

Article-5 (2001 년 씀)
(필자의 승인 없는 인용을 금함)

Hangul Speedy Input for ECJ and other languages

- Shorthand input method of Hangul for M17N environment –

By Matthew Y. Ahn

Introduction

Unicode is an interoperable code system for all languages. Any writing system based on Unicode can be processed without re-booting or restarting your computer system.

AhnMaTae Phonetic Hangul Keyboard Input System (APHKIS) is a phonetically arranged Hangul (Korean writing system) keyboard system. The input method is similar to the English stenographic shorthand machine. The difference is that it does not require a separate machine but rather utilizes existing computer hardware. APHKIS can be accessed directly from a computer without any hardware changes.

The purpose of this paper is to demonstrate how it is possible and to lay a practical guide for how to use this system for anyone who is interested in short hand input system for their own respective writing systems. Currently APHKIS is developed to use English version of Microsoft Word with IME with Korean (see: <http://office.microsoft.com/downloads/2002/imekor.aspx>). APHKIS is being developed by the Center for Artificial Intelligence Research (CAIR) of KAIST (Korea Advanced Institute of Science and Technology) and is available free of charge to anyone who request for it at <http://ai.kaist.ac.kr/ahnmatae>.

There are various speedy input systems invented for various languages, but phonetic Hangul keyboard simultaneous input system is an efficient and economical processing of any languages. For instance, English processing on computer needs stenographic shorthand machine, separate from the computer, to input English words simultaneously, while APHKIS can input words directly into the computer. It is approximately three times faster in input speed in comparison to serial input system in Korean writing system, and it is expected to have the same speed of input in other languages.

Hangul as a Scientific Writing System

There are phonetic symbols, such as International Phonetic Alphabet (IPA) and Bopobopo (Chinese phonetic symbols, name of the first four alphabets). But they are for limited languages and not for universal languages: IPA for Latin languages and Bopomofu for Chinese language only.

Hangul is a perfect phonetic symbol, which can apply to any languages. It was invented in 1443 by King Sejong and eight court scholars after more than half a dozen years of toiling efforts of inventing most scientific and most versatile writing system in the world. Despite the fact that most of the world writing systems is developed by hundreds or thousands years of evolutions, Hangul was invented deliberately as phonetically accurate symbols to depict all the sound in nature.

Some Western scholars believe that Hangul writing system was adopted by neighboring countries' writing systems, such as Mongolian Phags-pa writing system and or Sanskrit, but it is a result of complete misunderstanding of Hangul. *Hunminjeongeum Haerae* (Explanation to Teaching People with Correct Sound, 訓民正音解例, published in 1446), which explains how it was invented and how it should be used, clearly indicates its own unique invention. One of the authors, *Chung Inji (1396-1478)*, explains that Hangul was deliberately invented to depict the shapes of sound creating organs for consonants, and

Hangul is so amazingly easy to learn, that it requires only one hour of practice to read and write Hangul. The reasons for such easiness are: (1) No single Jamo looks alike visually, (2) Consonants and Vowels are distinguishable by its shape, (3) Each Jamo has only one sound value, (4) Each Jamo has a unique frequency of sound.

(1) No Jamo looks alike.

English has confusing typography. For example, capital letter ‘I’ (ai) and small letter ‘i’ (el) look so much alike, that a non-native English user confuses in reading and writing. However, there is no single Hangul Jamo typography, which looks similar.

(2) Consonants and Vowels are visually different.

In Hangul, Vowels are dot(.), horizontal line (—) and vertical line (|) and or with combinations of it. Consonants are combinations of above geometric signs plus slush (/); reverse slush (\) and or a circle (○). Therefore, it is easy to distinguish which Jamo is consonant and which is vowel.

(3) Only one sound value for each Jamo.

In English, alphabet ‘a’ can be pronounced differently, such as o (ball), eo (another), ei (baby) and a (cola). On the other hand, Hangul Jamo vowel ‘ㅏ’ is pronounced as ‘a’ in the case of cola, and nothing else. Only the exception is ‘ㅑ’. When it is placed before the vowel(s) as a leading consonant, it does not have sound value and acts only as a fill code. When it is used after vowel(s) as a trailing consonant, it has a sound value of ‘ng’, a nasal sound. When it comes after another consonant, it softens the sound of combining consonant, as in the case of V (ㅓㅑ) and F or PH (ㅕㅑ) sound.

