AUTOMATIC GENERATION OF KOREAN PRONUNCIATION VARIANTS BY MULTISTAGE APPLICATIONS OF PHONOLOGICAL RULES

Jehun Jeon, Sunhwa Cha*, Minhwa Chung*, Jun Park**, and Kyuwoong Hwang***

* Department of Computer Science, Sogang University
1 Shinsoo-Dong, Mapo-Ku, Seoul 121-742, Korea

** Human Interface Department, Electronics and Telecommunications Research Institute
Yusung P.O. Box 106, Taejon 305-600, Korea

E-mail: mchung@ccs.sogang.ac.kr

ABSTRACT

Phonetic transcriptions are often manually encoded in a pronunciation lexicon. This process is time consuming and requires linguistic expertise. Moreover, it is very difficult to maintain consistency. To handle these problems, we present a model that produces Korean pronunciation variants based on morphophonological analysis. By analyzing phonological variations frequently found in spoken Korean, we have derived about 800 phonemic contexts that would trigger the applications of the corresponding phonemic and allophonic rules. In generating pronunciation variants, morphological analysis is preceded to handle variations of phonological words. According to the morphological category, a set of finite state automata tables reflecting phonemic context is looked up to generate pronunciation variants. Our experiments show that the proposed model produces mostly correct pronunciation variants of phonological words consisting of several morphemes.

1. INTRODUCTION

A pronunciation lexicon is an important resource for speech recognition and text-to-speech systems. Many speech recognition systems have used a pronunciation lexicon with multiple pronunciation variants of each word for better recognition accuracy. Text-to-speech systems also require multiple pronunciation variants corresponding to various utterance conditions for generating more precise and natural speech.

Phonetic transcriptions are often manually encoded in a pronunciation lexicon. This process is time consuming and requires linguistic expertise. Moreover, it is very difficult to maintain consistency. To handle these problems, many researches have proposed methods for generating pronunciation variants automatically [2][4][8]. In this paper, we present a model that produces Korean pronunciation variants based on morphophonological analysis.

A Korean syllable consists of two or three phonemes: an initial consonant, a vowel, and an optional final consonant. Since Korean is a syllable-based phonetic language, an orthographic transcription may be considered as a phoneme sequence. However, pronunciation is not always the same as the orthographic transcription. According to an utterance context, phonological variations may occur in a letter-to-pronunciation conversion, resulting in a range of phonetic transcription possibilities.

Phonemic configurations and morphological properties of an orthographic transcription affect phonological variations. For example, an *eojeol*[1] "*kam ki*" may be a noun or a combination of an *eogan*[2] "*kam*" and an *eomi*[3] "*ki*". When "*kam ki*" is a noun, its pronunciation is /K AA M G IY/. When it is a combination of an *eogan* and an *eomi*, its pronunciation is /K AA M KK IY/[4]. That is, the first consonant "*k*" of the second syllable "*ki*" is pronounced as either /G/ or /KK/ depending on the morphological properties. The phonemic configuration also affects pronunciation. The first consonant "*k*" of the first syllable "*kam*" would be pronounced as /G/ if the preceding phoneme is a voiced sound. Therefore, we have 4 possible pronunciations for the word "*kam ki*".

Since Korean is an agglutinative language, thousands of *eojeols* may be generated from a given root word by combining substantial and formal morphemes. Furthermore, in spoken dialogue, two or more *eojeols* may be merged to form an *eonjeol*[5]. Pronunciation lexicon is required to cope with these variations. Since it is impossible to record all combinations of morphemes as lexical entries of pronunciation lexicon, morphological analysis is necessary for generating Korean pronunciation variants for large vocabulary continuous speech recognition systems and text-to-speech systems.

This paper describes a pronunciation generation model that produces a range of possible phonetic transcriptions according to phonemic configurations and morphological properties of an input orthographic transcription. By analyzing phonological

* This work was supported by Grant No. KOSEF 961-0904-025-2 from Korea Science and Engineering Foundation.

[1] *Eojeol* is a spacing unit of Korean orthography and it corresponds to a word or a phrase in English.

[2] *Eogan* is a root of a verb or adjective in Korean.

[3] *Eomi* is a verb-ending or adjective-ending in Korean.

[4] /KK/ is a tensified sound of /K/.

[5] *Eonjeol* is a phonological word or a unit of pause in spoken Korean. More than one *eojeol* are often uttered without pause to form an *eonjeol*.

variations frequently found in spoken Korean, we have derived phonemic context that would trigger the application of the corresponding phonemic and allophonic rules. In generating pronunciation variants, morphological analysis is preceded to handle variations of Korean *eojeols*. According to the morphological category, a set of finite state automata tables reflecting phonemic context is looked up to generate pronunciation variants. In an experiment the pronunciation generator automatically produces mostly correct pronunciation variants of *eojeols* consisting of several morphemes.

