A Corpus-Based Analysis of Korean Segments Produced by Japanese Learners

Hyejin Hong¹, Sunhee Kim², Minhwa Chung¹

¹Department of Linguistics, Seoul National University, Seoul, Republic of Korea
²Speech Synthesis Research Lab, NHN Corporation, Seoul, Republic of Korea

souble1@snu.ac.kr, sunhkim@nhn.com, mchung@snu.ac.kr

Abstract

This paper examines variations of Korean segments produced by Japanese learners of Korean. For corpus-based statistical analysis, we have used Korean read speech corpus produced by Japanese learners. Contrastive analysis of the target language and the source language is performed to provide information for interpreting the results of corpus analysis. Segmental variations are analyzed by aligning canonical phonetic transcriptions with auditory phonetic transcriptions of the corpus. The results show that (1) Japanese learners tend to demonstrate substitutions due to differences in the phonemic systems of the two languages; and (2) they are likely to omit a consonant or insert a vowel to deal with the different syllable structures. These results with detailed statistical data are useful for designing a computer-assisted pronunciation training and assessment system for Japanese learners of Korean.

Index Terms: Korean language education, Japanese learners, segmental variations, computer-assisted language learning

1. Introduction

Recent years have seen that computer-assisted language learning (CALL) systems have been developed in line with advances in spoken language technology [1]. Due to a large variability observed in non-native speech, it is difficult to automatically process non-native learners’ speech and to provide corrective feedback [2]. Therefore, analyzing variations in non-native learners’ speech is essential to lay the groundwork for developing a CALL system. In general, a contrastive analysis [3], which compares learners’ native language with their target foreign language, is used to predict the general patterns of variations in non-native speech. In previous studies [4][5], the general patterns of Korean speech produced by Japanese learners are described based on the contrastive analysis. However, there are some drawbacks in the contrastive analysis such as (1) it is difficult to predict all possible variations, (2) it is uncertain that the predicted variations are totally matched with the ones which are found in learners’ real speech, and (3) it is difficult to quantify which variations are more frequently produced by non-native learners. Statistical analysis of a spoken corpus produced by non-native learners is needed to compensate for drawbacks of the contrastive analysis.

In order to develop a computer-assisted pronunciation training and assessment system for learners of Korean as their foreign language, as a preliminary study, we have examined segmental variations of Korean produced by non-native speakers using both contrastive analysis and corpus-based statistical analysis. Considering that interest in learning Korean has been growing with the spread of Korean popular culture [6], CALL systems for teaching Korean can provide a useful learning environment to many foreign learners. In this paper, focusing on Japanese learners of Korean, the salient segmental variations produced by Japanese learners are presented and the factors which influence the variations are examined, based on a systematic analysis of speech data from 34 Japanese learners.

The remaining part of this paper is organized as follows. Section 2 presents a contrastive analysis of Korean and Japanese, which provides a ground for interpreting the results of corpus analysis. Section 3 describes details of materials and methods for analyzing Korean segments produced by Japanese learners. Analysis results are presented in Section 4, which is followed by conclusions in Section 5.

2. Contrastive Analysis

Contrastive analysis assumes that learners’ speech production is mainly influenced by their native language, and difficulties learners encounter can be predicted by comparison of their native language and the target language. A more recent model of foreign language learning, the Speech Learning Model (SLM), claims that the phonetic system of the native language and that of the target language are mutually influenced [7]. Considering that learners’ native language influences the foreign language speech production to a large extent, it is helpful to compare the native language and the target language.

Differences in the Korean and Japanese phonemic systems which are expected to be related to substitutional variations are discussed in this section, and differences in syllable structures which are likely to lead to insertions or deletions of segments are presented as well.

2.1. Phonemic systems

The focus of the analysis is more on phonemes than on phones, since our corpus analysis is performed at a level close to the phonemic level. The phonemic consonantal inventory of Korean and that of Japanese are presented in Table 1 and Table 2, respectively.

Table 1. Korean consonants. Adapted from [8].

