Acoustic Analysis of Korean L2 Learners in the Acquisition of English and Japanese Stop Voicing Contrasts

Katsumasa Shimizu
Faculty of Foreign Studies, Nagoya Gakuin University, Nagoya, Japan
kshimizu@ngu.ac.jp

Abstract

The present study examines phonetic characteristics of Korean L2 learners in their acquisition of English and Japanese stop voicing contrasts. Voice onset time (VOT) was measured on their Korean, English and Japanese utterances, and it was found that VOT is functional in distinguishing a three-way contrast of Korean stops, and that Korean L2 learners use their tense, unaspirated stops for voiced stops in English and Japanese, but it was not clearly shown how they use other types of categories for voiceless stops in English and Japanese. Phonological features /VOICE, ASPIRATED, TENSE/ were proposed to specify the laryngeal characteristics of Korean, English and Japanese.

Index terms: Korean L2 learners, VOT, stop voicing, English, Japanese

1. Introduction

This paper examines phonetic characteristics which Korean L2 learners exhibit in the process of acquiring stop voicing contrasts in English and Japanese. Korean is known to have a three-way contrast in its stop consonants, and they differ in manner and point of articulation. They are usually classified as voiceless, tense, unaspirated stops (Type 1 stops), voiceless lax, weakly aspirated stops (Type 2 stops), and voiceless, strongly aspirated stops (Type 3 stops). Some phoneticians simply call them fortis, lax, and aspirated stops, respectively, while others call them strong, weak and aspirated. The disagreements on the terms to apply to the stops reflect some difficulty in characterizing their phonetic features on purely phonetic grounds. There are no voiced stops in their phonemic inventory in Korean, except in a few cases [1].

The contrast in Korean stops has been of interest among phoneticians, since it involves several modes of vocal fold adjustments, and has been studied from physiological, acoustic and phonological points of view. First, physiologically, Kim (1965) carried out an electromyographic test of bilabial stops and reported greater muscle activity of the lips for Type 1 stops and found that these muscle serve to differentiate the three types of stops, and Type 1 stops are specifically characterized by the activity of the thyroarytenoid muscle [3]. Further, Kagaya (1974) examined the glottal width and the timing of the closure by using a fiberscope, and reported that there is a considerable difference among the three types of stops in glottal width; the smallest glottal width for Type 1 stops, the intermediate for Type 2 stops, the largest for Type 3 stops [4].

Next, acoustically, there have been several studies, and Han and Weitzman (1970) and Shimizu (1996) examined such features as VOT, F0 curve and other acoustic characteristics for distinguishing the three types of stops, and reported that VOT and F0 curve are both functional for characterizing the three-way contrasts [5]. Furthermore, aerodynamically, Dart (1987) measured the air pressure and oral flow of Type 1 and Type 2 stops, and found that Type 1 stops are produced with higher intra-oral pressure before release, but a lower flow after release [7]. Through these studies, it can be stated that the three stops in Korean differ in laryngeal muscle activity, glottal width, and timing of glottal closing relative to articulatory release and these differences in articulation are reflected in such acoustic features as VOT, F0, F1 and intensity. The problem is the lack of examination on how these features are coordinated with each other, and how these features are reflected when Korean L2 learners acquire the two-way contrast of stop consonants in English and Japanese as a second language.

The present study was undertaken as part of a research project to examine voice–voiceless contrast of stop consonants, and was aimed at examining how Korean L2 learners exhibit the phonetic characteristics in learning English and Japanese stop voicings. Japanese and English have two types of voicing contrasts for stops: voiced /b, d, g/ and voiceless /p, t, k/, and are two-category languages in the classification of voicing contrast. The difference between the two categories in Japanese is one of voicing, but the situation in English is a bit different from those in Japanese. Voiced stops in English are fully voiced in the middle position, while they tend to be voiceless unaspirated in the initial one. Voiceless stops in English are aspirated in the syllable-initial position but are unaspirated in other positions. Under such situations, English is generally considered to be one of aspiration-oriented languages. As is generally known, the stops in English occur in the initial, medial and final positions, while those in Japanese do in the initial and medial positions, and they are contextually influenced in phonetic environments. It will be significant to examine how a three-way contrast of stops in Korean is correlated to a two-way contrast in English and Japanese, since it will clarify the laryngeal relationship among these languages, and the study will reveal how a three-way contrast of Korean stops is manifested in a two-way contrast in the acquisition of English and Japanese as a second language.

