

Input Method for Foreign Language Based on the Hunminjeongeum Korean Input Keyboard

Soon-Bum Lim¹, Da-Som Lee² and Eun-Joo Sin^{3*}

^{1,2}Dept. of IT Engineering, Sookmyung Women's University, Cheongpa-ro 47-gil
100, Yongsan-gu, Seoul, 140-742, Korea

³(04310) 511, Myungshin, Sookmyung Women's University, Cheongpa-ro 47-gil
100 (Cheongpa-dong 2(i)-ga), Yongsan-gu, Seoul, Korea

TEL : +82-10-3064-3536, FAX: +82-

¹sblim@sookmyung.ac.kr, ²lee_ds@daum.net, ^{3*}kiki75@gmail.com

Abstract

Hangul, which is the Korean alphabet, features excellent phonetic principles. Hangul can be an efficient alternative for languages that do not have their own alphabet, or which have their own language, but for which it is still difficult because of their inefficient system for use in the digital environment. However, the Hangul alphabet system currently in use is standardized only for the pronunciations of the Korean language. In reality, it is therefore incapable of notating foreign languages with different phonetic systems. In order to solve this problem, in this paper, we analyze the input problems that can occur in the extended Hangul, and we propose a suitable input system.

Keywords: Hangul, Korean alphabet, Input System

1. Introduction

Hunminjeongeum is a writing system with an excellent phonetic representation, and it is capable of notating any language based on the principles of its phonemes. The project of Hangul adoption for the Cia-Cia tribe is a good example that shows its potential for supporting languages that have no alphabet, and to notate such foreign languages as pronounced with a high efficiency. However, the Hangul system that is currently in use is standardized to represent the pronunciations mainly used in Korean phonetics, which makes it impossible to represent the phonetic systems of foreign languages. In order to solve this problem, there is a need for additional graphemes that are based on Hunminjeongeum.

However, these extra graphemes cannot be input with the existing Korean keyboard using the current input system. In most cases, a separate input interface is therefore used, or the expression of additional graphemes incurs additional tedious steps and time delay. Several studies and patent cases have aimed to solve this problem, but to date, there has been no reported solution to inputting double consonants based on a dictionary with the existing keyboard input system.

In this paper, we propose an input system based on Hunminjeongeum for the notation of foreign languages. The proposed system is convenient system as it does not require an additional interface. In addition, it will be useful as it allows inputs with a high efficiency by suggesting input words using a dictionary database.

Received (August 8, 2017), Review Result (October 26, 2017), Accepted (November 9, 2017)

* Corresponding Author

2. Related Works

In recent times, there have been many patent applications and studies that are related to Hangul as there is a growing interest in Hangul for its potential applications. In this paper, we address the input problems for improvement, and we examine the cases for past patents and studies, with a particular focus on the Chinese language and the old Hangul input system.

First, we review the Chinese Word Input Apparatus and Method, in which Chinese words and the notation written in Hangul of the corresponding Chinese pronunciations are paired and saved in a database. When the user inputs Hangul letters, the Apparatus extracts and displays corresponding Chinese characters from the database [1]. In order to solve the problem of Byeongseo² and Yeonseo³, which occurs when typing Hangul letters for the Chinese notation, using this technology, graphemes with such problems were categorized into the 1st consonant, 2nd consonant, 1st vowel, and 2nd vowel. The method automatically combines graphemes either in Yeonseo or Byeongseo, and inputs them if the 2nd consonant is inputted within a given time after the input of the 1st consonant.

This is processed in the “input decision section” and “combination input section.” The input decision section decides whether the 1st and 2nd graphemes are inputted within a given time, and if they are input within the time limit, the combination input section combines them either in Yeonseo or Byeongseo. However, this method is problematic in that it is impossible to input other combinations of graphemes, with the exception of the selected sets of the 1st and 2nd consonants and vowels. Moreover, the efficiency is affected as the user has to wait for a specific time when the input is not diagraphs.

Second, it is an input apparatus for Chinese pronunciation notation, in which consonants and vowels are combined according to the order of the input, and are converted into Chinese phonetic notation signals. When a number of consonants or vowels are consecutive in the input letters, the conversion mechanism included the conversion of each of the consonant and vowel set into either Yeonseo or Byeongseo notation [2]. This method uses the special characters * or # for vertical or horizontal combinations of input letters. As a result, the method generates an additional input of keys for special characters on top of the existing input of Hangul keys, and the problem is therefore that it increases the number of key inputs for the user.