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Dental/Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>ū bʰ pʰ pʷ</td>
<td>d̥ tʰ tʷ</td>
<td>ŧ kʰ kʷ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>*</td>
<td>* wʰ wʷ</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>*</td>
<td>* sʰ sʷ</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>η</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-vowel</td>
<td>wʰ</td>
<td>j</td>
<td>wʰ uů</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* classified as both bilabial and velar due to double articulation
Korean has a three-way distinction of stops and affricates: lenis, aspirated and fortis. The lenis stops /b, ɗ, ɡ/ and the lenis affricate /ɗ/ are realized as slightly aspirated, the aspirated stops /pʰ, tʰ, kʰ/ and the affricate /ʨʰ/ as heavily aspirated, and the fortis stops /p, t, k/ and the affricate /ʨ/ as laryngealized and unaspirated [10]. They are all voiceless. In contrast, Japanese shows voicing contrasts in stops and affricates. Korean shows a two-way distinction between the lenis fricative /s/, the fortis fricative /ʃ/, whereas there is a two-way voicing contrast in Japanese fricatives. For nasals, Korean has the velar nasal /ŋ/ which is only permitted as the final consonant, while Japanese has a moraic nasal ‘N’ which has various phonetic realizations including [m, n, ŋ, ɴ]. The Korean velar semi-vowel /ɰ/ is a phoneme which does not exist in the Japanese phonemic system.

Regarding vowels, Korean has more phonemes than Japanese: the Korean vowel system has 8 vowels (excluding the two vowels [y], [w] since they are mostly pronounced as diphthongs rather than monophthongs), whereas there are 5 vowels in the Japanese vowel system.

<table>
<thead>
<tr>
<th>Stop</th>
<th>Bilabial</th>
<th>Dental/Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alveolar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ɸ</td>
<td>s</td>
<td>ɕ</td>
<td>d</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td>N*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semi-vowel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* moraic nasal

Table 2. Japanese consonants. Based on [9].

Table 3. Korean vowels. Adapted from [8].

<table>
<thead>
<tr>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>Close-mid</td>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>Open-mid</td>
<td>ə</td>
<td>A</td>
</tr>
</tbody>
</table>

Table 4. Japanese vowels. Adapted from [9].

Although there are predicted detailed phonetic differences, the Korean vowel system has counterparts for all the Japanese vowels. However, the Japanese vowel system lacks the three vowels /i/, /ɯ/, /u/, which exist in the Korean vowel system.

The comparative study shows that the two languages differ in both consonants and vowels. These differences in the phonemic systems can lead to the prediction that Japanese learners show variations when the target phoneme is missing in Japanese or shows different distribution.

2.2. Syllable structures

A syllable in Korean is composed of an optional consonant, an optional semi-vowel and a vowel followed by an optional final consonant: (C)(j/w/ɰ)V(C) [8]. Japanese allows a syllable composed of an optional onset such as a consonant or a consonant and a semi-vowel /ʝ/, and a nucleus such as a vowel or a vowel followed by the moraic nasal ‘N’ or the first half of a geminate consonant ‘Q’: (C)j(V/N/Q) [11]. Basically, the canonical Japanese syllables are open syllables except when either a moraic nasal or a geminate consonant occurs as a coda. In contrast, Korean permits both open syllables and closed syllables. Differences in their syllable structures can lead to the prediction that Japanese learners tend to simplify Korean syllable structure in order to make it similar to the structure permitted in Japanese.

The comparative study shows that Korean and Japanese significantly differ in both phonemic systems and syllable structures, which poses problems in Japanese learners’ Korean segmental production.

3. Method

In order to examine Korean segments produced by Japanese learners, we adopt a corpus-based statistical analysis using a Korean read speech corpus produced by Japanese learners.

3.1. Speech material

The speech material for Korean speech produced by Japanese learners is taken from the Korean read speech corpus uttered by 100 learners with various native languages such as Japanese, Chinese, English, Russian, Turkish and Vietnamese.

