Irregular Pronunciation Detection for Korean Point-of-Interest Data Using Prosodic Word*

Sunhee Kim(SNU), Je Hun Jeon(SNU), Minsoo Na(SNU), Minhwa Chung(SNU)

<Contents>

<Abstract>

Irregular Pronunciation Detection for Korean Point-of-Interest Data Using Prosodic Word

Sunhee Kim, Je Hun Jeon, Minsoo Na, Minhwa Chung

This paper aims to propose a method of detecting irregular pronunciations for Korean POI data adopting the notion of the Prosodic Word based on the Prosodic Phonology (Selkirk 1984, Nespor and Vogel 1986) and Intonational Phonology (Jun 1996). In order to show the performance of the proposed method, the detection experiment was conducted on the 250,000 POI data. When all the data were trained, 99.99% of the exceptional prosodic words were detected, which shows the stability of the system. The results show that similar ratio of exceptional prosodic words (22.4% on average) were detected on each stage where a certain amount of the training data were added. Being intended to be an example of an interdisciplinary study of linguistics and computer science, this study will, on the one hand, provide an understanding of Korean language from the phonological point of view, and, on the other hand, enable a systematic development of a multiple pronunciation lexicon for Korean TTS or ASR systems of high performance.

* Keywords: Irregular pronunciation, Point-of-interest, Prosodic word, TTS, ASR.

* This research was funded in parts by the Electronics and Telecommunications Research Institute, Daejeon, Korea and by Ministry of Commerce, Industry and Energy's Brain Neuroinformatics Research Program (Project No. 10023913-2005-32).

1. Introduction

Point-of-Interest (POI) data consist of personal names, place names including administrative districts, and names of business and recreational locations, and they are known to enable applications such as Telematics, wireless location-based services (LBS), real estate multiple listing services (MLS), retail site analysis, competitive and market research, sales territory analysis, business and tourism, etc.. As most of POI data are proper nouns and morphologically complex words, they include a quantity of phonologically irregular pronunciations, which are in general referred to as exceptions in the Grapheme-to-Phoneme (GTP) conversion.

Being a major component of a Text-to-Speech (TTS) system and an automatic speech recognition (ASR) system, a GTP conversion maps the orthography to its corresponding sequence of phonemes [1][2][3][4]. A Korean Grapheme-to-Phoneme conversion system is composed of a set of regular rules and an exceptions dictionary [5][6][7][8]. Until the proposals by [9] and [10], the irregular pronunciations were regarded as exceptions, which should be recorded in a simple and random manner and, consequently, it was almost impossible to detect them in a systematic way. [9][10] showed that, contrary to general expectations, the exceptions for the GTP conversion are characterized by certain phonological alternations with their limited contexts. According to [8], Korean ordinary text corpus includes 6.67% of irregular pronunciations from 53,750 sentences.

As most irregular pronunciations are supposed to be found in nouns [9][10], the POI data are presumed to comprise higher percentage of irregular pronunciations than ordinary text corpus. Furthermore, as the POI data include abundant names or proper nouns, the same words showing irregular pronunciations in the sentences might not be found in the POI data. That is, compared to ordinary text corpus, the POI data are supposed to include the words showing irregular pronunciations of different type and in increased number. Based on [9] and [10], [8] proposed how to extract the exceptions from added text corpus. The departure point of this study is to apply the exceptions extraction method proposed by [8] to the Korean POI data, which consist of 250,000 entries.

From the linguistic point of view, the location enriched POI data include a quantity of complex words or compound words, which need to be divided into their constituents preliminarily before deriving their pronunciations along with their exceptions.[1] In order to divide each word into its constituents, we use its prosodic information rather than morphological information based on the Prosodic Phonology

[11][12][13] and Intonational Phonology [14]. Research on the application of the prosodic information to the area of the spoken language processing was mainly presented during the conference Prosody 2001 [15], but most papers presented there were more focused on the acoustic characteristics such as pitch, duration and energy than on the structural aspect of the prosody [16][17].