2. ANALYSIS OF PHONOLOGICAL VARIATIONS IN KOREAN

Orthographic transcription is converted into phonetic transcription through phonological process. Phonological variations occur first at a phonemic level, and then at an allophonic level. Phonemic variations can be characterized by phonemic rules that describe how a phoneme is inserted, deleted, or substituted by other phoneme. Allophonic variations can be characterized by allophonic rules that describe how a phoneme is realized as various allophones in utterance.

Based on literature survey [1][5][7], we have identified 18 major phonemic rules, as shown in Table 1, that explain phonemic variations frequently found in spoken Korean. From this analysis, we have derived 761 phonemic contexts that would trigger the application of the corresponding phonemic rule. In most cases, a phonemic context is defined as an ordered set of two adjacent consonants: the final consonant of a syllable and the first consonant of the next syllable. Whenever a defined phonemic context exists in an input phoneme configuration, the corresponding phonemic rule is applied. As a result, the input phoneme configuration changes, which may trigger another phonemic rule.

Among the 18 rules shown in Table 1, rules 1 to 11, 17 and 18 are obligatory. These rules must be applied, whenever the corresponding phonemic context exists in an input phoneme sequence. Three obligatory rules have more constraints. Rule 11 is applied only to phonemes across a morpheme boundary in a compound noun or an *eojeol* boundary in an *eonjeol*. Rules 17 and 18 are applied only to phonemes across a morpheme boundary between an *eogan* and an *eomi*. On the contrary, rules 12 to 16 are optional. These rules produce nonstandard pronunciations which actually happen frequently in normal conversation. We need to consider them together with standard pronunciations to improve the accuracy of speech recognition systems.

A set of basic phonemes together with allophones defines phone-like units (PLUs) used for phonetic transcription. There are many allophonic rules in Korean such as devoicing of semivowels, devoicing of vowels, palatalization, uvularization, voicing, consonant weakening, labialization, nasalization, implosion, flapping, and so on. Since some allophones do not have a significant acoustic difference from other allophones, distinguishing all the allophones may just increase the number of PLUs and have negative effect on the accuracy of speech recognition systems. Therefore, unlike phonemic rules, we do not have to reflect all the allophonic rules in generating pronunciation variants. We have considered only 3 allophonic rules that cause significant acoustic difference in the resulting allophones. They are voicing, implosion, and palatalization that have 26, 3, and 23 phonemic contexts, respectively. Based on this analysis, we have defined 43 PLUs for phonetic transcription of Korean. Among these, 17 PLUs are not defined in ARPAbet and thus can be considered as characteristics of Korean pronunciation.

	Rule Name	Number of Phonemic Contexts
1	Neutralization of final consonant	170
2	Simplification of final consonant cluster	244
3	Aspirationalization	22
4	Liaison	41
5	Lateralization	10
6	Nasalization of obstruent	34
7	Nasalization of liquid	19
8	d-palatalization	3
9	Tensification	136
10	Final consonant h-deletion	1
11	n-insertion	30
12	Umraut	5
13	Deletion of final consonant with the same place of articulation as the next consonant's	6
14	Insertion of final consonant with the same place of articulation as the next consonant's	6
15	Conversion into bilabial or velar sound	17
16	Initial consonant h-deletion	5
17	Conversion into semivowel	5
18	Vowel deletion	7

Table 1: Phonemic rules and the number of phonemic contexts that would trigger the application of the corresponding phonemic rule.

3. AUTOMATIC GENERATION OF KOREAN PRONUNCIATION VARIANTS

In this section, we describe how the analysis in Section 2 is reflected in generating Korean pronunciation variants.

As shown in Figure 1, we have separated phonological process into three stages: (1) obligatory phonemic process, (2) optional phonemic process, and (3) allophonic process. By applying phonological rules following this order, pronunciation variants are successfully generated. For example, we get two standard pronunciations /K AA M G IY/ and /K AA M KK IY/ from an input *eojeol* "*kam ki*" according to its morphological properties. In addition, we also get two more frequently happening nonstandard pronunciations by optional phonemic process.

