Statistical Analysis of Korean Pronunciation Variations

Kyong-Nim Lee and Minhwa Chung
Department of Computer Science, Sogang University, Korea
{knlee, mchung}@sogang.ac.kr

ABSTRACT

In this paper, we present a statistical analysis of Korean pronunciation variations using a Grapheme-to-Phoneme (GTP) system. The GPT system generates pronunciation variants by applying rules modeling obligatory and optional phonemic changes and allophonic changes in spoken Korean. Experimental results using a PBS (Phonetically Balanced Sentence) Speech DB of 60,000 sentences show that the most frequently happening obligatory phonemic variations are in the order of *liaison, tensification, aspirationalization*, and *nasalization of obstruent*, and the most frequently happening optional phonemic variations are in the order of *initial consonant /h/-deletion, insertion of final consonant with the same place of articulation as the next consonant's*, and *deletion of final consonant with the same place of articulation as the next consonant's*. These statistics can be used for improving the performance of speech recognition systems.

1. INTRODUCTION

Statistics of pronunciation variations can be used for improving the performance of continuous speech recognition. For example, pronunciation variations across morpheme boundary need to be reflected into the search network for providing all possible pronunciation contexts. In order to reduce the network size and search time, we can selectively reflect some major pronunciation variations based on the statistics. Since it is very difficult to estimate the statistics from a large corpus of auditory transcriptions, we present a statistical analysis of Korean pronunciation variations based on the phonetic transcriptions automatically obtained by a GTP system [5].

Pronunciation variations occur first at phonemic level, and then at allophonic level. The GPT system used for our experiments generates pronunciation variants by applying rules modeling obligatory and optional phonemic changes and allophonic changes in Korean. Experiments using a PBS DB of 60K sentences and 44K Korean morphemes show that the most frequently happening obligatory phonemic variations are in the order of *liaison, tensification, aspirationalization*, and *nasalization of obstruent*, and the most frequently happening optional phonemic variations are in the order of *initial consonant /h/-deletion, insertion of final consonant with the same place of articulation as the next consonant's*, and *deletion of final consonant with the same place of articulation as the next consonant's*.

This paper is organized as follow. Section 2 introduces the Korean phonological rules. Section 3 describes our model for automatic generation of Korean pronunciation variations. In Section 4 and 5, we describe the task material and analysis results, followed by conclusions.

2. KOREAN PHONOLOGICAL RULES

As shown in Table 1, we have identified 20 major phonemic rules based on the literature survey [1][6][7]. These rules explain phonemic variations frequently found in spoken Korean. We have also derived 816 phonemic contexts, which can trigger the application of the corresponding phonemic rule.

Table 1. List of Korean phonemic rules and the number of corresponding phonemic contexts

No	Rule Name	# of contexts
1	Neutralization of final consonant	117
2	Simplification of final consonant cluster	256
3	Aspirationalization	21
4	Liaison	42
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	Deletion of final consonant with the same place of articulation as the next consonant's*	7
13	Insertion of final consonant with the same place of articulation as the next consonant's*	6
14	Conversion into bilabial or velar sound*	17
15	Initial consonant /*h*/-deletion*	5
16	Vowel shortening of '*ui*' with concatenation first consonant	18
17	Vowel shortening of predicative conjugated form	3
18	Vowel shortening of '*ye*'*	17
19	Vowel shortening of '*ui*' except first syllable*	2
20	Vowel lengthening of *eomi* '*eo*'*	1

The first 15 rules are consonant-related and the last 5 rules are vowel-related. These rules are further classified as obligatory and optional. Obligatory rules are applied to every underlying form which meets the corresponding phonemic context, producing a single surface form. Optional rules are marked by * in Table 1. They produce a non-canonical pronunciation, which is not a standard pronunciation, however, happens frequently in spoken