(4) Different frequency for each Jamo.

Professor Han, Taedong, a retired theologian at Yonsei University in Seoul, describes in his recent book, *Phonology of Sejong Period*, (written in Korean and published by Yonsei University Press, 1998. A must read book by those who study phonetics and those who pursue scientific research on artificial intelligence area), that each basic consonant has different frequency. When he has tested with sonogram, on five basic consonants, following results were obtained (on page 50 of his book):

<u>Name of</u> <u>Jamo</u>	<u>Distance from</u> <u>Voice Box</u>	<u>Frequency of</u> <u>Sound</u>	<u>First</u> <u>Octave</u>	<u>Second</u> <u>Octave</u>
-------------------------------	--	-------------------------------------	-------------------------------	--------------------------------

ㄱ	15.5 cm	2, 270 Hz	4, 500 Hz	7, 000 Hz
ㅋ	13.5cm	2, 500 Hz	5, 000 Hz	7, 500 Hz
ㆁ	13.5 cm	3, 500 Hz	7, 000 Hz	
ㄷ	?	?		
ㅇ	11.5 cm	4, 000 Hz	8, 000 Hz	

He found that the five consonants had same frequencies and distance ratio as flutes, which was based on five musical notes of fifteenth century. He writes: "Hangul is only the scientific writing system in the world, which is based on the shapes of sound creating organs of human being, and based on the frequencies of musical instrument of the time".

Why APHKIS is so fast?

Even though Hangul was invented five and a half centuries ago, a long time before computer was invented, it is designed as if it was invented for the computer age. The reason is that it was designed for all the sounds in the nature (*declaration of King Sejong in Hunminjeongeum*), and APHKIS was designed to use it on the computer most faithfully what the inventors intended. How fast one can type is not known yet, however, it is predictable that one can type as fast as one can speak, as most of the shorthand systems are.

Followings are the reasons for fast input by APHKIS:

(1) Syllabic Character input system.

Hangul is designed to write as a syllabic character. Order of writings of syllabic characters is: Leading consonant(s) and medial Vowel(s), and or without Trailing consonant(s). Another word, Hangul syllabic characters are conjoined L + V, or L + V + T. Most of Hangul characters are consisted of Leading consonants and Vowels (approximately 80 %) only, however, there are approximately 20 % of syllabic characters, which are consisted of Leading consonant(s) and Vowels with Trailing consonants. Averaging two and half of Jamo make one syllabic character. Therefore, theoretically, simultaneous input system is two and half times faster than serial input system.

(3) No shift keys are used.

APHKIS can input Hangul without using any shift keys. Shift key use slows down the input speed and promotes typist's fatigue due to the fact that it uses weakest fingers of both hands.

Input of double consonants ㅃ, ㅆ, ㅈ, ㅊ, ㅌ use simultaneous pushing of next right side keys with ㅅ, ㅍ, ㅊ, ㅌ, ㅎ.

Input of harsh sound Jamo of ㅋ, ㅌ, ㅍ use simultaneous pushing of ㅍ, ㅊ, ㅌ, ㅎ with ㅎ. Multiple vowels are automatically semi-composed when they are pushed simultaneously.

Trailing consonants harsh sound Jamo of ㅋ, ㅌ, ㅍ use simultaneous pushing of ㅍ, ㅊ, ㅌ, ㅎ, with ㅎ. (Instruction on how to use APHKIS is available on line at <http://ai.kaist/ahnmatae>)

(4) Shortest distance of finger traveling.

The longer the finger travels, the slower input speed occurs. APHKIS is designed to minimize the finger traveling by, (a) concentrated layout of most frequently used Jamo on home rows (second row from the bottom), (b) not assigning any Jamo on the fourth row (top row where the Roman numerals are located), (c) use of both thumbs for the bottom row's trailing consonants to free other eight fingers from traveling to the bottom row.

(5) Ergonomic design.

Inner fingers are stronger and faster than outer fingers of human hands. By assigning more frequently used Jamo inside, and less frequently used Jamo outside, the speed of input can be accelerated while fatigues are minimized.

(6) Space bar and sentence marks are simultaneously inputted as trailing.

Hangul uses spaces after every word, averaging about three syllabic characters, and sentence marks such as comma, period, question mark, etc. By simultaneously input them as trailing to the last syllabic character saves stroke numbers, and accelerate the speed of input.

