longest matching and Myanmar word segmentation using syllable level longest matching was proposed by Hla Hla Htay and Kavi Narayana Murthy [99, 100]. This approach needs dictionaries, lexical resources, corpus, etc. in advance. The mentioned two authors made a detail morphological analysis of Myanmar text documents, cleaned up to remove hypertext markup, collected stop words such as prepositions/post-positions, conjunctions, particles, inflections, etc. (e.g. "သဖြင့်", "သူ", "သည်"), collected n-grams of syllables up to 5-grams (e.g. bigram "ဖန်တီး", trigram "လက်ခနဲ", 4-grams "စဉ်းစဉ်းစားစား") and developed Myanmar corpus including more than 17 lakh sentences [101]. From their studies, syllabification obtained 99% accuracy and word segmentation obtained 80% accuracy [101].

Another possible way of syllable segmentation is to use defined syllable breaking rules based on the characteristics of Myanmar syllable formations. In general, syllable breaking point can be determined by matching pairs of characters for putting break point between them or not. However, it is not sufficient or efficient for the process of syllable segmentation. Here is an introduction to existing segmentation rules for Myanmar language.

Segmentation rules related to Devowelising (e.g. "မင်း", "ကော်", "မြန်"), Syllable Chaining (e.g. "လက္ခဏာ", "အတ္တုပ္ပတ္တိ", "တက္ကသိုလ်", "စက္ကူ") , Kinzi (e.g. "သင်္ကန်း", "အင်္ကျီ", "စင်္ကြံ"), Loan Words (e.g. "မားစ်", "ဂျော့ချ်", "မစ္စတာ", "ဩဂုတ်", "အိန္ဒိယ"), Great Sa (e.g. "ဣဿာ", "အဝိဿဇ္ဇိက", "အသဿတ", "ဝေဿန္တရ") and Contractions (e.g. "ယောက်ျား", "ကျွန်ုပ်", "သို့", "ပိန္နဲပင်", "သင်္ဘော") were proposed by Zin Maung Maung and Yoshiki Mikami [102]. A test corpus containing 32,238 Myanmar syllables was tested with their developed rule-based syllable segmentation program, and it was reported that 99.96% accuracy was achieved.

Similar rule-based approach was proposed by Tun Thura Thet, Jin-Cheon Na and Wunna Ko Ko, and their rules were defined for "Single Character", "Special Ending Characters", "Second Consonant", "Last Character", "Next Starter" and "Miscellaneous". Here, "Miscellaneous" covers non-Myanmar characters, numeric characters, punctuation marks, spaces and similar characters [103]. Their proposed strategy can be divided into two parts; one is rule-based syllable segmentation and the other is dictionary based statistical syllable merging. They reported that their word segmentation achieved precision: 98.94%, recall: 99.05% and F-measure: 98.99%.

Here is a brief explanation on an overall process of rule-based syllable segmentation. First, the system converts the input string into equivalent category (e.g. "ကျောင်းသားများ" to "CMVVCAFCVFCMVF, here C=Consonant, M=Medial, V=Vowel, A=Athat, F=Final"), second, breaks the syllable boundary based on defined rules (e.g. CMVVCAF | CVF | CMVF), and then prints segmented string as a human readable format (e.g. "ကျောင်း|သား|များ"). Figure 12-1 shows the overall process of rule-based syllable segmentation flowchart.

Other applicable approaches or analysis methods on Myanmar syllable and word segmentation are Finite State Automata (FSA), morphological rules, rule-based stemmer, Statistical approach, etc. [104] [105]. Encoding methods such as ASCII and Unicode should be considered for defining rules and program development [106]. Syllable based dual weight algorithm for line braking in Myanmar Unicode was proposed by Keith Stribley who mentioned that comparing two Myanmar characters only to determine line breaking point is insufficient and requires complicated context analysis to solve ambiguity cases [107].

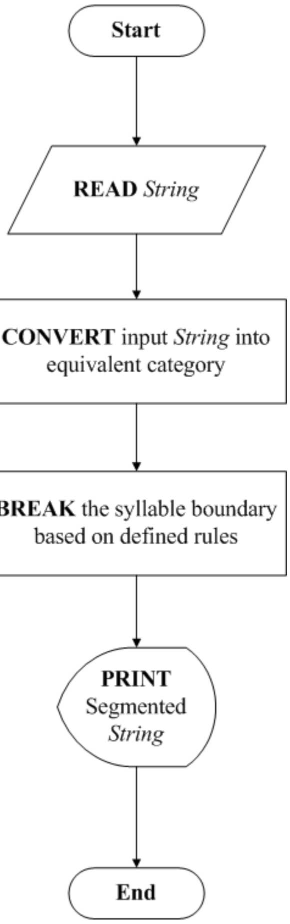


Figure 12-1. Flowchart of rule-based syllable segmentation overall process

Image taken from my published paper [135]

12.1.2 Segmentation Rule based on Positional Prediction

In general, Myanmar words are formed by one or more syllables. As mentioned in Chapter 1, Myanmar syllables are constructed with base consonant and surrounded vowels. All vowels have their defined writing position or combination patterns. My interests are those four directional writing positions (Left, Right, Up and Down) or combination patterns for developing Positional Prediction text input interfaces. And thus, the definition of a grapheme cluster in my research is not equivalent to that of a Unicode grapheme cluster in the meaning [108] [109]. My definition is as follows:

A grapheme cluster is a base character (typically a consonant) without or with vowel(s), medial(s), final(s) various sign(s) and stacked consonant(s). It is not necessary to be pronounceable or meaningful.

For example, there are "three" grapheme clusters in a Myanmar word "ကျောင်းသား" (student in English), which are "ကျော", "င်း" and "သား". However, there are only "two" grapheme clusters in Unicode, which are "ကျောင်း" (school) and "သား" (son).

Rule-based grapheme cluster segmentation method is a simple approach used for character grouping. There are only 4 rules in my system as follows:

Rule 1: *put break point in front of a character*

A break point is put in front of all Myanmar Consonants from "က" (Ka) to "အ" (A), including "ဉ" (Nya) and "ဿ" (Great Sa), all Myanmar Independent Vowels "ဣ" (I), "ဤ"(II), "ဥ"(U), "ဦ" (Uu), "ဩ" (O), "ဪ" (Au), four Various Signs "၌" (Locative), "၍" (Completed), "၎" (Aforementioned), "၏" (Genitive), all Myanmar Digits "၀" (Zero), "၁" (One), "၂" (Two), "၃" (Three), "၄" (Four), "၅" (Five), "၆" (Six), "၇" (Seven), "၈" (Eight), "၉" (Nine) , and all Myanmar Punctuations "၊" (Little Section) and "။" (Section).

Rule 2: *put break point at the end of a character*

A break point is put at the end of all Myanmar Independent Vowels except "ဦ" UU, four Myanmar Various Signs, all Myanmar Digits and all Myanmar Punctuations if they are still mixed with other characters.

Rule 3: *relating to non Myanmar characters*

A break point is put in front of and at the end of all non Myanmar characters including space character if they are still mixed with other characters.

Rule 4: *cleaning the mixed grapheme clusters*

Although Rule 1, Rule 2 and Rule 3 can change their working order (e.g. first, run Rule 2, second, run Rule1 and third, run Rule 3) or be combined as a function in the program, Rule 4 should start after checked and applied by other rules. The main purpose is to clean the mixed grapheme clusters and remove unnecessary break points (e.g. remove "|" break point or pipe character from "U+1039|consonant"). Some facilities are planned to be added based on system requirements (e.g. a new rule is added for breaking old Myanmar syllables) in this rule.

Note: Rule 1 to Rule 4 above are basically considered for Unicode encoding and correct typing order.

Here is an explanation with an example Myanmar sentence "အင်္ဂလိပ်လူမျိုးဂျော့ချ်"George"ကစက္ကူပေါ်၌၁*?!လို့ရေးခဲ့သည်။" (English man George wrote down ၁*?! on a piece of paper) to show how segmentation rule is based on PP work. This example sentence contains not only consonant and vowels but also English word "George", English loan word "ဂျော့ချ်" (i.e. spelling of "George" in Myanmar language), KinZi "င်္", subscripted character "္က", various sign "၌", Myanmar digit "၁", three English symbols "*", "?" and "!", etc., and Myanmar punctuation sign "။". "|" (pipe character) character is used for grapheme cluster separator symbol.

Input string:

အင်္ဂလိပ်လူမျိုးဂျော့ချ်"George"က စက္ကူပေါ်၌ ၁*?!လို့ရေးခဲ့သည်။

After Rule1 applied:

|အ|င ် ္|ဂ|လ ိ|ပ ်|လ ူ|မ ျ ိ ု း|ဂ ျ ေ ာ ် ့ |ခ ျ ်"George"|က space|စ|က ္ |က ူ |ပ ေ ါ ်
|၌space|၁*?!|လ ိ ု ့ |ရ ေ း|ခ ဲ ့ |သ|ည ်|။

After Rule2 applied:

|အ|င ် ္|ဂ|လ ိ|ပ ်|လ ူ|မ ျ ိ ု း|ဂ ျ ေ ာ ် ့ |ခ ျ ်"George"|က space|စ|က ္ |က ူ |ပ ေ ါ ်
|၌|space|၁|*?!|လ ိ ု ့ |ရ ေ း|ခ ဲ ့ |သ|ည ်|။|

After Rule3 applied:

|အ|င ် ္|ဂ|လ ိ|ပ ်|လ ူ|မ ျ ိ ု း|ဂ ျ ေ ာ ် ့ |ခ ျ ်|"|G|e|o|r|g|e|"|က|space |စ|က ္ |က ူ |ပ
ေ ါ ် |၌|space |၁|*|?|!|လ ိ ု ့ | |ရ ေ း|ခ ဲ ့ |သ|ည ်|။|

After Rule4 applied:

|အ|င ် ္ ဂ|လ ိ|ပ ်|လ ူ|မ ျ ိ ု း|ဂ ျ ေ ာ ် ့ |ခ ျ ်|"|G|e|o|r|g|e|"|က| space|စ|က ္ က ူ |ပ ေ

ါ ် |၌| space|၁|*|?|!|လ ိ ု ့| |ရ ေ း|ခ ဲ ့ |သ|ည ်|။|

Output as human readable format:

(|အ|င်္ဂ|လိ|ပ်|လူ|မျိုး|ဂျော့|ချ်|"|G|e|o|r|g|e|"|က| |စ|က္ကူ|ပေါ်|၌| |၁|*|?|!|လို့|ရေး|ခဲ့ |သ|ည်|။|)

Note: Some spaces are put so that Unicode encoding of Myanmar characters can be seen clearly. On the other hand, the space actually included in the input string is mentioned with a word "space".

12.1.3 Implementation

For the implementation, Visual Basic programming language was chosen, which is simple coding and suitable for rapid development. Although the segmentation tools were designed mainly to create Positional Prediction database for all Myanmar consonants, there are three main functions in total; which are 1) Positional Prediction combination pattern segmentation 2) character segmentation and 3) grapheme cluster segmentation based on Positional Prediction text input concept. User Interface of the prototype can be seen in Figure 12-2.

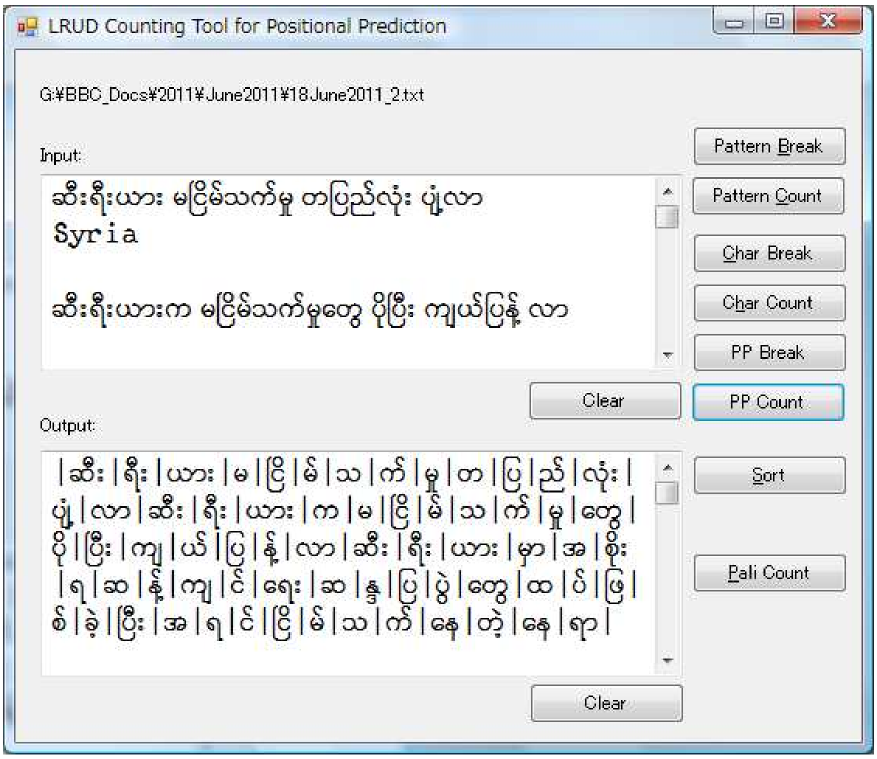


Figure 12-2. Grapheme cluster segmentation tool of Myanmar language

12.1.4 Positional Prediction Combination Patterns Segmentation

All dependent Myanmar characters such as dependent vowels, medial consonants and various signs were divided into four groups based on their four directional writing positions (i.e. Left, Right, Up or Down). Here, Left characters are "ေ" and "ြ", Right characters are "ျ", "ာ", "ါ" and "း", Up characters are "ိ", "ီ", "်", "ံ" and "ဲ" and Down characters are "ု", "ူ", "ွ", "ှ" and "့". Subscripted characters such as "္က" (subscripted Ka), "္ခ" (subscripted Kha), "္ဂ" (subscripted Ga), "္ဃ" (subscripted Gha) and "္စ" (subscripted Ca) were also considered and defined as down characters. The process of producing Positional Prediction patterns is not complex and they are working as character matching based on the defined patterns. The following is an example of input string and output:

Input String:	*Output:*
တက္ကသိုလ်ကကား	|C|PS|CUD|CU|C|CRR|
အင်္ကျီပေးမယ်။	|C|KS|CRU|CLR|C|CU|SB|
သားသမီးများ	|CRR|C|CUR|CRRR|
မန္တလေးကမမ	|C|PS|CLR|C|C|C|
ပါမောက္ခသတ္တိ	|CR|CLR|PS|C|PSU|
တွေမနေနဲ့	|CDL|C|CL|CUD|

Number of Combination Patterns:

C=11, PS=3, CUD=2, CU=2, CRR=2, KS=1, CRU=1, CLR=3, SB=1, CUR=1, CRRR=1, CR=1, PSU=1, CDL=1, CL=1

Here:

C = Consonant

L = Left character

R = Right character

U = Up character

D = Down character

PS = Consonant cluster (Padsint)

KS = Kinzi character

SB = Symbol

12.1.5 Character Segmentation

The process of character segmentation is to read Myanmar characters one by one from the input plain text file and put break point for every Myanmar character. This is considered as a necessary process for statistical analysis like character frequencies for Myanmar. The following is an example of input string and output:

Input String:

အင်္ကျီသင်္ကြန်ဂင်္ဂါမြစ်အင်္ဂါသားမင်းသိင်္ခ

Output:

|အ|င်္|က|ျီ|သ|င်္|က|ြ|န|်|ဂ|င်္|ဂ|ါ|မ|ြ|စ|်|အ|င်္|ဂ|ါ|သ|ား|မ|င|
်း|သ|ိ|င်္|ခ|

Number of Characters:

အ=2, င်္=5, က=2, ျီ=1, သ=3, ြ=2, န=1, ်=2, ဂ=3, ါ=2, မ=2, စ=1, ား=1, င=1, ်း=1, ိ=1, ခ=1

12.1.6 Grapheme Cluster Segmentation

Grapheme cluster segmentation process is segmentation of Myanmar word into grapheme clusters based on Positional Prediction text input concept [38] [39]. It is rule-based segmentation, and grapheme cluster boundary is determined by four rules. Explanation about four rules in details is shown in section 5. The following is an example of input string and output:

Input String:

မေမေမေမေပါပါပါပါကိုကိုကြေကြေနပ်နပ်လှလှကောကြေနပ်ရဲ့လား

Output:

|မေ|မေ|မေ|မေ|ပါ|ပါ|ပါ|ပါ|ကို|ကို|ကြေ|ကြေ|န|ပ်|န|ပ်|လှ|လှ|ကော|
ကြေ|န|ပ်|ရဲ့|လား|

Number of Grapheme Clusters:

မေ=4, ပါ=4, ကို=2, ကြေ=3, န=3, ပ်=3, လှ=2, ကော=1,
ရဲ့=1, လား=1

12.1.7 Results

This section discusses preprocess and difficulties of grapheme cluster segmentation by using existing Myanmar electronic documents like e-books. It also presents output results from experiments with correctly typed (i.e. correct typing order without spelling errors)

Myanmar-English dictionary in terms of Positional Prediction combination patterns and distribution of Myanmar characters based on the grapheme cluster segmentation.

Using Existing Electronic Document

Input text file used for grapheme cluster segmentation initial studies is the content of a Myanmar book named "အိန္ဒိယနိုင်ငံ၏ဗုဒ္ဓသာသနာတော်သမိုင်း" (history of Buddha religion in India) written by U Myint Swe and downloaded from "MyanmarISP" e-book website [110]. Its format is a plain text file of 62 pages in total. The reason for selecting the book relating to Buddha religion is that it contains not only Myanmar words but also Myanmar Pali words as well as loan words. Myanmar Pali words are also considered very important for text analysis and text input method interface of Myanmar language. This is because Myanmar Pali words use various consonant clusters such as "ဗုဒ္ဓ", "ဓမ္မ" and "အနုပ္ပဗန္ဓနပ္ပစ္စပြ္ပဋ္ဌာန", and play a very important role in Myanmar language.

Some preparations are required before starting segmentation process. Original PDF format e-book uses old Myanmar fonts "WinResearcher", "WinResearcherA", "WinTaungyiA", "WinPhaAn", "WinMonotype", etc., and all of them are ASCII fonts and do not follow Unicode encoding of Myanmar language. And thus, as a first step, a PDF e-book was converted to a plain text file by using freely available document converter "DocCharConvert Version 1.3" [111]. It was developed by ThanLwinSoft and supports conversion between different font encodings. "WinInnwa => Padauk" conversion option was used, and "Windows-1252" input encoding and "UTF-8" output encoding were selected to get Unicode plain text. For all three segmentation processes, "Myanmar3" Unicode font was used, which is Unicode encoding and recommended by Myanmar Unicode and NLP Research Center [29]. The second step is to fix conversion errors manually. There are four types of conversion errors in total. Examples of conversion errors and possible reasons to cause errors are as follows:

Typing mistake:

မစ္စိျမထေရ ← မဇ္ဈိမထေရ် (pg. 13)
(typed စ (Ca) + ျ (Ya) instead of ဈ (Jha))
ပဥ္စရဋ္ဌာနိ ← ပဉ္စရဋ္ဌာနိ (pg. 20)
(typed "ဥ" (U) character instead of "ဉ" (Nya Lay) character)
ခရစ်ှစ် ၄၅၅–ခုနှစ်၌ ← ခရစ်နှစ် ၄၅၅–ခုနှစ်၌ (pg. 43)
တော်င်စခန်းမြို့ဖြစ်၏။ ← တောင်စခန်းမြို့ဖြစ်၏။ (pg. 35)
ဟိန္ဒိူ ← ဟိန္ဒူ (pg. 47)

Typing order mistake:

ပငို ်ကြ၏။ ← ပိုင်ကြ၏။ (pg. 23)
ပါဠေိဝါဟာရအရ ← ပါဠိဝေါဟာရအရ (pg. 32)
တူးဖေါ်ရရသို မျှသော ← တူးဖေါ်ရရှိသမျှသော (pg. 52)

English to Unicode encoding conversion error:

ွမနနု ← Greek (pg. 23)
ဗီ ၈၀–၌ ← AD ၈၀–၌ (pg 40)
အမှတ် ှ–၅ စေတီ ← အမှတ် S-5 စေတီ (pg. 55)

Symbols to Unicode encoding conversion error:

▪ မှတ်ချက်။ ။ ← [မှတ်ချက်။ ။ (pg. 16)
္ထဂန္ဓာရတိုင်း ← *ဂန္ဓာရတိုင်း (pg. 20)
တွထံ ← တွင်– (pg 27)

Note: Left side of the arrow "←" is the conversion error and right side of the arrow is the original form. "pg. xx" refer the page number in original e-book.

Other interesting facts were also found that the Myanmar3 font does not support some character combinations and cannot display them smoothly. Although most of them are rarely used in the current Myanmar language, some are used frequently (e.g. ၅၀၀ိ, 500 Myanmar Kyat). This font display problem can be neglected if the combination order is correct, because it does not affect the segmentation process. The followings are examples of character combinations that cannot display smoothly with the Myanmar3 font:

ဂန္တွာ (cannot display န+္+တ+ွ combination) (pg. 20)
ဂိ့လာ (cannot display ဂ+ိ+့ combination) (pg. 20)
ဆုငွေ ၅၀၀ိ–ပေးမည် (cannot display ၀ (Digit Zero) + ိ (Vowel Sign I) (pg. 39)

Advantages of using existing electronic documents are 1) Myanmar e-books are available freely on the internet, 2) it is possible to collect large amount of data for creating PP database if existing Myanmar electronic contents are converted to Unicode encoding smoothly and 3) not only e-books but also many Myanmar blogs and websites use old ASCII based or partial Unicode fonts [28]. On the other hand, grapheme cluster segmentation from existing electronic documents such as plain text (.txt), Portable Document Format (.pdf), OpenDocument Text (.odt) and Word document (.doc, .docx) files requires checking conversion errors thoroughly.