ISBN 1-876346-48-5 © 2003 UAB

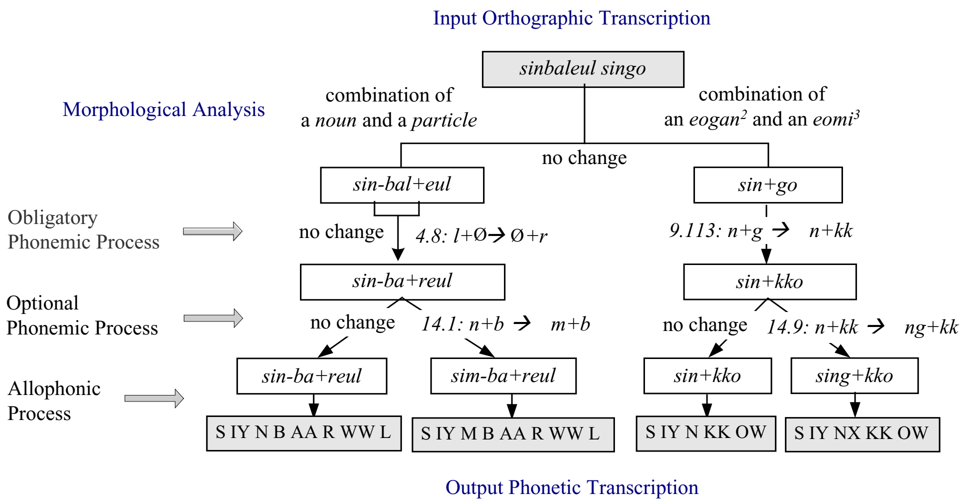

Figure 1. An example of the 3-stage phonological processes.

Korean. For explaining allophonic changes, we also have identified 3 major allophonic rules: *implosion, flapping,* and *voicing* [1].

3. AUTOMATIC GENERATION OF KOREAN PRONUNCIATION VARIATIONS

3.1 System overview

The GTP system produces a phonetic transcription from Korean *eojeols*[1] by using a set of finite state automata. We used the two-level model proposed by Koskenniemi [4], which transforms an underlying lexical representation to its surface representation. Here, the underlying representation corresponds to an orthographic transcription and the surface representation corresponds to a phonetic transcription. Our pronunciation generation algorithm goes through the following steps: (1) morphological analysis of input orthographic transcriptions, (2) applications of 15 obligatory phonemic rules (the rules 1-10, 11, 16, and 17 in Table 1) according to the morphological category to generate canonical pronunciation, (3) applications of 5 optional phonemic rules (the rules 12-15, and 18, 19, and 20 in Table 1) to augment a non-canonical phonemic transcription, and (4) applications of 3 allophonic rules to generate a phonetic transcription.

3.2 Modeling within-morpheme and cross-morpheme pronunciation variations

Since phonemic variations depend on both morphological category and phonemic context, we have augmented phonemic rule automata with the scope of the rule application such as 'E/M/C/I/O/m' flags. 'E' means that the rule can be applied to *eojeol* boundary, 'M' to morpheme boundary, 'C' to morpheme boundary of a compound noun, and 'I' to any location within a morpheme, respectively. 'O' stands for an optional rule and 'm' for a multiply applicable rule, respectively.

We have modeled 20 major phonemic changes using 816 subrules, as shown in Table 2. L3 stands for the last consonant of a syllable, R1 for the first consonant of the following syllable, respectively. As shown in Figure 1, we can get the right phonemic transcription "sin-ba+reul sin-kko" from an input text "sin-bal+eul # sin-go" by applying rules 4.8 and then 14.1 to "sim-bal+eul", and rules 9.133 and then 14.9 to "sing-kko", respectively. Here, '-' stands for a syllable boundary, '+' for a morpheme boundary, and '#' for an eojeol boundary, respectively. The symbol '-' in the scope means that since Rule #14 is an optional rule, it can be applied to any location.

