Proceedings of the 4th GLOW in Asia 2003

Generative Grammar in a Broader Perspective

Edited by
Hang-Jin Yoon

Published for
The Korean Generative Grammar Circle
&
Institute for Cognitive Science, Seoul National University

Hankook
Flaps in American English and Korean:
An Acoustic and Perceptual Study

William Idsardi and Eun-kyung Sung
University of Delaware and Cyber University of Foreign Studies

1. Introduction

Linguistic experience influences the mind of a listener. Listeners pick up the same acoustic signals and may interpret the signals differently according to their language backgrounds. Linguistic experience in a particular language decreases sensitivity for some phonetic contrasts that are not used in the language. One way to examine the effect of linguistic experience on speech perception is to compare the perceptual patterns of phonetically relevant segments by two groups of speakers who have different linguistic experiences. Most cross-language studies have focused on native language interference in the perception of non-native contrasts, and are restricted to the perception of certain phones such as stops or liquids. There has been little investigation of the perception of sounds that are classified as different categories such as sonorants and obstruents in two languages.

Flaps in AE and Korean provide an appropriate case since flaps are differently categorized in the two languages. That is, flaps in Korean and AE are phonetically very similar, but the functions of the flaps in the phonological systems of the two languages are crucially different. The flap in Korean is an allophone of the lateral liquid /ɾ/ found when /ɾ/ comes between vowels, e.g., puʔ [puɾi] ‘beak’. It contrasts with the alveolar lax stop /ɾ/, as we can see in a minimal pair, e.g., muli-ɾa [muɾiɾa] ‘unreasonable’ vs. muti-ɾa [muɾiɾa] ‘blunt, dull’. It should be noted that the Korean /ɾ/ further contrasts with /ɾ darken 3rd person singular. alveolar tense and aspirated stops, respectively, e.g., soɾo [sæɾo] ‘an apostle’ vs. soɾ o [sæɾo] ‘mayor (archaic)’ vs. soɾ o [sæɾo] ‘sandy soil’. By contrast, the flap in AE is an allophone of the alveolar stops /ɾ/ and /ɾ/, e.g., latter and ladder, both pronounced [læɾə].
contrasts with the liquids /l/ and /l/, e.g., Betty vs berry vs belly. The following figure shows the relationship of flaps and the main variants of liquid and alveolar stop sounds in Korean and AE:

(1) Phonological status of flaps, alveolar stops and liquids in Korean and AE:

<table>
<thead>
<tr>
<th>Korean</th>
<th>Alveolar Stops</th>
<th>Liquids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological level</td>
<td>/l/</td>
<td>[l]</td>
</tr>
<tr>
<td>Phonetic level</td>
<td>[l]</td>
<td>[l']</td>
</tr>
<tr>
<td>Phonological level</td>
<td>/l/</td>
<td>[l']</td>
</tr>
</tbody>
</table>

The present research explores the relationship between phonetics and phonology by comparing acoustic properties of flaps in AE and Korean, and investigates the role of the first language experience by comparing different perceptual patterns of a phonetically similar segment by two groups of speakers. The acoustic properties of AE and Korean flaps are analyzed on the basis of the data from a production experiment. In evaluating the flaps in the two languages, the measures of closure duration, presence of burst and voicing during the closure were compared. The effects of the phonological function of flaps are investigated through a series of perception experiments on phonetically relevant sounds in Korean and AE.

2. Experiment 1: Production

2.1. Hypothesis

There would be no systematic acoustic differences between AE and Korean flaps.

2.2. Methodology

2.2.1. Subjects

Ten AE speakers (five males and five females) who were undergraduate students at the University of Delaware (UD), and ten Korean speakers (five males and five females) who attended the English Language Institute (ELI) at UD, participated in this study.

2.2.2. Stimuli and Procedure

The twelve target items contained the flap [ɾ] between vowels, e.g., CV:V. All the tokens were real words in AE or Korean. Intervocalic contexts were chosen since these are the only environments where both AE and Korean flaps occur. The items in the stimuli included a variety of vocable contexts for flaps, and the vocalic contexts of the AE items were approximately parallel to those of the Korean items. Each item was produced inside a carrier phrase to avoid possible effects of citation form or utterance-final position (Klatt 1976). In total, 240 items for each group (12 tokens*2 repetitions*10 subjects) were analyzed.