The third technology combines Hangul letters, which correspond to both the consonants and vowels comprising Chinese pronunciations, and the tone markers that mark the tones. It then inputs the pronunciations of the Chinese characters in corresponding Hangul notation [3]. However, the output method of this technology differs from that of the general modern Hangul notation system. Therefore, it requires a separate keyboard that enables the input of Hangul consonants and vowels as well as the tone marking characters in order. In addition, it is not possible to use the existing modern Hangul keyboard as it is with this technology.

Some studies reported that the national standard KSX2016, where 117 more graphemes were added in 2007 to process old Hangul in order to represent 1.6 million syllables, worked outside the principles of Hunminjeongeum, and they proposed a new web input keyboard [4]. This keyboard was devised to receive grapheme information from the web-based Jeongeum keyboard, and save it in Jeongeum codes. These codes list the graphemes

² Byeongseo is a term in Korean grammar, which is a way of arranging consecutive consonant letters horizontally to fit into a Korean character block. For example, ㄱ, ㅋ, ㆁ or ㄱㄱ/ㄱㄱㄱ/ㄱㄱㄱ

³ Yeonseo is a counterpart term of Byeongseo in Korean grammar, which is a way of arranging consecutive consonant letters vertically to fit in a Korean character block. For example, ㄱ + ㅇ = ㄱㅇ

internally on the computer, and outputs the syllables on the screen according to Seongeumbeob⁴.

3. Input System of Additional Hangul Graphemes (Diagrams)

The additional graphemes that were added to enhance the phonetic property of Hangul are the extended graphemes based on Hunminjeongeum. The meaning of the diagraph used in this paper is a consonant cluster consisting of two or more consonants, and not double consonants used in modern Hangul. In the extended Hangul, the need for consonant clusters and vowel clusters was increased, but there is no problem with vowel clusters in the input principle. This is because multiple inputs of vowels do not result in an ambiguity in syllables.

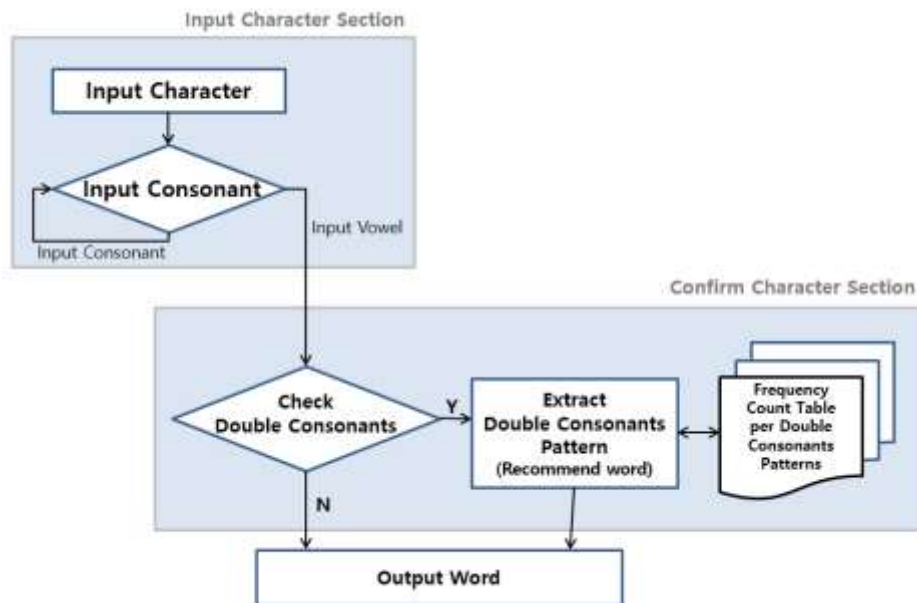


Figure 1. Expanded Hangul Input System

The input system proposed in this paper allows an input of extended graphemes (diagrams). Although it uses the existing keyboard, it is still expected to reduce the extra steps and to significantly reduce the time delay. When inputting diagraphs, a separation between the final consonant of a present syllable and the initial consonant of a following syllable becomes ambiguous, and an input is not possible.