2. Experimental Procedure

2.1. Subjects

Abstract

The present study examines phonetic characteristics of Korean L2 learners in their acquisition of English and Japanese stop voicing contrasts. Voice onset time (VOT) was measured on their Korean, English and Japanese utterances, and it was found that VOT is functional in distinguishing a three-way contrast of Korean stops, and that Korean L2 learners use their tense, unaspirated stops for voiced stops in English and Japanese, but it was not clearly shown how they use other types of categories for voiceless stops in English and Japanese. Phonological features /VOICE, ASPIRATED, TENSE/ were proposed to specify the laryngeal characteristics of Korean, English and Japanese.

Index terms: Korean L2 learners, VOT, stop voicing, English, Japanese

1. Introduction

This paper examines phonetic characteristics which Korean L2 learners exhibit in the process of acquiring stop voicing contrasts in English and Japanese. Korean is known to have a three-way contrast in its stop consonants, and they differ in manner and point of articulation. They are usually classified as voiceless, tense, unaspirated stops (Type 1 stops), voiceless lax, weakly aspirated stops (Type 2 stops), and voiceless, strongly aspirated stops (Type 3 stops). Some phoneticians simply call them fortis, lax, and aspirated stops, respectively, while others call them strong, weak and aspirated. The disagreements on the terms to apply to the stops reflect some difficulty in characterizing their phonetic features on purely phonetic grounds. There are no voiced stops in their phonemic inventory in Korean, except in a few cases [1].

The contrast in Korean stops has been of interest among phoneticians, since it involves several modes of vocal fold adjustments, and has been studied from physiological, acoustic and phonological points of view. First, physiologically, Kim (1965) carried out an electromyographic test of bilabial stops and reported greater muscle activity of the lips for Type 1 stops and found that these muscle serve to differentiate the three types of stops, and Type 1 stops are specifically characterized by the activity of the thyroarytenoid muscle [3]. Further, Kagaya (1974) examined the glottal width and the timing of the closure by using a fiberscope, and reported that there is a considerable difference among the three types of stops in glottal width; the smallest glottal width for Type 1 stops, the intermediate for Type 2 stops, the largest for Type 3 stops [4].

Next, acoustically, there have been several studies, and Han and Weitzman (1970) and Shimizu (1996) examined such features as VOT, F0 curve and other acoustic characteristics for distinguishing the three types of stops, and reported that VOT and F0 curve are both functional for characterizing the three-way contrasts [5]. Furthermore, aerodynamically, Dart (1987) measured the air pressure and oral flow of Type 1 and Type 2 stops, and found that Type 1 stops are produced with higher intra-oral pressure before release, but a lower flow after release [7]. Through these studies, it can be stated that the three stops in Korean differ in laryngeal muscle activity, glottal width, and timing of glottal closing relative to articulatory release and these differences in articulation are reflected in such acoustic features as VOT, F0, F1 and intensity. The problem is the lack of examination on how these features are coordinated with each other, and how these features are reflected when Korean L2 learners acquire the two-way contrast of stop consonants in English and Japanese as a second language.

