Korean Students' Pronunciation of Minimal Pairs of English Words with a Fricative

Byunggon Yang
Pusan National University
E-mail: bgyang@pusan.ac.kr

Annual Review of English Learning & Teaching No.10

The JACET Kyushu-Okinawa Chapter

November 30, 2005
Korean Students' Pronunciation of Minimal Pairs of English Words with a Fricative

Byunggon Yang
Pusan National University
E-mail: bgyang@pusan.ac.kr

Abstract
In this study, the author examined Korean students' pronunciation of minimal pairs of English words with a fricative in order to discover which fricatives were most difficult to distinguish. Thirty students attending a Korean university participated in a recording session. The author listened to each pair on a speech analysis software and judged whether there was a noticeable distinction in their pronunciation of the minimal pairs and whether each fricative was correctly produced. The decision was made based on auditory impression along with visual information from the analysis software. Results showed that almost 70% of the students produced the minimal pairs distinctively and 78% of the fricatives correctly. The minimal pair with the lingua-dental fricative came out with the lowest distinction rate while almost a perfect distinction was made between the voiceless lingua-alveolar and lingua-palatal fricatives. The voiced lingua-palatal fricative in the middle position scored the lowest followed by the lingua-dental fricative in the onset position while the lingua-alveolar fricatives both in the onset and coda position scored 100%. The majority of the fricative errors seemed to be attributed to the negative transfer of the uneven Korean fricative or plosives. Further studies will be desirable on the other English consonants in the future.

Introduction
Recently, English has become the international language by which people can communicate with each other. For a clearer communication in English, pronunciation skills seem to be very important. Among English vowels and consonants, the author believes that fricatives play a great role in clear pronunciation because they account for the majority of segmental sounds in English words. According to Maddieson (1984), the total number of fricatives found in a given language ranged between 0 to over 12. Only 6.3% of the world languages surveyed had more than eight fricatives, to which English belongs.

Since this paper will make most of the acoustic and auditory information for the judgment of distinct pronunciation, the author will review those features of English fricatives. English fricatives are produced by generating turbulence noise from a narrow constriction in the vocal tract. They can be identified (1) by the formation of a narrow constriction in the vocal tract, (2) by the development of turbulent airflow, and (3) by the generation of turbulent noise (Kent & Read, 1996, p. 121). Generally, they are produced at five places in the vocal tract: labio-dental--/f, v/, lingua-dental--/θ, ʒ/, lingua-alveolar--/s, z/, lingua-palatal--/ʃ, ʒ/, and glottal--/h/. Acoustic cues for fricatives are friction noise and transitions and voicing (Kent, Dembowski & Lass, 1996, p. 208-209). Listeners seem to divide fricatives by relative intensity into sibilants of higher intensity /s, z, ʃ, ʒ/ and the low intensity fricatives /f, v, θ, ʒ/. Acoustically, the output noise energy divides them into stridents with much noise and nonstridents with less noise. According to the voicing manner, they are grouped together into voiced fricatives--/v, ʒ, z, ʒ/ and voiceless ones--/θ, ʃ, f, h/. In Korean,
there are only three kinds of fricatives /s/, /s'/, /h/. Thus, we would expect many problems in Koreans’ production of the eight English fricatives. The first and second Korean fricatives are distinguished by their intensity: lenis versus fortis, respectively. The aspirated /h/ can be comparable to English /h/, which will not be considered to cause any problems for Korean students and will not be discussed in this paper.