How fast can it be?

As most other shorthand typing systems, this system can be as fast as one can speak. It is too early to predict exactly how fast one can type with this system. However it is reasonably predictable that it is much faster than the serial input system in various languages.

In the early part of 1980's, the test result of AhnMaTae keyboard and the Korean Standard keyboard in the USA, came to the conclusion that AhnMaTae keyboard was twice faster in memory of key positions in comparison with the Korean Standard keyboard. Input speed of serial typing at the initial stage was not much different. However, after half a year of typing experience one could type at least 50 % faster with AhnMaTae keyboard than the Korean Standard. Similar test was conducted in China in 1996 and 1967 and about the same result occurred. As of November 2001, a similar test is conducted at Pyongyang, North Korea, and preliminary result shows the same result as in US and China.

As this new simultaneous input system on Word was finalized in September of 2001, it is expected that the comparative study between simultaneous input method of AhnMaTae keyboard and serial input system of K. S. keyboard, will be available some time in 2002. As the typing is training of autonomous nerve system, it requires a long period of time to master a system and to come to the concrete conclusion of its input speed.

How fast APHKIS type for other languages would not be much different from Korean. Since current computer keyboards are designed for English usage and inherited from the typewriter age, characters are assigned only on the first three rows from the bottom. It has only 33 key positions available on three rows: 10 key positions on the first row, 11 on the second row (home row), and 12 on the third row. Most of world writing systems with basic alphabet far exceed these numbers and therefore it has to go up the upper key positions where numerical numbers are located. As in English typing, which requires shift key for Capital letters, most of other languages also require shift keys to type their alphabets. As keyboard has shift key bar on the far bottom of both ends and only reachable by small fingers of both hands, the more you use the little fingers, the slower typing speed occur, as our little fingers are weaker and slower than other fingers.

APHKIS is designed not to use shift keys at all, unless you need to use it for archaic Hangul. Hangul is so versatile in its use, that it can create millions of syllabic characters. It means any languages in the world can be written phonetically with Hangul.

Hangul syllabic characters are composed very regularly and predictably, and decomposed very easily. Ken Lunde, the author of CJKV (Chinese, Japanese, Korean and Vietnamese) Computer Processing, published in 2000, describes, "Hangul is considered to be one of the most scientific (or well-designed) writing systems due to its extremely regular and predictable structure". (p. 48)

Understanding of Hangul in Unicode

Most of average people misunderstand Hangul as an ideographic or a pictographic writing system when they see so many pre-composed Hangul characters (11,172 characters, from AC00 to D7A3) of modern usage in Unicode. Besides, there are also Compatibility Hangul Jamo (3130 to 318A) and Hangul Jamo (1100 to 11FF) in Unicode.

The confusion and misunderstanding of Hangul in the code system were caused by the historical misuse of Hangul on computer, and misplacement of Hangul Jamo on the Korean Standard Keyboard (KS Keyboard).

Current KS keyboard was adopted in 1969 as a keyboard for the teletypewriter, which was originally invented to use English. As English alphabet had 26 alphabet and Korean modern Hangul Jamo had 24 letters, that original KS Keyboard for this machine added two double vowels of Hangul (ㅢ, ㅣ) to match English. In the later years, it added five more consonants (ㅃ, ㅆ, ㅉ, ㅊ, ㅌ) and two more double vowels (ㅥ, ㅦ) to eliminate the confusion when conjoined into a composed syllabic character. Therefore, current KS Keyboard has 33 Jamo assigned on the computer keyboard. Advocates of current KS Keyboard insist that it was designed by a thorough frequency studies, but my research on Hangul Jamo frequency study revealed that it was not true. For example, the frequency rate of current Jamo 'ㅢ' on KS keyboard is less than 0.001 % of the total usage. There are many other double vowels, which have far higher frequency rates.

Double consonants of ㅃ, ㅆ, ㅉ, ㅊ, ㅌ was absolute necessity on KS Keyboard due to the confusion of whether ㅃ, ㅆ, ㅉ, ㅊ, ㅌ could be used as Leading consonants or Trailing consonants when inputting them serially. For instance, when ㅃ ㅃ ㅃ ㅃ ㅃ ㅃ was inputted serially, as KS keyboard does, it could compose differently as 각각 or 가깝, depending on where you end the trailing consonant on the first syllabic character. To eliminate such confusion it was necessary to have separate double consonants on the keyboard.