2.3. Results

2.3.1. Closure Duration

The duration of each flap was measured from the decrease of acoustic
energy associated with the beginning of the closure interval until the increase of energy associated with the following unstressed vowel if there was no identifiable release burst. If there was a clear release burst indicated by a high-energy peak, the closure duration was measured until the beginning of the burst portion. All markings were made at the positive zero crossings preceding the first positive-going peak of the first pitch pulse of each flap segment and preceding the first positive-going peak of the first pitch pulse of the following vowel.

The mean closure duration of AE flaps is 19 ms, and that of Korean flaps is 20 ms. Standard deviation is not large in either AE or Korean. An unpaired t-test shows that there is no significant difference in mean values of closure duration between AE and Korean flaps ($t(18) = -0.18, p > .05$). Three AE speakers (AM1, AM4, and AF1) provided the shortest mean values of closure duration since they often produced flaps whose closure duration was less than 10 ms. The following boxplots show the mean closure duration for each group separately.

(2) Boxplots of Mean Closure Duration for AE and Korean Flaps

As can be seen in (2) the Korean flaps have a more compact shape than the AE flaps, and the AE flaps have more variability than the Korean flaps. The large variability in the AE flaps came from the subgroup of AE speakers who often produced very short flaps.

2.3.2. Voicing

The main source of voicing information around the closure for stops is the presence or absence of a voicing bar which is seen in the spectrograms (Lisker 1978). If a voicing bar was not clear in the spectrograms, the waveforms were also examined in an expanded view. If there were regular periodic wave shapes during the closure, the flap token was considered a voiced one, although the waveform of the flap was weaker with a simpler shape than that of adjacent vowels. The percentage of voicing occurrence during the closure duration is high in both AE and Korean. For the AE speakers, 82% of the flaps are voiced. This average is compatible with the frequency of AE flaps (86%) given by Fox and Tuerk (1977) and that (81%) given by Price (1981).

The AE speakers fall into two groups in terms of voicing occurrence: One group of five speakers (AM1, AM3, AM4, AF1 and AF3) almost always produced voiced flaps (96% - 100% of flaps are voiced), and the other group of speakers (AM2, AM5, AF2, AF4 and AF5) often produced flaps without voicing (only 46% - 79% of the flaps are voiced). The Korean speakers' flap production was more homogeneous. Eight out of the ten Korean speakers produced voiced flaps most of the time (92% - 100%). Thus, the Korean flaps are almost always voiced. The non-parametric Mann-Whitney two-independent samples test based on ranks shows that there is no significant difference in voicing occurrence between the two groups ($z = -1.468, p > .05$) despite individual variation in AE flaps. The following boxplots show the percentage of voicing occurrence for the two groups.

2 AM = American English male speaker
AF = American English female speaker
KM = Korean male speaker
KF = Korean female speaker
(3) Boxplots of Percentage of Voicing Occurrence in AE and Korean Flaps

The Korean flaps are more homogeneous than the AE flaps in terms of voicing occurrence as well as closure duration. The larger variability in the AE flaps than in the Korean flaps may be attributed to two subgroups of the AE speakers that were identified on the basis of the data.

2.3.3. Release Burst

The release burst is an aperiodic signal portion preceding the first glottal pulse of the following vowel. The burst is seen as a dark vertical line extending to the high frequencies, and the amplitude of a burst and the duration of frication noise vary across speakers. Release bursts are rare in both the AE and Korean flaps, though the Korean flaps are accompanied by release bursts more often than the AE flaps. The average percentage of release bursts in the AE flaps is 25%, and this frequency is compatible with that (29%) given by Price (1981). Four out of the ten AE speakers (AM1, AM4, AM5, AF1) seem to very rarely have bursts (less than 15%), whereas all the Korean speakers have bursts more than 20% of the time. Considering closure duration, voicing and bursts together, the flap production of the three speakers (AM1, AM4, AF1) often has a very short closure duration, and is almost always accompanied by voicing. A non-parametric Mann-Whitney two-independent samples test based on ranks shows that there is a marginal difference in the burst occurrence between the two groups \( z = -2.164, p < 0.05 \). The Korean flaps are more often accompanied by release bursts than the AE flaps. The following boxplots show the percentage of burst occurrence.

(4) Boxplots of Percentage of Burst Occurrence in AE and Korean Flaps

The boxplots also show the different distribution of percentage of burst occurrence between the AE and Korean flaps since the two boxes do not overlap. The following table shows the summary of the results in the production experiment.