Heinz and Stevens (1961) proposed a mathematically defined transfer function for fricative resonance as $T(f) = [P(f) Z(f)] R(f)$, in which $T(f)$ is the transfer function in the frequency domain and, $P(f)$ for formants or poles, $Z(f)$ for antiformants or zeros, and $R(f)$ for the radiation characteristic. The formula is an expanded form of the original transfer function for vowels with an additional zero term. Zeros indicate the antiresonance and are determined by the size of the back cavity and the size of the constriction itself. Those cavities are smaller than the whole vocal tract, thus zeros are more widely spaced than poles. Poles and zeros tend to occur close at low frequencies but separated at higher frequencies. The poles and zeros cancel each other out at low frequencies, which one can see in the lower spectral energy at low frequencies. For the labio-dental fricatives, the size of the front cavity is too small to form a strong resonance energy so that their spectra come out flat and diffuse. Johnson (1997, p. 118) mentioned that labial fricatives are made with turbulence at the lips so that the front cavity for them is very short. However, lingua-alveolar and lingua-palatal sounds have intense high energy in the spectra because the front cavity is large enough to produce a visible noise energy at the lower frequency range. For palato-alveolar fricatives, American English speakers usually produce them with lip rounding as well as with the sublingual cavity; thus, their spectral peaks appear lower because the front cavity becomes much longer than labial fricatives. For a man’s production of /s/ the lowest resonance or cutoff frequency is reported to be around 4 kHz with the cavity length of 2 cm. For /s/'/ it comes down to 3 kHz (Kent, Dembowski & Lass, 1996, p. 194-195). The spectral mean for the lingua-alveolar is usually lower than that of the lingua-palatal. Thus, the cutoff frequency can be a useful cue for telling the two kinds of fricatives apart. Ladefoged and Maddieson (1996) commented that the gesture forming the constriction requires a higher degree of articulatory precision than that for stops and nasals because just a little variation along the position of the vocal tract leads to a great deal of acoustic difference. In addition, Johnson (1997, p. 118) illustrated evenly spaced spectra for each fricative when he transformed the acoustic spectra for a series of Egyptian Arabic fricatives into the auditory scale Mel. That equidistant distribution on the spectra suggests that we may have to be careful about judging the fricative distinction in a manner that is solely dependent on the acoustic output because the physical frequency scale may not properly represent the auditory impression of a listener.

Finally, voicing is the cue to tell the voiced fricative from the voiceless one. Generally, a voice bar, the low frequency sound of glottal pulsing, on the speech spectrogram may be a salient cue of voicing. However, even without it, listeners can make judgments about the voicing of a syllable-final fricative by the duration of the preceding vowel (Denes, 1955; Borden & Harris, 1984). Denes found that people heard /s/ when the /z/ from /juz/ was cut and pasted onto the end of /jus/. Raphael (1972) also tested listeners’ perception of voicing distinctions in coda positions. He found that vowels of shorter duration provoked the perception of voiceless final consonants and vice versa. He pointed out that American English speakers did not always release final stops, and they distinguished the voicing of the final consonant by controlling the duration of the preceding vowel. Moreover, Johnson (1997, p. 115) pointed out that voiced fricatives are more difficult to produce than voiceless ones. The speaker should prepare high volume velocity to produce the turbulent noise with the vocal cords close together to produce the voicing. The narrow glottal opening to produce the voicing part of the consonant will not allow much air for the energy needed to produce the
turbulent noise. That may explain why many people tend to produce a voiceless fricative instead of a voiced one, especially in coda position.

The aim of this paper was to examine Korean students' pronunciation of minimal pairs of English words with a fricative in order to find out the most persistent pronunciation errors. Because the acoustic analysis of the pronunciation did not display all the information of each speaker's organic gestures, the author decided to rely on the auditory information along with the acoustic information in judging students’ pronunciation. That way, this paper may provide more practical information on the teaching of English fricatives to Korean students.

**Method**

Thirty subjects (fifteen males and fifteen females) without any hearing problems participated in the experiment. They were junior students attending a Korean university who enrolled in an English pronunciation class. They were instructed to produce 11 English minimal pairs with a fricative. The pairs were as follows: *belief-believe, class-clash, clothe-cloth, day-they, face-faith, ice-eyes, measure-major, sea-she, see-zee, think-sink, view-few*. The author asked them to pronounce each pair clearly. Their productions were recorded on a notebook computer (SAMSUNG SX 10) at a sampling rate of 22,050 Hz with a microphone using one of the folders of *Alvin, "Prompted Recording"* (Hillenbrand, 2005).

The recorded sounds were read and opened in an edit window on a speech analysis software, Praat (Boersma & Weenink, 2005). The edit window displayed the waveform, pitch contour, and intensity contour on the spectrogram. From the window, the author zoomed in and out or listened to each segment or the whole minimal pair to judge whether the two words in the pair sounded the same or different. Additionally, the author utilized its acoustic information to confirm the perceptual difference correctly. Then, the author identified each minimal pair as distinct or not distinct. The correct distinction rate was determined by obtaining the percent ratio of the word pair distinctively produced to the total pairs. The same procedure was applied to the pronunciation of each fricative and the correct ratio was collected. While listening to the student's pronunciation and observing his or her spectrogram, error sounds were noted for discussion.