Therefore, the system proposed in this paper maintains a database of foreign language notations and a frequency count table for each of the diagraph patterns. When a user inputs a diagraph, the system extracts and displays the words based on the frequency count table for each diagraph pattern in the database. A user can notate the foreign languages rapidly and precisely without setting an additional input interface for foreign language notation.

In the extended Hangul, the character input section identifies the input phonemes, and it also identifies the number of consonants until a vowel is input. Then, in the character check section, the number of consecutive consonants and the presence of double consonants are checked to decide whether double consonants were input. The system

⁴ Seongeumbeob is a set of rules for the first Hangul orthography provided in the Hunminjeongeum document, and is different from the modern day Korean orthography.

searches the pattern of corresponding double consonants, and inputs the most frequently suggested patterns from the frequency count table per each pair of consonant pattern. The frequency count is established through a pattern analysis of the words in the dictionary database, and the accuracy of the frequency is maintained through continued updates.

4. Cases According to Input Form of Diagraphs

In order to solve the problem arising from inputting additional graphemes that are required for the notation of foreign languages, we examined possible cases according to the input form of diagraphs. The diagraphs are as in Table 1, when all of the diagraphs in old Hangul are restored to enable the representation of foreign languages.

Table 1. Added diagraphs upon Restoration of Old Hangul

Diagraphs upon restoration of old Hangul
ㄴㄴ ㄴㄷ ㄴㄹ ㄴㅁ ㄴㅂ ㄴㅅ ㄴㅆ ㄴㅈ ㄴㅊ ㄴㅋ ㄴㆁ ㄴㅇ
ㅁㅁ ㅁㅂ ㅁㅅ ㅁㅆ ㅁㅈ ㅁㅊ ㅁㅋ ㅁㆁ ㅁㅇ
ㅂㅂ ㅂㅅ ㅂㅆ ㅂㅈ ㅂㅊ ㅂㅋ ㅂㆁ ㅂㅇ

Other than the diagraphs, the consonants, which can be a problem when restoring old Hangul, are bansiot (△) and old ieung (○), which cannot be input on the current keyboard. Bansiot has a softer pronunciation than siot (ㄴ), and was used only in between syllables. However, it used to be put in the initial place of a syllable when notating the Chinese language, and it is used when representing the voiced alveolar fricative [z]. The old ieung (○) has a velar sound, and it used to be a letter representing the velar nasal, and is the modern day final consonant ieung (ㅇ). In the initial era when the Hunminjeongeum was first used, the current ieung was used as a final consonant filler to Chinese characters that did not have a final. The use of these two old Hangul graphemes depends on the phonetic systems of foreign languages. If they are added, we should consider inputting bansiot using [ctrl+ㄴ] and old ieung by [shift+○], while considering the current keyboard system.

In old Hangul, diagraphs of a maximum of three single consonant combinations were used. If the consonants in a former syllable and a latter syllable are combined, a total of six multiple consonants can be input, and the cases for which the ambiguity arises when distinguishing syllable by syllable are as below.

- ① Vowel 1 + diagraph (2) + vowel 2
 - Consonant vowel 1 / diagraph (2) vowel 2
 - Consonant vowel 1 consonant 1 / consonant 2 vowel 2
- ② Vowel 1 + diagraph (3) + vowel 2 (example of input: ㄴ ㄱ ㅅ)
 - Consonant vowel 1 / diagraph (3) vowel 2
 - Consonant vowel 1 consonant / diagraph (2) vowel 2
 - Consonant vowel 1 diagraph (2) / consonant vowel 2
- ③ Vowel 1 + diagraph 4 + vowel 2 (example of input: ㄴ ㄱ ㅅ ㄱ)
 - Consonant vowel 1 / diagraph (4) vowel 2 (no occurrence)
 - Consonant vowel 1 consonant / diagraph (3) vowel 2 (no occurrence)
 - Consonant vowel 1 diagraph 1(2) / diagraph 2 (2) vowel 2
 - Consonant vowel 1 diagraph (3) / consonant 1 vowel 2

5. Decision-making Rules for Determining Input Words

In the above section, we examined types of diagraphs that were added for the notation of foreign languages, and we analyzed ambiguous cases in syllabic distinction for inputs. This ambiguity prevents the input of diagraphs in practice, and results in a process of selecting an additional interface or an input method using a time lag. The input system proposed in this paper increases the input efficiency by suggesting priority diagraph patterns that we expect to be preferred based on the patterns found in a dictionary database, which enables the existing Hangul keyboard for inputs without an additional interface.