(5) Mean Values of Flaps in AE and Korean
(Standard Deviation in parentheses)

<table>
<thead>
<tr>
<th></th>
<th>AE</th>
<th>Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closure Duration</td>
<td>19 ms (13)</td>
<td>20 ms (9)</td>
</tr>
<tr>
<td>Voicing</td>
<td>82% (19)</td>
<td>95% (6)</td>
</tr>
<tr>
<td>Burst</td>
<td>25% (6)</td>
<td>42% (4)</td>
</tr>
</tbody>
</table>

3. Discrimination Experiments: Flaps, Alveolar Stops and Liquids

3.1. Hypotheses

The hierarchy of the perceptual difficulty is hypothesized on the basis of the phonological status of sounds. A phonemic difference means that two
phones of a pair occur systematically and signal differences in meaning in L1. Thus, the two phones are categorized into two different phonemes. An allophonic difference indicates that two phones of a pair occur systematically, but do not signal differences in meaning in L1. Thus, the two phones are categorized into the same phoneme. A partially phonemic difference shows that two phones in a pair contrast meaning in certain contexts, but not in other contexts. Therefore, although the two phones are categorized as the same phoneme in the L1 phoneme inventory, the two phones are contrastive in a certain position. A non-native difference means that the pair represents neither phonemic nor phonetic difference since one of the phones in a pair does not occur in L1. Speakers of the language depend on the acoustic difference to discriminate between two phones in a pair. The amount of acoustic difference varies depending on each pair. The following shows the hypothesized hierarchy of the perceptual difficulty (from the least to the most):

- phonemic difference (separate phonemes in L1) least
- partially phonemic difference (phonemic contrast in a context) less
- allophonic difference (allophonic variants in L1) more
- non-native difference (native vs. non-native phones) most

3.2. Methodology

3.2.1. Subjects

The subjects were twenty native speakers of AE (eight males and twelve females) and twenty native speakers of Korean (ten males and ten females).

3.2.2. Stimuli and Procedure

The stimuli in the present experiment were constructed using edited natural speech that was derived from the productions of two female AE or Korean speakers. Each pair consisted of two very short pieces of words produced by the two speakers with standard pronunciation. For the AE stimuli, each token included an intervocalic flap [r], an alveolar stop [d] or [tʰ], or a liquid [ɾ] or [l] between the preceding vowel [u] and the following vowel [i], e.g., [ur i], [ud i]. In the stimuli including [r], [ɾ] and [l], the preceding vowel was always stressed. In the stimuli including [d] and [tʰ], however, the following vowel was always stressed, otherwise, the sounds in question would be pronounced as a flap. In each token, 40 ms of each vowel was included, and the other part of the vowel was cut off in order to reduce the effect of the stress patterns, and to obviate the different effects of vowels for native and non-native speakers. Only two high vowels, [u] and [i], were used in the present experiment because these two vowels were perceptually very similar in AE and Korean according to the author and several AE speakers. Also, both Korean and AE flaps, alveolar stops and liquids can occur in this environment, [uCi]. In total, 80 test pairs and 80 fillers were used. The test pairs consisted of 40 same pairs [5 consonants (r, d, t, l) x 4 repetitions x 2 speakers], and 40 different pairs [5 contrasts (r-d, r-t, r-l, l-r) x 4 repetitions x 2 speakers]. The fillers also consisted of 40 same pairs and 40 different pairs, and each item in a pair included [b], [p] or [s] between vowels. There was a 1500 ms interval between the two members of a pair, and a 3000 ms interval between the pairs.

For the Korean stimuli, The short pieces of words included intervocalic flap [ɾ], alveolar lax stop [d], alveolar aspirated stop [tʰ], alveolar tense stop [tʰ], and geminate lateral [ll] between the preceding vowel [u] and the following vowel [i], e.g., [ur i], [ud i]. The stimuli consisted of 80 test items and 80 fillers. The test items contained 40 same items [5 consonants (r, d, t, l) x 4 repetitions x 2 speakers] and 40 different items [5 contrasts ([r-d], [r-t], [r-l], [l-r], [l-t], [l-d]) x 4 repetitions x 2 speakers]. The fillers also consisted of 40 same items and 40 different items, and each item in a pair included [b], [p] or [s] between vowels. The pairs in the present experiments were chosen on the basis of the phonological or phonetic status of the phones in each pair. The following tables show the phonemic relations of pairs of segments in the AE and Korean stimuli.
3.3. Results

In order to investigate the subjects' performance in discrimination tasks, sensitivity was measured for each stimulus pair for each subject by applying signal detection theory (Green & Swets 1966, Macmillan & Creelman 1991, Macmillan 2002) Signal Detection Theory (SDT) is a framework for understanding accuracy that makes the role of decision clear by providing a single sensitivity index from any discrimination paradigm. The following figures display the mean discrimination sensitivity for each pair in the AE and Korean stimuli by the AE and Korean speakers.