The goal of this paper is to propose a method of detecting irregular pronunciations for Korean POI data using their prosodic information based on the Prosodic Phonology [11][12][13] and Intonational Phonology [14]. Being intended to be an example of an interdisciplinary study of linguistics and computer science, this study will, on the one hand, provide an understanding of Korean language from the phonological point of view, and, on the other hand, enable a systematic development of a multiple pronunciation lexicon for Korean TTS or ASR systems of high performance.

2. Phonological Rules and Exceptions for Korean GTP Conversion

Like other ordinary GTP conversion systems proposed in previous research [5][6][7][8], the one for the POI data is composed of a set of regular rules and an exceptions dictionary. When most POI data are compounds of nouns, any morphological information is irrelevant, and consequently, only phonological constrained rules are required for the pronunciation generation, such as: (1) Neutralization in Coda, (2) Simplification of consonant clusters, (3) Phonological tensification, (4) Aspiration, (5) Nasalization of obstruent, (6) Lateralization, (7) Double Nasalization, (8) h-Deletion, (9) h-Nasalization, (10) Liaison, (11) Palatalization, (11) Vowel Shortening [9][10].

The exceptions are the words for which alternative pronunciations are possible in the same context and the choice of the pronunciation depends on the words (idiosyncratic). Therefore, it is impossible to make rules for these words, which should be classified as exceptions. According to [9] and [10], the characteristics of the words classified as exceptions are related with certain limited phonological phenomena in Korean. There are three types of the phonetic contexts in which the exceptions appear: (I) when a sonorant is followed by a lenis obstruent; (II) when a nasal consonant is followed by a lateral consonant; (III) when a consonant in the coda position is followed by the vowel which is the nucleus of the following syllable containing a

1) Among 250,000 Korean POI data, very few of them take the form of a phrase or a sentence, and they do not include any irregular pronunciations.

silent consonant 'ㅇ'. And the phonological phenomena in each context are: the Lexical Tensification in (I); the Nasalization of the lateral consonant in (II); and, finally, the /n/-Epenthesis and the Liaison following the Neutralization in the coda or the Simplification of the consonant clusters in (III).

<TABLE 1> Exceptions and their phonetic contexts

<table>
<tr><th></th><th>p</th><th>t</th><th>s</th><th>c</th><th>k</th><th>l</th><th>V</th></tr>
<tr><td>m</td><td colspan="5" rowspan="5">(Ⅰ)</td><td></td><td></td></tr>
<tr><td>n</td><td>(Ⅱ)</td><td></td></tr>
<tr><td>N</td><td></td><td></td></tr>
<tr><td>l</td><td></td><td></td></tr>
<tr><td>V</td><td></td><td></td></tr>
<tr><td>C</td><td></td><td></td><td></td><td></td><td></td><td></td><td>(Ⅲ)</td></tr>
</table>

(C: a consonant of a consonant cluster; V: a vowel or diphthong)
(I) Lexical Tensification
(II) Nasalization of the lateral
(III) /n/-Epenthesis, Neutralization/ Simplification + Liaison

3. Prosodic Word

Given that most POI data are compound nouns, we compared the length of words in the POI data with that of general nouns found in an ordinary dictionary [18]. The length of most words in the POI data was more than twice longer than that of general nouns. The classification of the POI and that of the nouns of a general dictionary [18] shown in <Table 2>.