Table 2. Examples of phonemic change rules used in Figure 1

Phonemic Context			Change Code		Rule No	Scope
L3	R1	→	L3	R1		E/M/C/I/O/m
l	Ø	→	Ø	r	4.8	1 1 0 0 0 0
n	g	→	n	kk	9.113	0 1 1 0 0 1
n	b	→	m	b	14.1	-
n	kk	→	ng	kk	14.9	-

4. EXPERIMENTAL RESULTS

4.1 Task material

Several studies have been made on the statistics of orthographic transcriptions in the text and phonetic transcriptions in the pronunciation lexicon. However, it is difficult to know the phonological phenomena in intra-morpheme and inter-morpheme boundaries. In this study, we focused on pronunciation variations in the continuous speech. For this purpose, 60K phonetically balanced sentences with 44K Korean morphemes are collected from newspapers and textbooks. It is designed

[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.

ISBN 1-876346-48-5 © 2003 UAB

and constructed for common use on sentences that cover various phonological environments and balanced triphones.

Table 3. Statistics of the Samsung PBS DB[4]

Count	# Total	# Unique	Average
Sentence	60,000	60,000	9.2 eojeol
Tagged *eojeol*	551,820	170,419	2.1 morpheme
Tagged morpheme	1,160,597	44,303	1.9 syllable
Syllable	2,230,845	167,949	-
Morpheme boundary	608,777	-	-

4.2 Automatic extraction of applying phonemic rules

We have generated phonological variations automatically using the GTP system introduced in Section 3. Phonemic changes occur for about 25.4% of 1.2M morpheme inputs: 12.2% for obligatory and 13.2% for optional phonemic changes. Figure 2 shows the breakdown of rule applications according to the morphological category. We can see that pronunciation variations happen more frequently at contents words than at function words.

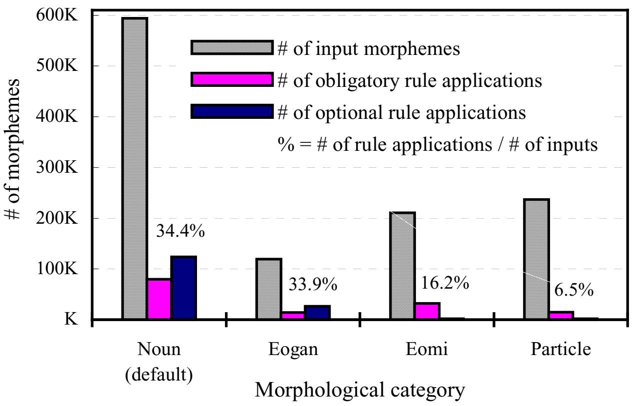

Figure 2. Distributions of the phonemic change rule applications.

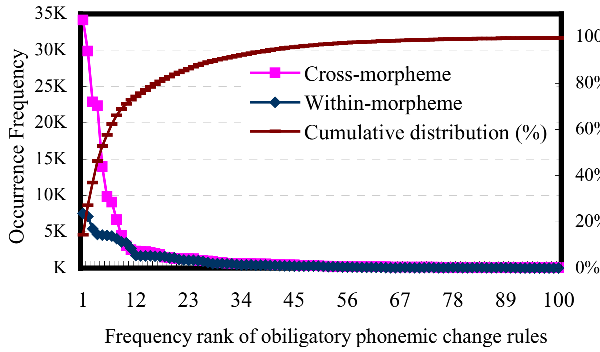

Figure 3. Distributions of obligatory phonemic change rule applications according to the location of rule application.

Figure 3 shows the cumulative distribution of the top 100 frequently applied obligatory phonemic change rules among 757 obligatory rules. In this experiment, 192 phonemic rules are actually applied to the corpus 289,169 times. The 36th rule was applied about 1000 times and the 82nd rule was applied about 100 times.

4.3 Analysis of obligatory phonemic variations

Obligatory phonemic variations are further classified by distinguishing the locations of the rule application as within-morpheme and cross-morpheme. As shown in Figure 4, the most frequently occurring obligatory phonemic variations are in the order of *liaison, tensification, aspirationalization*, and *nasalization of obstruent*.