Historically, those who were involved in the standardization on computer keyboard and codes were confused about the mechanical structure of Hangul composition and decomposition

systems that they have made Hangul input system as English alphabets, and codes were assigned as if they were ideographic characters. The first codes were given only with 2,350 pre-composed Hangul syllabic characters and then started adding it up. Finally, it ended up at 11,172 characters in Unicode, ISO/IEC 10646 and KS X 5700. However, this huge pre-composed pseudo-ideographic character numbers cannot satisfy to express other languages and all the sounds in the nature as intended by its inventors.

On the other hand, commercial vendors used their own keyboard, mostly Kong Byung Woo system, which had separate keys for Leading and Trailing consonants. They also used different code systems, such as one byte system using escape sequence and or double bytes with five bits assigned to each Jamo. To satisfy all the parties involved, in fact, another code standard in Korea, KS X 5601, was revised at least half a dozen times since its first inception.

Hangul was invented to scribe all the sound in the world, from whooping of cranes to blowing of winds, by not only using basic Jamo but by combining two or three consonants and multiple numbers of vowels, and conjoining them as a syllabic letter.

Hangul seems very confusing in conjoining system, but in reality, it is not. As Ken Lund described in his book, it has extremely regular and predictable structure. However, Hangul is written in rectangular shape and its composition of syllabic character is made up by different glyphs, that it looks very much alike confusing. For instance, leading consonant glyph of ‘ㄱ’ changes its size and shape for six times: three times by the shapes of following vowel(s), and three times when it has trailing consonants. Following syllabic character with leading consonant glyph of ‘ㄱ’ is self-explanatory.

가, 고, 과, 감, 곶, 팜,

APHKIS uses not only the pre-composed syllabic Hangul characters, but uses Hangul Jamo to conjoin them when it does not have pre-composed syllabic characters. It will compose with Hangul Jamo by following the Unicode Hangul composition rules. It will be economical to use the pre-composed syllabic character as it has only two bytes. However, if there is not exactly matching phonetic sound in foreign language, then it will assign a new code using the private use area code number and making a new font.

To understand how Hangul can be used on multi-lingual environment, I will explain only three writing systems; English, Chinese and Japanese. However, I want to emphasize that Hangul can be used on other languages as input method.

Hangul usage for English shorthand input system.

As already explained, English can not use the computer keyboard for shorthand purpose. It has to use the stenographic short hand machine. Stenographic shorthand machine's alphabet layout is very similar to AhnMaTae Keyboard. Only the difference is that, in stenographic shorthand machine, four vowels are located bottom row to be used with two thumbs. Left four fingers use about a dozen leading consonants, and right hand fingers use about the same number of the trailing consonants. On the other hand, AhnMaTae Phonetic Hangul keyboard has ten trailing consonants located bottom row to be used with two thumbs and or with any free fingers. Ten leading consonants are located on the second and third row on the left to be used by four finger of left hand. And ten vowels on the second and third row on right hand side are used by four finger of right hand.

Though the system may look very similar, Hangul shorthand system is far more effective to accurately write any phonetic values in any languages in the world. English is very limited in that sense, and it cannot even use its own language on computer for shorthand purpose without the help of other machine. Chinese and Japanese use English to extract their languages, but you have to remember English is not an efficient system to use as phonetic symbol.

The biggest advantage of Hangul shorthand system is that it does not require another machine. APHKIS is developed to input Hangul directly on the computer keyboard.

Since English and Korean belong to completely different language group, its phonetic values are different. Therefore, modern Hangul Jamo as it is used now with KS keyboard cannot adequately meet its need in English. For instance, in current Korean sound, there is no equivalent sound value of English F, V, TH and Z. But with APHKIS, one can easily make these sounds by combining two different consonants, in the form of ㅍㅇ , ㅂㅇ , ㄷㅅ , and ㅈㅅ . In this case, we have codes in Unicode. There may be some other combined consonants of leading and or trailing consonants for English extraction which does not have code. However, there is no need for vowels as Unicode already has 66 vowels in Hangul Jamo with codes (1161 to 11A2). They are far more than adequate to write all English vocabularies.