<TABLE 2> Comparison of the POI data and noun entries of YKD classified by the number of syllables

# of syllable	POI Data		YKD Data	
	# of words	%	# of words	%
1	192	0.08%	1,021	3.32%
2	5,825	2.33%	18,080	58.75%
3	24,375	9.77%	9,775	31.76%
4	84,141	33.72%	1,685	5.48%
5	54,381	21.79%	182	0.59%
6	40,120	16.08%	30	0.10%
7	21,663	8.68%	0	0.00%
8	11,115	4.45%	0	0.00%
Over 9	7,738	3.10%	0	0.00%
Avg. # of Syll.	4.98		2.42	

Another fact observed in the POI data is that most words composed of more that three syllables can be divided into its constituents, which are prosodically and morphologically free. This kind of compounds, formed in the syntax, called phrasal compounds, are distinguished from the lexical compounds, which are lexically fixed combinations, derived in the morphology [14]. Thus the compounds in the POI data are basically phrasal compounds which need to be separated by their constituents. The problem in question is how to separate the POI words into their constituents. To answer this question, we adopt the notion of the Prosodic Word based on the Prosodic Phonology [11][12][13] and Intonational Phonology [14].

Within the Prosodic Phonology [11][12][13], the Prosodic Word is a prosodic unit, which is a constituent of the prosodic structure of the given language. The hierarchical prosodic structure includes Syllable, Foot, Prosodic Word, Phonological Phrase, and Intonational Phrase [11]. However, assuming the Accentual Phrase as the lowest prosodic level based on the tonal pattern, Jun [14] was not sure about the role of the prosodic unit in her model of prosodic hierarchy within the framework of Intonational Phonology. She defines nevertheless the prosodic word as "the minimal sequence of segments which can be produced as one Accentual Phrase (AP)." Adopting this notion of prosodic word as a potential AP, the compounds in the POI data are composed of prosodic words, which are either simple nouns or lexical compounds. A phrasal compound consists of more than two prosodic words, whereas a lexical compound constitutes one prosodic word, which is realized as one AP.

Some examples of phrasal compounds from POI data were given in (1) and those of lexical compounds in (2). In addition, some phrasal compounds which include lexical compounds as constituents are shown in (3). ('-' indicates the boundary of prosodic word.)

(1) a. /hankuk-kyoyuk/ 'Korea-education: Korea Education'
b. /hankuk-kyoyuk-sinmunsa/ 'Korea-education-newspaper company: Korea Education Newspaper Company'
c. /hankuk-kyoyuk-munhwasa/ 'Korea-education- publishing company: Korea Education Publishing Company'
(The number of the words including /Hankuk/ as a constituent in the POI data is 3,517.)

(2) a. /kimpap/ 'laver-rice: Laver Sushi (rice rolled in dried laver)'
b. /tɯŋpul/ 'light-fire: light'

c. /s'alcip/ 'rice-house: rice store'

(3) a. /kimpap-nala/ 'Laver Sushi-nation: Laver Sushi Nation (a name of a store)
b. /kimpap-site/ 'Laver Sushi-Era: Laver Sushi Era (a name of a store)
c. /kimpap-canc^{h}i/ 'Laver Sushi-feast: Feast of Laver Sushi (a name of a store)
(The number of the words including /kimpap/ as a constituent in the POI data is 208.)

When each prosodic word is realized as an AP, no phonological rule applies between the adjacent prosodic words. However, when the two adjacent prosodic words are realized as one AP due to various conditions such as speech rate or speaker variation, any relevant phonological rule for this context applies. That is, the phonological rule for the given context applies optionally depending on the given conditions, resulting in pronunciation variations, which is directly related with the pronunciation modeling problem especially in the area of ASR. If each prosodic word is optionally realized as an AP, the examples in (1) will have following pronunciations as in (4). (The changed segments by the application of phonological rules (Phonological Tensification) were represented in bold.)

(4) a. /hankuk-kyoyuk/ [hankuk kyoyuk]
[hankukk'yoyuk]
b. /hankuk-kyoyuk-sinmunsa/ [hankuk kyoyuk sinmunsa]
[hankuk kyoyuks'inmunsa]
[hankukk'yoyuk sinmunsa]
[hankukk'yoyuks'inmunsa]
c. /hankuk-kyoyuk-munhwasa/ [hankuk kyoyuk munhwasa]
[hankuk kyoyuŋmunhwasa]
[hankukk'yoyuk munhwasa]
[hankukk'yoyuŋmunhwasa]