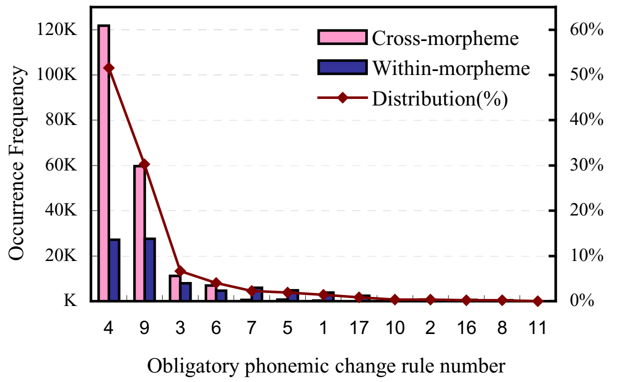

Figure 4. Distributions of obligatory phonemic rule application.

Table 4 shows examples for the five most frequently applied rules. The most frequently applied rule for within-morpheme variations is the rule 4.4 for *liaison*. For example, "un-yeong" (meaning 'management') changes to "u-nyeong" by this rule. The most frequently applied rule for cross-morpheme variations is the rule 9.72 for *tensification.* For example, "iss-da" (meaning 'exist') is realized as /it-tta/ by this rule. In this rule, *neutralization of final consonant* is embedded to reflect repeated applications of *neutralization of final consonant* and *tensification*.

Table 4. Top 5 obligatory phonemic change rules

Rank	Within-morpheme		Cross-morpheme	
	No	Change rule	No	Change rule
1	4.4	'n'+Ø→ Ø+/n/	9.72	'ss'+'d'→/t/+/tt/
2	4.8	'l'+Ø → Ø+/r/	4.4	'n'+Ø→Ø+/n/
3	9.4	'k'+'s'→/k/+/ss/	4.1	'k'+Ø→ Ø+/k/
4	9.1	'k'+'g'→/k/+/kk/	4.8	'l'+Ø→ Ø+/r/
5	5.1	'n'+'l'→/l/+/l/	4.17	'm'+Ø→ Ø+/m/

4.4 Analysis of optional phonemic variations

In order to improve the accuracy of speech recognition, the GTP system also produces alternate pronunciations by applying optional phonemic change rules. These rules are not influenced by morphological properties as contrasted with obligatory rules. However, lexical entries change under the influence of the morpheme boundary. Thus, except for the vowel-related rules 18 to 20, we classified optional rules as within-morpheme and cross-morpheme

[4] HCI Lab of Samsung Advanced Institute of Technology has funded our construction of the PBS DB.

ISBN 1-876346-48-5 © 2003 UAB

categories. As shown in Figure 5, the most frequently occurring optional phonemic variations are in the order of *initial consonant /h/-deletion, insertion of final consonant with the same place of articulation as the next consonant's*, and *deletion of final consonant with the same place of articulation as the next consonant's*.

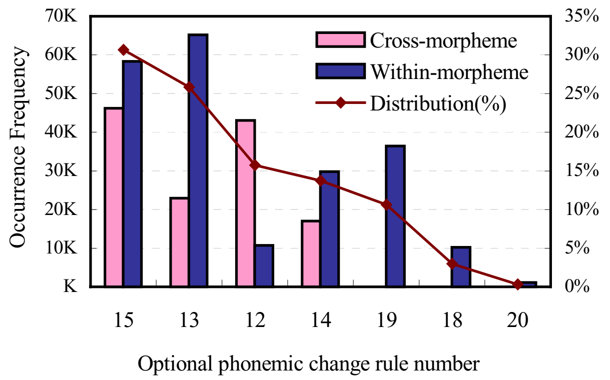

Figure 5. Distributions of optional phonemic rule applications.

The optional rule 13 explains pronunciation variations happening under carefully articulated slow speaking rate, whereas the optional rule 12 for variations under fast speaking rate [1]. We can reflect these characteristics as a parameter of the GTP system.

Table 5 shows the five most frequently applied optional phonemic change rules. The rule 19.2 for *vowel shortening* is the most frequent variation in within-morpheme, which changes 'ui' of adnominal case particle into /e/. The rule 15.1 for *initial consonant /h/-deletion* changes "ui-hak" (meaning 'medical science') to /ui-ak/.