To extract English with Hangul, it requires English vocabularies stored as dictionary. When equivalent phonetic Hangul shows up on monitor, and simultaneously space bar is pushed at the end of the last syllabic character, it will search for the right vocabulary and bring to the surface. It will be an easy part on programming because it will match one to one. However, English has a considerable number of vocabularies which has identical sound, such as 'sea' and 'see' (four strokes each with space), which is written in Hangul, '시' (only one stroke). Algorithm has been developed already to choose the right vocabulary in the sentence. Grammar and spelling check is a perfect example of such algorithm.

Here is an example of a sentence in English and its equivalent sound transcript in Hangul

This is the time for all good men to come to the aid of their country. = 68 strokes
딛 잇 더 타임 포 올 굿 맨 투 컴 투 더 에인 읍 테어 컨트리 = 21 strokes

Hangul will be displayed at the bottom of screen just like a TV screen with caption for the hearing impaired persons, so that the typist will have a chance to correct them in case of misspelling or any other errors. These Hangul characters are only for display, and not for printing.

Hangul shorthand for Ideographic Chinese character.

Hanzi (漢字, Japanese call it Kanji, Koreans call it Hanja) is as old as any other writing systems of the world. Its origin is not clear. What is clear is that it is developed as a pictographic writing system and each character has its own meaning. That is the reason why it is called ideographic character. Chinese use it as an official writing system and Japanese use it as a part of official writing system. South Koreans use

them occasionally, particularly among the older generations. North Koreans abandoned it several decades ago and only a very few scholars can read and write it.

The problem of its application on computer is the number of characters. No one knows exact number of it, but some estimate goes as many as 100,000 characters. In Unicode, there are 20,219 of them, in Unicode 3.0 and ISO/IEC 10646. This is called Uni-Han, as Chinese, Japanese and Koreans adopted as their common usage. However, there is proposed addition of more than 40,000 characters in addition to current numbers in the Unicode 3.1.

Another problem with Hanzi is that each character is pronounced differently. Pronunciation of same character by Chinese, Korean and Japanese are as different as English, French and German with same vocabulary. Even among Chinese, there are different pronunciations of same character by different dialects. Therefore, developing of input system by phonetics means development of several localized editions. For instance, Chinese in China as standard Chinese (普通話) is pronounced Zung Kwo Len, 中國人 (중꺁린), in Japanese it is pronounced Joo Kogu Jin (中國人) and in Korean it is pronounced as Jung Gook In, 중국인 (中國人).

As there are so many numbers of characters to be used on the computer, that there are so many input method developed so far. In each country it has developed its own unique systems. In China alone, there are so many different methods of input and different keyboard arrays that it is just impossible to explain them all in this paper. Only two most popular input systems are explained here.

One is called Woobi method (五筆法), which uses about 250 Chinese radicals, assigned half a dozen radicals on each key. It is the fastest way of inputting Chinese characters by keyboard. However, the drawback is that it takes so long time to train a person to operate it. Ordinary people cannot use it. The other problem is that it cannot extract 100 % characters. To choose the right character, one has to move prompt by mouse and then click, or by dragging prompt by arrows and then click to choose the right one, or pushing the numbered key, which is assigned on each character. It is a time consuming process. Microsoft recently developed a new system developed by artificial intelligent research area, such as voice recognition and written character recognition. However, it is better system than the current system of drawing clickers to the right character or choosing by number, but not as fast as APHKIS can process Chinese.

The other popular method is called Pinin method (拼音法), which uses English as means of medium. Most of ordinary people use this method, though it has so many of following problems using this method.

- (1) There are so many characters with the same sound value, sometimes several dozens of them, if not hundreds, that finding the right character and choosing them takes so much time.
- (2) Unless the sound value of Chinese does not match with English alphabet exactly, one has to strike different alphabets until it matches.

It is a really time consuming process. Most of the people prefer to use it, though it takes long time to process, because English has only 26 alphabets. As mentioned in the previous pages, English is not a good medium for processing other languages as each English alphabet has so many sound values.

Hangul is a perfect solution to process Hanzi. 238 Hangul Jamo in Unicode (1100 to 11F9) exactly matches with Chinese pronunciation. There is no need for creating any of combination consonants and vowels, or assigning new code numbers as in the case of English. Also Hangul is a part of its national standard in China, as approximately two million Sino-Koreans in Northern part of China use it. Using its own standard (Hangul) to process its own character (Hanzi) is a just normal process.