As was also proposed by [19][20] and [14], the prosodic word serves as a domain of the phonological rules in Korean. Within the framework of Prosodic Phonology, [19] and [20] argued that the Prosodic Word and Phonological Phrase were the domain of phonological and phonetic rules. On the other hand, [14] argued that the AP was the domain of Lenis Stop Voicing, Post Obstruent Tensing and Vowel Shortening, and the IP the domain of post-lexical phonological rules. Based on the phonetic experiments, [14] showed that her analysis was superior in dealing with the variability of the AP and IP domains due to non-linguistic factors such as semantic and pragmatic factors, speech rate and weight of the speech, while the Prosodic Phonologists had to restructure the domains of the Prosodic Word or Phonological Phrase within their syntax-based approach.

As this study considers only the phonological rules, the derivation of the phonetic rules will not be dealt with here. Adopting the notion of Prosodic Word as a potential AP, we assume that the Prosodic Word is the obligatory domain of phonological rules and, across the Prosodic Word, they apply in the optional manner. Thus our assumption not only explains the variability of the domains of the prosodic constituents but also provides the cause and effect of this variability, which is the key of the pronunciation modeling problem in the areas of ASR and TTS.

Unlike phrasal compounds, the lexical compounds are realized as one AP, which will not trigger any pronunciation variations. The lexical compounds in (2) are all examples which show exceptional pronunciations as in (5). (The changed segments caused by the exceptional pronunciations (Lexical Tensification) were represented in bold.)

(5) a. /kimpap/ [kim**p'**ap]
b. /twɯŋpul/ [twɯŋ**p'**ul]
c. /s'alcip/ [s'al**c'**ip]

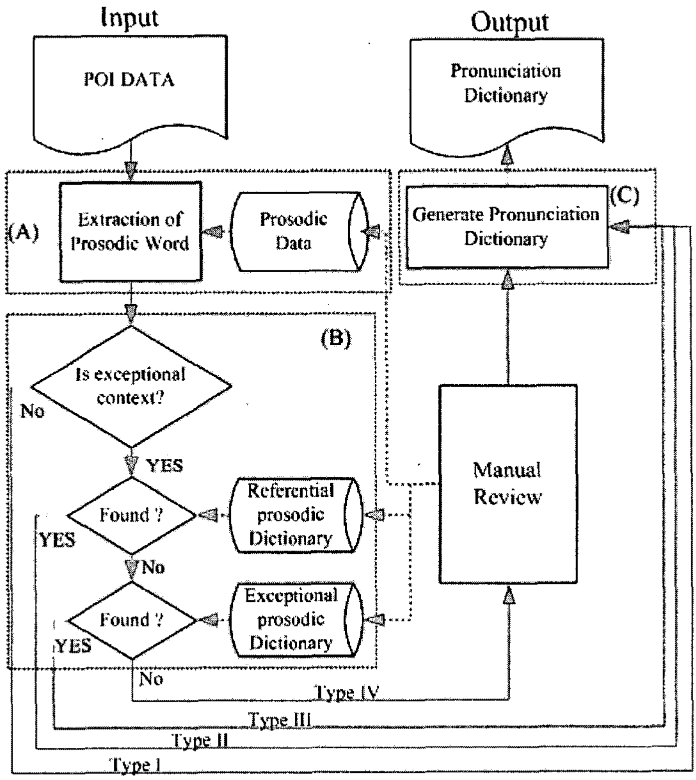
But, when a lexical compound is in itself a constituent of a phrasal compound, the phonological rule relevant for the context can apply optionally as in other phrasal compounds. The pronunciations of (3) are given in (6). (The changed segments caused by the application of phonological rules (Obstruent Nasalization, Phonological Tensification) were represented in bold.)

(6) a. /kimpap-nala/ [kim**p'**ap nala], [kim**p'**a**m**nala]
b. /kimpap-site/ [kim**p'**ap site], [kim**p'**ap**s'**ite]
c. /kimpap-canc^{h}i/ [kim**p'**ap canc^{h}i], [kim**p'**ap**c'**anc^{h}i]

As was shown in (5) and (6), the irregular pronunciations occur within the prosodic word, which means that, in order to extract the words with irregular pronunciations, it is required to extract in advance the prosodic words which show irregular pronunciations.