The data for the analysis consists of read speech produced by a total of 34 Japanese adult learners of Korean, aged from 22 to 52. Their proficiency in Korean language is distributed from novice to advanced levels. About 200 sentences are produced by each speaker. The speech data contains 6,877 sentences from textbooks for teaching Korean as a foreign language. Each sentence has 6.1 words on average. A total of 7,331 word types appear in the speech corpus.

3.2. Transcriptions

For the Korean speech corpus uttered by Japanese learners, at first, orthographic transcriptions at word level are manually created by 4 native transcribers who major in Korean linguistics and literature. The corresponding canonical phonetic transcriptions are automatically generated by using a Grapheme-to-Phoneme converter [12] and orthographic word transcriptions. Using canonical phonetic transcriptions as reference, each utterance is manually transcribed to get auditory phonetic transcriptions. Both canonical and auditory phonetic transcriptions include [ɾ] as an allophone of /ɾ/ and five allophones of syllable-final consonants [p, t, k, m, n] besides the Korean phonemic segments presented in Table 1 and Table 3. According to the canonical transcriptions, a total of 275,536 target segments (152,991 consonants and 122,545 vowels) are obtained.

3.3. Analysis of segmental productions

In order to generate variation matrices which can tabularize the relationship between the target segments and their actual realizations in the Japanese learners’ speech production, alignment of the auditory phonetic transcriptions with the
canonical phonetic transcriptions is performed. Based on a
dynamic programming algorithm, weighted distance measures
regarding the phonetic features are computed as in [13] are
introduced to acquire more accurate and consistent alignment.
Vowels are not aligned with consonants. Results of alignment are
manually checked when the alignments are suspicious.

Variation matrices are generated based on the alignment. In
each variation matrix, target segments and their realizations in the
learners’ speech are provided. A ‘variation’ occurs when a target
segment and its realized segment are not identical. For the analysis
of segmental variations which are expected to be salient patterns
of Japanese learners, only the target segments for which the
learners show below-average segment correctness are selected. In
this paper, among all the variations found in the analysis, the
variations which reach more than 2% are considered as salient
variations produced by Japanese learners. More discussions will
be given in the next section for only these variations.

4. Results and Discussion

4.1. Results

The Korean segments which do not reach the average segment
correctness of 96.13%, and their salient variations produced by
Japanese learners are listed in Table 5. The resultant variations
include 30 consonantal variations. Since no salient vocalic
variations occurred, only consonantal variations are discussed in
this paper.

4.2. Discussion

Most of the variations are the ones reported as the major
pronunciation problems of Japanese learners in previous research
[4][5]. However, variations such as deletion of /h/ and substitution
of /w/ for /uw/ are not captured in the previous knowledge-based
studies; they are newly introduced in this paper.

A closer look at the variations reveals that the target segments
are the ones which do not exist in the Japanese phonemic system
or the variations are the results of the learners’ strategies to deal
with unacceptable syllable structures in Japanese. Among the 30
variations, 22 variations are related to differences in phonemic
systems. These variations will be discussed in Section 4.2.1. There
are 8 variations which are due to differences in syllable structures.
As expected, these variations include the cases where the target
segment is deleted or an epenthetic vowel is inserted. The variations
due to differences in syllable structures will be discussed in Section 4.2.2.

4.2.1. Variations due to differences in phonemic systems

Japanese learners tend to realize the aspirated stops /pʰ, tʰ, kʰ/
and the aspirated affricate /tʃʰ/ as their fortis counterparts /p˭, t˭, k˭/
and /ts˭/, respectively. Substitutions of the lenis segments /b, d, g, v/
for the aspirated segments follow. In the case of the Korean
fortis segments /p, t, k, s/ of Japanese learners tend to produce
them more often as the lenis segments /b, d, g, v/ than as the
aspirated segments /pʰ, tʰ, kʰ, sʰ/. For the lenis segments, only two
segments – /b, v/ are selected as the ones to show the salient
variations. For the bilabial lenis stop /b/, Japanese learners
produce it more often as the aspirated stop /pʰ/ than as the fortis
stop /p/. Producing the fortis affricate /tʃ/ for the lenis affricate
/tʃʰ/ is selected as the salient variation, while the aspirated affricate
/tʃʰ/ is less likely to be produced as a substitution by Japanese
learners. For the alveolar fricatives which have the lenis and fortis
contrast, producing the alveolar lenis fricative /s/ as the alveolar
fortis fricative /sʰ/ is found to be salient, however, vice versa is not
selected as a salient variation. The variations related to the three-way
contrast in the Korean stops, affricates and fricatives are
illustrated in Figure 1.