Headers and footers of all pages have been removed in order to eliminate bias for words that occur in the header of every page. The output result summarized on Positional Prediction patterns is twenty eight and their number of distributions are “C=13996, CL=3211, CR=6027, CU=11220, CD=4361, CLR=2735, CLU=85, CLD=489, CRU=3501, CRD=942, CDU=6309, CLRU=1635, CLDR=198, CLDU=289, CLUDR=17, CRUD=603, I=245, IR=93, KS=273, PS=947, PSL=31, PSLR=19, PSR=343, PSU=317, PSD=125, PSDR=3, PSDU=23 and V=2597” (see Figure 11-3). Here, I = Independent Vowel, KS = Kinzi character, PS = Consonant + Subscripted consonant and V = Various Sign. The different combination orders are neglected (e.g. “L + R” and “R + L”, “C + L + D” and “C + D + L”) and considered as the same group. The number of occurrence of the same character group such as “CR”, “CRR” and “CRRR” is under a group “CR”. The followings are the Positional Prediction groups and their members:

Group:	***Member:***
C	C
CL	CL, CLL
CR	CR, CRR, CRRR
CU	CU, CUU
CD	CD, CDD
CLR	CLR, CLLR, CLRR, CRL, CRLR
CLU	CLU
CLD	CLD, CLDD, CLDL, CDL, CDDL, CDDLD, CDLD
CRU	CRU, CRUR, CUR, CURR, CURRR
CRD	CRD, CRDR, CRDDR, CDR, CDDR, CDDRR, CDRR
CDU	CDU, CUD, CUDD, CDDU, CDDUD, CDUD, CDUDD
CLRU	CLRU, CLUR, CRLRU
CDLR	CDLR, CLDLR, CLRD, CRDLR, CDDLR,
CLDU	CLDU, CLUD, CLUDD
CLUDR	CLDUR, CLUDR, CRDLRU, CDLRU
CRUD	CRDU, CRDUR, CRUD, CRUDD, CRUDR, CUDR, CDDUR, CDUDR, CDUR
I	I
IR	IR

KS	KS
PS	PS
PSL	PSL, PSLL
PSLR	PSLR
PSR	PSR, PSRR
PSU	PSU
PSD	PSD
PSDR	PSDR, PSRD
PSDU	PSDU, PSUD
V	V

To explain with more detail examples, CRU = C + R + U (e.g. “စီး”, “မ်း”, “ငိး”), CLUDR = C + L + U + D + R = (e.g. “ကြိုး”, “မော့်”, “ဖြိုး”), I = (e.g. “ဣ”, “ဤ”, “ဥ”), IR = I + R (e.g. “ဦး”), KS represents only Kinzi character “င်္” (e.g. C+KS+RU for “အင်္ကျီ”), PS (e.g. “က္က”, “ဍ္ဍ”, “ဋ္ဌ”), PSL = PS + L (e.g. “က္ခေ”, “န္ဒြ”, “စ္စေ”, “ဇ္ဇေ”), PSR = PS + R (e.g. “က္ကာ”, “ဏ္ဍား”, “န္ဒာ”, “ဇ္ဈာ”, “ပ္ပါ”, “ဗ္ဗာ”), PSDR = PS + D + R (e.g. “မ္ဘာ့”), etc. Among them, the top five occurrence patterns are “C = 13996, CU = 11220, CDU = 6309, CR = 6027 and CD = 4361”, and the last ten patterns are “PSDR=3, CLUDR=17, PSLR=19, PSDU=23, PSL=31” (see Figure 12-3).

Using Correctly Typed Content

Volunteer staff is hired for typing Myanmar words of the Myanmar-English dictionary that contains 26,485 Myanmar words [45]. Here are output results in detail.

The output result summarized on Positional Prediction patterns in the Myanmar-English dictionary is thirty one and their number of distributions are “C=22831, CL=3792, CR=8894, CU=26708, CD=7777, CLR=6792, CLU=313, CLD=1323, CRD=1853, CRU=8170, CUD=6426, CLUR=1390, CLRD=1205, CLUD=262, CRDU=2567, CLUDR=268, PS=595, PSL=41, PSR=303, PSU=117, PSD=92, PSLR=12, PSUR=1, PSUD=12, PSLRU=2, PSUDR=1, I=317, IR=190, IDU=3, V=12, KS=213” (see Figure 12-4). The followings are new Positional Prediction patterns found in the dictionary:

CLRL, CLLD, CLDLD, CRRD, CRDRR, CLLRU, CLLRD, CLDDR, CLDR, CLDRR, CRDL, CRDLRD, CDLRD, CRLRD, CLDUD, CRDUD, CRDUDR, CRDUDD, CLDDUR, PSUR, PSLRU, PSUDR, IDU

Here, PSUR = Consonant + Subscripted consonant + Up character + Right character (e.g. ဋ္ဌီး), PSLRU = Consonant + Subscripted consonant + Left character + Right character + Up character (e.g. မွှော်), PSUDR = Consonant + Subcripted consonant + Up character + Down character + Right character (e.g. ဋ္ဌိုး) and IDU = Independent Vowel + Down Character + Up Character (e.g. ဥုံ).

My segmentation tool can produce distributions of Myanmar consonants (see Figure 12-5), vowels, independent vowels (see Figure 12-10), consonant signs, various signs (see Fig. 12-10), symbols, athat or killer (see Figure 12-6), kinzi (see Figure 12-7), a consonant with stacked a consonant (see Figure 12-8), stacked consonants (see Figure 12-9), etc. Figure 12-11 clearly shows distribution of vowel tall AA or Mauk Cha. Making analysis of output results, it is possible to produce distributions of each Myanmar character and its combination patterns clearly. For example, the following are the numbers of Positional Prediction combination patterns for all Myanmar consonants including Nya Lay "ဉ" and Tha Gyi "ဿ".

က=104, ခ=77, ဂ=47, ဃ=5, င=41, စ=40, ဆ=29, ဇ=25, ဈ=3, ဉ=10, ည=30, ဋ=8, ဌ=7, ဍ=3, ဎ=2, ဏ=26, တ=59,ထ=28, ဒ=36, ဓ=10, န=87, ပ=79, ဖ=56, ဗ=51, ဘ=28, မ=112, ယ=30, ရ=66, လ=74, ဝ=25, သ=47, ဟ=28, ဠ=6, အ=24, ဿ=5

Here, consonant Ma "မ=112" is the highest number of Positional Prediction patterns, and its combinations are as follows:

မား, မြေး, မ, မျော, မ်း, မု, မာ, မူ, မူး, မော, မျ, မြော့, မြ, များ, မ်, မြော, မိ, မဲ့, မှ, မီ, မှူ, မ္ဘာ, မံ, မဲ, မိုး, မှော, မီး, မ့်, မြီး, မွ, မွေ, မှီ, မျိုး, မှာ, မြှော, မ္ပ, မ္ဗ, မ္ဘာ, မ္မ, မ္ဘ, မြွာ, မွဲ, မျှော်, မို, မြား, မြွေ, မြေ, မုံ, မှူး, မော့, မွေး, မြူ, မြိ, မှို, မြူး, မြီ, မျာ, မျှ, မြို့, မှီး, မြဲ, မြှ, မြှေ, မွေ့, မြှူ, မြု, မြော်, မွ, မှား, မေ, မွှေ, မျှား, မှူ, မြု, မေး, မြုံ့, မော်, မှဲ့, မ္ပာ, မြို, မွှေး, မွှား, မွု, မ္ဗူ, မြေး, မေ့, မုံ့, မြွှာ, မြုံ, မှေး, မ္မိ, မ္ဘိ, မို့, မှော်, မျှော့, မြုံး, မျှော, မျို, မြှို, မြွ, မြာ, မြံ, မြိုး, မှူ့, မျှို့, မွှေ့, မြွှ, မုံး, မ္မူ, မွှော်, မ္ဘူ, မ္ဗ

The most frequently used combinations of Ma "မ" consonants can be also found out from the results. For example, the top five frequently used grapheme clusters are 1) မ=1522, 2) မ်း=1224, 3) မြ=347 4) မ်=345 and 5) မျ=326.

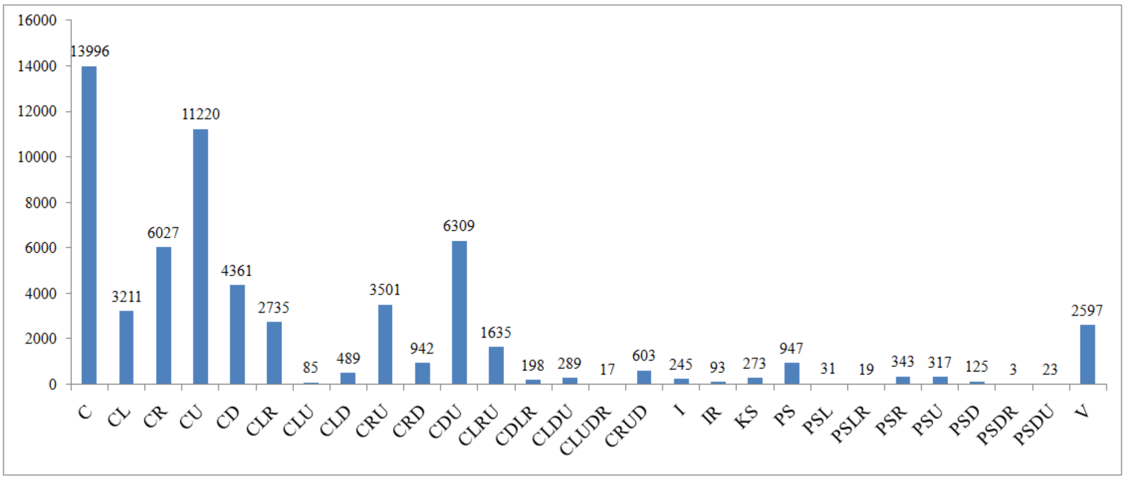


Figure 12-3. Distribution of Positional Prediction combination patterns in book titled "History of Buddha religion in India"

Image taken from my published paper [135]

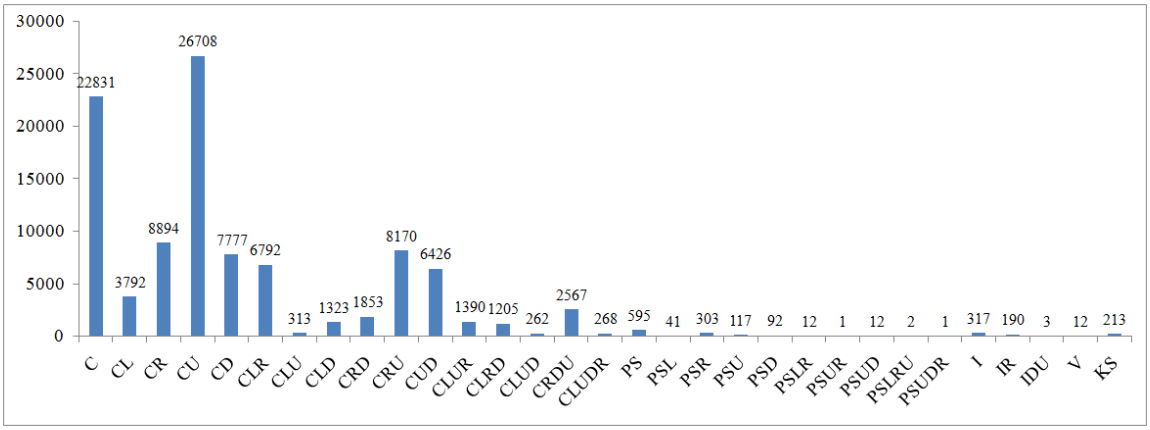


Figure 12-4. Distribution of Positional Prediction combination patterns in Myanmar-English dictionary

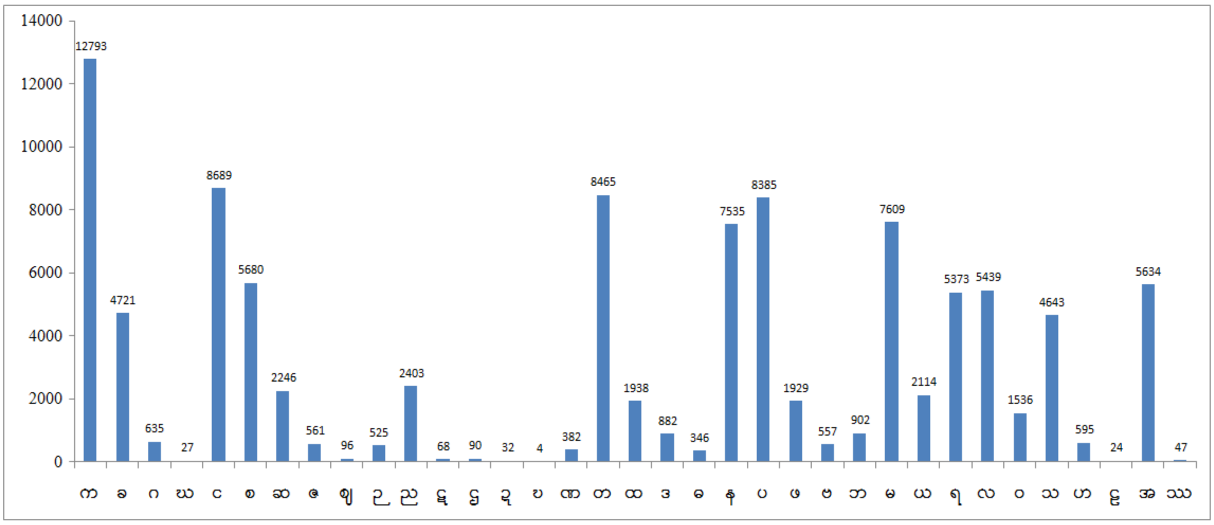


Figure 12-5. Distribution of Myanmar consonants in Myanmar-English dictionary

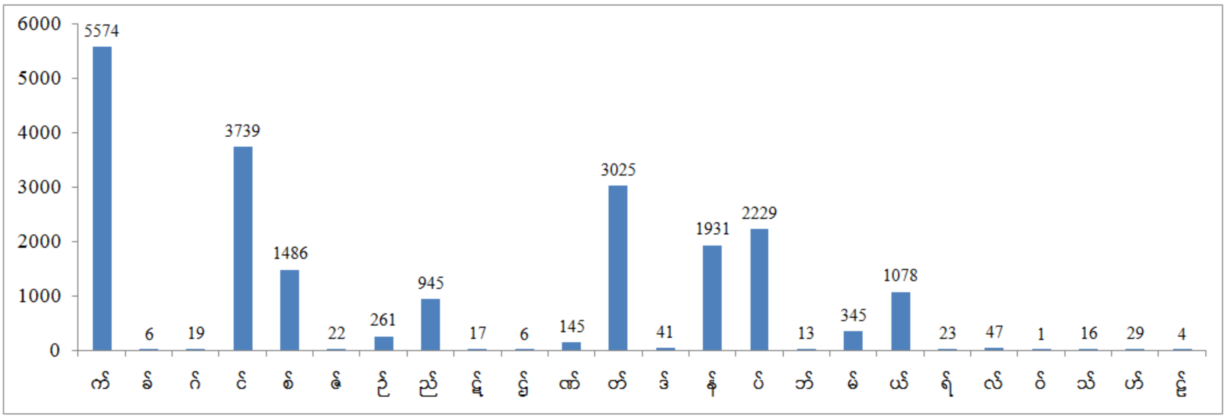


Figure 12-6. Distribution of athats or killers in Myanmar-English dictionary

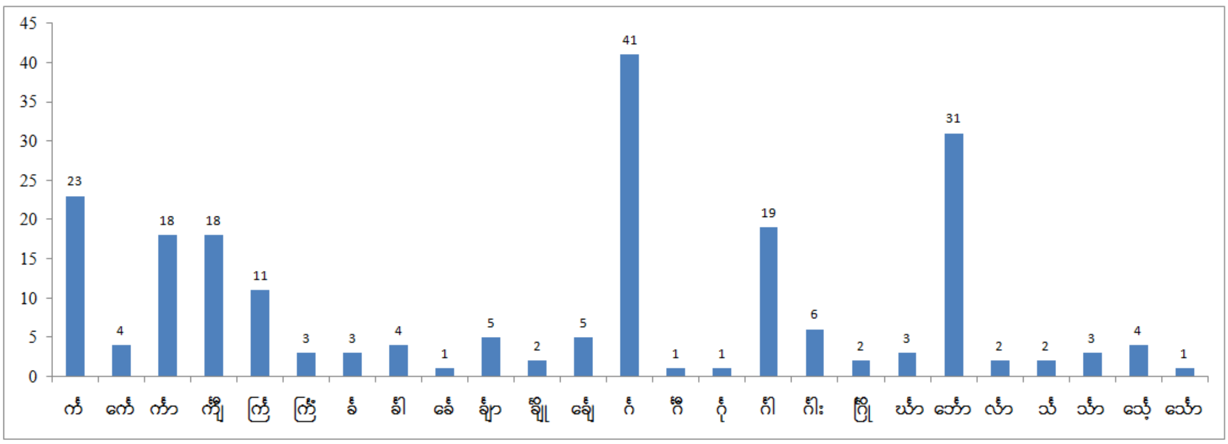


Figure 12-7. Distribution of kinzi in Myanmar-English dictionary

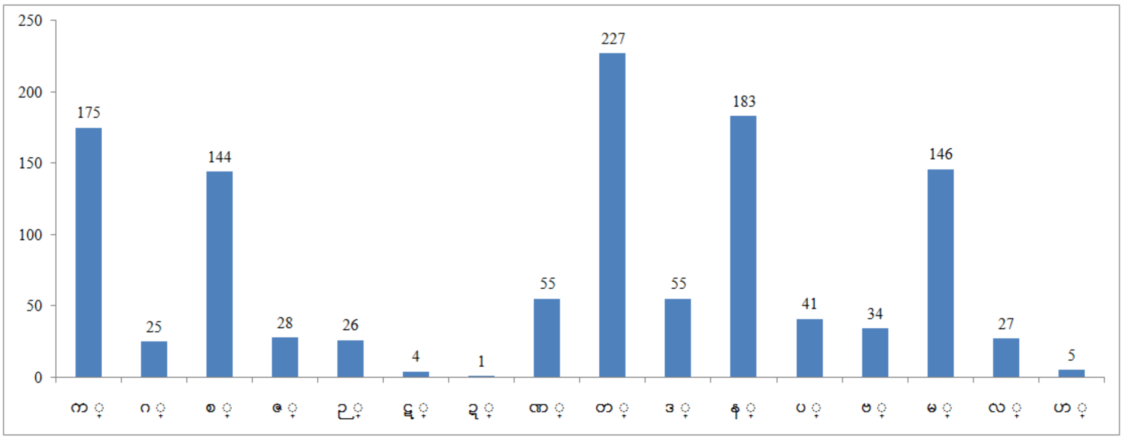


Figure 12-8. Distribution of consonants with stacked consonant in Myanmar-English dictionary

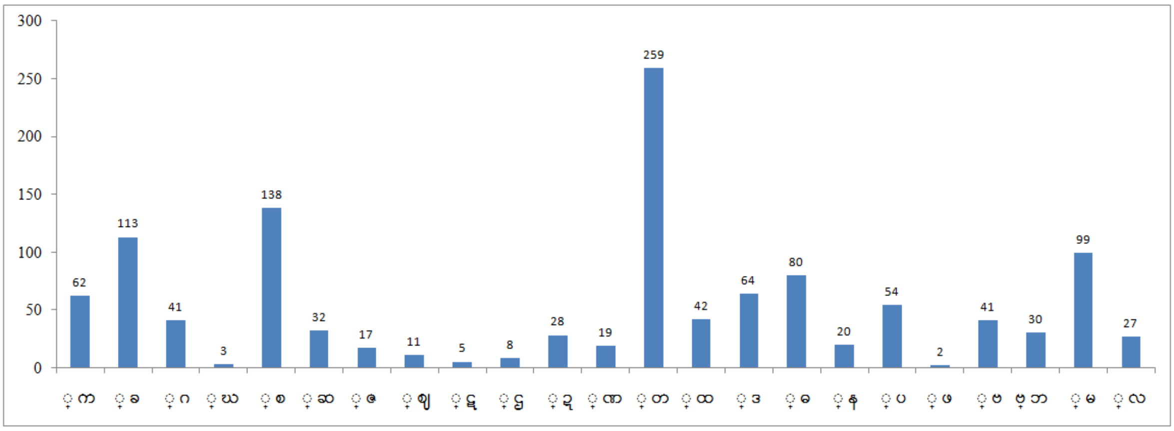


Figure 12-9. Distribution of stacked consonant in Myanmar-English dictionary

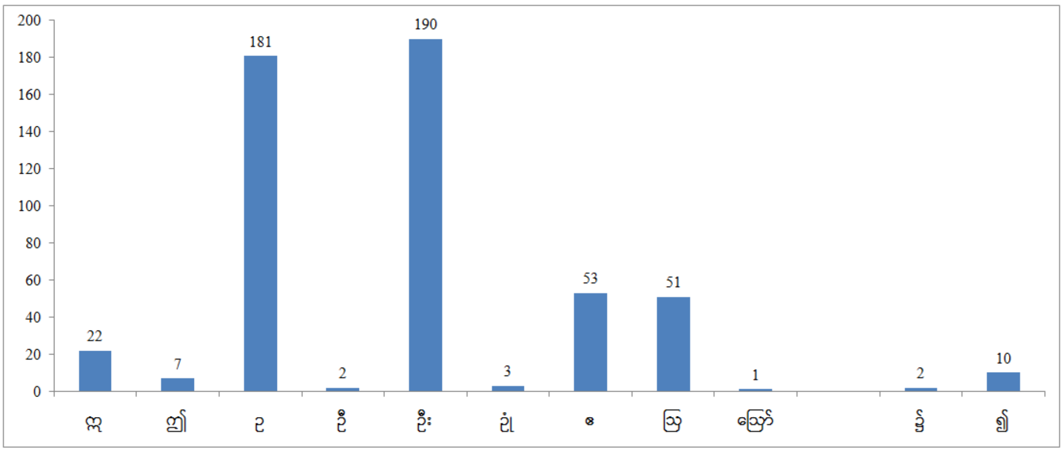


Figure 12-10. Distribution of independent vowels and various signs in Myanmar-English dictionary

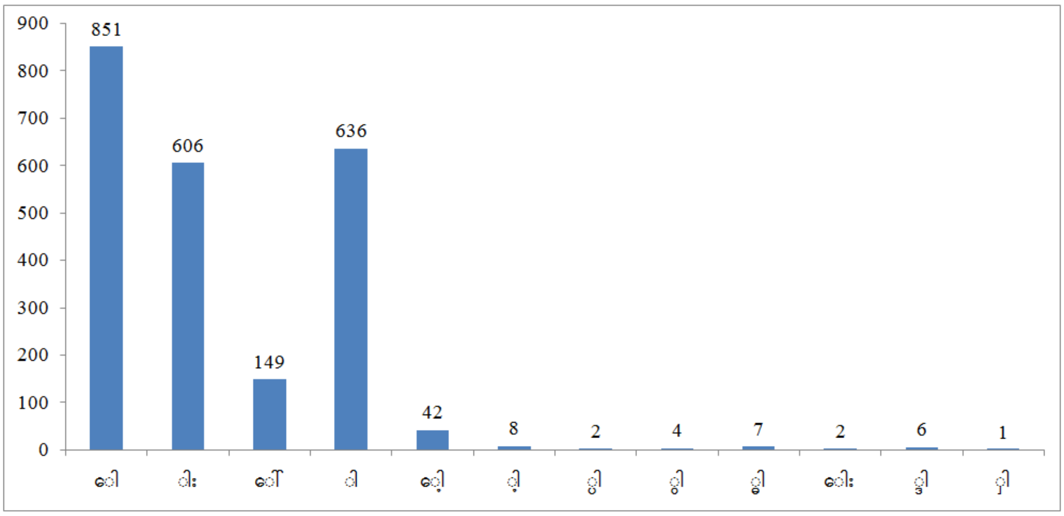


Figure 12-11. Distribution of vowel tall AA or Mauk Cha in Myanmar-English dictionary

Some regular expression patterns are used for searching and replacing of Myanmar text in my program [112]. For example, break point ("|", pipe character) is added between Kinzi and following consonant as follows:

```
Dim reg_exp As New Regex( _
        "(?<kinzi>င်္)(?<consonant>.)")
tempOutput = reg_exp.Replace( _
        tempOutput, "${kinzi}|${consonant}")
```

Round brackets group the regular expression between them. They capture the text matched by the regular expression inside them that can be referenced by the name between the sharp brackets [112]. If the input string is "အင်္ကျီသင်္ကြန်နင်္ဂါမြစ်", after running the regular expression mentioned above, the output string will be "အင်္|ကျီသင်္|ကြန်နင်္|ဂါမြစ်".