Hangul can be displayed as in the case of English process for correction purpose. Hanzi usually has one syllable sound value, and therefore, one Hangul syllabic character matches to one Chinese character. However, to shorten the number of similarly valued sound character, Chinese vocabulary should be stored in dictionary and extracted when the space bar is pushed at the end of last syllable of the

word. Most of vocabularies will show on screen. When there are several candidates, then choosing with the keyboard by certain characteristics of character will shorten input time.

As Hanzi has several characteristics, such as (1) stroke number(s), (2) number of radicals, (3) shape of first stroke, (4) shape of last stroke, (5) four different accents, any one or combinations of it will choose one vocabulary by one stroke. This is the fastest input method of Chinese characters and no other method will match its speed and accuracy. Let me explain these characteristics: (Alphabet assignments are temporary positions and used QWERTY keyboard)

- (1) Stroke number(s): Hanzi has stroke numbers, one to several dozens, and dividing them into five categories: 1-5 assigned on Q, 6-10 assigned on W, 11-15 on E, 16-20 on R and more than 20 on T.
- (2) Number of radical(s): Hanzi has number of radicals, one to several. It would be assigned one on A, two on S, three on D, four on F and five and or more on G.
- (3) Shape of first stroke: Hanzi has only five stroke shapes: Horizontal line on Y, vertical line on U, dot on I, right to left slash on O, and left to right slash on P.
- (4) Shape of last stroke: Same shapes as the first stroke but different keys are assigned for the last stroke: horizontal line on H, vertical line on J, dot on K, right to left slush on L and left to right slush on colon.
- (5) Four tone marks (四聲): Chinese language is a very musical language and has four tone marks. First one on V, second one on B, third one on N, and fourth one on M.

At the first trial, it looks very much confusing, but once the nerve systems are trained, the fingers will move automatically. Also, algorithm will be written to choose the most closely matching characteristic, so that one does not need to strike exactly right key to choose one.

Hangul application on Japanese.

Japanese writing systems itself is a multi-writing system and very interesting to learn how it is processed. It uses Ganji (漢字), Gatagana, Hiragana and Romanji (English alphabet). Sometimes, it writes from left to right, or top to bottom and right to left. Also it has a unique Japanese writing system, called 'ruby' type, a small Hiragana on the top of Chinese characters.

Ganji usage in Japan is very interesting as it has different way of pronouncing it. Originally Ganji was introduced through Korea and later years it came through different parts of China in different period, and therefore, Ganji has different readings by the period of importation and geographical route of its importation. In addition to it, it also read it by ideographic meanings. However, to minimize the confusion, this paper only deals with 'on' (音) readings or the way ordinary people pronounce it.

Gatagana is a simplified Hanzi of similar sounds, written in formal shapes, while Hiragana is another simplified Hanzi written in cursive way. All of them are monosyllabic writing system and it does not have trailing consonants, except nasal sound of 'n'. English is used very frequently in modern Japanese, and writing system is either spelled in English alphabet or

pronunciation is spelled in Katakana. Since Japanese are in lack of trailing consonants, the way it pronounced sounds very strange to foreigners. McDonald is pronounced as 'Mag-do-nar-do'.

Unlike Korean, which belongs to Altaic language group, such as Turkish, Finish and Mongolian, Japanese belongs to pacific islander's language group. This is the reason why it does not have a trailing consonant sound, except nasal sound of 'n'. However, order of sentence is identical with Korean and use of honorific is the same. Use of verb endings and grammatical rules are identical. So was the religious influence of Buddhism and teachings of Confucian ethics.

As Japanese use multi-writing system of Ganji, Hiragana, Gatagana and Romanji, it is extremely difficult and complicated in its application on computer. So, many of them use 'Eigo' (英語, English) as a means of medium of extraction of Japanese language. As already mentioned on English and Chinese, English is not a suitable medium for processing other languages in speedy way, as it can not extract itself as shorthand without the help of other machine. Hangul is a perfect solution to solve the problem of slow process of Japanese language, because it was invented not only for Korean language, but invented for all languages.

Ganji processing will be similar to what already described in Chinese processing method. Only the difference is that Japanese will use different pronunciation on each character and it cannot use the four tones as in the case of Chinese. Some of the glyphs of Japanese Ganji are different than that of same Chinese character.