![Graph showing the distribution of mean sensitivity (d') functions for the AE and Korean stimuli.](image)

For the AE stimuli, whereas the AE speakers show variation among the pairs, between the [r-d] pair and all other pairs in particular, the Korean speakers' performance is more or less consistent across the pairs in general, the AE speakers performed well on all the pairs except the [r-d] pair, and individual data show that all the AE speakers uniformly performed poorly on the [r-d] pair. Although the Korean speakers' performance on the AE stimuli
does not show much variation, their performance is not as good as that of the AE speakers on all the pairs except [r-d]. When the data were submitted to a repeated measures ANOVA with groups (2 levels: AE vs Korean) as the between-subjects factor, and pairs (5 levels: [r-d], [r-t'], [r-l], [r-r], [l-r]) as the within-subjects factor, the results confirmed that there was a significant effect of pair \[F(4, 152) = 31.04, p < 0.01\] Group was not significant \[F(1, 38) = 154, p > 0.05\], but there was a significant pair x group interaction \[F(4, 152) = 19.92, p < 0.01\] Group as a whole did not show a significant effect since although the AE speakers performed much worse than the Korean speakers on the [r-d] pair, the AE speakers performed better than Korean speakers on all other pairs. Thus, the group effect on one pair, [r-d], was obscured by the opposite group effect on the other pairs, [r-t'], [r-l], [r-r] and [l-r].

A one-way repeated measures ANOVA was conducted for the AE and Korean speakers separately with pair as the within-subjects factor. For the AE speakers, there was a significant effect of pair \[F(4, 76) = 51.11, p < 0.01\] Bonferroni post hoc tests showed that the AE speakers' mean d' scores of the [r-d] pair was significantly lower than those of all other pairs (p < 0.01 for all comparisons) There was no significant difference among the other pairs. For the Korean speakers, a one-way repeated measures ANOVA also showed that there was a significant effect of pair \[F(4, 76) = 4.198, p < 0.01\] Bonferroni post hoc tests showed a marginal difference between [r-t'] and [l-r] (p = 0.05) That is, the Korean speakers' performance on [l-r] was a little worse than that on [r-t']. There were no significant differences among the other comparisons.

In order to compare the groups for each pair, a between-subjects one-way ANOVA was conducted. When the Bonferroni correction was applied, there was a significant difference between groups in two pairs, [r-d] and [l-r] \[F(1, 38) = 24.74, p < 0.01; F(1, 38) = 9.614, p = 0.004 < 0.005, respectively\] The performance of the AE speakers on the [r-d] pair was worse than that of Korean speakers, while the performance of AE speakers on the [l-r] pair was better than that of Korean speakers Both the AE and Korean speakers performed well on the [r-t'] pair.

For the Korean stimuli, whereas the Korean speakers did not show variation between the pairs in the Korean stimuli, the AE speakers showed considerable variation between the pairs. In general, the Korean speakers performed uniformly well on all the pairs. The AE speakers showed low sensitivity to the [r-d] pair in the Korean stimuli as they did in the AE stimuli. The AE speakers also showed low sensitivity to the [d-t'] pair, although their sensitivity to [r-d] was much lower than that to [d-t']. The data were submitted to a repeated measures ANOVA with groups (2 levels: AE vs Korean speakers) as the between-subjects factor and pairs (5 levels: [r-d], [r-t'], [r-l], [r-r], [d-t']) as the within-subjects factor. The results confirmed that there was a significant effect of pair \[F(4, 152) = 44.69, p < 0.001\], group \[F(1, 38) = 41.95, p < 0.001\], and pair x group interaction \[F(4, 152) = 28.75, p < 0.001\]