This research work is in progress, and I have been making an analysis of the outputs and updating the program but have not made a formal evaluation yet. The evaluation process for the mentioned three segmentation methods was done manually by checking all the Positional Prediction patterns and comparing the results between character segmentation and grapheme cluster segmentation. After updating conversion errors and typing mistakes of the contents of e-book and Myanmar-English dictionary several times, the final output was 100% correct for the three segmentation methods by the proposed rule. The main difference with existing syllable segmentation methods is that my segmentation rules mainly focus on a base consonant and other characters directly combined with the base consonant (i.e. Other characters are divided as separate syllables such as “ကြ|က်”, “မြ|င်း”, “သ|ည်”, “ချ|ဉ်”, “ဝိ|ဇ္ဇာ” and “ဟိ|မ|ဝ|န္တာ|တော|င်|တ|န်း”). Although the current results are based only on two Myanmar books, with large amount of data, the output results would be useful for statistical analysis on distributions of grapheme clusters, characters and Positional Prediction patterns of Myanmar language.

12.1.8 Summary

This section has reported my developing grapheme cluster segmentation tool based on Positional Prediction text input concept. Presentation was made on the segmentation rules based on Positional Prediction text input concept and the output results of initial study with a Myanmar e-book and follow-up study with a Myanmar-English dictionary in detail. This tool can provide an analysis of Positional Prediction patterns, distribution of Myanmar consonants, dependent vowels, independent vowels, stacked characters, etc. The current version of segmentation tool only provides the segmentation results by the defined format. There is also a plan to add new user interface for searching Myanmar character distributions of inputted text file by Positional Prediction patterns such as “CL” (i.e. any Consonant + Left character) for “ရေ”, “လေ”, “နေ”, “ကျေ” and “ကြေ”, and “မLR” (i.e. Ma Consonant + Left character + Right character) for “မော”, “မျော”, “မြော”, “မေး” and “မြေး” in the near future.

12.2 Development for Khmer Language

A prototype of grapheme cluster segmentation tool for Khmer language was developed by one of my colleagues Channa VAN with Python version 3.2. For a test data, a simple xml format of Khmer word list created by Chenda NOU was used, which was collected from Khmer spelling dictionary (published by National Language Institute in Phnom Penh in 2005). The xml data format is as follows:

```
  <?xml version="1.0" encoding="UTF-8" standalone="no" ?>
- <KhmerLexicon>
- <Entity>
  <HeadWord>ក</HeadWord>

  <POS>ន.</POS>

  <POS>កិ.</POS>

  </Entity>
```

- <Entity>

<HeadWord>ກໍ່</HeadWord>

<POS>ກິ.ວິ.</POS>

<POS>ບ.</POS>

</Entity>

- <Entity>

CHAPTER 13
Conclusion

13.1 Discussion

This dissertation presented three main proposals for text entry in Asian syllabic languages called "Positional Mapping", "Positional Gesture" and "Positional Prediction". These three proposals derived from the same core concept that is "designing text input interfaces based on character writing positions or vowel combination with a consonant". It is a simple but useful concept that enables to create a common and user-friendly text input interfaces for Asian syllabic languages. Moreover, based on this concept, it is possible to create new interaction techniques especially for mobile devices. Among these three methods, Positional Gesture is designed to work with simple gesture command and it targets physically-challenged users. Numerous Positional Mapping, Positional Gesture and Positional Prediction prototypes have been designed, developed and evaluated for Myanmar, Bengali, Khmer, Nepali and Thai languages. Although the results vary depending on the version, overall results proved that the mentioned three methods are user-friendly, usable and suitable for text entry in Asian syllabic languages.

13.1.1 Specification of Prototypes

Table 13-1 shows the specification visualization table of the developed prototypes of Positional Mapping, Positional Gesture and Positional Prediction. This table classifies text input techniques based on the number of software keys and hardware keys, input area size (small, medium, big), input devices (e.g. a mouse, a stylus, a keyboard, a joystick, etc.), input language (i.e. native language or Romanization) and input method (e.g. tapping, gesture, handwriting etc.)

Table 13-1. Specification of prototypes

Prototype	No of Software Keys	No of Hardware Keys	Input Area Size	Input Device	Input Language	Input Method
Positional Mapping (Mobile Phone)	20 Keys	0 Keys	Small	Mouse, Stylus, Touchpad	Native Language	Tapping
Positional Mapping (Software Keyboard)	(Myanmar), 55 Keys, 120 Keys (Bengali), 67 Keys, 122 Keys (Khmer)	0 Keys	Big	Mouse, Stylus, Touchpad	Native Language	Tapping
Positional Mapping (PDA)	17 Keys	0 Keys	Small	Mouse, Stylus, Touchpad	Native Language	Tapping
Positional Mapping (Ergodex DX1)	9 Keys	17 Keys (on Dx1)	Small	Ergodex DX1	Native Language	Tapping
Positional Mapping (Dual Joystick)	9 Keys	Dual Joystick Keys	Small	Dual Joystick Touchpad	Native Language	Tapping
Positional Gesture (PG)	8 Commands (Myanmar), 9 Commands (Khmer and Bengali)	Left and Right Click (in case of Mouse and Stylus)	Medium	Mouse, Stylus, Trackball, Touchpad and Wii Remote Control	Native Language	Gesture & Tapping
Positional Prediction (Software Keyboard)	55 (Myanmar and Nepali), 52 (Khmer and Bengali)	0 Keys	Medium	Mouse, Stylus and Touchpad	Native Language	Tapping
PP_Clickwheel	55 (Myanmar and Nepali), 52 (Khmer), 62 (Thai)	5 Keys (+ Clockwise and Anti-Clockwise Scrolling)	Small	Clickwheel	Native Language	Scrolling & Tapping
Roman_HW_PP	11 Keys (+ Handwriting Area)	0 Keys	Medium	Mouse, Stylus and Touchpad	Romanization	Handwriting
DKM Fingerspelling	38 Keys (Unshift) and 23 Keys (Shift) (including two Shift, Space, Backspace and Enter Keys)	38 Keys (Unshift) and 23 Keys (Shift)	Medium	Mouse, Stylus, Touchpad and Keyboard	Native Language	Tapping
RomanPP_Fingerspelling	31 Keys (including two Shift, Space, Backspace and Enter Keys)	31 Keys (including two Shift, Space, Backspace and Enter	Medium	Mouse, Stylus, Touchpad and Keyboard	Romanization	Tapping

The table indicates all of the prototypes and their main specifications for text entry described in this dissertation. Similarities as well as great diversities are apparent among the prototypes. Although text input interfaces of all the prototypes are based on the same concept, there are also great diversities when input devices are considered such as between a mobile phone keypad, a stylus and a clickwheel, a trackball and a touchpad, and this fact proves that my three proposals are applicable to various mobile devices.

The number of hardware and software keys usage differences of prototypes is also an important factor. It highlights that the mentioned three prototypes are not only usable with mobile keypads but also applicable to use with various input devices such as a large touch screen monitor and an electronic whiteboard.

Another important accomplishment of this work is that it covers three different areas; keyboard mapping (i.e. Positional Mapping), gesture interface (i.e. Positional Gesture) and predictive text input method (i.e. Positional Prediction).

13.1.2 Major Results

Beyond the specification of prototypes, this dissertation has presented empirical studies of Positional Mapping, Positional Gesture and Positional Prediction. In summary, the significant empirical results that highlight the strengths of my three approaches are as follows:

Positional Mapping (mobile phone): 31% faster than normal multi-tap for Myanmar language, and 46% faster than AKTEL key mapping in Bengali language.

Positional Mapping (software keyboard): 41% faster than Acharya and 58% faster than Akkhor with a tablet PC, and 86% faster than Acharya and 73% faster than Akkhor with an electronic whiteboard.

Positional Gesture: Average CPM was 14.29 with a pen and 19.27 with a mouse respectively for Myanmar language, and average CPM was 17.72 with a trackball and 17.69 with a mouse for Khmer language. Comparing Positional Gesture typing speed with a software keyboard, it is slower for both Myanmar and Khmer languages. The reason is that very few gesture commands are used in my prototypes to enable easier typing but software keyboards use one to one key mapping (e.g. 58 keys + shifted mode + right-alt-mode for Khmer software keyboard).

PP_Clickwheel: Average CPM was 18.9 with a clickwheel mouse for Khmer language. With the current PP_Clickwheel prototype, the fastest typing speed was 28.9 CPM and the slowest typing speed was 8.2 CPM.

Roman_HW_PP: Average CPM was 27.2 CPM by natives and 26.9 CPM by non-natives for Myanmar language. With the current Roman_HW_PP prototype, the fastest typing speed was 39.6 CPM by natives and 35.5 CPM by non-natives, and the slowest typing speed was 16.8 CPM by natives and 12.4 CPM by non-natives.

From the results of Likert scale evaluation on all the prototypes by users (see Table 4-3 and Table 5-1 for Positional Mapping, Table 6-3 and Table 6-4 for Positional Gesture, Table 7-1 for Positional Prediction, Table 8-2 for PP_Clickwheel, Table 9-2 for Roman_HW_PP and Table 10-5, Table 10-6 and Table 10-7 for Positional Prediction Fingerspelling), it can be said that my approach is a possible solution to find out common and user-friendly text input interfaces for Myanmar, Bengali, Khmer, Nepali and Thai languages.

13.2 Contributions

The contributions of the current work are divided into four categories; concepts, artifacts, methods and tools. Concepts are general proposals that may be applied to other text entry methods or interaction designs. Artifacts are developed prototypes useful for designing a real product and making simulation for other Asian syllabic languages. Methods are new empirical approaches used for designing and developing prototypes. Tools are developed font and utility software useful to other researchers.

13.2.1 Concepts

Keyboard or keypad mapping concept was proposed based on characters writing position of Asian syllabic languages. It was presented that this concept called *Positional Mapping (PM)* is a user-friendly keyboard mapping for native users of Myanmar, Khmer and Bengali languages, which is applicable for various computing devices such as a mobile phone, a PDA and an electronic whiteboard.

Gesture text input concept was proposed based on characters writing position of Asian syllabic languages. It was presented that this concept called *Positional Gesture (PG)* is a learnable gesture text input for native users of Myanmar and Khmer languages, which is usable with various input devices such as a mouse, a stylus, a trackball, a touchpad and a tablet PC.

Predictive text input concept was proposed based on characters writing position of Asian syllabic languages. It was presented that this concept called *Positional Prediction (PP)* is an effective and user-friendly predictive text input for native users of Myanmar and Khmer languages, which is applicable for various mobile devices such as a mobile phone, a PDA and the iPod.

13.2.2 Artifacts

Different versions of Positional Mapping prototype were developed for a mobile phone, a PDA, a software keyboard, the Ergodex DX1 input system and a joystick. All these prototypes support Myanmar, Bengali and Khmer languages, and can be extended to other Asian syllabic languages. Keypad mapping configuration file is used for a mobile phone prototype (see Figure 3-9 and Figure 3-10), and keypad layout can edit and apply on-the-fly.

Positional Gesture prototypes for Myanmar, Khmer and Bengali languages were developed and evaluated in conformity to the study results on the gesture text input interface based on characters writing position for Asian syllabic languages. Positional Gesture prototype requires no special recognition techniques or hardware development, and is usable with existing input devices such as a mouse, a trackball, a touchpad and a trackpoint. Study results with this prototype can guide future innovative gesture text input interfaces for Asian syllabic languages.

Different versions of Positional Prediction text input prototype were developed, some of which are not presented in this dissertation like Positional Prediction prototype for a Myanmar mobile phone [96]. All the prototypes are designed mainly for a mobile phone, a PDA and a touch screen computing device. This dissertation presented the basic prototype of Positional Prediction text input concept for Myanmar and Khmer languages, Positional Prediction for clickwheel mobile devices (PP_Clickwheel) for Khmer language, Romanized handwriting and Positional Prediction prototype (Roman_HW_PP) for Myanmar language and Romanized Positional Prediction prototype for Myanmar fingerspelling (RomanPP_Fingerspelling). These prototypes strongly demonstrated the concept of Positional Prediction text input, its extendibility not only to Asian syllabic languages but also to their fingerspelling and applicability for mobile devices.

Most of the prototypes mentioned above include mechanism to collect keystrokes for typing error and KSPC analysis. Some of the logged data are shown in Appendix A, Appendix B and Appendix C.

13.2.3 Methods

Positional Prediction text input method is required to build database of possible combinations of vowels with a consonant for each language (see “Appendix I” for Myanmar language). Based on this database, only possible candidate list of vowels with a selected consonant by users can be provided. Possible or impossible combinations were estimated by native users and already built for Myanmar, Nepali, Khmer and Thai languages. This approach and database should be useful for other predictive text input research of those languages.

Appendix G presents *PicMobi prototyping,* and Appendix H introduces pictures of Myanmar language PP text input interface designed for a mobile phone. This prototyping method is suitable for rapid development, which is similar to paper prototyping and easily implemented with a mobile phone and a tablet PC today. If user interface pictures are well-organized based on the scenario, user study can be held without a facilitator. This method is applicable for user interface design and development of various mobile applications.

"Romanized Handwriting" for "consonant" and "tapping" for "vowel" is a new and combined approach for text input interface research. Roman_HW_PP was designed, developed and evaluated as a good example (see Chapter 9). Touch screen technology has matured in the mobile phone application and this kind of "handwriting and tapping" or "combined" text input method is useful for future research.

Rule-based grapheme cluster segmentation method was proposed for Myanmar language (see Chapter 12). There are only 4 rules, which are Rule 1: putting break point in front of a character, Rule 2: putting break point at the end of a character, Rule 3: relating to non Myanmar characters and Rule 4: cleaning the mixed grapheme clusters. This grapheme cluster segmentation method may be applicable for other Asian syllabic languages.

13.2.4 Tools

Grapheme cluster segmentation tool (version 1.0) was developed for Myanmar language (Figure 12-2). It provides "Positional Prediction combination patterns segmentation", "character segmentation" and "grapheme cluster segmentation". The output is Unicode text, and it can be easily pasted into any text editor and statistics package. This tool would be applicable for Myanmar language NLP analysis.

Fingerspelling font was developed for the evaluation of DKM and RomanPP fingerspelling prototypes, and it is the first and only fingerspelling font for standard Myanmar fingerspelling (see Appendix F). This font should be useful for creation of fingerspelling educational contents.

13.3 Future Work

This dissertation has clearly presented that Positional Mapping, Positional Gesture and Positional Prediction are user-friendly and applicable for mobile devices as well as that they are common text input interfaces for Asian syllabic languages with Myanmar, Bengali, Khmer, Nepali and Thai languages. However, there is still a need to improve or update them. Some of the work in progress and future projects are highlights in this section.

Positional Mapping prototypes in this dissertation have focused only on keyboard mapping. Although positive feedbacks were received from users and appropriate typing speed was achieved (i.e. 42.59 CPM with a tablet PC and 50.12 CPM with an electronic whiteboard in Bengali, 34.62 CPM with a tablet PC using a stylus and 52.60 CPM with a tablet PC using a mouse for Khmer and 54.50 CPM with a notebook using a mouse for Myanmar), there is still a possibility of typing speed improvement if predictive text entry technique is used. Keyboard mapping by Positional Mapping and grapheme cluster prediction by Positional Prediction may enable better typing speed.

Positional Gesture with an eye tracking input device could be applicable for accessibility input. There is a plan for further refinements on the current prototype and follow-up analysis on the combination of Handwriting and Positional Gesture etc. in the near future.

The proposed Positional Prediction prototypes only provide the possible combinations of vowels with a consonant or grapheme cluster. Positional Prediction of word level may be successful, and development in the near future is taken into consideration.

Currently, there are three projects in progress relating to Positional Prediction. They are Positional Prediction with Handwriting, Positional Prediction for PDA, and Positional Prediction and LRUDC (Left, Right, Up, Down, Center) button. Positional Prediction with Handwriting interface allows users to write consonants by handwriting and predict the possible combinations with a consonant by four directional arrow keys. A prototype for Khmer was developed by Yin Huotely for her master thesis in Global Information and Telecommunication Studies, Waseda University. Based on the results of her research, average typing speed of Positional Prediction with Handwriting is 29.47 CPM, which is slightly higher than 28.12 CPM achieved with her proposed character based handwriting. The pilot study results of Positional Prediction for PDA with the current prototype are: KSPC (normal mode) is 1.91, KSPC (long press mode) is 1.45 and average CPM of three first-time users is 22.90 CPM [118]. The main idea of Positional Prediction and LRUDC interfaces is to apply Positional Prediction text input concept to each Myanmar consonant button or key. A software button is logically divided into 5 areas (i.e. Left, Right, Up, Down and Center), and users' average typing speed of the current LRUDC prototype is 34.61 CPM with a mouse and 31.14 CPM with a pen [119].

Myanmar character distribution analysis was made with grapheme cluster segmentation tool, and the results of experiments with a Myanmar-English dictionary were presented (Chapter 12). There is still a need for detail statistical analysis with a huge corpus to produce Positional Prediction text input candidate list sorted by frequently used grapheme clusters. And this kind of analysis will be required for other Asian syllabic languages.

New evaluation methods for Positional Mapping, Positional Gesture and Positional Prediction are also interesting as future work. Especially, operations during actual walking and with text-creation instead of text-copy task etc. may change users' typing speed and acceptance.

Finally, development of Positional Mapping, Positional Gesture and Positional Prediction prototypes for other Asian syllabic languages such as Hindi, Tibetan, Lao and Sinhala and implementation of user study may explore more these proposals.

13.4 Closing Remarks

This dissertation has attempted to demonstrate the following thesis:

> *A common text input interfaces such as Positional Mapping (PM), Positional Gesture (PG) and Positional Prediction (PP), Which are based on characters writing position, are user-friendly, effective on mobile devices and applicable for all Asian syllabic languages.*

Mobile devices are empowering millions of people in developed countries by giving them access to information, educational contents and services etc. in convenient ways. I predict that this wave will spread to Asian developing countries like Myanmar in the near future. However, there are still many difficulties in text messaging and information access and search in Asian syllabic languages, and overcoming of these difficulties is a challenge in the NLP research area. The user study results with Positional Mapping, Positional Gesture and Positional Prediction have indeed proven that common text input interfaces are available to Asian syllabic languages, and this approach might be a possible solution. I hope this dissertation will lead to the establishment of text input methods that are suitable for practical use in Asian syllabic languages.

APPENDIX A:

A Part of PG User Study Log File

User command operation for 4th Time:

Consonants, Right Vowels, Left Vowels, BACKSPACE, Right Vowels, BACKSPACE, Right Vowels, BACKSPACE, Right Vowels, Consonants, Left Vowels, Up Vowels, Right Vowels, Consonants, Left Vowels, Left Vowels, Consonants, Right Vowels, Consonants, Left Vowels, Up Vowels, Right Vowels, ENTER, Enter, Consonants, Consonants, Right Vowels, Consonants, Down Vowels, BACKSPACE, Down Vowels, Left Vowels, Right Vowels, Consonants, Up Vowels, Down Vowels, Right Vowels, Consonants, Down Vowels, Consonants, Consonants, Consonants, Up Vowels, Down Vowels, Right Vowels, ENTER, Enter, Consonants, BACKSPACE, Down Vowels, Consonants, Down Vowels, Left Vowels, Right Vowels, Consonants, Up Vowels, Down Vowels, Consonants, Up Vowels, Right Vowels, Consonants, Up Vowels, Consonants, Consonants, Consonants, BACKSPACE, BACKSPACE, Consonants, Up Vowels, Consonants, Consonants, Up Vowels, Down Vowels, Consonants, Right Vowels, Left Vowels, Right Vowels, Consonants, Consonants, Up Vowels, Down Vowels, Right Vowels, BACKSPACE, BACKSPACE, BACKSPACE, Down Vowels, Down Vowels, Up Vowels, Right Vowels, Consonants, Up Vowels, Consonants, Consonants, Left Vowels, Down Vowels, Consonants, Up Vowels, Consonants, Up Vowels, Down Vowels,

Transcribed text:

ကျားကြီးခြေရာကြီး

ပညာရွှေအိုးလူမခိုး

ရှေးထုံးလဲမပယ်ယနဲ့စျေးသုံးလဲမကြွယ်နဲ့

= =

Here,

Consonants = Command for displaying consonants

(i.e. double-click)

Left Vowels = Left gesture command for “left vowels”

(i.e. dragged mouse pointer to left direction)

Right Vowels = Right gesture command for “right vowels”

(i.e. dragged mouse pointer to right direction)

Up Vowels = Up gesture command for “upper vowels”

(i.e. dragged mouse pointer to up direction)

Down Vowels = Down gesture command for “down vowels”

(i.e. dragged mouse pointer to down direction)

BACKSPACE = Backspace command for deleting one character

(i.e. right-click drag operation)

APPENDIX B:

A Part of PP_Thumb User Study Log File

User Command Operation for 5th Time:

, B-Thu, L-Candidates, B-Ka, B-5, B-SelectedChar, B-Ya, B-1, B-Up, L-Candidates, B-Ka, B-2, B-Right, L-Candidates, B-Ka, B-5, B-Up, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Pa, B-5, B-SelectedChar, B-Ta, B-1, B-Left, B-Down, L-Candidates, B-Ya, B-2, B-SelectedChar, B-Ta, B-1, B-Right, L-Candidates, B-Ka, B-1, B-Left, B-Right, L-Candidates, B-Pa, B-1, B-Left, B-Up, L-Candidates, B-Ta, B-5, B-Left, B-Right, B-Up, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Ta, B-5, B-Left, L-Candidates, B-Ka, B-1, B-Left, B-Right, L-Candidates, B-Ka, B-5, B-Up, B-Right, L-Candidates, B-Ya, B-2, B-SelectedChar, B-LRUD, B-Up, B-Down, L-Candidates, B-LRUD, B-Up, L-Candidates, B-Ya, B-3, B-SelectedChar, B-LRUD, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Ka, B-5, B-Right, B-Down, L-Candidates, B-Pa, B-2, B-Down, L-Candidates, B-Ta, B-5, B-Up, L-Candidates, B-Pa, B-1, B-Right, L-Candidates, B-Ta, B-1, B-Up, L-Candidates, B-Ha, B-3, B-SelectedChar, B-Ya, B-5, B-SelectedChar, B-Ca, B-1, B-Up, L-Candidates, B-Ka, B-1, B-SelectedChar, B-ModeForward, B-Pa, B-Thu, B-Thu, B-Ha, B-Ta, B-Pa, B-Genitive, B-ModeBackward, B-PtPm, L-Candidates, B-Enter, B-Ha, B-3, B-Right, L-Candidates, B-Ta, B-1, B-Up, B-Down, L-Candidates, B-Ha, B-3, B-SelectedChar, B-Ka, B-2, B-Right, L-Candidates, B-Pa, B-2, B-Down, L-Candidates, B-Ta, B-5, B-Up, B-Right, L-Candidates, B-Pa, B-1, B-Left, L-Candidates, B-Ta, B-5, B-Up, L-Candidates, B-Ca, B-2, B-SelectedChar, B-Ka, B-1, B-Up, L-Candidates, B-Ka, B-1, B-Down, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Ta, B-3, B-Right, L-Candidates, B-Pa, B-1, B-Up, L-Candidates, B-Ta, B-5, B-Left, B-Right, B-Up, L-Candidates, B-PtPm, L-Candidates

ks:198

Transcribed text:

သူငယ်ချင်း၊

မတွေ့ရတာကြာပြီနော်။

နေကောင်းရဲ့လား။

ငါ့ဖုနံပါတ်အသစ်က၅၀၀၇၄၅၉၊

Appendix

အားတဲ့အခါဖုန်းပြန်ဆက်ကွာ။

ဒါပဲနော်။

= =

User Command operation for 6th Time:

, B-Thu, L-Candidates, B-Ka, B-5, B-SelectedChar, B-Ya, B-1, B-Up, L-Candidates, B-Ka, B-2, B-Right, L-Candidates, B-Ka, B-5, B-SelectedChar, B-LRUD, B-Up, L-Candidates, B-LRUD, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Pa, B-5, B-SelectedChar, B-Ta, B-1, B-Left, B-Down, L-Candidates, B-Ya, B-2, B-SelectedChar, B-Ta, B-1, B-Right, L-Candidates, B-Ka, B-1, B-Left, B-Right, L-Candidates, B-Pa, B-1, B-Left, B-Up, L-Candidates, B-Ta, B-5, B-Left, B-Right, B-Up, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Ta, B-5, B-Left, L-Candidates, B-Ka, B-1, B-Left, B-Right, L-Candidates, B-Ka, B-5, B-Up, B-Right, L-Candidates, B-Ya, B-2, B-Up, B-Down, L-Candidates, B-Ya, B-3, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-Ka, B-5, B-Right, B-Down, L-Candidates, B-Ka, B-Pa, B-2, B-Down, L-Candidates, B-Ta, B-5, B-Up, B-Right, L-Candidates, B-Ta, B-5, B-Up, L-Candidates, B-Pa, B-1, B-Right, L-Candidates, B-Ta, B-1, B-Up, L-Candidates, B-Ha, B-3, B-SelectedChar, B-Ya, B-5, B-SelectedChar, B-Ca, B-1, B-Up, L-Candidates, B-Ka, B-1, B-SelectedChar, B-ModeForward, B-Pa, B-Thu, B-Thu, B-Ha, B-Ta, B-Pa, B-Genitive, B-PtPm, B-PtPm, B-BackSpace, B-BackSpace, B-ModeBackward, B-PtPm, L-Candidates, B-Enter, B-Ha, B-3, B-Right, L-Candidates, B-Ta, B-1, B-Up, B-Down, L-Candidates, B-Ha, B-3, B-SelectedChar, B-Ka, B-2, B-Right, L-Candidates, B-Pa, B-2, B-Down, B-SelectedChar, B-LRUD, B-Down, L-Candidates, B-Ta, B-5, B-Up, L-Candidates, B-LRUD, B-Right, L-Candidates, B-Pa, B-1, B-Left, L-Candidates, B-Ta, B-5, B-Up, L-Candidates, B-Ca, B-2, B-SelectedChar, B-Ka, B-1, B-Up, L-Candidates, B-Ka, B-1, B-Down, B-Right, L-Candidates, B-PtPm, L-Candidates, B-Enter, B-3, B-Right, L-Candidates, B-Pa, B-1, B-Up, L-Candidates, B-Ta, B-5, B-Left, B-Right, B-Up, L-Candidates, B-PtPm, L-Candidates

ks: 209

Transcribed text:

သူငယ်ချင်း၊

မတွေ့ရတာကြာပြီနော်။

နေကောင်းရဲ့လား။

ငါ့ဖုန်းနံပါတ်အသစ်က၅၀၀၇၄၅၉၊

အားတဲ့အခါဖူန်းပြန်ဆက်ကွာ။

ဒါပဲနော်။

= =

Here,

B-Thu = Pressed “Thu” button

L-Candidates = Choose a Listed Candidate

B-5 = Pressed “Five” button

B-SelectedChar = Pressed “Selected Character” button (i.e. consonant)

B-LRUD = Pressed “LRUD” button

(this button is designed for editing dependent vowel and medial characters,

LRUD means Left, Right, Up and Down)

B-PtPm = Pressed “PtPm” button

(this button is designed for Myanmar symbols “၊” (U+104A) and “။” (U+104B))

B-BackSpace = Pressed “BackSpace” button

B-Enter = Pressed “Enter” button

ks: 209

(ks stand for “total keystrokes”. Here, 209 is the total number of keystrokes typed by users)

APPENDIX C:

A Part of PP_Clickwheel User Study Log File

(18 Feb 2009)

Ko Thein (1st time)

သူငယ်ချင်း၊မတွေ့ရတာကြာပြီနော်။နေကောင်းရဲ့လား။ငါ့ဖုန်းနံပါတ်အသစ်က၅၀၀၇၄၅၉၊အားတဲ့အခါဖုန်းပြန်ဆက်ကွာ။ဒါပဲနော်။

, B-Left, B-Center, B-Center, B-Down, B-Right, B-Right, B-Center, B-Center, B-Center, BackSpace, B-Center, BackSpace, B-Center, B-Center, B-Center, B-Center, B-Down, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Up, B-Right, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Left, B-Down, B-Down, B-Left, B-Up, B-Up, B-Up, B-Up, B-Center, B-Center, B-Center, BackSpace, B-Center, B-Center, B-Center, B-Center, B-Left, B-Down, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Left, B-Right, B-Center, B-Center, B-Center, B-Left, B-Up, B-Center, B-Center, B-Center, B-Left, B-Right, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Left, B-Center, B-Center, B-Center, B-Left, B-Right, B-Center, B-Center, B-Center, B-Up, B-Right, B-Center, B-Center, B-Center, B-Up, B-Down, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, BackSpace, B-Center, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Center, B-Center, PgDn, B-Center, B-Center, B-Center, B-Center, B-Right, B-Down, B-Center, B-Center, B-Center, B-Down, B-Center, B-Center, B-Center, B-Up, B-Right, B-Center, B-Center, B-Right, B-Center, B-Up, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Up, B-Down, B-Center, B-Center, B-Center, B-Center, B-Center, B-Right, B-Center, B-Right, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, BackSpace, B-Center, B-Center, B-Center, B-Center, B-Down, B-Center, B-Center, B-Center, B-Up, B-Right, B-Center, B-Center, B-Center, B-Left, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, BackSpace, B-Center, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Down, B-Right, B-Center, B-Center, B-Center, B-Center, B-Center, B-Right, B-Center, B-Center, B-Center, B-Up, B-Center, B-Center, B-Center, B-Left, B-Right, B-Up, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Center, B-Up

B-Center = 88

B-Left = 11

B-Right = 21

B-Up = 23

B-Down = 11

BackSpace = 6

PgDn = 1

Here,

B-Center = Pressed “Center” button

B-Left = Pressed “Left” button

B-Right = Pressed “Right” button

B-Down = Pressed “Down” button

BackSpace = Pressed “Backspace” button

APPENDIX D:

An Example of Likert Scale Evaluation on PM

User Name: Kabir , Age: 32

Date: 25 December 2007

Tablet PC, Text Input Method: Akkhor,

Likert Scales (range 1-5)	1	2	3	4	5	avg
Difficult-Easy		x				
Painful-Enjoyable			x			
Slow-Fast			x			
Dislike-Like			x			

Text Input Method: Acharya

Likert Scales (range 1-5)	1	2	3	4	5	avg
Difficult-Easy			x			
Painful-Enjoyable				x		
Slow-Fast				x		
Dislike-Like				x		

Text Input Method: Positional Mapping

Likert Scales (range 1-5)	1	2	3	4	5	avg
Difficult-Easy				x		
Painful-Enjoyable				x		
Slow-Fast					x	
Dislike-Like					x	

APPENDIX E:

An Example of Likert Scale Evaluation on DKM and RomanPP_Fingerspelling

[illegible]

Date: 18 . 3 . 2010

User Name: nyeinmahaw

Age: 28

Sex: Femal

DTP Work Experience: 7 years

Education: B.A Philosophy

Direct Keyboard Mapping (DKM)

Likert Scales	1	2	3	4	5
Difficult-Easy					✓
Painful-Enjoyable				✓	
Slow-Fast					✓
Dislike-like					✓

(1) 0 . 12 . 68
(2) 0 . 21 . 13
(3) 0 . 19 . 89
(4) 0 . 15 . 56
(5) 0 . 19 . 20
(6) 0 . 14 . 70
(7) 0 . 11 . 56
(8) 0 . 15 . 48
(9) 0 . 13 . 25
(10) 0 . 14 . 47.

Comment and Suggestion:

Romanized Positional Prediction (RomanPP)

Likert Scales	1	2	3	4	5
Difficult-Easy					✓
Painful-Enjoyable			✓		
Slow-Fast				✓	
Dislike-like					✓

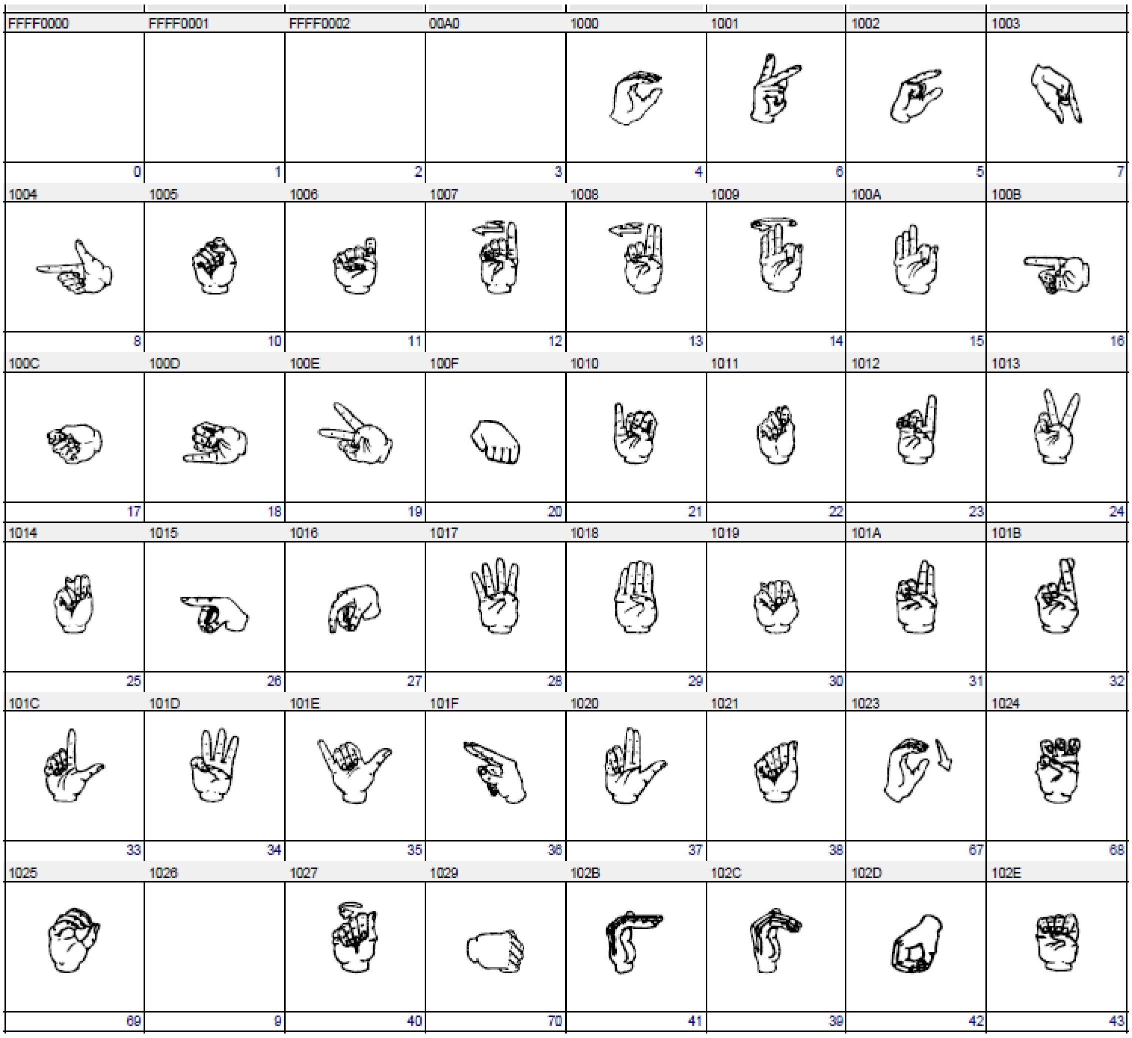
(1) 3 . 26 . 41
(2) 3 . 53 . 96
(3) 3 . 12 . 74
(4) 3 . 22 . 63
(5) 3 . 18 . 41
(6) 3 . 40 . 03
(7) 2 . 58 . 06
(8) 2 . 54 . 91
(9) 2 . 48 . 29
(10) 2 . 50 . 71.

Comment and Suggestion:

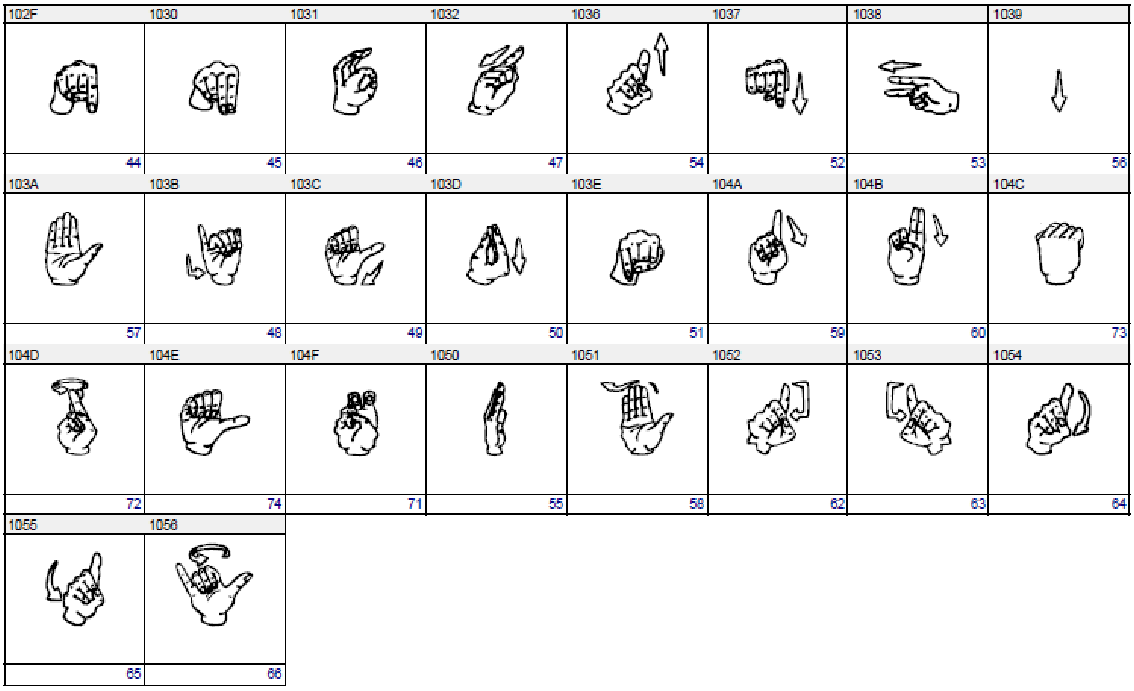
[illegible] Fin gerspelling [illegible]
[illegible] Fingerspelling [illegible]
[illegible]
[illegible]

APPENDIX F:

Font Table of mmFingerspelling.ttf (Ver. 2)



FFFF0000	FFFF0001	FFFF0002	00A0	1000	1001	1002	1003
0	1	2	3	4	6	5	7
1004	1005	1006	1007	1008	1009	100A	100B
8	10	11	12	13	14	15	16
100C	100D	100E	100F	1010	1011	1012	1013
17	18	19	20	21	22	23	24
1014	1015	1016	1017	1018	1019	101A	101B
25	26	27	28	29	30	31	32
101C	101D	101E	101F	1020	1021	1023	1024
33	34	35	36	37	38	67	68
1025	1026	1027	1029	102B	102C	102D	102E
69	9	40	70	41	39	42	43


102F
44
1030
45
1031
46
1032
47
1036
54
1037
52
1038
53
1039
56
103A
57
103B
48
103C
49
103D
50
103E
51
104A
59
104B
60
104C
73
104D
72
104E
74
104F
71
1050
55
1051
58
1052
62
1053
63
1054
64
1055
65
1056
66

APPENDIX G:

PicMobi Prototyping

Theoretically, low- and high-fidelity prototyping have their roles, and both of these techniques are useful for the product development life cycle. However, it is sometimes necessary to create a text input interface prototype rapidly to explore my new idea and provide visual and physical experiences that far exceed traditional paper prototyping. Are there any possible solutions? One is “prototyping with pictures and real mobile devices”.

Today pictures can be uploaded to mobile devices in various ways easily. Most mobile devices support a number of transfer techniques such as memory stick, USB cable, Infrared (IrDA), Bluetooth, http upload and email. Using picture uploading feature for mobile interface prototyping, this prototyping method can be applied to any mobile device and suitable for rapid prototyping. My approach is very simple and similar to paper prototyping technique, but it draws digital pictures by using drawing software and uploads them to mobile devices. The followings are three key steps for prototyping with pictures and mobile devices:

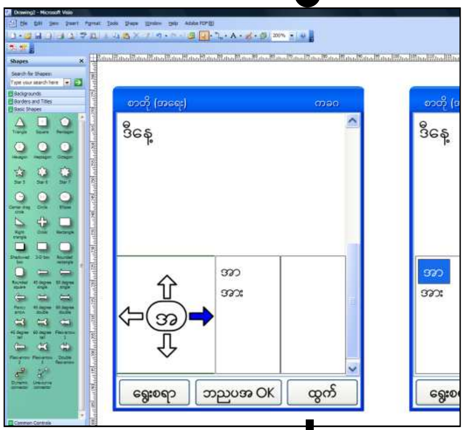
Step 1: Design or draw key screens of user interface with drawing software (e.g. Microsoft Office Visio) based on the user study scenario, and save each screen as picture format (e.g. JPEG, GIF, PNG, TIFF) with the resolution (e.g. 480x320 pixels for Apple iPhone, 800x480 pixels for Nokia N900) supported by mobile devices for user study.

Step 2: Arrange the sequence of representative screen pictures according to the user study scenario (e.g. save ordered alphabetically such as “a, b, c …” or numerically such as “1, 2, 3 …”).

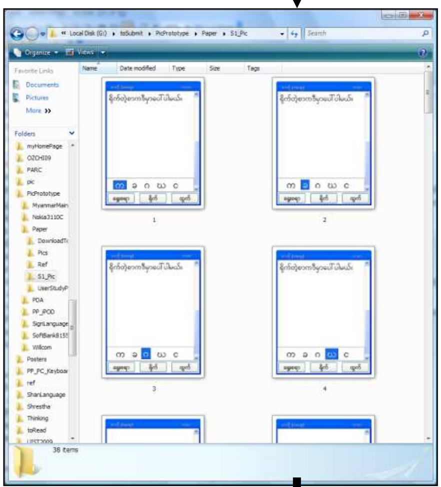
Step 3: Upload the screen pictures to the mobile device, and confirm the display sequence and resolution. Now, the process of “prototyping with picture and mobile devices” is already finished.

Key steps for prototyping with pictures and mobile devices:

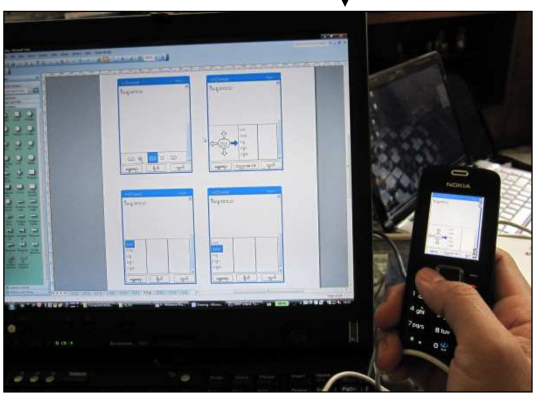
Step1: (Designing UI)



Step 2: (Arrange the sequence of display)



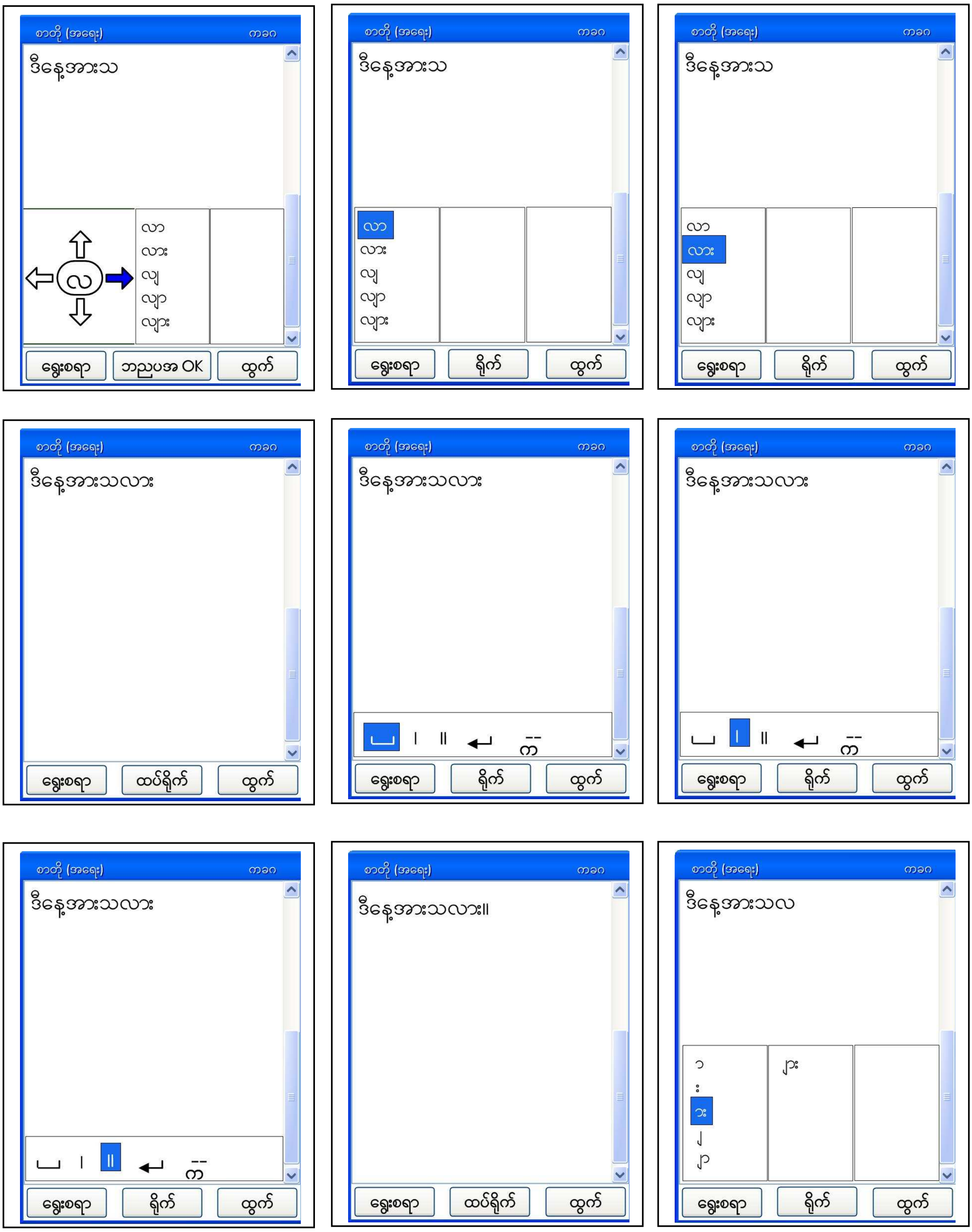
Step 3: (Upload pictures to mobile device)