5.1. Pattern Frequency-Count Table for Extraction of Input Words

We used a dictionary database containing the phonetics of foreign languages in order to extract input words. Based on the dictionary database, we composed frequency count tables of diagraph patterns. A frequency-count table increases the input efficiency by suggesting priority patterns of the highest frequency when ambiguous diagraph patterns are input.

The below Table 2 is a frequency-count table of patterns “consonant vowel 1 / diagraph (2) vowel 2” and “consonant vowel 1 consonant 1 / consonant 2 vowel 2,” which can occur when a cluster of two consonants is input. Then, Table 3 and Table 4 are frequency-count tables for each pattern that can occur with a cluster of three consonants and a cluster of 4 consonants, and the tables exist respectively according to the form of a consonant cluster.

Table 2. Frequency-count table of Diagraph Patterns with a Cluster of Two Consonants

a = frequency of “consonant vowel 1 / diagraph (2) vowel 2” pattern

b = frequency of “consonant vowel 1 consonant 1 / consonant 2 vowel 2” pattern

		Vowel1 (Vowel of the former syllable)				Total
		ㅏ	ㅑ	ㅓ	...	
Vowel2 (Vowel of the latter syllable)	ㅏ	(a ₁₁ ,b ₁₁)	(a ₂₁ ,b ₂₁)	(a ₃₁ ,b ₃₁)		(∑a _{i1} ,∑b _{i1})
	ㅑ	(a ₁₂ ,b ₁₂)	(a ₂₂ ,b ₂₂)	(a ₃₂ ,b ₃₂)		(∑a _{i2} ,∑b _{i2})
	ㅓ	(a ₁₃ ,b ₁₃)	(a ₂₃ ,b ₂₃)	(a ₃₃ ,b ₃₃)		(∑a _{i3} ,∑b _{i3})
	ㅕ	(a ₁₄ ,b ₁₄)	(a ₂₄ ,b ₂₄)	(a ₃₄ ,b ₃₄)		(∑a _{i4} ,∑b _{i4})
	ㅗ	(a ₁₅ ,b ₁₅)	(a ₂₅ ,b ₂₅)	(a ₃₅ ,b ₃₅)		(∑a _{i5} ,∑b _{i5})
	ㅛ	(a ₁₆ ,b ₁₆)	(a ₂₆ ,b ₂₆)	(a ₃₆ ,b ₃₆)		(∑a _{i6} ,∑b _{i6})
	ㅜ	(a ₁₇ ,b ₁₇)	(a ₂₇ ,b ₂₇)	(a ₃₇ ,b ₃₇)		(∑a _{i7} ,∑b _{i7})
	ㅠ	(a ₁₈ ,b ₁₈)	(a ₂₈ ,b ₂₈)	(a ₃₈ ,b ₃₈)		(∑a _{i8} ,∑b _{i8})
	ㅡ	(a ₁₉ ,b ₁₉)	(a ₂₉ ,b ₂₉)	(a ₃₉ ,b ₃₉)		(∑a _{i9} ,∑b _{i9})
	ㅣ	(a ₁₁₀ ,b ₁₁₀)	(a ₂₁₀ ,b ₂₁₀)	(a ₃₁₀ ,b ₃₁₀)		(∑a _{i10} ,∑b _{i10})
Total		(∑a _{1j} ,∑b _{1j})	(∑a _{2j} ,∑b _{2j})	(∑a _{3j} ,∑b _{3j})		(∑a _{ij} ,∑b _{ij})

Based on Table 3, the probability of “안다” based on this frequency-count table is

$$P(a_{11}) = \frac{a_{11}}{a_{11} + b_{11}} \quad , \quad P(b_{11}) = \frac{b_{11}}{a_{11} + b_{11}}$$

, and that of “아 ㄴ ㅏ” is . Patterns with higher probability are suggested for input.