This section will illustrate some waveforms and spectrograms of four minimal pairs produced by a native Australian female instructor in the university so that the reader can better understand how the author employed acoustic analyses in his decision.

**Figure 1** shows the production of the pair *belief-believe* produced by the Australian speaker. The upper window displays the waveform while the lower window does the spectrogram, pitch contour, and intensity contour. The analysis settings for the spectrogram were as follows: the view range from 0 Hz to 8000 Hz, the window length to 0.005 seconds, and the dynamic range to 40 dB. The view range was set wider than the default one with 5000 Hz in order to catch the high frequency noise of the fricative /s/. For pitch, the range was set between 75 Hz to 500 Hz and optimized for the intonation contour (Autocorrelation method). The standard settings were used for intensity contour.
From the figure one can note that the Australian speaker produced the coda /v/ in the second word believe of the pair as voiceless as was pointed out by Harris (1958). The author zoomed in the last coda section of the word "believe" but found not any visible periodic segment. We may judge that the speaker produced the two words the same if we rely on the acoustic information only. However, the author could tell a perceptual difference easily just because the preceding vowel duration was quite different. The duration of the second syllable peak of the vowel /i/ was measured from the rising second formant to the onset of the fricative noise /f/: 0.144 seconds for the word belief; 0.248 seconds for the word believe. Even the intensity of the final fricative appeared different. The voiceless /f/ had a stronger pattern than its counterpart. Thus, the author decided to apply two steps in judging whether a speaker pronounced a minimal pair distinctively or a given fricative onset or coda correctly. First, the author listened to each pair or word several times and expanded any ambiguous section from the acoustic analysis window to listen to its segment again. Second, some acoustic cues on the spectrogram, duration, intensity or pitch contour were employed to judge their distinction rather consistently across the stimuli. Figure 2 illustrates the word pair face-faith produced by the female speaker.
The author could hear a distinctive noise which must indicate the contact between the upper teeth and tongue tip of the speaker for the fricative /θ/ even though the spectrogram displayed only a nebulous trace for the weak frication. It might be related to the nonlinear auditory impression along the physical frequency range. Human beings can notice subtle changes in the frequency range around 3,500 Hz (See Figures 5,11, Denes & Pinson, 1970, p. 8). Moreover, whether the two coda sections were cut and pasted onto one window, the energy difference in the intensity curve seemed noticeable. In other words, the energy of the frication noise for the lingua-alveolar fricative appeared stronger than that of the lingua-dental fricative.

For the lingua-alveolar and lingua-palatal fricative pair she-see, the cutoff frequency was employed as a cue along with its auditory impression. Figure 3 shows its production by the female speaker.

![Figure 3](image)

*Figure 3. A waveform and acoustic analysis of the pair, she-see.*

The cutoff frequency for the fricative in the word *she* was around 2,227 Hz while that of the word *see* was around 4,332 Hz. The intensity also appeared stronger for the lingua-palatal fricative. Perceptually, they sounded quite different.

Figure 4 shows the pair *major-measure* produced by the same speaker. Here again, no voicing can be seen in the production of the speaker for the /dʒ/. For the judgment of the minimal pair distinction, the silent gap between the preceding and the following vowels of the word *major* can be a useful cue, which can be easily checked on the spectrogram.
Finally, when the subject produced each pair at a very low intensity with an ambiguous spectrographic analysis, the author relied on the auditory impression after amplifying the speech segment for the final judgment.

Results and Discussion

Table 1 reveals the results of minimal pair distinction by the Korean students. Ratio indicates the percent ratio of distinctly produced pairs to the total number of the pairs.