APPENDIX H:

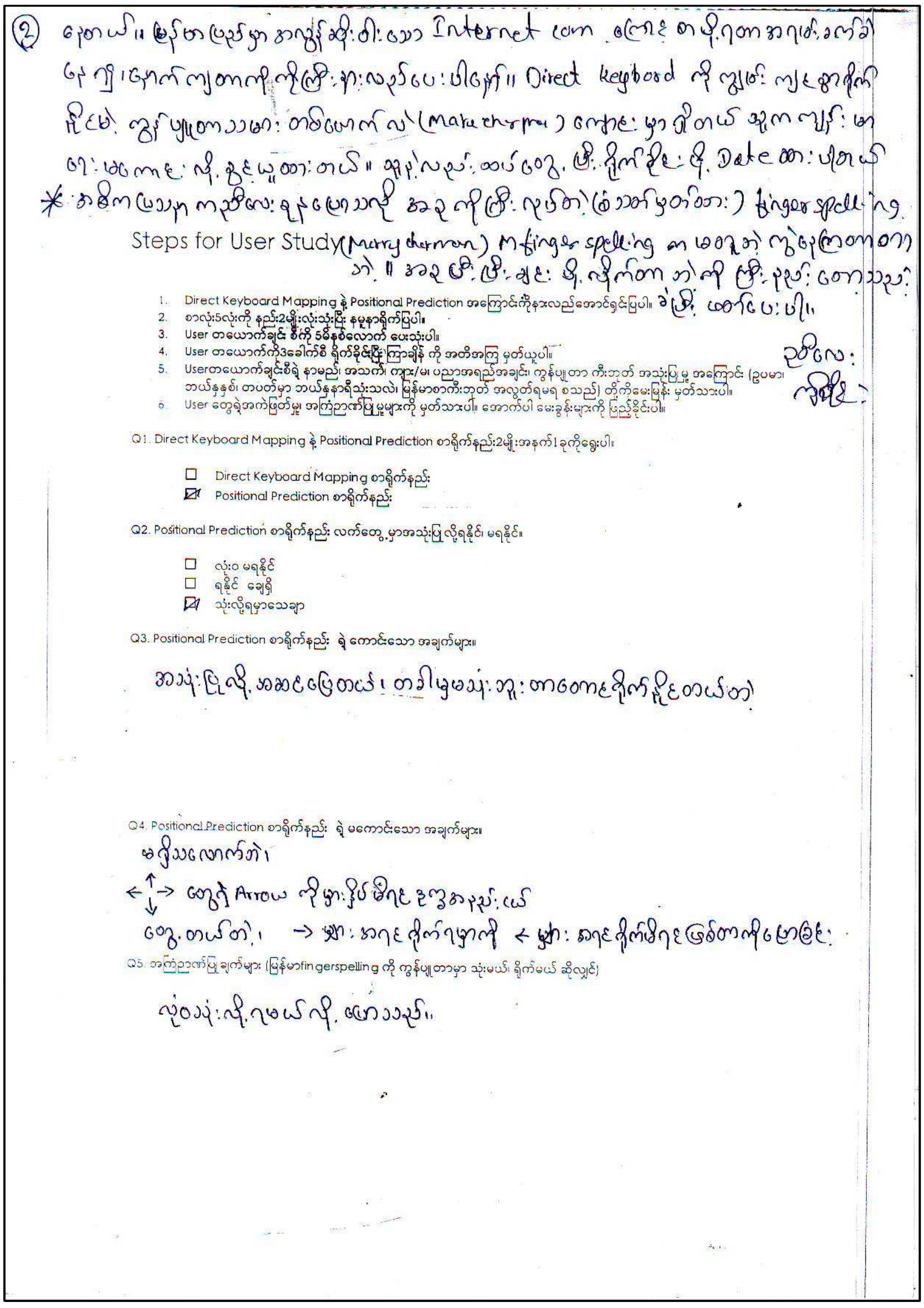
Pictures of PicMobi Prototyping

(PP Text Input Interface for Mobile Phone)



APPENDIX I:

An Example of User Study Manual of RomanPP Fingerspelling


Steps for User Study

1. Direct Keyboard Mapping နဲ့ Positional Prediction အကြောင်းကိုနားလည်အောင်ရှင်းပြပါ။
2. စာလုံး5လုံးကို နည်း2မျိုးလုံးသုံးပြီး နမူနာရိုက်ပြပါ။
3. User တယောက်ချင်း စီကို 5မိနစ်လောက် ပေးသုံးပါ။
4. User တယောက်ကို3ခေါက်စီ ရိုက်ခိုင်းပြီး ကြာချိန် ကို အတိအကျ မှတ်ယူပါ။
5. Userတယောက်ချင်းစီရဲ့ နာမည်၊ အသက်၊ ကျား/မ၊ ပညာအရည်အချင်း၊ ကွန်ပျူတာ ကီးဘုတ် အသုံးပြုမှု အကြောင်း (ဥပမာ၊ ဘယ်နှနှစ်၊ တပတ်မှာ ဘယ်နှနာရီသုံးသလဲ၊ မြန်မာစာကီးဘုတ် အလွတ်ရမရ စသည်) တို့ကိုမေးမြန်း မှတ်သားပါ။
6. User တွေရဲ့အကဲဖြတ်မှု အကြံဥာဏ်ပြုမှုများကို မှတ်သားပါ။ အောက်ပါ မေးခွန်းများကို ဖြည့်ခိုင်းပါ။

Q1. Direct Keyboard Mapping နဲ့ Positional Prediction စာရိုက်နည်း2မျိုးအနက်1ခုကိုရွေးပါ။

- [ ] Direct Keyboard Mapping စာရိုက်နည်း
- [x] Positional Prediction စာရိုက်နည်း

Q2. Positional Prediction စာရိုက်နည်း လက်တွေ့မှာအသုံးပြုလို့ရနိုင်၊ မရနိုင်။

- [ ] လုံးဝ မရနိုင်
- [ ] ရနိုင် ချေရှိ
- [x] သုံးလို့ရမှာသေချာ

Q3. Positional Prediction စာရိုက်နည်း ရဲ့ ကောင်းသော အချက်များ။

Q4. Positional Prediction စာရိုက်နည်း ရဲ့ မကောင်းသော အချက်များ။

Q5. အကြံဥာဏ်ပြုချက်များ (မြန်မာfingerspelling ကို ကွန်ပျူတာမှာ သုံးမယ်၊ ရိုက်မယ် ဆိုလျှင်)

APPENDIX J:

Possible Combinations of Myanmar Vowels, Medials with a Consonant

C က

CL ကြ, ကေ, ကြေ

CR ကျ, ကား, ကာ, ကျာ, ကျား

CU က်, ကီ, ကံ, ကိ, ကဲ

CD ကွ, ကူ, ကု

CLR ကြော, ကြေး, ကော, ကျော, ကျေ, ကျေး, ကြား, ကေး, ကြာ

CLU ကြီ, ကြိ, ကြံ, ကြဲ

CLD ကြွေ, ကွေ့, ကြွ, ကြူ, ကြု, ကွေ

CRD ကျူ, ကူး, ကွာ, ကျွ, ကျု, ကျူး

CRU ကျိ, ကျီ, ကျီး, ကျဲ, ကျံ, ကျ်ား, ကီး

CUD ကဲ့, ကံ့, ကို , ကုံ, ကွဲ, ကွီ, ကွဲ့

CLUR ကြီး, ကော်, ကြော်, ကျော်

CLRD ကော့, ကျွေး, ကျော့, ကြူး, ကြော့, ကွေး, ကြွား, ကြွေး

CLUD ကြွံ, ကြို, ကြို့, ကြုံ, ကြုံး

CRDU ကျိုး, ကိုး, ကုံး, ကျို, ကျို့, ကျုံး, ကျုံ့, ကျွဲ, ကျွံ, ကျံ့, ကျုံ

CLUDR ကြိုး

PS က္က, က္ခ

PSL

PSR က္ကား, က္ခာ, က္ကာ

PSU က္ခိ

PSD က္ကူ, က္ကု, က္ခု

PSLR က္ခော, က္ကော

PSUR

PSUD က္ကို, က္ခို

PSLRU

PSUDR

KS င်္ကာ, င်္ကျိ, င်္က, င်္ကေ, င်္ကြံ, င်္ကြ

C ခ

CL ခြေ, ခေ, ခြ

CR ချ, ခါး, ချား, ချာ, ခါ

CU ခံ, ခဲ, ခိ, ခ်

CD ခု, ခွ, ခူ

CLR ချေ, ချော, ခေါ, ခြား, ခြော, ချေး, ခြာ

CLU ခြံ, ခြိ

CLD ခွေ, ခြု, ခြွ, ခြွေ, ခွေ့

CRD ချွ, ခွား, ချု, ခွာ, ခူး, ချူ, ခွါ

CRU ချိ, ချီး, ချီ, ချဲ

CUD ခို, ခွဲ, ခုံ, ခွံ, ခဲ့, ခံ့, ခွံ့, ခို့

CLUR ချော်, ခေါ်, ခြီး

CLRD ချော့, ခြူး, ခွေး, ချွေ, ချွေး

CLUD ခြုံ, ခြို

CRDU ချုံ, ချွံ, ခိုး, ခုံး, ချိုး, ချို, ချဲ့, ချို့, ချုံ့, ချုံး, ချွံ

CLUDR ခြိုး

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

Appendix

PSUDR

KS င်္ချေ, င်္ချာ, င်္ခ, င်္ခါ, င်္ခေ, င်္ချေ, င်္ချို

C ဂ

CL ဂေ

CR ဂါ, ဂါး, ဂျာ, ဂျ

CU ဂ်, ဂိ, ဂီ, ဂံ, ဂဲ

CD ဂု, ဂူ, ဂွ

CLR ဂေါ, ဂျော, ဂေး

CLU

CLD

CRD ဂျူ, ဂျူး, ဂျု, ဂျွ

CRU ဂျီ, ဂျိ, ဂျီး

CUD ဂုံ, ဂို

CLUR ဂေါ်, ဂျော်

CLRD ဂေါ့, ဂွေး

CLUD ဂြို

CRDU ဂိုး, ဂျို, ဂျုံ, ဂျွံ, ဂုံး

CLUDR

PS ဂ္ဂ, ဂ္ဃ

PSL

PSR

PSU ဂ္ဂီ, ဂ္ဂိ

PSD

PSLR

PSUR

PSUD ဂ္ဂို

PSLRU

PSUDR

KS င်္ဂါ, င်္ဂ, င်္ဂါး, င်္ဂြို, င်္ဂီ, င်္ဂု

Appendix

C ဃ

CL

CR ဃာ

CU ဃိ

CD

CLR ဃော

CLU

CLD

CRD

CRU

CUD

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS င်္ဃာ

C င

CL ငြ

CR ငါ, ငါး

CU င်, ငံ

CD ငှ, ငု, ငူ

CLR ငြာ, ငေး, ငေါ, ငြား, ငြော

CLU ငြိ

CLD ငွေ, ငွေ့, ငြူ

CRD ငါ့, ငွား, ငှာ, ငှား, ငူး

CRU င်း

CUD င့်, ငို, ငုံ့, ငဲ့, ငုံ, ငံ့, ငို့, ငှဲ့

CLUR ငေါ်, ငြီး

CLRD ငေါ့, ငွေး

CLUD ငြို, ငြှိ

CRDU ငိုး, ငုံး

CLUDR ငြိုး

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဓ

CL ဓေ

CR ဓား, ဓာ

Appendix

CU စ်, စိ, စီ, စံ, စဲ

CD စွ, စု, စူ

CLR စော, စေး

CLU

CLD စေ့, စွေ, စွေ့

CRD စွာ, စူး

CRU စီး

CUD စို, စွဲ, စုံ, စို့, စုံ့, စွံ

CLUR စော်

CLRD စော့, စွေး

CLUD

CRDU စိုး, စုံး

CLUDR

PS စ္စ, စ္ဆ

PSL စ္ဆေ, စ္စေ

PSR စ္စာ, စ္ဆာ

PSU

PSD စ္စူ, စ္စု

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဆ

CL

CR ဆာ, ဆား

CU ဆီ, ဆဲ, ဆိ, ဆံ

CD ဆု, ဆွ

Appendix

CLR ဆေး, ဆော

CLU

CLD ဆွေ, ဆွေ့

CRD ဆူး

CRU ဆီး

CUD ဆွဲ, ဆို, ဆူ, ဆုံ, ဆံ့, ဆွံ, ဆို့, ဆုံ့, ဆွံ့

CLUR ဆော်

CLRD ဆွေး, ဆော့

CLUD

CRDU ဆိုး, ဆုံး

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဓ

CL ဓေ

CR ဓာ

CU ဓ်, ဓိ, ဓီ

CD ဓွ, ဓူ, ဓု

CLR ဓော

CLU

Appendix

CLD

CRD ဒူး, ဒွာ

CRU ဒီး

CUD ဒွဲ, ဒို

CLUR ဒေါ်

CLRD

CLUD

CRDU ဒိုး

CLUDR

PS ဒ္ဒ

PSL

PSR ဒ္ဓ, ဒ္ဒာ, ဒ္ဒါ, ဒ္ဓာ

PSU ဒ္ဒိ

PSD ဒ္ဒု

PSLR

PSUR ဒ္ဓိ

PSUD

PSLRU

PSUDR

KS

C ဈ

CL

CR ဈာ

CU

CD

CLR ဈေး

CLU

CLD

CRD

Appendix

CRU

CUD

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C	ဥ
CL	
CR	ဥာ
CU	ဦ
CD	
CLR	
CLU	
CLD	
CRD	ဍာ
CRU	ဦး
CUD	ဦ့

CLUR

CLRD

CLUD

CRDU

CLUDR

PS ဋ္ဌ

PSL

PSR ဋ္ဌာ

PSU

PSD ဋ္ဌူ

PSLR

PSUR

PSUD

PSLRU

PSUDR ဋ္ဌိုး

KS

C ည

CL

CR ညာ, ညား

CU ည်, ညီ, ညိ, ညံ, ညဲ

CD ညွ, ညု, ညှ, ညွှ, ညူ

CLR ညော

CLU

CLD

CRD ညှာ, ညူး

CRU ည်း, ညီး

CUD ညို့, ညှ့်, ညို, ညှိ, ညှို့, ညှံ့, ညှီ, ညှို

CLUR

CLRD ညှော

Appendix

CLUD

CRDU ညှိုး, ညိုး

CLUDR ညှော်

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဋ္ဌ

CL ဋ္ဌေ

CR ဋ္ဌာ

CU င်္ဋ္ဌ, ဋ္ဌီ, ဋ္ဌိ

CD ဋ္ဌု

CLR

CLU

CLD

CRD

CRU

CUD

CLUR

CLRD

CLUD

CRDU

Appendix

CLUDR

PS ဍ

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဒ

CL

CR ဒာ

CU ဒီ, ဒိ, ဒ်

CD

CLR ဒေး, ဒော

CLU

CLD

CRD

CRU

CUD

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဍ

CL

CR ဍ္ဍ

CU ဍိ

CD

CLR

CLU

CLD

CRD

CRU

CUD

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

Appendix

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ပ

CL

CR

CU

CD

CLR

CLU

CLD

CRD

CRU

CUD ပို

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

PSU

PSD

Appendix

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C	ဏ
CL	
CR	ဏာ
CU	ဏ်, ဏီ, ဏိ, ဏ္ဏိ
CD	ဏု, ဏူ
CLR	
CLU	
CLD	
CRD	ဏူာ
CRU	ဏ်း
CUD	ဏို
CLUR	
CLRD	
CLUD	
CRDU	
CLUDR	
PS	ဏ္ဍ, ဏ္ဌ, ဏ္ဋ, ဏ္ဏ
PSL	ဏေ
PSR	ဏ္ဏား, ဏ္ဋာ, ဏ္ဋား, ဏ္ဏာ, ဏ္ဌာ
PSU	ဏ္ဌိ, ဏ္ဋိ
PSD	ဏ္ဋု
PSLR	ဏ္ဏေး
PSUR	

PSUD ဏ္ဍို

PSLRU

PSUDR

KS

C တ

CL တေ, တြ

CR တာ, တျာ, တျား, တျ, တား

CU တ်, တိ, တံ, တီ, တဲ, တိံ

CD တု, တွ, တူ

CLR တော, တေး, တျော, တြာ

CLU တြီ, တြိ

CLD တွေ့, တွေ, တေ့

CRD တူး, တွား, တွာ

CRU တီး

CUD တို, တို့, တွဲ, တဲ့, တုံ့, တုံ, တံ့, တွီ, တွဲ့

CLUR တော်

CLRD တော့, တွေး

CLUD

CRDU တိုး, တုံး

CLUDR

PS တ္တ, တ္ထ

PSL တ္တေ

PSR တ္တား, တ္တာ, တ္ထာ

PSU တ္တီ, တ္တိ, တ္ထိ

PSD တ္တူ, တ္ထု, တ္တု

PSLR

PSUR

PSUD တ္ထုံ, တ္တုံ့, တ္တို

PSLRU

Appendix

PSUDR

KS

C	ထ
CL	ထေ
CR	ထား, ထာ
CU	ထိ, ထံ, ထဲ, ထီ
CD	ထွ, ထူ, ထု
CLR	ထော, ထေး
CLU	
CLD	ထွေ, ထေ့
CRD	ထူး, ထွား, ထွာ
CRU	ထီး
CUD	ထို, ထုံ, ထုံ့, ထို့
CLUR	ထော်
CLRD	ထွေး, ထော့
CLUD	
CRDU	ထိုး, ထုံး
CLUDR	
PS	
PSL	
PSR	
PSU	
PSD	
PSLR	
PSUR	
PSUD	
PSLRU	
PSUDR	
KS	

C	ဒ
CL	ဒေ, ဒြ
CR	ဒါ, ဒျ
CU	ဒ်, ဒိ, ဒီ, ဒံ
CD	ဒူ, ဒု, ဒွ
CLR	ဒေါ, ဒြာ, ဒေး
CLU	
CLD	
CRD	ဒူး, ဒွါ
CRU	ဒီး
CUD	ဒို, ဒွိ, ဒို့, ဒုံ
CLUR	ဒေါ်
CLRD	ဒွေး, ဒေါ့
CLUD	
CRDU	ဒိုး, ဒုံး
CLUDR	
PS	ဒ္ဓ, ဒ္ဒ
PSL	ဒ္ဒေ, ဒ္ဓေ
PSR	ဒ္ဓါ, ဒါး, ဒ္ဒါ
PSU	ဒ္ဓိ, ဒ္ဒိ
PSD	
PSLR	
PSUR	
PSUD	
PSLRU	
PSUDR	
KS	
C	ဓ

Appendix

CL

CR ဓာ, ဓား

CU ဓိ, ဓံ, ဓီ

CD ဓူ, ဓု, ဓွ

CLR

CLU

CLD

CRD

CRU

CUD ဓို

CLUR

CLRD

CLUD

CRDU

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဋ

CL ဋေ

Appendix

CR	နား, နာ, နျာ
CU	န်, နီ, နံ, နိ, နဲ
CD	နု, နွ, နူ, နှ, နွှ, နု
CLR	နော
CLU	
CLD	နေ့, နွေ
CRD	နူး, နှာ, နွား, နာ့, နျူ, နှူး, နွှာ
CRU	န်း, နီး
CUD	န့်, နွဲ့, နို့, နှံ, နှဲ, နှိ, နွံ, နို, နိူ, နံ့, နု်, နဲ့, နုံ, နွှဲ, နှံ့, နုံ့, န့်, နှဲ့, နွဲ, နှီ
CLUR	နော်
CLRD	နှော, နွေး, နှေး, နွှေး, နော့
CLUD	
CRDU	နိုး, နှီး, နိူး, နုံး
CLUDR	
PS	န္န, န္ထ, န္တ, န္ဓ, န္ဒ
PSL	န္ဒြေ, န္ဒြ, န္တေ, န္ဓေ
PSR	န္တာ, န္ဓာ, န္ဒာ, န္ဒား
PSU	န္တိ, န္တီ, န္နိ, န္နဲ, န္ဓိ, န္ဒီ, န္တံ
PSD	န္ဒု, န္တု, န္ဒူ
PSLR	န္ဓော, န္တေး, န္ဓေး, န္တော
PSUR	
PSUD	န္ဒို, န္ဓို
PSLRU	
PSUDR	
KS	