4. Irregular Pronunciation Detection

In this section, we propose a method to detect the irregular pronunciations for the POI data based on [8] and [21] for the generation of the pronunciation lexicon in ASR systems. This method includes reiterations to increase the data which are newly added. The flow of detection of irregular pronunciation lexicon generation is shown below in <Figure 1>.



<FIGURE 1> Overall flow of irregular pronunciation detection and Dictionary generation

The process was composed of 3 sub processes including a manual review by experts. First sub process (A) divides input POI data into prosodic words, which are stored in the Prosodic Data with their additional information. Each prosodic word in the Prosodic Data in (A) will have its location information which it occupies within the word such as the initial, the middle, and the final of the word as well as its frequency information. On this stage, the proper boundary of each prosodic word will be demarcated based on the decision function using location and frequency information, and syllable length of each prosodic word. When a word is not divided into any prosodic words, it will be recorded as a prosodic word by itself, which will eventually help to reduce error rate. When any predefined Prosodic data do not exist, this process will not take place (This is the case of the experiment without any training data described below in Section 5.2.).

In the second sub process (B) each prosodic word will be decided whether it is exceptional or regular (not exception) with reference to predefined referential and exceptional prosodic words. Four types of output will be derived through this process. Type I indicates regular prosodic words, Type II referential prosodic words, Type III exceptional prosodic words, and Type IV possible exceptional prosodic words. Some examples of the regular prosodic words drawn from (4) are listed in (7). The Type III exceptional prosodic words were already given in (2) and (5), and the Type II referential prosodic words which appear in the same contexts of the exceptional prosodic words in (2) and (5) are given in (8).

(7) a. /hankuk/ 'Korea'
  b. /kyoyuk/ 'education'
  c. /sinmunsa/ 'newspaper company'
  c. /munhwasa/ 'publishing company'

(8) a. /hwaŋkɯmpi/ 'extreme and mean ratio'
  b. /toŋpaŋ/ 'oriental'
  c. /culca/ 'measuring tape'

The Type IV possible exceptional prosodic words are those who appear in the contexts of exceptional prosodic words but are not yet determined as such, and are then subject to the next process of the manual review by experts. On this stage, it will be verified whether it can constitute a prosodic word or it needs to be divided into other prosodic words, and will be determined whether it is a referential (Type II) or exceptional (Type III) prosodic word. After the manual review, prosodic data extracted in sub process (A) and referential and exceptional prosodic dictionary in sub process (B) will be updated. Finally, once all prosodic words of the input POI data are assigned by their correct types, sub process (C) generates pronunciation variants of them. The variants of prosodic words of all four types will be generated by applying predefined phonological rules.

5. Experiments and Results

5.1. Analysis of POI Data

We divided the total data of 250,000 into five subsets ordered by their frequency range as in <Table 3>. Each word in the data was separated manually by the prosodic words which constitute the word. The average number of the syllables of the words was almost 5 whereas the average number of syllables of the prosodic words was about 2. This difference of length indicates the possibility that each word was separated on average by more than two prosodic words.

<TABLE 3> Analysis of POI data ordered by their frequency

Freq. Range	Avg. # of Syllable	Avg. # of Prosodic Word	Prosodic Word			
			Total	Unique	Accum.	New
1~50k	4.4049	1.8176	90,249	21,458	21,458	-
50~100k	4.8697	2.0459	101,609	17,908	32,965	11,507
100~150k	5.0253	2.0994	104,229	20,146	43,751	10,786
150~200k	5.3157	2.2076	109,713	19,475	54,027	10,276
200~250k	5.3221	2.1796	108,338	19,715	64,465	10,438
Total	4.9875	2.0700	514,138	64,465		

(* Accum : the number of accumulated total of prosodic words,
*New : the number of unseen prosodic words in the previous data range.)