Table 1

<table>
<thead>
<tr>
<th>Pairs</th>
<th>belief-believe</th>
<th>class-clash</th>
<th>clothe-cloth</th>
<th>day-they</th>
<th>face-faith</th>
<th>Ice-eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Correct Distinction</strong></td>
<td>22</td>
<td>30</td>
<td>13</td>
<td>12</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td><strong>Ratio (%)</strong></td>
<td>73</td>
<td>100</td>
<td>43</td>
<td>40</td>
<td>70</td>
<td>87</td>
</tr>
<tr>
<td><strong>Pairs</strong></td>
<td>measure-major</td>
<td>sea-she</td>
<td>see-zee</td>
<td>think-sink</td>
<td>view-few</td>
<td><strong>Total</strong></td>
</tr>
<tr>
<td><strong>Number of Correct Distinction</strong></td>
<td>7</td>
<td>30</td>
<td>27</td>
<td>24</td>
<td>19</td>
<td><strong>231</strong></td>
</tr>
<tr>
<td><strong>Ratio (%)</strong></td>
<td>23</td>
<td>100</td>
<td>90</td>
<td>80</td>
<td>63</td>
<td>70</td>
</tr>
</tbody>
</table>
From the table we can see that the Korean students produced 70% of all the minimal pairs distinctively. The fricative distinction of the five minimal pairs in the onset positions (view-few, day-they, see-she, see-zee, think-sink) or coda positions (belief-believe, class-clash, clothe-cloth, face-faith, ice-eyes) was made around 75% (112 pairs out of 150) correctly. It is interesting to find the same correct ratio for the fricatives in the two different positions. When the data were grouped into the place of articulation, the lowest (23%) distinction came out for the pair measure-major followed by 58% (70 pairs out of 120) for lingua-dental fricatives; 68% (41 pairs out of 60) for labio-dental fricatives in both onset and coda positions; 88% (53 pairs out of 60) for lingua-alveolar fricatives; 100% for lingua-palatal fricatives. Only seven subjects out of thirty distinguished the voiced fricative /s/ from the voiced affricate /dʒ/ in the pair measure-major. Figure 5 shows a female student’s production of the pair measure-major. There was not much difference in the auditory impression along with the acoustic analysis. A slight difference in the duration and pitch of the second syllable could be observable but those are negligible enough to be heard. Thus, the author judged that the subject did not clearly distinguish the pair.

![Waveform and spectrogram of a female student’s production of the minimal pair, measure-major](image)

Figure 5. A waveform and spectrogram of a female student’s production of the minimal pair, measure-major

On the other hand, all the subjects distinguished the lingua-palatal pairs (sea-she and class-clash) correctly, which might be attributed to the sufficient acoustic cue by the different cut-off frequency.

Table 2 shows the distinction of each fricative. The percent correct ratio was determined to display in the chart so that general comparison among the fricatives could be made.

From the table, we found that the subjects distinguished 78% of all the fricatives clearly. When the data were grouped into the onset and coda fricatives, results showed that 72% (152 fricatives out of 210) of the onset sounds were pronounced correctly while 83% (200 fricatives out of 240) of the coda sounds were spoken correctly. The voiced fricative /s/ scored 73% (22 out of 30) in the middle position. The data were grouped into the place of production: 69% (83 out of 120) for labio-dental fricatives /f-v/; 66% (99 fricatives out of 150) for lingua-dental fricatives; 92% (110 fricatives out of 120) for lingua-alveolar fricatives; 91% (82 cases out of 90) for lingua-palatal fricatives. Individually, the lowest distinction (43%) was made in the voiceless lingua-dental fricative in the coda position followed by the voiceless labio-dental fricative.
(57%). The low score might be related to the different consonantal sound system of English and Korean. Here again, the voiceless lingua-alveolar and lingua-palatal fricatives in the onset or coda positions scored 100%, as was seen from the perfect distinction of the minimal pairs in the previous table.

Table 2
Number and ratio of the fricatives correctly distinguished by the Korean students