C	ပ
CL	ပြ, ပေ, ပြေ
CR	ပျ, ပါး, ပျာ, ပါ, ပြာ, ပျား
CU	ပ်, ပီ, ပိ, ပဲ, ပံ

CD ပွ, ပု, ပူ

CLR ပေါ, ပေး, ပြော, ပြား, ပျော, ပြေး, ပေါး

CLU ပြဲ, ပြီ, ပြိ

CLD ပြု, ပွေ့, ပြွ, ပြွေ့, ပြူ, ပေ့, ပွေ, ပြွေ

CRD ပွား, ပျူ, ပူး, ပါ့, ပျူး

CRU ပျံ, ပီး

CUD ပွဲ, ပုံ, ပို, ပို့, ပဲ့, ပံ့, ပုံ့, ပွဲ့

CLUR ပျော်, ပြီး

CLRD ပြူး, ပေါ့, ပွေး, ပျော့, ပြော့

CLUD ပြို, ပြုံ, ပြွဲ

CRDU ပုံး, ပျိုး, ပျို, ပျုံ့, ပျို့

CLUDR ပြိုး, ပြုံး

PS ပ္ပ

PSL

PSR ပ္ပါ, ပ္ပာ

PSU ပ္ပိ, ပ္ပီ, ပ္ပံ

PSD ပ္ပု

PSLR

PSUR

PSUD ပ္ဖို, ပ္ပို

PSLRU

PSUDR

KS

C ဖ

CL ဖြ, ဖြေ, ဖေ

CR ဖျ, ဖား, ဖျား, ဖာ, ဖျာ

CU ဖီ, ဖိ, ဖဲ

CD ဖွ, ဖု, ဖူ

CLR ဖော, ဖျော, ဖြော, ဖြာ, ဖေး, ဖြား

Appendix

CLU ဖြဲ

CLD ဖြူ, ဖြု, ဖွေ

CRD ဖူး, ဖွား, ဖွာ, ဖျူး

CRU ဖီး, ဖျံ

CUD ဖဲ့, ဖို, ဖွဲ့, ဖုံ, ဖို့, ဖွဲ, ဖွို, ဖွီ, ဖွံ့

CLUR ဖော်, ဖြီး, ဖျော်, ဖြော်

CLRD ဖော့, ဖြူး, ဖျော့, ဖွေး

CLUD ဖြို့, ဖြို, ဖြုံ

CRDU ဖိုး, ဖုံး, ဖျိုး

CLUDR ဖြုံး, ဖြိုး

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ဗ

CL ဗေ, ဗြ

CR ဗျာ, ဗျား, ဗျ, ဗာ, ဗား

CU ဗိ, ဗီ

CD ဗွ, ဗု

CLR ဗြာ, ဗြား, ဗြော

CLU ပြဲ, ပြိ

CLD ပြု, ပွေ

CRD ပျူ, ပူး, ပျာ့, ပျု

CRU ပီး

CUD ပုံ, ပို, ပို့, ပွိ

CLUR ပော်

CLRD ပော့, ပွေး, ပြူး

CLUD ပြို့, ပြို

CRDU ပွိး, ပိုး, ပုံး, ပျို့, ပျိုး

CLUDR ပြုံး

PS ပ္ပ, ပ္ဘ

PSL ပ္ပေ

PSR ပ္ပာ

PSU ပ္ဘိ, ပ္ပိ

PSD ပ္ဘု

PSLR ပော

PSUR ပ္ဘိး, ပ္ဘာ

PSUD

PSLRU

PSUDR ပျို

KS

C ဘ

CL ဘေ

CR ဘာ, ဘား, ဘျ

CU ဘီ, ဘဲ, ဘ်, ဘိ

CD ဘု, ဘူ, ဘွ

CLR ဘော, ဘေး

CLU

CLD

CRD ဘူး, ဘွား, ဘွာ

CRU ဘီး

CUD ဘို, ဘုံ, ဘွဲ့, ဘို့, ဘဲ့, ဘွို

CLUR ဘော်

CLRD ဘော့

CLUD

CRDU ဘိုး

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS ဘော်

C မ

CL မြ, မြေ, မေ

CR မား, မာ, မျ, များ, မျာ, မျှား

CU မ်, မိ, မီ, မံ, မဲ

CD မု, မူ, မှ, မှူ, မွ, မွှ

CLR မျော, မော, မြော, မြား, မေး, မြေး, မြာ

CLU မြိ, မြီ, မြဲ, မြံ

CLD မွေ, မြွေ, မြူ, မြှ, မြွှေ, မွေ့, မြှူ, မြှု, မွှေ, မြု, မေ့, မြွ, မွေ့, မြွှ

CRD မူး, မှာ, မှူး, မျှ, မှား, မွှား

CRU　မ်း, မီး

CUD　မဲ့, မ့်, မှီ, မွဲ, မို, မုံ, မှိ, မှို, မှဲ့, မုံ့, မို့, မှို့

CLUR　မြီး, မြော်, မော်

CLRD　မြှေး, မြှော့, မှော, မြှော, မြွှာ, မော့, မွေး, မြူး, မွှေး, မြွှာ, မှေး, မျှော့, မျှော

CLUD　မြို့, မြုံ့, မြို, မြုံ, မြှို

CRDU　မိုး, မျိုး, မှိး, မျို, မျှို့, မုံး

CLUDR　မျှော်, မှော်, မြှုံး, မြိုး

PS　မ္ပ, မ္ဗ, မ္မ, မ္ဘ, မ္ဖ

PSL

PSR　မ္မာ, မ္ဘာ, မ္ပာ

PSU　မ္မိ, မ္ဘိ

PSD　မ္ပု, မ္ဗူ, မ္မု, မ္ဘူ

PSLR

PSUR

PSUD

PSLRU　မ္မော်

PSUDR

KS

C　ယ

CL　ယေ

CR　ယာ, ယား, ယျာ, ယျ

CU　ယိ, ယီ, ယ်, ယံ

CD　ယု, ယူ, ယွ, ယှ

CLR　ယော, ယျော

CLU

CLD　ယှေ

CRD　ယူး

CRU　ယီး

CUD　ယ့်, ယုံ, ယို, ယဲ့, ယို့, ယွဲ့, ယှိ

Appendix

CLUR ယော်

CLRD

CLUD

CRDU ယိုး

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C ရ

CL ရေ

CR ရာ, ရား

CU ရံ, ရဲ, ရိ, ရီ, ရ်

CD ရွ, ရု, ရူ, ရှ, ရှု, ရွှ, ရှူ

CLR ရော, ရေး

CLU

CLD ရွေ, ရှေ့, ရွေ့, ရွှေ, ရှေ, ရွေ, ရေ့, ရွှေ့

CRD ရွား, ရူး, ရှာ, ရှူး, ရွာ, ရှား, ရာ့, ရွှာ

CRU ရီး, ရ်း

CUD ရုံ, ရို, ရို့, ရှို, ရဲ့, ရှို, ရွဲ, ရွံ့, ရံ့, ရွှဲ, ရှို့, ရှုံ့, ရွဲ့, ရွံ, ရွှို, ရှီ, ရှဲ, ရှုံ, ရွှံ့

CLUR ရော်

Appendix

CLRD ရော့, ရွေး, ရှော, ရှေး, ရှော့

CLUD

CRDU ရိုး, ရုံး, ရှုံး, ရှိုး

CLUDR ရှော်

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C လ

CL လေ

CR လာ, လျား, လျ, လား, လျာ

CU လိ, လီ, လ်, လံ, လဲ

CD လု, လွ, လွ, လှ, လှူ, လူ, လူ့

CLR လေး, လော, လျော

CLU

CLD လွေ, လေ့, လှေ, လှေ့, □, လွေ့

CRD လူး, လွာ, လှူ, လျှာ, လွား, လျှ, လျှူ, လှာ

CRU

CUD လဲ့, လို, လို့, လှို, လုံ, လွဲ, လှဲ, လှီ, လွဲ, လှုံ, လွံ, လှိ, လှံ, လုံ့, လှုံ့, လွံ့

CLUR လော်, လျော်

CLRD လျှော, လွေး, လှေး, လျှော့, လှော, လျော့, လော့

CLUD

Appendix

CRDU လုံး, လိုး, လီူး, လျှိုး, လျှုံ, လျှို, လျှို့

CLUDR လုော်, လျှော်

PS လွ

PSL

PSR လွာ

PSU လွိ

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS င်္လာ

C ဝ

CL ဝေ

CR ဝါး, ဝါ

CU ဝိ, ဝီ, ဝံ, ဝဲ, ဝ်

CD ဝူ, ဝု

CLR ဝေး, ဝေါ

CLU

CLD ဝေူ့, ဝေ့

CRD ဝှူး, ဝူာ, ဝူါ

CRU

CUD ဝို, ဝံ့, ဝဲူ

CLUR ဝေါ်

CLRD ဝေူး

CLUD

CRDU ဝိုး, ဝီူး

CLUDR

PS

PSL

PSR

PSU

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C	သ
CL	သေ
CR	သား, သာ
CU	သီ, သံ, သိ, သဲ, သ်
CD	သူ, သု, သွ, သူ့
CLR	သော, သေး
CLU	
CLD	သေ့, သွေ့, သွေ
CRD	သွား, သျှ, သျှူး, သျှု, သွာ, သျှား
CRU	သီး
CUD	သို, သို့, သုံ့, သွဲ့, သုံ, သဲ့, သံ့, သွံ
CLUR	သော်
CLRD	သွေး, သျှော, သော့
CLUD	
CRDU	သုံး, သိုး, သျှိ, သျှီ, သျှို, သျှိုး
CLUDR	

Appendix

PS	
PSL	
PSR	
PSU	
PSD	
PSLR	
PSUR	
PSUD	
PSLRU	
PSUDR	
KS	င်္သေ့, င်္သော, င်္သ, င်္သာ
C	ဟ
CL	ဟေ
CR	ဟာ, ဟား
CU	ဟ်, ဟိ, ဟီ, ဟဲ, ဟံ
CD	ဟု, ဟူ, ဟွ
CLR	ဟော, ဟေး
CLU	
CLD	ဟေ့
CRD	ဟူး
CRU	ဟီး
CUD	ဟို, ဟုံ, ဟဲ့, ဟို့
CLUR	ဟော်
CLRD	ဟော့
CLUD	
CRDU	ဟုံး, ဟိုး
CLUDR	
PS	ဟ္မ
PSL	

Appendix

PSR	ပ္မာ
PSU	ပ္မီ
PSD	
PSLR	
PSUR	
PSUD	
PSLRU	
PSUDR	
KS	

C	ဋ
CL	
CR	ဋာ
CU	ဋိ, ဋ်, ဋီ
CD	ဋု
CLR	
CLU	
CLD	
CRD	
CRU	
CUD	
CLUR	
CLRD	
CLUD	
CRDU	
CLUDR	
PS	
PSL	
PSR	
PSU	

Appendix

PSD

PSLR

PSUR

PSUD

PSLRU

PSUDR

KS

C	အ
CL	အေ
CR	အား, အာ
CU	အိ, အံ, အီ, အဲ
CD	အု, အူ
CLR	အော, အေး
CLU	
CLD	အေ့
CRD	
CRU	အီး
CUD	အို, အုံ, အို့, အံ့, အဲ့, အုံ့
CLUR	အော်
CLRD	အော့
CLUD	
CRDU	အုံး, အိုး
CLUDR	
PS	
PSL	
PSR	
PSU	
PSD	
PSLR	

PSUR

PSUD

PSLRU

PSUDR

KS

C	ဿ
CL	
CR	ဿာ
CU	ဿီ, ဿံ
CD	
CLR	
CLU	
CLD	
CRD	
CRU	
CUD	ဿုံ
CLUR	
CLRD	
CLUD	
CRDU	
CLUDR	
PS	
PSL	
PSR	
PSU	
PSD	
PSLR	
PSUR	
PSUD	

PSLRU

PSUDR

KS

APPENDIX K:

Some Example of Khmer Subscript Consonants

ក

ក្ក   កក្កដ, កក្កដា, តុក្កតា, រតក្កសិលា, នេបក្កបណ្ឌិត

ក្ខ   ចក្ខុ, បេក្ខជន, បេក្ខនារី, អារក្ខ, អារក្ខា, រុក្ខជាតិ, គរុក្ខ,

លក្ខការណ៍, លក្ខខណ្ឌ, អនក្ខរកម្ម, យក្ខ, បដិក្ខ,   នបក្ខ, វិរូបក្ខ, លក្ខណា, អក្ខរា

ក្យ   ពាក្យ, វាក្យខ័ណ្ឌ, បាក្យសំដី, រាក្យ, រាក្យសា

ខ

គ

គ្គ   វគ្គ, មគ្គុទេសន៍, បដិវគ្គ, វច្ឆមគ្គ, មហាវគ្គ, សមគ្គ, មគ្គផល, អគ្គិសនី

គ្ឃ

គ្យ អគ្យន្តរាយ, អគ្យាគារ, អគ្យសនី

ឃ

ង

ង្ក កង្កាញ់, កង្កួច, កង្កែប, បង្ក, បង្កើន, បង្កើត, បង្កា, រូបមង្គ, សរីរាង្ក, សង្កសី

ង្ខ កង្ខា, កង្ខើញ, សង្ខារ

ង្គ អង្គ, អង្គធាតុ, មេគង្គ, គង្គា, ព្យាង្គ, ឧត្តុង្គ, ប្រាង្គប្រាសាទ, មង្គល, រូបរាង្គ, សិរីរាង្គ

ង្ឃ សង្ឃ, ភិក្ខុសង្ឃ, សង្ឃទាន

ង្រ កង្ក្រង, ពង្រ, ពង្រៀង, ចង្រិត, ចង្រៃ

ង្វ កង្វល់, កង្វះ, ពង្វៀង, វង្វេង, រង្វាន់, ចង្វាក់

ង្ស វង្ស, គង់វង្ស, រង្ស, រាជវង្ស

ង្ហ កង្ហារ, សិង្ហ, សង្ហា, សិង្ហបុរី, បរង្ហ, ចង្ហូរ, បង្ហៀរ, បង្ហិន, បង្ហាញ

ង្អ ចង្អេរ, ចង្អូរ

ច

ច្ច កិច្ច, កិច្ចការ, មច្ចវប្បកម្ម, សច្ចៈ, សច្ចធម៌, អរិយសច្ច, អាទិច្ចគាម, អាទិច្ចវង្ស, អាទិច្ចវណ្ណ,
មច្ចុរាជ, បច្ចេកទេស,

ច្ឆ មច្ឆវប្បកម្ម, មច្ឆា, ស្វេតច្ឆត្រ, តិរិច្ឆាន

ច្យ

ឆ

ជ

ជ្ជ រជ្ជ, រជ្ជកាល, ភេសជ្ជៈ, វេជ្ជបណ្ឌិត, វិជ្ជមាន, អនជ្ជកម្ម, កម្មជ្ជវាត

ជ្ឈ មជ្ឈមណ្ឌល, មជ្ឈដ្ឋាន

ជ្ញ អ្នកប្រាជ្ញ, ប្រាជ្ញា

ជ្យ រាជ្យ, រាជ្យសម្បត្តិ, សោយរាជ្យ

ឈ

ញ

ញ្ច　　កញ្ចាស់, កញ្ចប់　　កញ្ចែ, បញ្ចប់, បញ្ចវគ្គី, បញ្ចុក

ញ្ឆ　　កញ្ឆក់, កញ្ឆែត, កញ្ឆា

ញ្ជ　　កតញ្ជលី, កញ្ជរ, កញ្ជើ

ញ្ឈ　　បញ្ឈរ, បញ្ឈប់

ញ្ញ　　កញ្ញា, បញ្ញា, ធម្មនុញ្ញ, កតញ្ញូ, សាមញ្ញ, ព្រហ្មញ្ញ, អរញ្ញ, សញ្ញា, ហិរញ្ញវត្ថុ,

បូជាយញ្ញ, មនុញ្ញផល

ញ្ជ្រ　　កញ្ជ្រែង, កញ្ជ្រោក, កញ្ជ្រៀវ, កញ្ជ្រោល

ញ្ហ　　បញ្ហា

ដ

ដ្ត　　បថ្យាវដ្ត, សេដ្តកិច្ច

ដ្ឋ　　សេដ្ឋី, សេដ្ឋកិច្ច, រដ្ឋ, បសេដ្ឋ, រដ្ឋមន្ត្រី, រដ្ឋការ, រដ្ឋសភា, រដ្ឋធម្មនុញ្ញ,

កោដ្ឋ, រដ្ឋបាល, ធម្មបទដ្ឋ, អូដ្ឋអរិដ្ឋ, ពិសិដ្ឋ, ម៉ដ្ឋ, ហ្មដ្ឋ

ដ្ដ　　វដ្ដ, វដ្ដរដូវ, វដ្ដ

ឋ

ឌ

ឌ្ឍ	វឌ្ឍ, យសវឌ្ឍិនី, អភិវឌ្ឍ, អភិវឌ្ឍន៍, ឋានបឌ្ឍ

ឍ

ណ

ណ្ដ	សណ្ដ, និលកណ្ដ, បណ្ដាល, កណ្ដាល, កណ្ដុរ

ណ្ឋ	សណ្ឋាន, ទសកណ្ឋ

ណ្ឌ	មជ្ឈមណ្ឌល, ឥណ្ឌា, ឥណ្ឌូចិន, បិណ្ឌ, លក្ខខណ្ឌ, មណ្ឌលគិរី, ខណ្ឌ

ណ្ណ	វណ្ណ, សុវណ្ណ, វណ្ណកម្ម, វណ្ណៈ, បណ្ណាគារ

ណ្យ	បុណ្យ

ត

ត្ត	ព្រះបរិត្ត, វត្ត, គោត្តនាម, ភោគោត្តសាស្ត្រ, អវត្តមាន, ឯកត្ត, អនុវត្ត,

ប្រព្រឹត្ត, ចិត្ត, យោត្តមុត្តី, នរោត្តម, សត្តឃាដ, ខេត្ត, មេត្តា, ទស្សវត្ត, កត្តា

ត្ថ សមត្ថភាព, សមត្ថកិច្ច, វត្ថុ, សេត្ថវិថី, ហត្ថកម្ម, ហត្ថលេខា, បរមត្ថ, ធម្មត្ថចរិយា, ឧបត្ថម

ត្យ អាទិត្យ, អាមាត្យ

ត្វ សត្វ

ត្ស សតវត្ស, ទសវត្ស

ថ

ទ

ទ្ទ ឧទ្ទេសនាម

ទ្ធ លទ្ធភាព, លទ្ធផល, ឥទ្ធិពល, កម្មសិទ្ធ, យុទ្ធ, យុទ្ធជន, ស្មិទ្ធស្នាល, មហិរិទ្ធ

ទ្យ វិទ្យា, ជាអាទ្យ, បាទ្យ

ធ

ន

ន្ត	និមន្ត, រថយន្ត, ភាពយន្ត, ក្សាមក្សាន្ត, កំសាន្ត, ទិស្សឥសាន្ត, រូបមន្ត

ចរន្ត, សូត្រមន្ត, ស្រាន្ត, សម្រាន្ត, សូប្រកាន្ត, នព្វន្ត, យន្តហោះ, បន្ត,

បន្តិច, បន្តួច, កន្តឹប, អន្តរជាតិ

ន្ថ	បន្ថយ, បន្ថែម

ន្ទ	ចន្ទ, ព្រះច័ន្ទ, ភូមិន្ទ, សន្ទនា, កន្ទេល, កន្ទុំរុយ, បន្ទរ, អានន្ទ, សុរិន្ទ, ផលានន្ទា

ន្ធ	រន្ធ, គន្ធពិដោរ, កន្ធែក, ពន្ធ, ប្រពន្ធ, ប្រព័ន្ធ

ន្ន	អាសន្ន, អាសន្នរោគ, សន្និបាត

ន្ម	ដំបូន្មាន

ន្យ	សូន្យ, សន្យា, សាមាន្យ

ន្ត្រ	កន្ត្រៃ, កន្ត្រោង, កន្ត្រើយ, កន្ត្រាក់

ន្ល	ទន្លេ, ពន្លក, បន្លិច, បន្លាច, កន្លះ, កន្លង់, កន្លង

ន្ស	កន្សែង, កន្សោម, អន្សម, កន្សៃ, បន្សាប

ន្ឋ	ន្ឋាលនឹង

ប

ប្ប	សប្បាយ, កប្បាស, វប្បកម្ម, វប្បធ៌ម, អកប្បកិរិយា, រដ្ឋប្បវេណី,

អាយុកប្ប, សិប្បនិម្មិត, សិប្បកម្ម, គុណប្បសំសិនី, អនេកប្បការ

ប្ស អប្សរ, អប្សរា

ផ

ព

ព្ទ ស្វរសព្ទ, ទូរស័ព្ទ, សព្ទសាធុការ, កិរិយាសព្ទ

ព្វ សព្វ, កាតព្វកិច្ច, ប៉ិនប្រសព្វ, សព្វសាធុការ, ស្រោចស្រព្វ,

សំណព្វ, បព្វជិត

ភ

ម ម្ប, កាយសម្បទា, សម្បទាន, សម្បូរ

ម្ផ សម្ផស្ស

ម្ព គម្ពីរ, គុម្ព

ម្ភ បារម្ភ, ស្តុកស្តម្ភ, អារម្ភកថា, ចតុស្តម្ភ, ម្ភៃ, សម្ភព, កម្ភៈ

ម្ម សកម្ម, សកម្មជន, សកម្មភាព, កម្មករ, កម្មាភិបាល, ពលកម្ម,

ហាតកម្ម, កសិកម្ម, ពលកម្ម, អាជីវកម្ម, សេវាកម្ម, កម្មវិធី, ស្ថាបត្យកម្ម, ឧស្សាហកម្ម

ម្យ សមរម្យ, ប្រណម្យ, រម្យទម

យ

យ្យ, ទេយ្យទាន, កោសេយ្យសិរី, ឧបមេយ្យ, ទេយ្យធម៌, អាយ្យការ

រ

រ្ត រាមកេរ្ត៍, កេរ្តិ៍ឈ្មោះ

រ្ក

រ្ម វរ្ម័ន

រ្យ អារ្យធម៌, អាចារ្យ, សិអារ្យមេត្រី, រាមសូរ្យ, ពិទូរ្យសូរ្យកាន្ត, បាដិហារ្យ

រ្ស រណសិរ្ស

ល

ល្យ កល្យាណ, កល្ប

ល្ល អាល្លជី

វ

វ្រ

វ្ហ វ្ហី, វ្ហើយ, វ្ហោះវ្ហើយ

ស

ស្ស មនុស្ស, សិស្ស, រហស្សកម្ម, បុស្ស

ហ

ហ្គ ហ្គាណា

ហ្ន ហ្នឹង, ហ្ន

ហ្ម ហ្ម, ហ្មង, ហ្មត់ចត់, ហ្មឺន

ហ្ល ហ្ល័កហ្ល័ង, ហ្លាក់ខាំ, ហ្លួង

ហ្វ ហ្វឹកហ្វឺន, ហ្វូស៊ីល, កាហ្វេ

ហ្ស ហ្សឺណែវ

ឡ

អ

អូ អឿង

BIBLIOGRAPHY

[1] Twitter (Wikipedia)
http://en.wikipedia.org/wiki/Twitter
(Accessed date: October 16, 2011)

[2] Twitter
http://twitter.com/
(Accessed date: October 16, 2011)

[3] Facebook (Wikipedia)
http://en.wikipedia.org/wiki/Facebook
(Accessed date: October 16, 2011)

[4] Facebook
http://www.facebook.com/
(Accessed date: October 16, 2011)

[5] Mon Language
http://en.wikipedia.org/wiki/Mon_language
(Accessed date: March 20, 2011)

[6] Unicode Code Chart of Myanmar
http://www.unicode.org/charts/PDF/U1000.pdf
(Accessed date: March 20, 2011)

[7] Unicode Code Chart of Khmer
http://www.unicode.org/charts/PDF/U1780.pdf
(Accessed date: June 23, 2011)

[8] Nepali Language
http://en.wikipedia.org/wiki/Nepali_language
(Accessed date: December 20, 2008)

[9] Unicode Code Chart of Devanagari
http://www.unicode.org/charts/PDF/U0900.pdf
(Accessed date: December 20, 2008)

[10] Pan Localization Working Papers 2004 to 2007

http://www.panl10n.net/english/TOC.htm#Nepali

(Accessed date: January 10, 2009)

[11] Nepali Alphabets and Pronunciation

http://www.omniglot.com/writing/nepali.htm

(Accessed date: October 16, 2011)

[12] Banu Oja and Shambhu Oja, Nepali: A Beginner's Primer Conversation and Grammar, Department of Asian Studies, South Asia Program, Cornell University, 3rd edition 2004

[13] Abugida

http://en.wikipedia.org/wiki/Abugida

(Accessed date: October 15, 2011)

[14] IEEE Annals of the History of Computing

January-March 2009 (vol. 31 no. 1)

ISSN: 1058-6180

[15] Baraha (Free Indian Language Software)

http://www.baraha.com

(Accessed date: October 15, 2011)

[16] AAA+

http://www.aa.tufs.ac.jp/~kmach/gicas/ASTI/AAA/info/info_e.htm

(Accessed date: October 15, 2011)

[17] Khmer language

http://en.wikipedia.org/wiki/Khmer_language

(Accessed date: October 16, 2011)

[18] Huffman, F. E. 1970. Cambodian System of Writing and beginning reader with Drills and Glossary. Yale University Press.