The present study was undertaken as part of a research project to examine voice–voiceless contrast of stop consonants, and was aimed at examining how Korean L2 learners exhibit the phonetic characteristics in learning English and Japanese stop voicings. Japanese and English have two types of voicing contrasts for stops: voiced /b, d, g/ and voiceless /p, t, k/, and are two-category languages in the classification of voicing contrast. The difference between the two categories in Japanese is one of voicing, but the situation in English is a bit different from those in Japanese. Voiced stops in English are fully voiced in the middle position, while they tend to be voiceless unaspirated in the initial one. Voiceless stops in English are aspirated in the syllable-initial position but are unaspirated in other positions. Under such situations, English is generally considered to be one of aspiration-oriented languages. As is generally known, the stops in English occur in the initial, medial and final positions, while those in Japanese do in the initial and medial positions, and they are contextually influenced in phonetic environments. It will be significant to examine how a three-way contrast of stops in Korean is correlated to a two-way contrast in English and Japanese, since it will clarify the laryngeal relationship among these languages, and the study will reveal how a three-way contrast of Korean stops is manifested in a two-way contrast in the acquisition of English and Japanese as a second language.

2. Experimental Procedure

2.1. Subjects
The subjects in the present study are four native speakers of Korean, three male and one female, and are speakers of the Seoul dialect. They are undergraduate students of Osaka University, and are studying Japanese as a foreign language. They have studied English for eight years in Korea. Their ages are 22 – 23 years. They have resided in Japan for the period from one to two years, and their Japanese proficiency is in advanced level, while their English proficiency is in the intermediate one.

2.2. Linguistic Materials

Linguistic materials for recording in the present study were shown below. They include two minimal triplets, differing in the initial stops, nine minimal pairs in English, and six minimal pairs in Japanese. Korean materials were written in Korean characters and were read in a carrier sentence. The recording was made in a sound-proof recording room at Osaka University.

Table 1. Mean VOT values of Korean Stops (ms)
(N=24 for Type 1 and 2 stops, N=12 for Type 3 stops) (s.d. in parenthesis)

<table>
<thead>
<tr>
<th>Type 1 stops</th>
<th>Type 2 stops</th>
<th>Type 3 stops</th>
</tr>
</thead>
<tbody>
<tr>
<td>p*</td>
<td>19.0(5.7)</td>
<td>p</td>
</tr>
<tr>
<td>t*</td>
<td>17.8(9.6)</td>
<td>t</td>
</tr>
<tr>
<td>k*</td>
<td>33.6(12.1)</td>
<td>k</td>
</tr>
</tbody>
</table>

It can be seen from Table 1 that VOT values increase in the order from Type 1 stops to Type 3 stops, and there is a difference among three categories of Korean stops. It is apparent that Type 3 stops are produced with considerable delay of voicing and are strongly aspirated. The result and tendencies are in general agreement with previous studies [6], [8]. In my own study [6], Type 1 stops showed the shortest value of VOT, Type 2 stops did the intermediate value between Type 1 and Type 3 stops, and Type 3 stops showed a characteristically longer value of VOT than those of other two types. Although there is a tendency that velar stops show longer delay of voicing among three places of articulation, this was not clearly observed in three types of stops, though Type 2 stops seemed to show this trend slightly.

Table 2 shows the mean VOT values of English stops uttered by Korean speakers. From Table 2, it can be said that voiced stops in English are produced with voicing delay, and that voiceless stops in English are produced with a long delay of voicing and the VOT values are closer to those of Type 3 stops in Korean.

Table 2. Mean VOT values of English Stops by Korean speakers (ms) (N= 16) (s.d. in parenthesis)

<table>
<thead>
<tr>
<th>Voiced</th>
<th>Voiceless</th>
</tr>
</thead>
<tbody>
<tr>
<td>b 24.3(14.3)</td>
<td>p 79.1(30.8)</td>
</tr>
<tr>
<td>d 13.8(28.7)</td>
<td>t 70.8(20.9)</td>
</tr>
<tr>
<td>g 39.3(16.3)</td>
<td>k 101.0(27.5)</td>
</tr>
</tbody>
</table>

Table 3 shows the mean VOT values of Japanese stops uttered by Korean L2 learners, and it can be said that voiced stops in Japanese are produced with a voicing delay, as found in the case of English voiced stops.