A one-way repeated measures ANOVA was conducted for the AE and the Korean speakers separately with pair as the within-subjects factor. On the one hand, for the AE speakers, the results showed a significant effect of pair \[F(4, 76) = 47.49, p < 0.01\] Bonferroni post hoc tests showed that the AE speakers' mean d' scores of the [r-d] pair were significantly lower than those of all the other pairs (p < 0.01 for the comparisons of [r-d] vs [r-t'], [r-d] vs [r-l], [r-d] vs [l-r]); p < 0.01 for the comparison of [r-d] vs [d-t']). In addition, the mean d' scores of the [d-t'] pair was significantly lower than those of three other pairs, [r-t'], [r-l] and [l-r] (p < 0.01 for all comparisons), and significantly higher than those of the [r-d] pair (p < 0.01). On the other hand, for the Korean speakers, the results of a one-way repeated measures ANOVA revealed that pair was marginally significant \[F(4, 76) = 3.85, p = 007 < 0.05\] The follow-up Bonferroni post hoc test, however, did not reveal any significant difference between pairs (p > 0.05 for all comparisons). In order to compare the groups for each pair in the Korean stimuli, a between-subjects one-way ANOVA was conducted. The results showed that the Korean speakers performed better than the AE speakers on three pairs, [r-d], [r-t'], and [d-t'] \[F(1, 38) = 68.17, F(1, 38) = 17.76, F(1, 38) = 24.62, respectively, p < 0.001 for all three pairs\] The AE speakers' performance on [r-t'] and [r-l] was as good as the Korean speakers'
performance. The AE speakers, however, did not perform better than the Korean speakers on any pairs.

In order to compare the AE pairs and the Korean pairs for each group separately, paired t-tests were conducted. Only three pairs, [r-d], [r-tʰ], and [r-l], were compared between the AE and the Korean stimuli since both the AE and the Korean stimuli included these three pairs. The AE speakers did not show any difference between the AE and Korean stimuli regarding these three pairs (p > 0.05 for all three pairs). Their sensitivity to the [r-d] pair was low in both the AE and Korean stimuli, and their sensitivity to the [r-tʰ] and [r-l] pairs was high in both sets of stimuli. By contrast, the Korean speakers showed a significant or marginal difference between the AE and Korean stimuli regarding these pairs (p = 0.03 < 0.05 for [r-d], p = 0.02 < 0.05 for [r-tʰ], and p = 0.02 < 0.05 for [r-l]). Their sensitivity to these pairs in the Korean stimuli was higher than in the AE stimuli.

4. Categorization Experiments: Flaps, Alveolar Stops and Liquids

4.1. Hypotheses

The AE speakers categorize the flaps as alveolar stops, whereas the Korean speakers would categorize the flaps as liquids. In addition, the influence of the LI phonological system is greater in the categorization tasks than in the discrimination tasks.

4.2. Methodology

4.2.1. Subjects

The subjects were the same as those in the discrimination experiments. Twenty native speakers of AE (eight males and twelve females) who were undergraduate students at UD, and twenty native speakers of Korean (ten males and ten females) who were ELI students at UD, took part in this experiment. It should be noted that among the AE speakers who were originally recruited, six speakers who had substantial experience in Spanish (more than four years during high school and college) were eliminated from the previous and present experiments. Spanish has a flap (tap) which is acoustically very similar to AE flaps, and the Spanish flap is spelled as “r”, so the subjects with significant Spanish experience may be influenced by Spanish spelling.

4.2.2. Stimuli and Procedure

The stimuli in the present study consisted of 40 test words and 40 fillers which were all nonce words in AE. The stimuli were produced by a female AE or Korean speaker with standard pronunciation. Nonce words were used to prevent the subjects' perceptions from being influenced by their familiarity with particular lexical items.

For the AE stimuli, the 40 test words included AE flap [r], alveolar stops [d] and [tʰ], and liquids [l] and [l] between various vowels, e.g., [ni[r]o], [sotʰu]. The preceding vowels were chosen among /i, a, o, u/ and the following vowels among /i, o, a/. The vowel [a] was excluded from the following vowels since words which end with [a] are very rare in AE. The onset of each nonce word was chosen among the AE consonants with the exception of a flap, alveolar stops and liquids. Since the stimuli were AE nonce words, the AE stress pattern was employed. In other words, the AE flaps were produced between a preceding stressed vowel and a following unstressed vowel, whereas the AE alveolar stops [d] and [tʰ] were produced between a preceding unstressed vowel and a following stressed vowel. The AE liquids [l] and [l] were produced either with the preceding or following vowel stressed. The filler items included [s, z, b, p, g, k] between various vowels.