[19] Displaying Khmer Script

http://www.khmeros.info/download/DisplayingKhmerScript.pdf

(Accessed date: October 16, 2011)

[20] KhmerOS, 2008

http://www.khmeros.info/drupal/
(Accessed date: October 15, 2011)

[21] Thai Language
http://en.wikipedia.org/wiki/Thai_language
(Accessed date: August 21, 2009)

[22] Thai Writing System
http://thaiarc.tu.ac.th/thai/thaiwrt.htm
(Accessed date: October 15, 2011)

[23] Richard Gillam, "Unicode Demystified", September 2002, Addison Wesley, Page 331-332

[24] Thai Technical Characteristics
http://www01.ibm.com/software/globalization/topics/thai/tech.jsp
(Accessed date: October 15, 2011)

[25] Unicode
http://en.wikipedia.org/wiki/Unicode
(Accessed date: October 15, 2011)

[26] Input System and Encoding Myanmar Language (Presented by Mr. Zaw Htut)
http://www.myanmarnlp.net.mm/doc/index.htm,
(Accessed date: October 16, 2011)

[27] WinMyanmar Keyboard
http://www.myanmars.net/winmyanmar/
(Accessed date: October 16, 2011)

[28] Zawgyi Myanmar Keyboard
http://www.zawgyi.net/
(Accessed date: October 15, 2011)

[29] Myanmar3 Keyboard (Myanmar Unicode and NLP Research Center)
http://www.myanmarnlp.net.mm/
(Accessed date: October 15, 2011)

[30] Nepali Language
http://en.wikipedia.org/wiki/Nepali_language

(Accessed date: December 20, 2008)

[31] Pan Localization Working Papers 2004 to 2007
http://www.panl10n.net/english/TOC.htm#Nepali
(Accessed date: January 10, 2009)

[32] Nepali Alphabets and Pronunciation
http://www.omniglot.com/writing/nepali.htm
(Accessed date: October 16, 2011)

[33] Banu Oja and Shambhu Oja, Nepali: A Beginner's Primer Conversation and Grammar, Department of Asian Studies, South Asia Program, Cornell University, 3rd edition 2004

[34] Nepali Unicode Keyboard Layout Standardization based on Genetic Algorithm
http://dl.dropbox.com/u/547169/jwalanta.com.np/nepalikeyboardlayout_report_draft.pdf

[35] Download Page of Nepali Unicode Keyboard Layout
http://www.madanpuraskar.org/downloads.php
(Accessed date: October 15, 2011)

[36] Market share of mobile handset in Thailand
http://www.zdnetasia.com/news/communications/0,39044192,62050412,00.htm
http://www.zdnetasia.com/news/communications/0,39044192,62050010,00.htm
(Accessed date: October 15, 2011)

[37] I. Scott MacKenzie, 2002, KSPC (Keystrokes per Character) as a Characteristic of Text Entry Techniques, Proceedings of the 4th International Symposium on Mobile Human-Computer Interaction with Mobile Devices, Heidelberg, Germany, Springer-Verlag, pp. 195 - 210

[38] Ye Kyaw Thu, "Positional Prediction: Consonant Cluster Prediction Text Entry Method for Burmese (Myanmar Language)", Extended Abstracts of the 26th ACM Conference on Human Factors in Computing Systems (CHI 2008), Florence, Italy, April 5-10, 2008, pp. 3783-3788

[39] Ye Kyaw Thu, Ouk Phavy, Yoshiyori URANO and Mitsuji MATSUMOTO, "Positional Prediction for Khmer Language (Cluster Predictive Text Entry Method in Mobile Devices)", IIEEJ 4th Mobile Image Research Meeting, March

4, 2008, International Conference Center, Waseda University, Tokyo, Japan, Page 5-10

[40] MySM

http://www.mysmonline.com/

(Accessed date: October 15, 2011)

[41] Landolt C or Landolt Broken Ring

http://en.wikipedia.org/wiki/Landolt_C

(Accessed date: October 15, 2011)

[42] Home page of Forum.Nokia.Com

http://www.forum.nokia.com/

(Accessed date: October 15, 2011)

[43] MacKenzie, I. S. and Kumiko Tanaka-Ishii, Text Entry Systems (Mobility, Accessibility, Universality), Morgan Kaufmann Press (2007).

[44] Yamada, A historical study of typewriters and typing methods: From the position of planning Japanese parallels, Journal of Information Processing, 2, 175-202

[45] Department of Myanmar Language Commission, Myanmar-English Dictionary, Department of the Myanmar Language Commission, pp.xi (1993).

[46] Soukoreff, R. W. and MacKenzie, I. S., "Measuring errors in text entry tasks: An application of the Levenshtein string distance statistic." Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI '01). Seattle, Washington (March 31-April 5, 2001). New York: ACM Press, pp. 319-320.

[47] Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology, 140, 1–55.

[48] A. Pavlovych, W. Stuerzlinger (2003). Less-Tap: A Fast and Easy-to-learn Text Input Technique for Phones, Graphics Interface 2003, AK Peters, ISBN 156881207-8, 97-104.

[49] Saied B. Nesbat, Ph.D., A System for Fast, Full-Text Entry for Small Electronic Devices. Proceeding of the Fifth International Conference on Multimodal Interfaces, ICMI 2003, Vancouver, (Nov 5-7, 2003), 4-11

[50] Rick W. Ha, Pin-Han Ho, and Xuemin (Sherman) Shen, SIMKEYS: An efficient keypad configuration for mobile communications. IEEE Communications Magazine, November 2004

[51] FasTap Keypad Layout

http://www.digitwireless.com

(Accessed date: October 15, 2011)

[52] DeltaII Keypad Layout

http://www.chicagologic.com

(Accessed date: October 15, 2011)

[53] Ye Kyaw Thu and Yoshiyori URANO, “Positional Mapping Multi-tap for Myanmar Language”, the 12th International Conference on Human–Computer Interaction, July 22~27, 2007, Beijing International Convention Center, Beijing, P.R.China, (Proceedings Volume 2, LNCS_4551, ISBN: 978-3-540-73106-1) Page 486–495

[54] Acharya Multilingual Editor for Indian Languages

http://imli.sourceforge.net/index.html

(Accessed date: October 15, 2011)

[55] Download link of NiDA Khmer Software Keyboard

http://www.tavultesoft.com/keyman/downloads/keyboards/details.php?KeyboardID=401&FromKeyman=0

[56] Information of SMART Electronic Whiteboard

http://www.smartboard.co.jp/

(Accessed date: October 15, 2011)

[57] Ye Kyaw Thu, Ouk Phavy and Yoshiyori URANO, “Positional Gesture: Simple Gesture Text Input for Khmer”, the 127th Human Computer Interaction Symposium 2008, January 31 to February 1, 2008, Faculty Club (Gakushi Kaikan), Hiroshima University, Higashi Hiroshima, Japan, SIGHCI, Information Processing Society of Japan (IPSJ) SIG Technical Reports, 2008 – HCI – 127, Page 115–122

[58] Goldberg, D., Richardson, C., “Touch Typing With a Stylus”, Proc. of the ACM Conference INTERCHI '93, Amsterdam, The Netherlands, 1993, pp. 80-87

[59] MacKenzie, I.S., Soukoreff, R.W., “Text entry for mobile computing: Models and Methods, Theory and Practice”, Human-Computer Interaction 17, 2002, pp. 147-198

[60] Graffiti (Palm OS)

http://en.wikipedia.org/wiki/Graffiti_%28Palm_OS%29

(Accessed date: October 15, 2011)

[61] Butter, Andrea; Pogue, David (2002). Piloting Palm. New York: John Wiley & Sons. pp. 62–66. ISBN 0471089656

[62] Jacob O. Wobbrock, Brad A. Myers, John A. Kembel, “EdgeWrite: A Stylus-Based Text Entry Method Designed for High Accuracy and Stability of Motion”, Proc. of the ACM Symposium on User Interface Software and Technology (UIST '03), Vancouver, British Columbia, 2003, pp. 61-70

[63] R. Balaji, V. Deepu, Sriganesh Madhvanath and Jayasree Prabhakaran, “Handwritten Gesture Recognition for Gesture Keyboard”, Proc. of the Tenth International Workshop on Frontiers in Handwriting Recognition (IWFHR-10), La Baule, France, 2006.

[64] HP Labs India

http://www.hpl.hp.com/india/research/aad-gkb.html

(Accessed date: October 15, 2011)

[65] Win Myanmar Visual Keyboard

http://winmyanmar.myanmars.net/other_features.htm

(Accessed date: February 1 2008)

[66] Chinese Input Methods for Computer

http://en.wikipedia.org/wiki/Chinese_input_methods_for_computers

(Accessed date: October 15, 2011)

[67] Masui, T. POBox: An efficient text input method for handheld and ubiquitous computer. Proc. HUC ’99, 289-300

[68] MacKenzie, I. S., Kober, H., Smith, D., Jones, T., & Skepner, E. LetterWise: Prefix-based disambiguation for mobile text input. Proc. UIST 2001. ACM Press (2001), 111-120.

[69] Ye Kyaw Thu and Yoshiyori URANO, “Positional Mapping Software Keyboard for Syllabic Scripts (Study with Bangla and Khmer Languages)”, Proc. ICCA2008, February 14~15, 2007, Yangon, Myanmar, Page 150-157

[70] MacKenzie, I. S., Chen, J., & Oniszczak, A. (2006). Unipad: Single-stroke text entry with language-based acceleration. Proceedings of the Fourth Nordic Conference on Human-Computer Interaction - NordiCHI 2006, pp. 78-85. New York: ACM.

[71] John Okell, 1971, A Guide To The Romanization Of Burmese, The Royal Asiatic Society of Great Britain and Ireland.

[72] Ye Kyaw Thu and Yoshiyori URANO, 2006, Development of Romanized Input Methods for Myanmar Language Short Messaging Service (SMS): Comparison with Multitap Input Method, Simulation and Analysis, Technical report of the 117th Human Interface Symposium 2006, Tsukuba, Japan, pp. 33-40.

[73] Burglish Myanmar Text Input Engine (Ver. 1.8.17.080919).

http://burglish.googlepages.com/testarea.htm (Accessed date: 12 December 2008)

(Accessed date: October 15, 2011)

[74] Textbook of Speaking and Myanmar Sign Communication, Mary Chapman School for the Deaf, 1988

[75] American Manual Alphabet

http://en.wikipedia.org/wiki/American_manual_alphabet

(Accessed date: October 15, 2011)

[76] Myanmar Sign Language Dictionary (Vol:1), Mandalay School for the Deaf, 2007

[77] British Deaf Association (BDA)

http://www.bda.org.uk/

(Accessed date: October 15, 2011)

[78] American Sign Language (ASL)

http://wiki.medpedia.com/American_Sign_Language

(Accessed date: October 15, 2011)

[79] Article of SignWriting in Wikipedia

http://en.wikipedia.org/wiki/Sutton_SignWriting

(Accessed date: October 15, 2011)

[80] SignWriting Site
http://www.signwriting.org/read.html
(Accessed date: October 15, 2011)

[81] James Ohene-Djan and Andrew Mould, "Finger-Chat: A Multi-mode Fingerspelling Chat System", in Proceedings of the Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment, June 29~July 2, 2004, CDROM

[82] Ye Kyaw Thu and Yoshiyori URANO, "A Comparison of Myanmar PC Keyboard Layouts", in Proceedings of the 8th International Symposium on Natural Language Processing, October 20~22, 2009, Bangkok, Thailand, pp. 15-20

[83] Ye Kyaw Thu and Yoshiyori URANO, "RomanPP: Myanmar Text Typing with Consonants Romanization and Possible Vowel Combinations Prediction with 4 Directional Arrow Keys", Proceedings of the 71st IPSJ National Conference, March 10~12, 2009, Ritsumeikan University, Shiga, Japan, pp. 4-17~4-18

[84] Yannis Haralambous (Translated by P. Scott Horne), "Font & Encodings", O'Reilly Media, Released: September 2007

[85] Pango engine
http://www.pango.org/
(Accessed date: October 15, 2011)

[86] Microsoft's Uniscribe Engine
http://www.microsoft.com/typography/developers/uniscribe/intro.htm
(Accessed date: October 15, 2011)

[87] IBM's ICU Engine
http://site.icu-project.org/
(Accessed date: October 15, 2011)

[88] QT framework
http://en.wikipedia.org/wiki/Qt_(framework)
(Accessed date: October 15, 2011)

[89] FreeType font engine

http://www.ascendercorp.com/services/font-engines/freetype/

(Accessed date: October 15, 2011)

[90] Tavultesoft Keyman

http://www.tavultesoft.com/keyman/

(Accessed date: October 15, 2011)

[91] Grover, D. L., King, M. T., and Kuschler, C. A. Patent No. US5818437, Reduced keyboard disambiguating computer. Tegic Communications, Inc., Seattle, WA (1998)

[92] Home Page of FontLab Ltd.

http://www.fontlab.com/

(Accessed date: October 15, 2011)

[93] Notable features of Thai alphabet

http://www/omniglot.com/writing/thai.htm

(Accessed date: October 15, 2011)

[94] Ye Kyaw Thu and Yoshiyori URANO, "Positional Mapping Myanmar Text Input Scheme for Mobile Devices", in proceeding of the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI 07), September 9~12, Singapore Polytechnic Auditorium, Singapore, Page 153 – 160

[95] Ye Kyaw Thu and Yoshiyori URANO, "Positional Mapping: Keyboard Mapping Based on Characters Writing Positions for Mobile Devices", in proceeding of the 9th International Conference on Multimodal Inter-faces (ICMI 2007), November 12~15, Nagoya, Japan, Page 110 – 117T9

[96] ACM Student Research Competition Grand Finals 2009, "Positional Prediction: Common and User Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices", June 27, 2009

http://www.acm.org/src/thu/YeKyawThu/thu.htm

(Notice: Myanmar and Khmer fonts need to install)

[97] Ye Kyaw Thu, Shrestha PRADYUMNA and Yoshiyori URANO, "PP_Clickwheel: Positional Prediction Nepali Text Input Interface with Clickwheel", Proceedings of the 2009 IEICE General Conference, March 17~20, 2009, Ehime University, Matsuyama, Japan, Page S-49~S-50

[98] T.H. Cormen, C.E. Leiserson, R.L. Rivest and C. Stein., “Introduction to Algorithms”, MIT Press/McGraw-Hill, ISBN: 0-262-03293-7, pp. 988.

[99] Hla Hla Htay and Kavi Narayana Murthy, “Myanmar Word Segmentation”, Proceedings of the 4th International Conference on Computer Applications (ICCA2006), February 23~24, 2006, Yangon, Myanmar, pp. 353-359

[100] Hla Hla Htay and Kavi Narayana Murthy, "Myanmar Word Segmentation using Syllable level Longest Matching", to appear in the proceedings of the IJCNLP-2008 Workshop on Asian Language Resources, 11-12 January 2008, Hyderabad, India

[101] Hla Hla Htay, G Bharadwaja Kuma and Kavi Narayana Murthy, “Statistical Analyses of Myanmar Text Corpora”, Creation of Lexical Resources for Indian Language Computing and Processing LRIL-2007, C-DAC, Mumbai, 26-28 March 2007

[102] Zin Maung Maung and Yoshiki MIKAMI, “A Rule-based Syllable Segmentation of Myanmar Text”, Proceedings of the IJCNLP-2008 Workshop on Asian Language Resources, 11-12 January 2008, Hyderabad, India, pp.51-58

[103] Tun Thura Thet, Jin-Cheon Na and Wunna Ko Ko, “Word segmentation for the Myanmar language”, Journal of Information Science, Volume 34 Issue 5, October 2008, pp. 688-704

[104] San Ko Oo and Yoshiki MIKAMI, “A Proposal for Myanmar Language Stemmer”, Workshop on Natural Language Processing for Asian Languages (a part of GITS/GITI Research Festival 2009), October 8, 2009, Waseda University, Honjo, Japan, pp. 70-75

[105] Aye Myat Mon, Myint Myint Thein, Su Su Htay, Soe Lai Phyue and Thinn Thinn Win, “Analysis of Myanmar Word Boundary and Segmentation by Using Statistical Approach, Proceeding of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE) 2010, Chengdu, China, August 20-22, 2010, pp. V5-233-V5-237

[106] Martin Hosken, “Unicode Technical Note #11, Representing Myanmar in Unicode: Details and Examples Version 3”, October 8, 2009

http://www.unicode.org/notes/tn11/UTN11_3.pdf

[107] Syllable Based Dual Weight Algorithm for Line Breaking in Myanmar Unicode

http://www.thanlwinsoft.org/ThanLwinSoft/MyanmarUnicode/Parsing/my2weightLineBreakAlg1_1.pdf

[108] Glossary of Unicode Terms
http://unicode.org/glossary/#grapheme_cluster
(Accessed date: October 15, 2011)

[109] Grapheme Cluster Boundaries
http://www.unicode.org/reports/tr29/
(Accessed date: October 15, 2011)

[110] Myanmar ISP e-books
https://www.myanmarisp.com/burmese_ebooks
(Accessed date: October 15, 2011)

[111] Document Character Converter
http://www.ThanLwinSoft.org
(Accessed date: October 15, 2011)/

[112] Jan Goyvaerts, "Regular Expressions (The Complete Tutorial)", ISBN: 1-4116-7760-9

[113] iPod
http://en.wikipedia.org/wiki/IPod
(Accessed date: October 15, 2011)

[114] Text Input Methods for iPod
http://www.ipodlinux.org/wiki/Text_Input_Methods
(Accessed date: 5 January 2009)

[115] Keypad Text Input Method
http://www.ipodlinux.org/wiki/Keypad
(Accessed date: 5 January 2009)

[116] Cursive Text Input Method
http://www.ipodlinux.org/wiki/Cursive
(Accessed date: 5 January 2009)

[117] Thumbscript Text Input Method
http://www.thumbscript.com/
(Accessed date: October 15, 2011)

[118] Ye Kyaw Thu and Yoshiyori URANO, “Positional Prediction Software Keyboard Layout for Myanmar PDA”, Proceedings of the 2008 IEICE Society Conference, Kawasaki, September 16-19, 2008, pp. S-155~S-156

[119] Ye Kyaw Thu and Yoshiyori URANO, “Analysis on Four Directional Arrow Keys (4DAK) and Left, Right, Up, Down and Center (LRUDC) Button Text Input Interfaces for Myanmar Language”, Proceedings of the Forum on Information Technology (FIT2008), Fujisawa, September 2-4, 2008, pp. 31-34

[120] Ye Kyaw Thu, “Romanization Based Asian Syllabic Languages Text Input Interfaces for Non-Natives”, Proceedings of the 10th APRU Doctoral Students Conference, Kyoto, Japan, July 6-10, 2009, (CD-ROM)

[121] Ye Kyaw Thu and Yoshiyori URANO, “PP_Clickwheel: Positional Prediction Khmer Text Input Interface with Clickwheel for Small Mobile Devices”, The Journal of the Institute of Image Electronics Engineers of Japan (IIEEJ), Vol.39, No.5, 2010, pp. 654-662, September, 2010

[122] Jackchai Butsrikui, Ye Kyaw Thu, Mitsuji MATSUMOTO and Yoshiyori URANO, “Analysis on Possible Combinations of a Consonant with Vowels and Tone Marks for Thai”, Workshop on Natural Language Processing for Asian Languages (a part of GITS/GITI Research Festival 2009), Waseda University, Honjo, Japan, October 8, 2009, pp. 31-38

[123] Phavy Ouk, Ye Kyaw Thu, Mitsuji MATSUMOTO and Yoshiyori URANO, “A Word-Based Predictive Text Entry Method for Khmer Language”, Proceedings of the 2008 International Conference on Information Reuse and Integration, Las Vegas, USA, July 13-15, 2008, pp. 214-219

[124] Ye Kyaw Thu, Jackchai Butsrikui, Mitsuji MATSUMOTO and Yoshiyori URANO, "A Comparison of Thai Mobile Phone Keypad Mappings", Proceedings of the Forum on Information Technology (FIT2009), Sendai, September 2-4, 2009, pp. 461-464

[125] Ye Kyaw Thu and Yoshiyori URANO, “MyTap and M9: Comparison of Two Existing Methods for Myanmar Mobile Phone Text Input”, the 133rd Human Computer Interaction Symposium, Ishikawa, May 15-16, 2009, IPSJ SIG Technical Reports, 2009-HCI-133, (online publication)

[126] Ye Kyaw Thu, “Common and User-Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices”, Adjunct Proceeding of the 21st annual

ACM Symposium on User Interface Software and Technology, Monterey, California, USA, October 19-22, 2008, pp. 27-30

[127] Ye Kyaw Thu, “Positional Mapping, Positional Gesture and Positional Prediction for Syllabic Languages: Study with Myanmar, Bangla and Khmer”, Proceedings of the 9th APRU Doctoral Students Conference, Vladivostok, Russia, July 14-18, 2008, (CD-ROM)

[128] Ye Kyaw Thu, “Common and User-Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices”, International DADD-PhD Summer Academy on Multimodal Human-Machine Interaction in Different Application Scenarios, University of Chile, Santiago, Chile, August 25-September 10, 2008, pp. 191-217

[129] Ye Kyaw Thu, Ouk Phavy and Yoshiyori URANO, “Positional Gesture for Advanced Smart Terminals: Simple Gesture Text Input for Syllabic Scripts Like Myanmar, Khmer and Bangla”, Proceedings of the first ITU-T Kaleidoscope Academic Conference, Geneva, Switzerland, May 12-13, 2008, pp. 77-84

[130] Ye Kyaw Thu and Yoshiyori URANO, “Romanized Myanmar Language Handwriting Text Input Interface for Non-natives”, Proceedings of the Interfaces and Human Computer Interaction 2009 and Game and Entertainment Technologies 2009, Algarve, Portugal, June 17-23, 2009, pp. 121-128

[131] Ye Kyaw Thu and Yoshiyori URANO, “Romanization with Vowel Positional Information: New Romanized Text Input Concept for Burmese (Myanmar Language)”, Proceedings of the 2008 IEICE General Conference, Kitakyushu, March 18-21, 2008, BS-3-8, pp. S-24~S-25

[132] Ye Kyaw Thu, “Development of Fingerspelling Font for Myanmar Deaf Education”, Proceedings of the 11th APRU Doctoral Students Conference, Jakarta, Indonesia, July 12-16, 2010, (CD-ROM)

[133] Ye Kyaw Thu, Sai Aung Win Maung and Yoshiyori URANO, “Direct Keyboard Mapping (DKM) Layout for Myanmar Fingerspelling Text Input (Study with Developed Fingerspelling Font “mmfingerspelling.ttf”)”, GITS/GITI Research Bulletin 2009-2010, Waseda University, pp. 127-135, July 31, 2010