2.3 Acoustic Analysis

Acoustic analysis of the recorded materials was made through AcousticCore 8 of Arcadia Co. The linguistic materials were digitized and analyzed at a sampling rate of 44.1 kHz, and the digitized materials were stored for reviewing and listening. The measurement of VOT was made by manually positioning two cursors in the display of the wide-band spectrograms.

3. Results

The measurement of voice onset time (VOT) was made for the interval between the consonant release and the onset of voicing for three types of Korean stops, two types of English and Japanese stops.
known that the VOT value increases as the point of articulation and this is in accordance with previous studies. Although it is students.

they can be considered to show voicing lead in the EFL speakers (ms)

show a greater value than those of bilabial and alveolar stops, strongly aspirated. In connection with this, it can be noted that voiceless Japanese and English, and it can be noted that voiceless stops in

Japanese show a medium length of voicing lag, while those in

English are distinguished by a timing dimension of VOT, and 

In comparing the VOT values of Korean stops with those of English and Japanese stops in Tables 2 and 3, it can be said that Korean L2 learners closely align their Type 1 stops with English and Japanese voiced stops, but they do not show consistent alignment of their voiced stops with English and Japanese voiceless ones.

Several studies have been done on VOT values in English and Japanese stops produced by the natives speakers of both languages, and according to my own study [9], they can be shown, as in Table 4.

Table 3  Mean VOT values of Japanese Stops by Korean speakers (ms) 
(N= 12)  (s.d. in parenthesis)

<table>
<thead>
<tr>
<th>Voiced</th>
<th>Voiceless</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>19.5(9.0)</td>
</tr>
<tr>
<td>d</td>
<td>23.5(12.0)</td>
</tr>
<tr>
<td>g</td>
<td>34.6(8.7)</td>
</tr>
</tbody>
</table>

It can be said that Korean L2 learners show voicing delay in the production of English and Japanese voiced stops, though they can be considered to show voicing lead in the EFL students.

4. Discussion

The examination of voice onset time (VOT) reveals that a three-way stop in Korean can be distinguished by VOT dimension, i.e., three domains such as short range, medium range and long range on the scale of VOT value. From Table 1, it can be seen that the lax unaspirated stops are weakly aspirated in word-initial position. Besides VOT values, it is known that other features such as Fo and Fo contour immediately after the release of closure and F1 values may be relevant for distinction, and it can be said that such a distinction involving tensing of laryngeal muscles can be manifested in several acoustic features.

In comparing the VOT values of Korean stops with those of English and Japanese stops in Tables 2 and 3, it can be said that Korean L2 learners closely align their Type 1 stops with English and Japanese voiced stops, but they do not show consistent alignment of their voiced stops with English and Japanese voiceless ones.

Several studies have been done on VOT values in English and Japanese stops produced by the natives speakers of both languages, and according to my own study [9], they can be shown, as in Table 4.

Table 4  Mean VOT values of Japanese and English Stops (ms)(s.d. in parenthesis)

<table>
<thead>
<tr>
<th>Japanese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/  41(17.1)</td>
<td>68(15.3)</td>
</tr>
<tr>
<td>/t/  30(12.7)</td>
<td>82(18.6)</td>
</tr>
<tr>
<td>/k/  66(12.1)</td>
<td>85(20.1)</td>
</tr>
<tr>
<td>/b/ -89(28.5)</td>
<td>-88(18.1)</td>
</tr>
<tr>
<td>/d/ -75(32.7)</td>
<td>-74(28.0)</td>
</tr>
<tr>
<td>/g/ -75(27.0)</td>
<td>-85(14.4)</td>
</tr>
</tbody>
</table>