Table 3. Example Frequency-Count Table of Diagraph Patterns with A Cluster of Two Consonants

		Vowel1 (Vowel of the former syllable)			Total
		ㅏ	...		
Vowel2 (Vowel of the latter syllable)	ㅏ	(a ₁₁ = The frequency of ‘ㅏ ㄴ ㅏ’, b ₁₁ = The frequency of ‘ㄴ ㅏ’)			(∑a ₁₁ , ∑b ₁₁)
	ㅑ	(a ₁₂ = The frequency of ‘ㅏ ㄴ ㅑ’, b ₁₂ = The frequency of ‘ㄴ ㅑ’)			(∑a ₁₂ , ∑b ₁₂)
	ㅓ	(a ₁₃ = The frequency of ‘ㅏ ㄴ ㅓ’, b ₁₃ = The frequency of ‘ㄴ ㅓ’)			(∑a ₁₃ , ∑b ₁₃)
	ㅕ	(a ₁₄ = The frequency of ‘ㅏ ㄴ ㅕ’, b ₁₄ = The frequency of ‘ㄴ ㅕ’)			(∑a ₁₄ , ∑b ₁₄)
	ㅗ	(a ₁₅ , b ₁₅) (a ₂₅ , b ₂₅) (a ₃₅ , b ₃₅)			(∑a ₁₅ , ∑b ₁₅)
	ㅛ	(a ₁₆ , b ₁₆) (a ₂₆ , b ₂₆) (a ₃₆ , b ₃₆)			(∑a ₁₆ , ∑b ₁₆)
	...				
Total		(∑a _{1j} , ∑b _{1j})			(∑a _{ij} , ∑b _{ij})

An example of a cluster of three consecutive consonants is “ㄷ+ㅈ+ㅊ.” In the case where “ㅏ” is the vowel before and after the cluster, the cases are a₁₁ = “ㅏ ㄷㅈㅊ ㅏ,” b₁₁ = “알 ㅈㅊ ㅏ” 그리고, and c₁₁ = “알 ㅏ.” according to the Table 4.

Table 4. Frequency-Count Table of Diagraph Patterns with A Cluster of Three Consonants

a = frequency of “consonant vowel 1 / diagraph (3) vowel” pattern
 b = frequency of “consonant vowel 1 consonant / diagraph (2) vowel 2” pattern
 c = frequency of “consonant vowel 1 diagraph (2) / consonant vowel 2” pattern

		Vowel1 (Vowel of the former syllable)				Total
		ㅏ	ㅑ	ㅓ	...	
Vowel2 (Vowel of the latter syllable)	ㅏ	(a ₁₁ , b ₁₁ , c ₁₁)	(a ₂₁ , b ₂₁ , c ₂₁)	(a ₃₁ , b ₃₁ , c ₃₁)		(∑a ₁₁ , ∑b ₁₁ , ∑c ₁₁)
	ㅑ	(a ₁₂ , b ₁₂ , c ₁₂)	(a ₂₂ , b ₂₂ , c ₂₂)	(a ₃₂ , b ₃₂ , c ₃₂)		(∑a ₁₂ , ∑b ₁₂ , ∑c ₁₂)
	ㅓ	(a ₁₃ , b ₁₃ , c ₁₃)	(a ₂₃ , b ₂₃ , c ₂₃)	(a ₃₃ , b ₃₃ , c ₃₃)		(∑a ₁₃ , ∑b ₁₃ , ∑c ₁₃)
	ㅕ	(a ₁₄ , b ₁₄ , c ₁₄)	(a ₂₄ , b ₂₄ , c ₂₄)	(a ₃₄ , b ₃₄ , c ₃₄)		(∑a ₁₄ , ∑b ₁₄ , ∑c ₁₄)
	ㅗ	(a ₁₅ , b ₁₅ , c ₁₅)	(a ₂₅ , b ₂₅ , c ₂₅)	(a ₃₅ , b ₃₅ , c ₃₅)		(∑a ₁₅ , ∑b ₁₅ , ∑c ₁₅)