Gatagana and Hiragana processing will be easy as it has only five vowels: A, I, U, E and O. Others are syllabic characters; combination of consonants and vowels. i.e., ga, gi, gu, ge, go, na, ni, nu, ne, no, etc.

Since Japanese writing systems use ganji, gatagana, hiragana and English, and it does not use trailing consonants, except 'n', any keys on the bottom rows can be used as the selection toggle key. For the convenience's sake, following bottom keys are assigned. Ganji toggle key is V, Gatagana is B, Hiragana is N, and English is V. All of these keys are easy to reach by both thumbs. By pushing these keys will change to a respective writing system, until other key is pushed for another writing system.

To type Japan (Nihon) in Ganji, simultaneously push V and 日本. It will show 日本(Japan) and 二本(two each) on the screen. To choose one right vocabulary, the characteristic of the first character ni(日) should be used. In this case of Characteristics of Ni(日) are: (1) Stroke number is four, and therefore it is Q. (2) Number of radical is one, and therefore it

is A. (3) Shape of first stroke is horizontal, therefore it is Y. (4) The shape of last stroke is horizontal, and therefore it is H. Typing QAYH simultaneously with Ganji choosing toggle (i.e. C) with left hand thumb will choose the correct character.

Benefits of APHKIS and Hangul encoding for M17N.

AhnMaTae Phonetic Hangul Keyboard Input System (APHKIS) was invented for all the language system. It is not only faster than conventional input method, but also will contribute to the world peace. How this input system could contribute for the peace of the world, can be explained by how it can prevent the international criminal activity.

Since Hangul is the best phonetic writing system in the world, it will become a medium for computer input system for all the languages, and the APHKIS is developed for shorthand input system for every language, a brand new area of its usage will be opened. That is the ability of computer to sort out different language by auto detection method.

The process will be following:

1. Almost all international criminal activities are carried out by group of people and their voice and or written messages will be sipped into the computer and the computer will encode this communication media into Hangul codes.
2. The encoded Hangul will be decoded into different languages.
3. Certain vocabulary will alert the operator and whole content will be recorded or printed out.

For instance, some one pronounced ai kil iu. This is encoded into Hangul, .아이, 킬, 유. In Korean, 아이 means child, 킬 does not mean anything, 유 means existence. Since 킬 does not mean anything and this sentence does not have a verb, it cannot be Korean. In Japanese, kil does not have any meaning, as the trailing consonant only use 'n', and it cannot be Japanese. In Chinese, there is no sound value with 'k' in the leading consonant and it cannot be a Chinese. Only the remaining possibility is English. In English, 아이 can be spelled as 'I' or 'eye'. 킬 can be spelled 'kill', 유 can be spelled 'U' or 'you'. In English, algorithms are written to identify which vocabulary fits into where by the grammar rules. In this case, it is not 'eye kill you', nor 'I kill U'. It is definitely 'I kill you', because it is grammatically correct. If the word 'kill' is chosen to raise the red flag, then the operator of this computer will look into it more thoroughly. Any language of descriptive nature of violent action can be flag raising words.

Conventional system of gathering information from the prospective international criminals was depended on the language experts. These experts listen to the conversations of criminals and or intercept their correspondence. This method is very expensive, inaccurate and slow. It is expensive because such bi-lingual expert require long period of time to be trained. It is also inaccurate as human can make some mistake in understanding some vocabularies. It is a slow process as most of the international criminals move fast and use different identifications.

References:

Ahn, Matthew, Super Speedy Input system of Hangul and Globalization. 2001,
International Phonetic Alphabet, London University
King Sejong & Jung Inji: Hoonminjungeum Haerae, 1446 \$1447(?)
Unicode Consortium: Unicode, 3.0, 2000
Han, Taedong, Phonology of King Sejong Era, Yonsei University press, 1998
Lund, Ken, CJKV, 2000
ISO/IEC 10646, 1993, Geneva
KSX 5700, Korean Standard Bureau, 1995
KSX 5601: Korean Standard Bureau. 1991
ASCII:
EUC-CN (GB 1988-89 and GB 2312-80), Chinese Standard:
EUC-JP (JIS X 0201-1997, JIS X 0208-1997, and JIS X 0212-1990) Japanese Standard:

(Rev. Apr. 3, 2002)