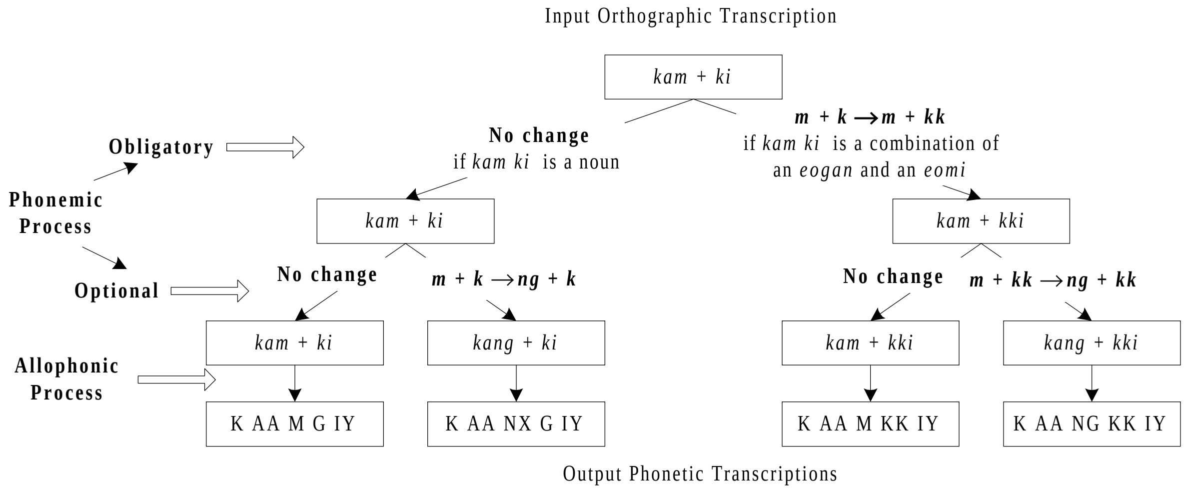

Figure 1: Three levels of phonological variations.

If the input consists of more than two *eojeols* and the preceding phoneme of the first consonant "*k*" of the first syllable "*kam*" is a voiced sound, the corresponding PLU for the consonant "*k*" would change from /K/ to /G/.

In general, a phonemic context triggers only one phonemic rule. In this case, a correct result is always generated and we call it an intrinsic rule ordering. Sometimes, a special phonemic context may trigger more than one phonemic rule. In this case, depending on the order of rule applications, some rules may generate incorrect results. In order not to overgenerate incorrect ones, we need a priority in selecting the right rule to apply. We call this priority an extrinsic rule ordering [5]. The idea of finite state transducers in a standard two-level morphology model [2][3][6] is applied to our system. Obligatory and optional phonemic rules are implemented as a set of finite state automata table that reflects both intrinsic and extrinsic rule orderings. In order to generate pronunciation variants, our model applies phonemic and allophonic rules to an input orthographic transcription by table look-up procedure.

Obligatory phonemic variations depend on both morphological categories and phonemic contexts. Therefore, we have distinguished obligatory phonemic rules that can be applied to phonemes (1) within a morpheme, (2) across a morpheme boundary in a compound noun, (3) across a morpheme boundary in an *eojeol*, or (4) across an *eojeol* boundary in an *eonjeol*.

To generate Korean pronunciation variants, the algorithm goes through the following steps:

1. Perform a morphological analysis of an input orthographic transcription, which may be a morpheme, *eojeol*, *eonjeol*, or sentence.
2. Generate standard pronunciation by applying obligatory phonemic rules according to the morphological category and phonemic context of the input.
3. Add nonstandard pronunciation by applying optional phonemic rules.
4. Generate the output phonetic transcription by applying allophonic rules.

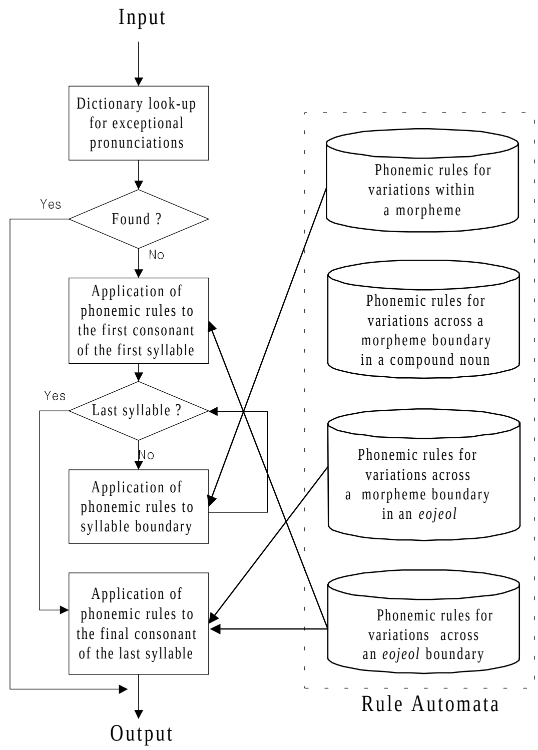

Figure 2: An obligatory phonemic rule processing module for a noun, adverb or adnoun input.