As shown in Table 4, there is a clear-cut distinction in VOT values for voiced and voiceless stops in both Japanese and English, and it can be noted that voiceless Japanese show a medium length of voicing lag, while those in English show a medium to long voicing lag. Since voicing lag is related to aspiration, it can be said that voiceless Japanese stops are moderately aspirated, while those in English are strongly aspirated. In connection with this, it can be noted that the VOT value for the voiceless velar stops in both languages show a greater value than those of bilabial and alveolar stops, and this is in accordance with previous studies. Although it is known that the VOT value increases as the point of articulation goes backward in the oral cavity, this was not observed in between bilabial stop and alveolar stops in Japanese. For the voiced stops there is a considerable voicing lead before the release of consonant closure, and there doesn’t seem to be any correlation between their values and the points of articulation.

5. Phonological Features for Stop Consonants in Three Languages

There has been a great deal of discussion on laryngeal features in phonology; what are the features for, what are the physical bases for the features, and what are the conventions in phonological descriptions, and so on. The feature is a basic unit in phonology and is considered to show the relation among the segments. In order to deal with cross-language voicing contrasts, it is necessary to have three phonological features, such as /VOICE/, /ASPIRATED/ and /TENSE/.

The features are considered to be abstract in nature and are to have binary values of plus and minus. It is assumed that other specific features can be added to the list of the features to make a more accurate description, if necessary. For the present study, three features are sufficient. The stops /p, t, k/ in Japanese are specified as [-voice], and the same stops in English are specified as [-voice, -aspirated, -tense]. This means that voicing contrasts in Japanese involve one feature, whereas those in Korean are different from Japanese because they require more feature specifications.

Although Japanese and English use two series of stops, they differ in physical features, and as shown in Table 4, English /p, t, k/ show a long voicing delay, which means a strong aspiration, but Japanese /p, t, k/ show a medial voicing delay. As shown in the above, the same phonetic symbols are used for voiced and voiceless stops in the two languages, but their physical characteristics of the segments differ from one another. Even if the two languages are said to have a voicing contrast, the stop voicing features are different from each other in a phonological framework.

That Korean L2 learners use Type 1 stops for English and Japanese voiced sounds /b, d, g/ implies that they take voicing features rather than tensing ones, and the phonological features such as voicing are more relevant for distinction than tensing. Further, as noted in Tables 2 and 3, they pronounce English and Japanese stops with a voicing delay and this means that a voicing lead is not necessarily a condition for voiced feature and voiced sounds are pronounced with a long duration from a considerable voicing lead to a voicing delay.

VOT is considered to be an efficient feature to specify voiced/voiceless distinction, and voicing lead may be closely related with voicelessness, while voicing delay may be related with aspiration and voicelessness. But this specification is not appropriate, since a number of voiced categories in Japanese and English uttered by Korean L2 learners are in the range of voicing delay. It may be necessary to have revised features to account for voiced/voiceless distinction in case of Korean L2 learners. Along with this proximity in VOT values, it is can be said that Type 1 stops are used in Korean loan words from English to represent voiced stops in English, and this may be due to the phonetic similarity between Type 1 stops and voiced ones in both languages.

6. Summary

The present study shows that a three-way contrast of Korean stops is distinguished by a timing dimension of VOT, and Korean L2 learners use their tense unaspirated stops for voiced
category in English and Japanese, and they use strongly aspirated stops for voiceless stops in English. However, it is not clear whether they use weakly or strongly aspirated stops for voiceless category in Japanese. Proximity in VOT values can be considered to be one of the main reasons of using a Korean stop category in learning L2 categories. Three phonological features such as /VOICE, ASPIRATED, TENSE/ can be proposed and used to specify the phonetic nature of stop consonants in three languages.

7. Acknowledgements

We would like to thank Professor Masayoshi Kakudo of Osaka University for arranging the subject in the present study. The present study was in part supported by a Grant-in-Aid for Scientific Research (C) from Japan Society for the Promotion of Science (22520593) in 2010.

8. References