[134] Ye Kyaw Thu, Shrestha PRADYUMNA and Yoshiyori URANO, “Analysis on Possible Combinations of Vowels with a Consonant of Nepali”, Proceedings of the 7th International Conference on Computer Applications (ICCA2009), Yangon, Myanmar, February 26-27, 2009, pp. 17-23

[135] Ye Kyaw Thu, and Yoshiyori URANO, “Grapheme Cluster Segmentation Tool for Myanmar Text Based on Positional Prediction Text Input Concept”, Proceedings of the 9th International Conference on Computer Applications (ICCA2011), Yangon, Myanmar, May 5-6, 2011, pp. 67-76

LIST OF PUBLICATIONS

Articles in refereed journals	Ye Kyaw Thu and Yoshiyori URANO, "PP_Clickwheel: Positional Prediction Khmer Text Input Interface with Clickwheel for Small Mobile Devices", The Journal of the Institute of Image Electronics Engineers of Japan (IIEEJ), Vol.39, No.5, 2010, pp. 654-662, September, 2010 Ye Kyaw Thu, Sai Aung Win Maung and Yoshiyori URANO, "Direct Keyboard Mapping (DKM) Layout for Myanmar Fingerspelling Text Input (Study with Developed Fingerspelling Font "mmfingerspelling.ttf")", GITS/GITI Research Bulletin 2009-2010, Waseda University, pp. 127-135, July 31, 2010
Presentations at International conferences	Ye Kyaw Thu, and Yoshiyori URANO, "Grapheme Cluster Segmentation Tool for Myanmar Text Based on Positional Prediction Text Input Concept", Proceedings of the 9th International Conference on Computer Applications (ICCA2011), Yangon, Myanmar, May 5-6, 2011, pp. 67-76 Nu Nan Dar Lin, Ye Kyaw Thu, Mitsuji MATSUMOTO and Yoshiyori URANO, "Survey on Ontology Languages", Proceedings of the 9th International Conference on Computer Applications (ICCA2011), Yangon, Myanmar, May 5-6, 2011, pp. 260-267 Ye Kyaw Thu, Sai Aung Win Maung and Yoshiyori URANO, "Direct Keyboard Mapping (DKM) for Myanmar Fingerspelling Text Input", Proceedings of the IADIS International Conferences, Interfaces and Human Computer Interaction 2010 and Game and Entertainment Technologies 2010, Freiburg, Germany, July 28-30, 2010, pp. 267-275 Ye Kyaw Thu and Yoshiyori URANO, "A Comparison of Myanmar PC Keyboard Layouts", Proceedings of the 8th International Symposium on Natural Language Processing (SNLP 2009), Bangkok, Thailand, October 20-22, 2009, pp. 15-20 Ye Kyaw Thu and Yoshiyori URANO, "Romanized Myanmar Language Handwriting Text Input Interface for Non-natives", Proceedings of the Interfaces and Human Computer Interaction 2009 and Game and Entertainment Technologies 2009, Algarve, Portugal, June 17-23, 2009, pp. 121-128 Ye Kyaw Thu, Shrestha PRADYUMNA and Yoshiyori URANO, "Analysis on Possible Combinations of Vowels with a Consonant of Nepali", Proceedings of the 7th International Conference on Computer Applications (ICCA2009), Yangon, Myanmar, February 26-27, 2009, pp. 17-23 Ouk Phavy, Ye Kyaw Thu and Mitsuji MATSUMOTO, "Khmer Word-based Predictive Non-QWERTY Soft Keyboard in Stylus-based Devices", Proceedings of the 2008 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2008), Herrsching am Ammersee, Germany, September 15-19, 2008, pp. 225-232 Phavy Ouk, Ye Kyaw Thu, Mitsuji MATSUMOTO and Yoshiyori URANO, "A Word-Based Predictive Text Entry Method for Khmer Language", Proceedings of the 2008 International Conference on Information Reuse and Integration, Las Vegas, USA, July 13-15, 2008, pp. 214-219

	Ye Kyaw Thu, Ouk Phavy and Yoshiyori URANO, “Positional Gesture for Advanced Smart Terminals: Simple Gesture Text Input for Syllabic Scripts Like Myanmar, Khmer and Bangla”, Proceedings of the first ITU-T Kaleidoscope Academic Conference, Geneva, Switzerland, May 12-13, 2008, pp. 77-84 [Young Author Recognition Certificate] Ye Kyaw Thu, "Positional Prediction: Consonant Cluster Prediction Text Entry Method for Burmese (Myanmar Language)", Extended Abstracts of the 26th ACM Conference on Human Factors in Computing Systems (CHI 2008), Florence, Italy, April 5-10, 2008, pp. 3783-3788 [2nd Prize for CHI08 Student Research Competition] Ye Kyaw Thu, Ouk Phavy and Yoshiyori URANO, “Positional Mapping Software Keyboard for Syllabic Scripts (Study with Bangla and Khmer Languages)”, Proceedings of the 6th International Conference on Computer Applications (ICCA2008), Yangon, Myanmar, February 14-15, 2008, pp. 150-157 MD. Monzur Morshed, Ye Kyaw Thu and Yoshiyori URANO, “Frequency Based Two–Layer Multitap Bangla Input Method for Mobile Phones”, Proceedings of the 10th International Conference on Computer and Information Technology (ICCIT 2007), Dhaka, Bangladesh, December 27-29, 2007, pp. 313-319 Ye Kyaw Thu and Yoshiyori URANO, "Positional Mapping: Keyboard Mapping Based on Characters Writing Positions for Mobile Devices", Proceedings of the 9th International Conference on Multimodal Interfaces (ICMI 2007), Nagoya, Japan, November 12-15, 2007, pp. 110-117 Ye Kyaw Thu and Yoshiyori URANO, “Positional Mapping Myanmar Text Input Scheme for Mobile Devices”, Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI 07), Singapore, September 9-12, 2007, pp. 153-160 [Best Paper Award] Ye Kyaw Thu and Yoshiyori URANO, “Positional Mapping Multi–tap for Myanmar Language”, Proceedings of the 12th International Conference on Human–Computer Interaction (HCI International 2007), Beijing, China, July 22-27, 2007, Part 2, pp. 486-495 Ye Kyaw Thu and Yoshiyori URANO, “Selection Based Myanmar Text Input Interfaces: (Proposal of Dividing All Myanmar Characters into 6 Groups)”, Proceedings of the 2nd IASTED International Conference on Human Computer Interaction (IASTED-HCI 2007), Chamonix, France, March 14-16, 2007, pp. 154-159 Ye Kyaw Thu and Yoshiyori URANO, “Selection Based Myanmar Text Input Interfaces for Small Screen Computing Devices”, Proceedings of the 5th International Conference on Computer Applications (ICCA2007), Yangon, Myanmar, February 8-9, 2007, pp. 215-222 Ye Kyaw Thu and Yoshiyori URANO, “Query Interfaces for Retrieving Myanmar Language Digital Resources”, Proceedings of the 2nd International Conference on Digital Libraries (ICDL 2006), New Delhi, India, December 5-8, 2006, pp. 53-65 Ye Kyaw Thu and Yoshiyori URANO, “Text Entry for Myanmar Language SMS: Proposal of 3 Possible Input Methods, Simulation and Analysis”, Proceedings of the 4th International Conference on Computer Applications (ICCA2006), Yangon, Myanmar, February 23-24,

	2006, pp. 199-207
Presentations at domestic academic meetings held by study groups	Ye Kyaw Thu and Yoshiyori URANO, “MyTap and M9: Comparison of Two Existing Methods for Myanmar Mobile Phone Text Input”, the 133rd Human Computer Interaction Symposium, Ishikawa, May 15-16, 2009, IPSJ SIG Technical Reports, 2009-HCI-133, (online publication) Ye Kyaw Thu and Yoshiyori URANO, “A Comparison of Myanmar PC Keyboard Layouts”, the 131st Human Computer Interaction Symposium, Oita, January 26-27, 2009, IPSJ SIG Technical Reports, 2009-HCI-131, pp. 15-21 Ye Kyaw Thu and Yoshiyori URANO, “LRUDC (Left, Right, Up, Down and Center) Button for Myanmar Language Software Keyboard”, the 128th Human Computer Interaction Symposium, Kobe, May 28-29, 2008, IPSJ SIG Technical Reports, 2008-HCI-128, pp. 159-164 Ye Kyaw Thu, Ouk Phavy, Yoshiyori URANO and Mitsuji MATSUMOTO, “Positional Prediction for Khmer Language (Cluster Predictive Text Entry Method in Mobile Devices)”, IIEEJ 4th Mobile Image Research Meeting, Tokyo, March 4, 2008, pp. 5-10 Ye Kyaw Thu, Ouk Phavy and Yoshiyori URANO, “Positional Gesture: Simple Gesture Text Input for Khmer”, the 127th Human Computer Interaction Symposium, Higashi Hiroshima, January 31-February 1, 2008, IPSJ SIG Technical Reports, 2008-HCI-127, pp. 115-122 Ouk Phavy, Ye Kyaw Thu and Yoshiyori URANO, “Two Keystrokes Text Entry Method for Khmer Mobile Phones”, the 127th Human Computer Interaction Symposium, Higashi Hiroshima, January 31-February 1, 2008, IPSJ SIG Technical Reports, 2008-HCI-127, pp. 123-130 Ye Kyaw Thu and Yoshiyori URANO, “Lookup Interfaces for Myanmar Electronic Dictionaries”, the 121st Human Interface Symposium, Kyoto, November 21-22, 2006, IPSJ SIG Technical Reports, 2006-HI-121, pp. 75-82 MD. Monzur Morshed, Ye Kyaw Thu, Yoshiyori URANO, “A Proposal of Reduced Keystroke Multitap Input Model for Bangla Short Message Service”, the 121st Human Interface Symposium, Kyoto, November 21-22, 2006, IPSJ SIG Technical Reports, 2006-HI-121, pp. 43-49 Ye Kyaw Thu and Yoshiyori URANO, “Development of Romanized Input Methods for Myanmar Language Short Messaging Service (SMS): Comparison with Multitap Input Method, Simulation and Analysis”, the 117th Human Interface Symposium, Tsukuba, January 13, 2006, IPSJ SIG Technical Reports, 2006-HI-117, pp. 33-40
Presentations at domestic conferences	Nu Nan Dar Lin, Sai Aung Win Maung, Ye Kyaw Thu, Mitsuji Matsumoto and Yoshiyori Urano, “Implementation of Japanese to Myanmar Online-Dictionary By Using Web Ontology Language (OWL)”, Proceedings of the 2011 IEICE Society Conference, Sapporo, September 13-16, 2011, pp. S-78~S-79 Huotely YIN, Ye Kyaw Thu, and Yoshiyori URANO, “New Khmer Handwriting Text Input

Interface: Handwriting and Positional Prediction", Proceedings of the 2009 IEICE Society Conference, Niigata, September 15-18, 2009, pp. S-133~S-134

Ye Kyaw Thu, Jackchai Butsrikui, Mitsuji MATSUMOTO and Yoshiyori URANO, "A Comparison of Thai Mobile Phone Keypad Mappings", Proceedings of the Forum on Information Technology (FIT2009), Sendai, September 2-4, 2009, pp. 461-464

Ye Kyaw Thu, Shrestha PRADYUMNA and Yoshiyori URANO, "PP_Clickwheel: Positional Prediction Nepali Text Input Interface with Clickwheel", Proceedings of the 2009 IEICE General Conference, Matsuyama, March 17-20, 2009, pp. S-49~S-50

Chinzorig Sashka, Ye Kyaw Thu, and Yoshiyori URANO, "Personalized review based recommendation system for tourism domain", Proceedings of the 2009 IEICE General Conference, Matsuyama, March 17-20, 2009, pp. 95

Huotely YIN, Ye Kyaw Thu, and Yoshiyori URANO, "Online Khmer Handwriting Character Recognition by employing Nearest Neighbor and DTW Algorithms", Proceedings of the 2009 IEICE General Conference, Matsuyama, March 17-20, 2009, pp. S-47~S-48

Ye Kyaw Thu and Yoshiyori URANO, "RomanPP: Myanmar Text Typing with Consonants Romanization and Possible Vowel Combinations Prediction with 4 Directional Arrow Keys", Proceedings of the 71st IPSJ National Conference, Shiga, March 10-12, 2009, pp. 4-17~4-18

Ye Kyaw Thu and Yoshiyori URANO, "Positional Prediction Software Keyboard Layout for Myanmar PDA", Proceedings of the 2008 IEICE Society Conference, Kawasaki, September 16-19, 2008, pp. S-155~S-156

Ye Kyaw Thu and Yoshiyori URANO, "Analysis on Four Directional Arrow Keys (4DAK) and Left, Right, Up, Down and Center (LRUDC) Button Text Input Interfaces for Myanmar Language", Proceedings of the Forum on Information Technology (FIT2008), Fujisawa, September 2-4, 2008, pp. 31-34

Ye Kyaw Thu and Yoshiyori URANO, "Romanization with Vowel Positional Information: New Romanized Text Input Concept for Burmese (Myanmar Language)", Proceedings of the 2008 IEICE General Conference, Kitakyushu, March 18-21, 2008, BS-3-8, pp. S-24~S-25

Ouk Phavy, Ye Kyaw Thu, Mitsuji Matsumoto and Yoshiyori URANO, "Cluster Predictive Text Entry Method in Mobile Phone for Khmer Language", Proceedings of the 2008 IEICE General Conference, Kitakyushu, March 18-21, 2008, BS-3-9, pp. S-26~S-27

Ye Kyaw Thu and Yoshiyori URANO, "Myanmar Text Typing with Numeric Keypad (Applying the concept of Positional Mapping)", Proceedings of the 70th IPSJ National Conference, Tsukuba, March 13-15, 2008, pp. 4-1~4-2

Ye Kyaw Thu, MD. Monzur Morshed and Yoshiyori URANO, "Positional Mapping for Bangla Mobile Phone", Proceedings of the Forum on Information Technology (FIT2007), Toyota, September 5-7, 2007, pp. 427-430

Ye Kyaw Thu and Yoshiyori URANO, "Myanmar Text Input Interface for Personal Digital

	Assistant (PDA): Applying the concept of Dividing All Myanmar Characters into 6 Groups", Proceedings of the 2007 IEICE General Conference, Nagoya, March 20-23, 2007, pp. S-105 ~S-106 Ye Kyaw Thu and Yoshiyori URANO, "Developing Efficient Input Methods for Myanmar Language SMS", Proceedings of the 2005 IEICE Communications Society Conference, Sapporo, September 20-23, 2005, B-17-5, pp. 624
Workshop	Ye Kyaw Thu and Yoshiyori URANO, "Software Assisting Tools In Teaching And Learning for Myanmar Students With Visual Disabilities", GII-GITS Workshop on NLP and ICT 2011 (a part of GITS/GITI Research Festival 2011), October 7, 2011, Nagaoka University of Technology, Nagaoka, Japan, pp. 57-63 Ye Kyaw Thu and Yoshiyori URANO, "Software Assisting Tools In Teaching And Learning for Myanmar Students With Visual Disabilities", "Language and Speech Science Research Laboratories" Technical Report (a part of GITS/GITI Research Festival 2011), October 5, 2011, Waseda University, Tokyo, Japan Ye Kyaw Thu, "Direct Keyboard Mapping (DKM) Layout for Myanmar Fingerspelling Text Input (Study with Developed Fingerspelling Font "mmFingerspelling.ttf")", Joint Workshop of Natural Language and Speech Recognition, the 9th International Conference on Computer Applications (ICCA2011), Yangon, Myanmar, May 5-6, 2011 Ye Kyaw Thu, "Software Assisting Tools in Teaching and Learning for Myanmar Students with Hearing and Visual Disabilities", Tsinghua University & GITS Joint Workshop, Tsinghua University, Beijing, China, March 12, 2011 Ye Kyaw Thu and Yoshiyori URANO, "Positional Prediction Text Input Interfaces for Asian Syllabic Languages", Language and Speech Science Research Laboratories Technical Report (a part of GITS/GITI Research Festival 2010), Waseda University, Tokyo, Japan, October 13, 2010 Ye Kyaw Thu and Yoshiyori URANO, "Prototyping with "Pictures & Mobile Devices": a Rapid Prototyping Technique for Mobile User Interfaces", Workshop on Natural Language Processing for Asian Languages (a part of GITS/GITI Research Festival 2009), Waseda University, Honjo, Japan, October 8, 2009, pp. 46-53 Jackchai Butsrikui, Ye Kyaw Thu, Mitsuji MATSUMOTO and Yoshiyori URANO, "Analysis on Possible Combinations of a Consonant with Vowels and Tone Marks for Thai", Workshop on Natural Language Processing for Asian Languages (a part of GITS/GITI Research Festival 2009), Waseda University, Honjo, Japan, October 8, 2009, pp. 31-38 Ye Kyaw Thu and Yoshiyori URANO, "Introduction to Concepts of Positional Mapping (PM) and Positional Prediction (PP) for Asian Syllabic Languages Mobile Devices", Global Information and Telecommunication Open Research Paper Reading (GITS/GITI Workshop 2008), Waseda University, Tokyo, Japan, October 8, 2008 [written in Japanese] Ye Kyaw Thu and Yoshiyori URANO, "Selection Based Myanmar Text Entry Using Game Pads", Global Information and Telecommunication Open Research Paper Reading

	(GITS/GITI Workshop 2006), Waseda University, Honjo, Japan, October 18, 2006, pp. 12-13 Ye Kyaw Thu and Yoshiyori URANO, "Development of Efficient Input Methods for Myanmar Language SMS", Global Information and Telecommunication Open Research Paper Reading (GITS/GITI Workshop 2005), Waseda University, Honjo, Japan, October 19, 2005, pp. 4-5
Article	Ye Kyaw Thu, "Prototyping with "Pictures and Mobile Devices" (PicMobi): A Rapid Prototyping Technique for Mobile User Interfaces", The Transactions of the Asiatic Society of Japan, Fifth series, volume 2, 2010, pp. 129-136
Doctoral Symposium and PhD Academy	Ye Kyaw Thu, "Software Assisting Tools in Teaching And Learning For Myanmar Students with Visual Disabilities (Research in Progress)", Proceedings of the 12th APRU Doctoral Students Conference, Beijing, China, July 4-7, 2011, (CD-ROM) Ye Kyaw Thu, "Development of Fingerspelling Font for Myanmar Deaf Education", Proceedings of the 11th APRU Doctoral Students Conference, Jakarta, Indonesia, July 12-16, 2010, (CD-ROM) Ye Kyaw Thu, "Development of Myanmar Fingerspelling Text Input Methods (Work in Progress)", International DADD-PhD Academy on Multimodal Human-Machine Interaction in Different Application Scenarios, Waseda University, Tokyo, Japan, February 15-27, 2010 Ye Kyaw Thu, "Romanization Based Asian Syllabic Languages Text Input Interfaces for Non-Natives", Proceedings of the 10th APRU Doctoral Students Conference, Kyoto, Japan, July 6-10, 2009, (CD-ROM) Ye Kyaw Thu, "Common and User-Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices", Adjunct Proceeding of the 21st annual ACM Symposium on User Interface Software and Technology, Monterey, California, USA, October 19-22, 2008, pp. 27-30 Ye Kyaw Thu, "Common and User-Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices", International DADD-PhD Summer Academy on Multimodal Human-Machine Interaction in Different Application Scenarios, University of Chile, Santiago, Chile, August 25-September 10, 2008, pp. 191-217 Ye Kyaw Thu, "Positional Mapping, Positional Gesture and Positional Prediction for Syllabic Languages: Study with Myanmar, Bangla and Khmer", Proceedings of the 9th APRU Doctoral Students Conference, Vladivostok, Russia, July 14-18, 2008, (CD-ROM)
Invited Talk and Others	Ye Kyaw Thu, "携帯機器におけるアジア音節言語共通のユーザーフレンドリーな文字入力インターフェース", Atsumi International Scholarship Foundation (AISF) Kenkyuu Hokokukai 2009, Tokyo, Japan, March 6, 2010 [presented in Japanese] Ye Kyaw Thu, "Prototyping with "Pictures & Mobile Devices", Young Scholars'

	Programme of the Asiatic Society of Japan, Tokyo, Japan, November 16, 2009
	Ye Kyaw Thu, "Introduction to Concepts of Positional Mapping (PM), Positional Gesture (PG) and Positional Prediction (PP) for Asian Syllabic Languages on Mobile Devices", Waseda University Doctoral Student Network Meeting, Tokyo, Japan, December 9, 2008
	Ye Kyaw Thu, "Keyboard Mapping and Text Input Interfaces of Asian Languages Based on Word Formation", Myanmar Language Processing Workshop, Nagaoka University of Technology, Nagaoka, Japan, June 9, 2008
Poster	Ye Kyaw Thu, "Font and Keyboard Layout for Myanmar Fingerspelling", Poster Session for 70th Anniversary of Research Institute for Science and Engineering, Waseda University, Tokyo, Japan, July 5-30, 2010
	Ye Kyaw Thu, YIN Huotely and Yoshiyori URANO, "Romanized Khmer Language Handwriting Interface for Non-Natives", Interaction 2009, Tokyo, Japan, March 5-6, 2009
Demo	Ye Kyaw Thu, Sai Aung Win Maung, Mitsuji MATSUMOTO and Yoshiyori URANO, "Software Assisting Tools in Teaching and Learning for Myanmar Students with Hearing and Visual Disabilities", 7th IEEE Tokyo Young Researcher Workshop, Tokyo, Japan, November 20, 2010
	Ye Kyaw Thu and Yoshiyori URANO, "RomanPP Handwriting: Myanmar Language Text Input Interface for Non-natives", 72nd National Convention of IPSJ, Tokyo, Japan, March 9-11, 2010
	Ye Kyaw Thu, Jackchai Butsrikui, Yoshiyori URANO and Mitsuji MATSUMOTO, "Romanization Based Asian Syllabic Languages Text Input Prototypes for Non-Natives", 3rd International Universal Communication Symposium (IUCS 2009), Tokyo, Japan, December 3-4, 2009
Award	Waseda University Grant for Special Research Projects (Project number: 2011B-215), 2011
	Waseda University Grant for Special Research Projects (Project number: 2009A-901), 2009
	3rd Prize for ACM Student Research Competition Grand Finals 2009, "Positional Prediction: Common and User Friendly Text Input Interfaces for Asian Syllabic Languages on Mobile Devices", ACM Awards Banquet, San Diego, USA, June 27, 2009
	Waseda University Expense Stipend, Student representative to the 9th APRU Doctoral Students Conference, Vladivostok, Russia, July 14-18, 2008
	Young Author Recognition Certificate of ITU-T Kaleidoscope Academic Conference, "Positional Gesture for Advanced Smart Terminals: Simple Gesture Text Input for Syllabic Scripts Like Myanmar, Khmer and Bangla", May 12-13, 2008
	2nd Prize for ACM Student Research Competition of CHI 2008, "Positional Prediction: Consonant Cluster Prediction Text Entry Method for Burmese (Myanmar Language)", April 5-10, 2008

	Best Paper Award for ACM MobileHCI 07, “Positional Mapping Myanmar Text Input Scheme for Mobile Devices”, September 9-12, 2007