Table 5. Top 5 optional phonemic change rules

Rank	Within-morpheme		Cross-morpheme	
	No	Change rule	No	Change rule
1	19.2	'ui'(particle)→/e/	12.3	't'+'tt'→ Ø+/tt/
2	15.1	Ø+'h'→ Ø+Ø	15.1	Ø+'h'→ Ø+Ø
3	13.4	Ø+'th'→/t/+/th/	15.5	'ng'+'h'→/ng/+Ø
4	15.5	'ng'+'h'→/ng/+Ø	13.2	Ø+'kh'→/k/+/kh/
5	13.7	Ø+'zz'→/t/+/zz/	15.2	'n'+'h'→ Ø+/n/

4.5 Analysis of allophonic variations

In Korean, there are 21 vowels and 46 consonants (19 for initial and 27 for final consonant) symbols. Generally, we use 40 phonemes because some consonants share the same pronunciation depending on the neighboring phonemes. Most speech recognition systems use phoneme-like units (PLUs) as a basic phone set, which are used for phonetic transcription. PLUs shows the influence of allophonic variations. In this work, we have considered three major consonant-related allophonic variations such as *implosion, flapping,* and *voicing*. Figure 6 shows the distribution of allophones for phonemes 'p', 't', 'k', 'z', 'h', and 'l'.

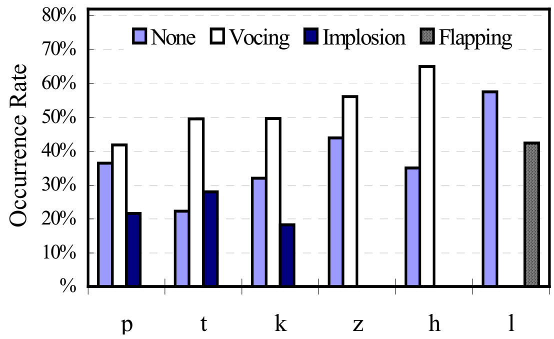

Figure 6. Distributions of allophonic variations for phonemes 'p', 't', 'k', 'z', 'h', and 'l'.

5. CONCLUSIONS

We have described statistical analysis of Korean pronunciation variations using a GTP system. The statistics are useful for improving the performance of speech recognition system and related researches [2][3]. If we use morphemes as lexical units for CSR systems, pronunciation variations across morpheme boundary need to be reflected in the search network in order to prepare for all possible contexts. Based on our statistical analysis, we can selectively apply the phonological variation rules to reduce network size and search time for speech recognition.

ACKNOWLEDGEMENTS

This work has been supported by Ministry of Science and Technology's Brain Neuroinformatics Research Program (Project No. M1-0107-01-0003). We are grateful to HCI Lab of SAIT for providing us the PBS DB.

REFERENCES

[1] Hoyoung Lee, *Korean Phonetics*, Teahak Press, 1996.

[2] H. Strik and C. Cucchiarini, "Modeling Pronunciation Variation for ASR: A Survey of literature," *Speech Communication*, 29(2-4): 225-246, 1999.

[3] Judith M. Kessens, Helmer Strik and Catia Cucchiarini, "Modeling pronunciation variation for ASR: Comparing criteria for rule selection," *Proc. of Workshop on Pronunciation Modeling and Lexicon Adaptation*, pp. 18-23, 2002.

[4] Kimmo Koskenniemi, "A general computational model for word form recognition and production," *Proc. of the 8th International Conference on Computational Linguistics*, 1984.

[5] Kyong-Nim Lee, Jehun Jeon and Minhwa Chung, "Automatic Generation of Pronunciation Variants for Korean Continuous Speech Recognition," *The Journal of the Acoustical Society of Korea*, 20(2):35-43, 2001.

[6] Korean Ministry of Education, *Korean Standard Rule Collections* (notification No. 88-2), 1988.

[7] Sang-Cheol Ahn, *An Introduction to Korean Phonology*, Hanshin Press, 1998.