For the Korean stimuli, The 30 test words included the Korean flap [r], alveolar plain stop [d], alveolar tense stop [tʰ] and alveolar aspirated stop [tʰ], and geminate lateral [ll] between vowels, e.g., [ni[r]a], [sa[tʰ]l]. The vowels preceding and following the flaps, alveolar stops and laterals were /i, a, o, u/. Fillers included [p, b, k, s, n] between vowels.
All the stimuli were recorded onto the IEAC RW-800 CD recorder, and the recordings were digitized at 44.1 kHz sampling frequency and 16 bit resolution using the EAC software program. All the items were automatically randomized for each subject. Each subject was seated in a sound-attenuated chamber. All the subjects heard 80 English nonce words through headphones in the PsyScope program on a Macintosh computer. The AE speakers were instructed to write down each word with normal English spelling on the response sheet. The Korean speakers were instructed to write down each word with English spelling for the AE stimuli, and with Korean spelling for the Korean stimuli. Each target word was repeated after a pause of 1000 ms.

4.3. Results

The following figure shows the percentage and frequency of responses for AE flaps by AE and Korean speakers.

(9) Percentage of responses in categorizing the AE flaps

(10) Percentage of responses in categorizing the Korean flaps of \( r(l) = \rightarrow \), others: ‘dr’ or ‘rd’

The AE speakers show a strong tendency to categorize the flaps as an AE alveolar stop ("d" or "t"), "d" in particular, while the Korean speakers most often categorize the AE flaps as an AE liquid ("r" or "l"), "r" in particular, the Korean flaps as the Korean liquid, "l". These results are consistent with those in previous perceptual experiments involving flaps (Monnet & Freeman 1972; Price 1981; Kim & Park 1995), indicating that the L1 phonological system strongly affects subjects’ performance.

The Korean speakers’ responses for the AE flaps are particularly interesting. The Koreans could distinguish between the AE flaps and AE liquids in the previous discrimination tests, showing that they were sensitive to the acoustic difference between AE \( r \) and \( t \). They were also familiar with the flapping of AE alveolar stops since they had studied AE for more than ten years in Korea, and were continuing their study of AE in the U.S. at the time they participated in the experiment. In addition, in a brief informal interview after the experiments, they told the experimenter that they were familiar with AE flaps used for "t" or "d". They, however, still show a strong tendency to categorize AE flaps as AE liquids when the AE flaps were presented in an isolated nonce word.

As expected, more "d" responses than "t" responses are found for the flaps. This result can be explained by the acoustic properties of flaps. In other words, flaps are acoustically more similar to "d" than "t" in terms of closure.
duration and VOT as was shown in the previous chapter.

5. Discussion

In the production experiment, the AE and Korean flaps did not show systematic acoustic differences. There are no statistically significant differences between the AE and Korean flaps in terms of the percentage of closure duration and the percentage of voicing occurrence. There is only a slight difference in terms of the percentage of burst occurrence. While canonical flaps are not thought to have bursts, in the present study, 25% of AE flaps, and 42% of Korean flaps contain a short burst. Although the Korean flaps manifest more bursts than the AE flaps overall, there is considerable individual variation within the AE group in terms of burst occurrence. One set of the AE subjects has very few bursts, whereas the other AE subjects are very similar to the Korean group. The greater number of bursts in the Korean flaps than in the AE flaps found in the present study is not able to account for the different categorization of the flaps, obstruents in AE, and sonorants in Korean. Obstruents, stops in particular, are sounds involving a closure of the oral tract followed by a sudden burst. Steriade (1993) also indicates that a major phonological difference between obstruents and sonorants is the existence of a release phase for the obstruents, and that stop releases are often accompanied by audible bursts. Thus, a burst can be interpreted as one of the main phonetic features of stops although it is not always required. The results of the present study lead rather in the opposite direction since Korean flaps manifest more bursts although they function as sonorants in the Korean sound system. Therefore, the present study shows that phonetic implementations of sounds are not always reflected in the phonological functions of the sounds.

The overall results of the discrimination experiments show that an abstract phonological system, rather than the numerical values of the acoustic parameters, most strongly affects listeners' perception of sounds. Listeners show difficulty in discriminating between two sounds if the two sounds are not categorized as two separate phonemes in their L1 phonological system. For example, the AE speakers had difficulty with the [r-d] pair in both the AE and Korean stimuli, and the [d-t'] pair in the Korean stimuli. In addition, the Korean speakers' performance was somewhat impaired with the [l-r] pair in the AE stimuli.