Two variations are related to the Korean final nasals. The Korean
syllable-final velar nasal [n] is confused by Japanese learners as [n]. Japanese learners’ difficulty in distinguishing [n] from [n] is in line with the study on
Japanese learners’ perception of English final nasals, which
reports that Japanese listeners have difficulties in distinguishing
[n] from [n] [14]. In our results, the Korean syllable-final alveolar
nasal [n] is confused by Japanese learners as [n] as well.

As expected, Japanese learners produce the target segments
that do not exist in the phonemic system of their native language
as incorrect ones.

Table 5. Korean segmental productions by Japanese
learners of Korean.

<table>
<thead>
<tr>
<th>Target segment</th>
<th>Freq.</th>
<th>%Corr</th>
<th>Realized Segment (%)</th>
<th>Error type</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>4210</td>
<td>95.13</td>
<td>pʰ 2.38</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>pʰ</td>
<td>1053</td>
<td>80.05</td>
<td>pʰ 11.02</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>pʼ</td>
<td>659</td>
<td>83.46</td>
<td>b 11.23</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>tʼ</td>
<td>755</td>
<td>75.63</td>
<td>b 11.23</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>tʰ</td>
<td>900</td>
<td>71.89</td>
<td>tʰ 22.67</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>tʼ</td>
<td>2184</td>
<td>89.74</td>
<td>dʰ 2.11</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>kʼ</td>
<td>1687</td>
<td>87.91</td>
<td>tʰ 2.47</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>kʰ</td>
<td>1282</td>
<td>73.24</td>
<td>kʰ 20.75</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>kʼ</td>
<td>3169</td>
<td>88.51</td>
<td>kʰ 2.90</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>kʼ</td>
<td>7671</td>
<td>95.80</td>
<td>wʰ 2.45</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>wʰ</td>
<td>2524</td>
<td>87.80</td>
<td>wʰ 2.45</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>wʼ</td>
<td>1076</td>
<td>91.36</td>
<td>wʼ 6.51</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>s</td>
<td>10022</td>
<td>93.33</td>
<td>sʼ 6.50</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>h</td>
<td>6434</td>
<td>93.43</td>
<td>- 6.31</td>
<td>Syllable structure</td>
</tr>
<tr>
<td>mʼ</td>
<td>6891</td>
<td>87.67</td>
<td>mV 9.48</td>
<td>Syllable structure</td>
</tr>
<tr>
<td>nʼ</td>
<td>13351</td>
<td>88.90</td>
<td>nʼ 7.59</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>nʼ</td>
<td>5630</td>
<td>77.25</td>
<td>nʼ 15.67</td>
<td>Phonemic system</td>
</tr>
<tr>
<td>l</td>
<td>8240</td>
<td>91.80</td>
<td>rV 5.44</td>
<td>Syllable structure</td>
</tr>
<tr>
<td>w</td>
<td>2937</td>
<td>95.34</td>
<td>- 4.66</td>
<td>Syllable structure</td>
</tr>
<tr>
<td>uʼ</td>
<td>84</td>
<td>65.48</td>
<td>- 19.05</td>
<td>Syllable structure</td>
</tr>
</tbody>
</table>

learners of Korean. For the alveolar fricatives which have the lenis and fortis
contrast, producing the alveolar lenis fricative /s/ as the alveolar
fortis fricative /sʰ/ is found to be salient, however, vice versa is not
selected as a salient variation. The variations related to the three-way
contrast in the Korean stops, affricates and fricatives are
illustrated in Figure 1.