	ㄹ	(a ₁₆ ,b ₁₆ ,c ₁₆)	(a ₂₆ ,b ₂₆ ,c ₂₆)	(a ₃₆ ,b ₃₆ ,c ₃₆)		($\sum a_{i6}, \sum b_{i6}, \sum c_{i6}$)
	ㅍ	(a ₁₇ ,b ₁₇ ,c ₁₇)	(a ₂₇ ,b ₂₇ ,c ₂₇)	(a ₃₇ ,b ₃₇ ,c ₃₇)		($\sum a_{i7}, \sum b_{i7}, \sum c_{i7}$)
	ㅑ	(a ₁₈ ,b ₁₈ ,c ₁₈)	(a ₂₈ ,b ₂₈ ,c ₂₈)	(a ₃₈ ,b ₃₈ ,c ₃₈)		($\sum a_{i8}, \sum b_{i8}, \sum c_{i8}$)
	ㅓ	(a ₁₉ ,b ₁₉ ,c ₁₉)	(a ₂₉ ,b ₂₉ ,c ₂₉)	(a ₃₉ ,b ₃₉ ,c ₃₉)		($\sum a_{i9}, \sum b_{i9}, \sum c_{i9}$)
	ㅕ	(a ₁₁₀ ,b ₁₁₀ ,c ₁₁₀)	(a ₂₁₀ ,b ₂₁₀ ,c ₂₁₀)	(a ₃₁₀ ,b ₃₁₀ ,c ₃₁₀)		($\sum a_{i10}, \sum b_{i10}, \sum c_{i10}$)
Total		($\sum a_{ij}, \sum b_{ij}, \sum c_{ij}$)	($\sum a_{2j}, \sum b_{2j}, \sum c_{2j}$)	($\sum a_{3j}, \sum b_{3j}, \sum c_{3j}$)		($\sum a_{ij}, \sum b_{ij}, \sum c_{ij}$)

In Table 5, an example of a cluster of four consecutive consonants is “ㄹ+ㅍ+ㅑ+ㅓ.” Considering only the consonants defined by Hunminjeongeum, the cases of a and b cannot occur. In other clusters of four consecutive consonant combinations, a and b cannot exist, and the syllabic ambiguity occurs with c and d, and the examples are c11 = “알 ㄹㅑㅓ” and d11 = “ㅑ ㅓ 가.”

Table 5. Frequency-Count Table of Diagraph Patterns with A Cluster of Four Consonants

a = frequency of “consonant vowel 1 / diagraph (4) vowel 2” pattern
 b = frequency of “consonant vowel 1 consonant / diagraph (2) vowel 2” pattern
 c = frequency of “consonant vowel 1 diagraph 1 (2) / diagraph 2 (2) vowel 2” pattern
 d = frequency of “consonant vowel 1 diagraph (3) / consonant 1 vowel 2” pattern

		Vowel1 (Vowel of the former syllable)			Total
		ㅏ	ㅑ	...	
Vowel2 (Vowel of the latter syllable)	ㅏ	(a ₁₁ ,b ₁₁ ,c ₁₁ ,d ₁₁)	(a ₂₁ ,b ₂₁ ,c ₂₁ ,d ₂₁)	...	($\sum a_{i1}, \sum b_{i1}, \sum c_{i1}, \sum d_{i1}$)
	ㅑ	(a ₁₂ ,b ₁₂ ,c ₁₂ ,d ₁₂)	(a ₂₂ ,b ₂₂ ,c ₂₂ ,d ₂₂)	...	($\sum a_{i2}, \sum b_{i2}, \sum c_{i2}, \sum d_{i2}$)
	ㅓ	(a ₁₃ ,b ₁₃ ,c ₁₃ ,d ₁₃)	(a ₂₃ ,b ₂₃ ,c ₂₃ ,d ₂₃)	...	($\sum a_{i3}, \sum b_{i3}, \sum c_{i3}, \sum d_{i3}$)
	...				
Total		($\sum a_{ij}, \sum b_{ij}, \sum c_{ij}, \sum d_{ij}$)	($\sum a_{2j}, \sum b_{2j}, \sum c_{2j}, \sum d_{2j}$)		($\sum a_{ij}, \sum b_{ij}, \sum c_{ij}, \sum d_{ij}$)

5.2. Process for Applying the Decision Tree of Input System

The technique involving decision trees is used when proposing patterns for the entered diagraphs in the input system; the process for the application in the input system is as shown in Figure 2 below:

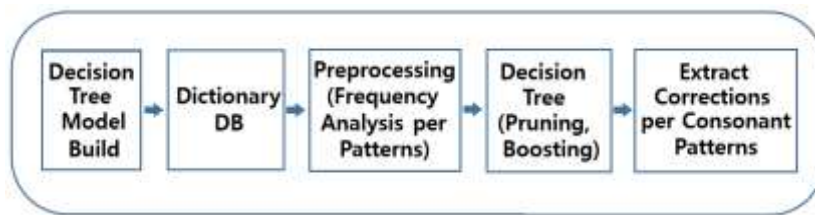


Figure 2. Decision-Tree Process for Input System

We built a decision tree model, which is the basis upon which the input words are proposed, and the model is preprocessed with suitable data using the dictionary database. Preprocessing involves analyzing the frequency of each pattern with syllabic ambiguity, and putting it in table form, and it should be possible to modify the table according to updates made to the dictionary database. Then, a decision is made by the decision tree on the input words, and based on this, proposed words are extracted for each grapheme pattern.

5.3. Decision-making Rules for Determining Input Words

The decision-making rules that are used for deciding input words are based on the pattern frequency-count table of the dictionary, and the decision-making rules for the proposed words requires that we select the patterns with the highest frequency of use. If one wants to input a pattern that is different from the primary suggestion, another set of probable patterns is proposed using the [F1] key. This can also be a time delay in the sense that an additional choice should be made when the primary suggestion is not suitable, but it reduces the delay up to a certain degree compared to the method that distinguishes the former and latter syllables every time, entailing a time delay.

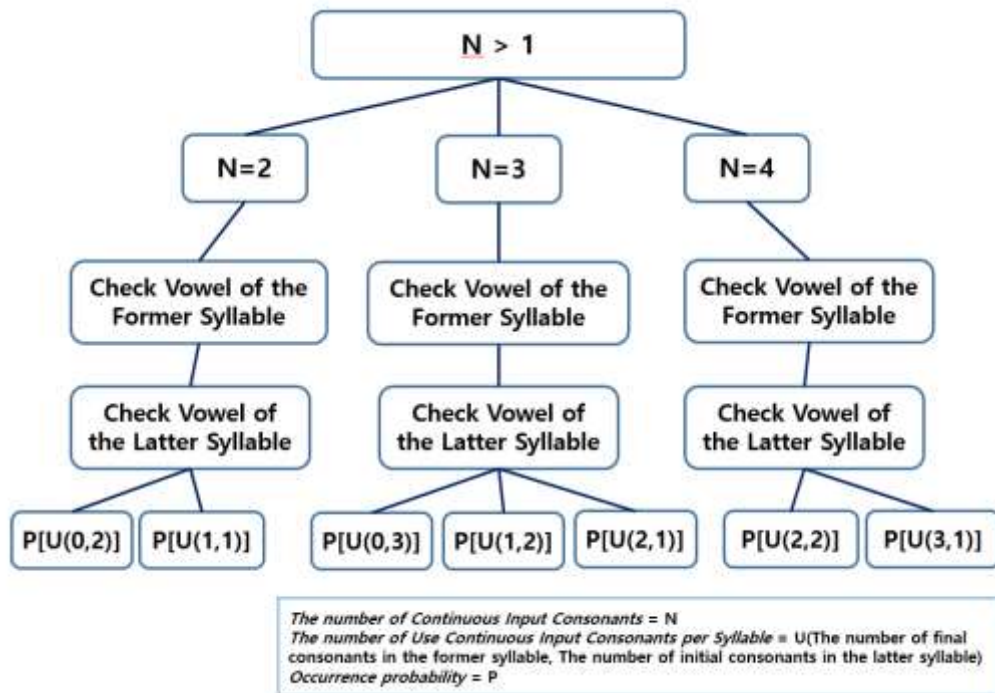


Figure 3. Decision-Making Rules for Determining Input Words

Figure 3 shows the decision-making rules employed to determine the input words. The process of determining diagraphs, which causes ambiguity of syllables, depends on the decision of the number of syllables that are input in advance. At the point when the input of consecutive consonants ends because of a vowel being input, the number of consecutive consonants (N) is decided, as well as whether it is a diagraph with syllabic ambiguity. If it is determined to be a diagraph, the types of patterns that can occur with the corresponding set of consecutive vowels is learned. Then, we obtain the probability of a comparison from the frequency-count table by checking each vowel that is placed before and after in order to suggest words.