Although the morphological analyzer used in step 1 provides 53 fine-grain morphological categories of Korean, we maintain only 12 coarse-grain morphological categories. They are divided into two groups:

1. simple noun, compound noun, adnoun, adverb, particle (noun ending), predicate (verb and adjective), and predicate-ending (verb-ending and adjective-ending)
2. suffix, interjection, number, special symbol, and foreign word.

Our model provides a processing module for each category in Group 1. As shown in Figure 2, according to the morphological category of an input, obligatory phonemic rules are applied by using the corresponding processing module. Group 2 contains words whose pronunciations are generally considered as exceptional and hard to process with rules. Our model uses an exceptional pronunciation lexicon for processing inputs whose morphological category belongs to Group 2. The lexicon also contains words that cannot be successfully processed due to semantic ambiguities and limitations of our rules.

4. EXPERIMENTAL RESULTS

We used three sets of data for performance evaluation. The first set consists of 123 *eonjeols* showing typical phonological variations. The second set consists of 1,219 *eojeols* taken from dialogues used for hotel and airline reservation. The third set consists of the first 2,881 *eojeols* taken from the 25,000-*eojeol* exceptional pronunciation lexicon used by ETRI's pronunciation generator, which, unlike our system, does not consider morphological properties.

Method	Without using morphological properties [2]	Proposed Model	
Test data set	Data 1	Data 1	Data 2
Input *eojeols*	123	123	1219
Output pronunciations	183	210	1353
Undergenerated pronunciations	40	0	10
Overgenerated pronunciations	13	2	23

Table 2: Result of pronunciation generation from the test input with typical phonological variations.

Since nonstandard pronunciations are also generated, we obtained more output pronunciations than inputs, as shown in Table 2. The previous version in [2] does not use morphological analysis, thereby resulting in lots of errors across morpheme boundaries. By employing morphological analysis and selectively applying phonemic rules according to morphological categories, the proposed model shows an improvement. Errors are classified as undergenerated and overgenerated pronunciations. Most errors still occur at morpheme boundaries due to failures of morphological analysis of spoken dialogues including interjections and mistakes in constructing the phonemic rule automata. In an experiment with Data 3, our system produces correct pronunciations for about 40% of the 2,881 entries used by ETRI's exceptional pronunciation lexicon. We are now conducting an extensive experiment using a test data consisting of all the entries of ETRI's 25,000-*eojeol* exceptional pronunciation lexicon and 100,000 *eojeols* taken from spoken dialogues. From this experiment, we expect an improvement in robustness of phonemic rule automata and morphological analyzer.

5. CONCLUSION

In this paper, we have presented a model that produces Korean pronunciation variants based on morphophonological analysis. Our model has the following contributions to the Korean spoken language processing: (1) Korean phonological variations can be easily and efficiently handled by the proposed set of finite state automata reflecting phonemic context. (2) By employing morphological analysis for generating pronunciation variants, more reliable and complete phonetic transcriptions are provided that can help to improve robustness and accuracy of Korean speech recognition systems.

6. REFERENCES

1. Joochae Bae, *Introduction to Korean Phonology*, Shinkoomoonhwa Press, 1995.
2. Jehun Jeon, Sunhee Wee, and Minhwa Chung, "Generating Pronunciation Dictionary by Analyzing Phonological Variations Frequently Found in Spoken Korean", *Proc. of International Conference on Speech Processing*, 1997.
3. Kimmo Koskenniemi, "A general computatioal model for word form recognition and production", *Proc. of the 8th International Conference on Computational Linguistics*, 1984.
4. Lori Lamel and Gilles Adda, "On Designing Pronunciation Lexicons for Large Vocabulary, Continuous Speech Recognition", *Proc. of International Conference on Spoken Language Processing*, 1996.
5. Hoyoung Lee, *Korean Phonetics*, Teahak Press, 1996.
6. Sungjin Lee, "A Two-level Morphological Analysis of Korean", M.S. Thesis, Department of Computer Science, KAIST, 1992.
7. Jungran Oh, *Modern Korean Phonology: Linear & Nonlinear Phonology*, Hyoungseol Press, 1995.
8. Maria-Barbara Wesenick, "Automatic Generation of German Pronunciation Variants", *Proc. of International Conference on Spoken Language Processing*, 1996.