Two variations are related to the Korean final nasals. The Korean
syllable-final velar nasal [n] is confused by Japanese learners as [n]. Japanese learners’ difficulty in distinguishing [n] from [n] is in line with the study on
Japanese learners’ perception of English final nasals, which
reports that Japanese listeners have difficulties in distinguishing
[n] from [n] [14]. In our results, the Korean syllable-final alveolar
nasal [n] is confused by Japanese learners as [n] as well.

As expected, Japanese learners produce the target segments
that do not exist in the phonemic system of their native language
as incorrect ones.
4.2.2. Variations due to differences in syllable structures

The results of the analysis on Korean segments in Japanese learners’ speech reveal that there are variations which are caused by differences in syllable structures. The prediction that Japanese learners encounter difficulties when Korean syllables have a final consonant is plausible because Japanese does not allow a coda except for the moraic nasal and the geminate consonant as presented in Section 2.2.

Japanese learners use two different strategies to deal with differences in syllable structures: either consonantal deletion or epenthetic vowel insertion. There are 1,839 occurrences of consonantal deletions (0.67%) and 1,196 occurrences of vocalic insertions (0.43%). Japanese learners omit the final Korean phonemic consonants /t/ and /k/ to avoid a syllable structure which is not permitted in Japanese. Japanese learners insert a vowel after the final consonant as well. This vowel insertion leads to re-syllabification making a new syllable as the case of the final consonant of /m/. For /l/, both consonantal deletions and vowel insertions are found, and Japanese learners tend to insert a vowel after /l/ more often than delete /l/.

For the deletion of /h/, 93.35% of /h/ is deleted when the preceding segment is a syllable-final consonant, especially a sonorant such as /m/, /n/ and /l/. This means that Japanese learners fail to produce /h/ followed by its preceding final consonant; instead, they omit the syllabic-initial consonant /h/ and then make a new syllable with the preceding syllable-final consonant as its onset. Note that /h/ is voiced when it is between two sonorants in Korean like this case; however, totally deleting /h/ between two sonorants is not acceptable as a standard pronunciation [8].

Korean permits a sequence of a consonant and a semi-vowel before a vowel; however, this is very limited in Japanese. Japanese learners omit /w/ or /u/ especially when there is a consonant which precedes them. These variations are the results of Japanese learners’ strategy to avoid a sequence of two consonants.

As expected, Japanese learners employ consonantal deletion or vowel insertion strategies to avoid syllable structure which is not allowed in their native language.

5. Conclusions

This paper examines Japanese learners’ Korean segmental variations by quantifying the patterns of learners’ variations occurring in the learners’ speech corpus. Based on contrastive analysis of differences in Korean and Japanese in terms of phonemic systems and syllable structures, the results of the corpus-based analysis on Korean segments produced by Japanese learners show that their segmental variations are related to differences in both phonemic systems and syllable structures. Firstly, Japanese learners are likely to show substitutional variations for Korean segments. Secondly, Japanese learners tend to omit a consonant or insert a vowel to meet the structural constraints of the native language’s syllable. The results confirm that learners’ segmental production in a foreign language is affected to a large extent by their native language as confirmed in a large amount of literature on foreign language acquisition.

The results of the analysis can provide background information which can be used when a computer-assisted Korean pronunciation training and assessment system is designed. In our future work, analysis of more detailed phonetic cues and effects of phonetic contexts will be investigated based on the preliminary results of this paper. In addition, analysis of Korean segments produced by learners with other languages will be examined.

6. Acknowledgements

This work was supported by the Industrial Strategic Technology Development Program, 10035252, development of dialog-based spontaneous speech interface technology on mobile platform, funded by the Ministry of Knowledge Economy of Rep. of Korea. The authors wish to thank the anonymous reviewers for their valuable comments and suggestions on an earlier version of this paper.

7. References