We categorized the prosodic words into three types; regular prosodic words, which did not include any contexts of irregular pronunciations; exceptional prosodic words, which show irregular pronunciations in the corresponding contexts <See Table 1>; and referential prosodic words, which did not show irregular pronunciations appearing in the contexts in which irregular pronunciations may occur. <Table 4> shows the statistics of the exceptional prosodic words and referential prosodic words. To extract the exceptional prosodic words, over 60 % of input prosodic words at every frequency range were reviewed to discriminate between exceptional prosodic words and referential prosodic words.

<TABLE 4> The percentage of exceptional and referential words

Frequency Range	Exceptional Prosodic Word				Referential Prosodic Word			
	% of Total	% of Unique	% of Accum.	% of New	% of Total	% of Unique	% of Accum.	% of New
1~50k	6.4%	11.2%	11.2%	-	51.4%	53.5%	53.5%	-
50~100k	5.6%	9.3%	11.0%	10.6%	50.3%	51.0%	53.0%	52.2%
100~150k	5.6%	8.4%	10.8%	10.2%	50.1%	50.7%	52.8%	52.2%
150~200k	5.4%	8.7%	10.6%	9.9%	50.7%	51.6%	52.9%	53.2%
200~250k	6.5%	9.7%	10.7%	10.9%	49.7%	51.7%	53.0%	53.5%

As shown in <Table 4>, the ratio of new prosodic words as well as that of the referential words was almost the same at every frequency range. The reason for this be explained with the results of the analysis of the frequency of each type of prosodic words as in <Table 5>. Over 70% of exceptional and referential prosodic words appeared only one time.

<TABLE 5> The frequency of each categorized prosodic words

	# of Unique	Avg. Freq	Single Appearance	
			#	%
Excp. Prosodic Word	6,871	4.3908	5,190	75.53%
Ref. Prosodic Word	34,152	7.5902	23,931	70.07%
Regular Prosodic Word	23,442	9.5875	14,845	63.33%

5.2. Irregular Pronunciation Detection

In this section, the detection experiment was performed based on the proposed method in <Figure 1> with 5 sets of POI data which are grouped by frequency range. As shown in <Table 6>, the experiment was conducted in 6 tests: the first test without training data, 4 tests with the increasing number of the training data, and the last test with all the data used for the training.

<TABLE 6> The results of detection experiments

Train	Test	# of Excep. in POI	Type III	Type IV		Missing
				Type III	Type II	Type III
None	250k	29,522	0	29,522	190,027	0
1~50k	50k~100k	5,569	3,577 (64.2%)	1,964	10,467	28 (0.50%)
1~100k	100k~150k	5,667	3,919 (69.2%)	1,737	9,120	11 (0.19%)
1~150k	150k~200k	5,783	3,994 (69.1%)	1,761	9,136	28 (0.48%)
1~200k	200k~250k	6,834	4,966 (72.7%)	1,832	8,765	36 (0.53%)
250k	250k	29,522	29,519 (99.99%)	3	16	0

This experiment is focused on detecting Type III and type IV prosodic words and the Missing Type III prosodic words. When the system was not trained, all possible exceptional prosodic words (Type IV) were in fact the exceptional words which were not separated at all by any prosodic words, because no training data were used in the experiment. In consequence, all the words in the exceptional contexts were detected as Type IV, which ranged more than 87% of 250,000 input data. Among these possible exceptional words, 15.14% (29,522) were finally extracted as exceptional words.

Each one of the next 4 tests was conducted with 50,000 words with the increasing number of training data of 50k, 100k, 150k and 200k. The test data with the previous group were used as training data for the experiment with the next group. When the 50,000 of the frequency range of 50k to 100k were tested with the training data of the frequency range of 0 to 50k, 3,577 Type III prosodic words were detected by the sub process (B) proposed in <Figure 1>. Also 12,431 Type IV prosodic words were detected and each of them was assigned by its correct type of prosodic word through the manual review. Finally, 1,964 new exceptional prosodic words of Type III were extracted with 28 missing. Missing Type III indicates the prosodic words of Type III, which are regarded as Type I or II by our proposed system and will eventually drive wrong pronunciations in the sub process (C).