<table>
<thead>
<tr>
<th>Words</th>
<th>believe</th>
<th>belief</th>
<th>clash</th>
<th>clothe</th>
<th>cloth</th>
<th>they</th>
<th>faith</th>
<th>ice</th>
<th>eyes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sounds</td>
<td>v</td>
<td>f</td>
<td>f</td>
<td>ð</td>
<td>ð</td>
<td>ð</td>
<td>s</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Number of Correct Distinction</td>
<td>17</td>
<td>27</td>
<td>30</td>
<td>20</td>
<td>28</td>
<td>13</td>
<td>21</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>Ratio (%)</td>
<td>57</td>
<td>90</td>
<td>100</td>
<td>67</td>
<td>93</td>
<td>43</td>
<td>70</td>
<td>100</td>
<td>90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Words</th>
<th>measure</th>
<th>she</th>
<th>see</th>
<th>zee</th>
<th>think</th>
<th>view</th>
<th>few</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sounds</td>
<td>3</td>
<td>f</td>
<td>s</td>
<td>z</td>
<td>ð</td>
<td>v</td>
<td>f</td>
</tr>
<tr>
<td>Number of Correct Distinction</td>
<td>22</td>
<td>30</td>
<td>30</td>
<td>23</td>
<td>17</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Ratio (%)</td>
<td>73</td>
<td>100</td>
<td>100</td>
<td>77</td>
<td>57</td>
<td>60</td>
<td>70</td>
</tr>
</tbody>
</table>

Perceptually, it was very difficult to judge whether the onset or coda fricatives were produced with the vocal cords vibrating. Acoustically, one can easily observe the voice bar on the spectrogram and the pitch values higher than zero. However, if we consider the fact that the pitch algorithms in most speech analysis software are still imperfect, it might not be good to depend solely on acoustic analysis for the judgment of correct pronunciation. The author found that only a few of the subjects actually produced voiced fricatives with two acoustic cues on the spectrogram. In particular, eight subjects realized the voiced labio-dental fricatives in the onset or coda positions and the lingua-dental fricative in the onset position with those acoustic cues. Ten subjects produced the voiced onset /z/ with the cues but only two did in the coda position. The voiced lingua-palatal and lingua-dental fricatives were realized as voiced by only four subjects out of a total of thirty. Further studies on the perceptual threshold of voicing by native English speakers may be desirable in order to establish guidelines for teaching Korean students voiced fricatives.

Finally, the author investigated the subjects’ error pattern. Many subjects substituted the voiceless plosive /p/ for the production of the labio-dental fricatives: Twenty cases out of sixty in the minimal pair view-few and thirteen out of sixty in the pair belief-believe. Figure 5 shows a female student’s production of the pair.
Figure 6. A waveform and spectrogram of a female student's production of the minimal pair, belief-believe

From the figure one can tell easily that the student produced a voiceless bilabial stop for the final fricative with a visible burst noise followed by a silent gap. It might be related to the Korean consonantal system without any equivalents of the English fricatives. Fourteen subjects produced the dental plosive /l/ for the words they and faith. The Korean affricate /lə/ was employed in the production of the two words: measure (six subjects) and zee (seven subjects). The lingua-alveolar fricative was used to produce the lingua-dental fricatives (twenty subjects). Eight subjects still produced the dental plosive for the onset of the word think.

Figure 6 clearly shows the dental plosive produced by a Korean subject. The voice onset time for the onset of the first word was 0.019 seconds, which falls on the stiff dental plosive in Korean (Ladefoged & Maddieson, 1996, p. 56). Those errors could be attributed to the negative transfer from the Korean consonantal system to the production of the English fricatives. If we had subjects at the beginning level of English proficiency, the percent correct ratio would have gone much lower.
Figure 7. A waveform and spectrogram of a female student's production of the minimal pair, think-sink

Conclusion

This study investigated the Korean students' pronunciation of minimal pairs of English words with a fricative onset or coda in order to determine which fricatives were most difficult to distinguish. The author listened to each pair produced by thirty Korean students on a speech analysis software and counted the number of the pairs and the fricatives distinctly pronounced based on the auditory impression along with acoustic information from a speech analysis software. Results showed that almost 70% of the students produced the minimal pairs distinctively and 78% of the fricatives correctly. The minimal pair with the lingua-palatal fricative came out with the lowest distinction rate while almost perfect distinction was made between the voiceless lingua-alveolar and lingua-palatal fricatives. Individually, the voiceless lingua-dental fricative in the coda position scored the lowest while the lingua-alveolar fricatives, both in the onset and coda positions, scored 100%. Error patterns of the fricatives seemed to reflect the negative transfer or the replacement of a similar Korean consonant for the English counterpart. Further studies will be desirable on the other English consonants and the native listener's ratings on the Korean student's production of English consonants.

References