6. Conclusion

In this paper, we proposed an input method of extended Hangul for the phonetic notation of Chinese and other foreign languages using the existing Korean keyboard. To express pronunciations that cannot yet be notated by the modern Hangul, we added the extended graphemes, and using suggestions that were based on the dictionary database, we solved the ambiguity that exists in the distinction of syllables, which occurs when inputting diagraphs among extended graphemes.

The expression of Chinese and other foreign languages close to their actual pronunciation in Hangul notation is absolutely necessary for persons who learn these languages, and it is also very important that this issue be addressed in order to reassure the potential of Hunminjeongeum for its excellent phonetic representation capacity.

We evaluated Hangul as a writing system that can be applied efficiently to languages that do not have their own alphabet for writing, and which require improvements for structural inefficiencies in their writing systems. However, there are not enough real examples of the practical development and application of Hangul into the extended Hangul based on the principles of creation of Hunminjeongeum. Based on the solution to the diagraph input of extended graphemes, the input system proposed in this paper is intended to address input problems that can occur in the future when expressing a variety of languages, as well as for further development in order to propose more efficient suggestion methods.

References

- [1] Sinmyeong Systems Ltd, "Chinese Word Input Apparatus and Method", publication number 10-1334154, (2013).
- [2] W. Jeong, "Apparatus for Inputting Korean Character as Chinese Phonetic Transcription and Method Thereof", publication number 10-867668, (2008).
- [3] Korea Advanced Institute of Science and Technology (KAIST), "Chinese Characters Pronunciation Displaying System using Hangul Phonetic Symbols", publication number 10-1243754, (2013).
- [4] J. Byeon and H. Lee, "A Web Input Method for Full Set Syllables Defined by Hunminjeongeum", Journal of Korean Institute of Information Scientists and Engineers: Computing Practice and Letters19, (2013), pp.371~375.
- [5] National Institute of Korean Language: <http://www.korean.go.kr/>.
- [6] G. Kim, "Extended Representation of Korean Characters for Worldwide Alphabets in the IT Age", Journal of Korean Language Informatics, vol. 14, no.1, (2012).
- [7] Wikipedia: <http://ko.wikipedia.org/wiki>.
- [8] G. Kim, H. Kim and H. Park "Transcription of Chinese Language in Hangul and keyboard standardization method", 27th Journal of Academic Symposium on Hangul and Korean Language Information Processing, (2015).
- [9] T. Jeong, "Proposal of syllabic expression method in foreign language transcription in Hangul", 27th Journal of Academic Symposium on Hangul and Korean Language Information Processing, (2015).
- [10] T. Jeong, "Syllable-based Romanization of Jeongeum", Kyung Hee University Press, (2005).
- [11] M. Ahn, "Zheng Yin Chinese Hanzi Input System, Is the First Step for Globalization of ZhengYin", Journal of Korean Language Informatics, vol. 14, no.1, (2012).
- [12] C. Yang, "Hunminjeongeum guide for transcription of world languages", Journal of Korean Language Informatics, vol. 11, no.2, (2009).

Authors



Soon-Bum Lim, he received his BS degree from Seoul National University, Korea, in 1982 and his MS and Ph.D degrees in computer science from KAIST(Korea Advanced Institute of Science and Technology), Korea, in 1983 and 1992, respectively. From 1989 to 1997, he was the engineering director for the font technology and printer division at Human Computer, Inc, and Trigem Computer, Inc. From 1997, he was an assistant professor in the Dept. of Computer Science at Konkuk University in Korea. From 2001, he is currently a professor in the Dept. of IT Engineering at Sookmyung Women's University in Korea. His main research interests are computer graphics, web and mobile multimedia contents, user interface, and electronic publishing such as font, eBook, and XML documents.



Da-Som Lee, she received her BS degree in Multimedia Science at Sookmyung Women's University in Korea. Currently, she is working on master's degree program in the Dept. of IT Engineering. She is interested in Multi-display environment, Web programming, and User interface.



Eun-Joo Sin, she received her Ph.D degrees in Computer Science from Sookmyung Women's University, Korea, in 2010. She is currently a professor in the Dept. of IT Engineering at Sookmyung Women's University. Her main research interests are Usability Evaluation, User interface and Sketch interface.