When the system was trained with different number of data, Type IV words with new input were detected with the ratio of 22.4% on average of 50,000 in four tests. This 22.4% of prosodic words of Type IV needs to be judged through the manual review whether they are adequate prosodic words, and then they are subject to be assigned each of them with Type III or Type II. The four tests show that similar ratio of Type III and Type IV prosodic words were detected even with the increase of the training data. These results are in fact related with the characteristics of the POI data shown in <Table 5>, which is, over 70% of prosodic words appear only one time.

When all the data were trained as shown in the results of the final test, 99.99% of the Type III without any missing were detected, which demonstrate the stability of the system. As the Missing Type III is closely related with the performance the pronunciation generation, it is critical to reduce its appearance in the system. Thus the percentage of its appearance ranging 0.43% on average reveals the good performance of the proposed system.

6. Conclusions and Future Work

In this paper, we proposed a method to detect the irregular pronunciations for the POI data, adopting the notion of Prosodic Word as the minimal sequence of segments which can be produced as one Accentual Phrase (AP). The compounds in the POI data are basically phrasal compounds which need to be separated by their constituents. Adopting the notion of prosodic word as the minimal sequence of segments, the compounds in the POI data are composed of prosodic words, which are either simple nouns or lexical compounds. A phrasal compound consists of more than two prosodic words, whereas a lexical compound constitutes one prosodic word, which is realized as one AP. We assumed in this paper that the Prosodic Word was the obligatory domain of phonological rules and, across the Prosodic Word, they applied in the optional manner. Thus our assumption not only explains the variability of the domains of the prosodic constituents but also provides the cause and effect of this variability, which is the key of the pronunciation modeling problem in the areas of ASR and TTS.

In order to show the performance of the proposed method, the detection performance experiment was conducted on the 250,000 POI data. Before conducting the detection experiment, the POI data were analyzed to depict their characteristics, which could affect the experiment. As a large number of the POI data are basically phrasal compounds, each word in the data was separated manually by the prosodic words which constitute the word. When all the data were trained, 99.99% of the Type III without any missing were detected, which eventually shows the stability of the system. Finally, the results show that similar ratio of Type III and Type IV prosodic words were detected even with the increase of the training data. These results are in fact related with the characteristics of the POI data, which is, over 70% of prosodic words appear only one time.

In the future work, the use of prosodic word in the GTP conversion can be extended to the area of ASR such as language modeling as well as the decoding process. We also intend to evaluate the recognition performance based on the lexicon constructed with the proposed method. Furthermore, we will apply the proposed method to the POI data or to the compound words in other languages.

REFERENCES

[1] T. Dutoit, *An Introduction to Text-to-Speech Synthesis,* Kluwer Academic Publishers, Norwell, MA, 1997.

[2] J. Allen, S. Hunnicutt et al., From text to speech: the MITalk system, *Cambridge University Press*, New York, NY, 1987.

[3] H. Strik, C. Cucchiarini, “Modeling pronunciation variation for ASR: a survey of literature”, *Speech Communication, 29(2-4)*: pp.225-246, 1999.

[4] M. Wester, Pronunciation modeling for ASR - knowledge-based and data-derived methods, *Computer Speech and Language*, 17, pp.69-85, 2003.

[5] B. Kim, G. Lee, J.-H. Lee, “Morpheme-based grapheme to phoneme conversion using phonetic patterns and morphophonemic connectivity information”, *ACM Transactions on Asian Language Information Processing*, Vol. 1, No. 1, pp.65-82, 2002.

[6] J. Jeon, S. Cha et al., “Automatic generation of Korean pronunciation variants by multistage applications of phonological rules”, *Proc. Interspeech 1998-ICSLP*, pp. 675-678, 1998.

[7] O.-W. Kwon and J. Park. “Korean large vocabulary continuous speech recognition with morpheme-based recognition units”, *Speech Communication*, Vol. 39 No. 3-4, pp. 287-300, 2003.

[8] S. Kim, J. Ahn et al., “A Korean grapheme-to-phoneme conversion system using selection procedure for exceptions”, *Proc. Interspeech 2004-ICSLP*, pp. 1905-1908, 2004.

[9] S. Kim, “Phonology of exceptions for Korean grapheme-to-phoneme conversion”, *Proc. Interspeech 2004-ICSLP*, pp.1285-1289, 2004.

[10] S. Kim. “Exceptional pronunciations based on text analysis”, *Malsori* Vol. 48. pp. 57-67, 2003.

[11] E. Selkirk, *Phonology and Syntax: The Relation Between Sound and Structure*, Cambridge: MIT Press, 1984.

[12] E. Selkirk, “On derived domains in sentence”, *Phonology 3*, pp. 371-405, 1986.

[13] M. Nespor and I. Vogel, *Prosodic Phonology*, Dordrecht: Foris, 1986.

[14] S.-A. Jun, The Phonetics and Phonology of Korean Prosody: intonational phonology and prosodic structure, Garland Publishing Inc., New York : NY, 1996.

[15] *Prosody in Speech Recognition and Understanding*, ISCA Tutorial and Research Workshop (ITRW), Molly Pitcher Inn, Red Bank, NJ, USA, Oct. 22-24, 2001, *ISCA* Archive, http://www.isca-speech.org/archive/prosody_2001.

[16] K. Hirose, K. Iwano, "Detection of prosodic word boundaries by statistical modeling of mora transitions of fundamental frequency contours and its use for continuous speech recognition", *Proc. Interspeech 2000-ICSLP*, Istanbul Vol. 3, pp. 1763-1766, 2000.

[17] K. Hirose, N. Minematsu et al., "Continuous speech recognition of Japanese using prosodic word boundaries detected by Mora transition modeling of fundamental frequency contours", *Proc. Prosody in Speech Recognition and Understanding, ISCA Tutorial and Research Workshop*, 2001, *ISCA* Archive, http://www.isca-speech.org/archive/prosody_2001.

[18] Center of Language and Information Development (ed.), Yonsei Korean Dictionary, Doosan Dong-Ah Publishing Co. 1998.

[19] O. Kang, *Korean Prosodic Phonology*, Ph. D. dissertation, University of Washington, 1992.

[20] E. Han, "The phonological word in Korean Japanese/Korean," *Linguistics 2, Ed.* P. Clancy, Stanford, CA: SLA/CSLI, pp.117-146, 1993.

[21] J. Jeon, M. Chung, "Automatic generation of domain-dependent pronunciation lexicon with data-driven rules and rule adaptation," *Proc. Interspeech-2005*, pp. 1337-1340, 2005.

접수일자: 2006년 2월 15일
게재결정: 2006년 3월 13일

▶ Sunhee Kim : Corresponding Author
Address: San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, South Korea
Affiliation: Center for Humanities Information, Seoul National University
Telephone: +82-2-880-7735
E-mail: sunhkim@snu.ac.kr

▶ Je Hun Jeon
Address: San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, South Korea
Affiliation: Center for Humanities Information, Seoul National University
Telephone: +82-2-880-9039
E-mail: jhjeon@snu.ac.kr

▶ Minsoo Na
Address: San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, South Korea
Affiliation: Interdisciplinary Program In Cognitive Science, Seoul National University
Telephone: +82-2-880-9039
E-mail: dix39@snu.ac.kr

▶ Minhwa Chung
Address: San 56-1 Sillim-Dong, Kwanak-Gu, Seoul 151-747, South Korea
Affiliation: Department of Linguistics, Seoul National University
Telephone: +82-2-880-9195
E-mail: mchung@snu.ac.kr