The absence or presence of a sound in the L1 phonological system, however, is not able to account for all the perceptual patterns observed. The hypothesis about perceptual difficulty on the basis of the phonological relation of sounds were not verified in the results of the discrimination tasks involving the AE and Korean stimuli. It was predicted that in the Korean stimuli, the AE speakers would have difficulty in discriminating between the sounds in both the [r-t'] and [d-t'] pairs since each pair includes a non-native sound. Although their performance of perception was impaired with the [d-t'] pair, they did not show any difficulty in the [r-t'] pair. Further, it was expected that the AE speakers would have more or less the same degree of difficulty in discriminating the sounds in the [r-d] and [r-t'] pairs since the two sounds in each pair are categorized as one phoneme. Their responses for these two pairs, however, are very contrastive. While they could not discriminate the sounds in the [r-d] pair at all, their performance on the [r-t'] pair was as good as that on the pairs involving a native phonemic contrast.

The hypothesis regarding the Korean speakers' performance of perception was also disconfirmed. That is, the Korean speakers' performance on the [r-t'] pair involving the partially phonemic difference was as good as that on the other pairs involving the full phonemic difference. In addition, the Korean speakers' performance on [l-r] in the AE stimuli was not impaired at all. Therefore, in both the AE and Korean stimuli, the phonemic or phonetic status of the sounds cannot predict the perceptual difficulty.

It seems that listeners do not hear a whole phoneme. They, rather, show sensitivity to a phonetic cue that is used for the contrastive features in their L1. The AE speakers could not detect the acoustic difference between [r] and [d] at all in both the AE and Korean stimuli since in AE the acoustic difference between these sounds is never used to contrast meaning. In other words, the AE speakers could not perceive the difference of closure duration and VOT between these two sounds since closure duration is not a primary cue to distinguish stops in their L1. Although VOT is an important cue to the
voicing contrast of AE stops, the AE speakers are not sensitive to the
difference of VOT values of [r] and [d]. It seems that the AE speakers are
not attuned to the difference of the VOT values if the VOT values of both
sounds are below 30 ms since in AE, the boundary that distinguishes between
voiced and voiceless stops is around 30 ms. The AE speakers also had
difficulty with the [d-t'] pair in the Korean stimuli since [d] and [t'] are
differentiated by the closure duration, and the closure duration is irrelevant in
their L1 phonology. The Korean speakers' difficulty with the [l-r] pair can
also be explained by their L1 phonological system since the acoustic cues to
distinguish between these two sounds, such as lowering F3, are not used to
contrast meaning in their L1.

The AE speakers' high sensitivity to the [r-t'] pair in both the AE and
Korean stimuli is also explained by their sensitivity to a phonetic cue. These
two sounds are acoustically different in terms of both VOT and closure
duration. It seems that AE speakers are able to perceive the difference of
VOT between [r] and [t'] since VOT is related to the phonemic distinction in
AE. In addition, we should reconsider the conception of allophones.
The relationship between allophones and underlying phonemes is not always the
same across pairs of sounds. For example, in the present study, the
relationship between [r] and /d/ is different from that between [r] and /t/;
Whereas /d/ is flapped by just reducing the closure duration, /t/ undergoes an
intermediate stage, neutralization, before the duration is reduced. Thus,
although [r] is an allophone of both /d/ and /t/, the process is different
between the /d/ flapping and the /t/ flapping. It is assumed that this different
process is also related to the different results in the perception of the [r-d]
and [r-t'] pairs. Further, we speculate that the Korean speakers' sensitivity to
the durational difference facilitates the Korean speakers' performance on the
[r-t'] pair.

The overall discrimination results of the present study are also consistent
with findings that have shown that partially contrastive L1 features can be
extended to new classes of sounds in the L2 (Brown 2000). Brown argues
that not phonemes, but features, guide the mapping process between the L2
input and the L1 grammar.

6 Conclusion

The present study empirically shows that AE and Korean flaps are
phonetically very similar although their functions in the phonological
systems of the two languages are crucially different. Thus, the findings of the
present study lead to the conclusion that a unified framework that fails to
make a distinction between phonetics and phonology cannot be maintained. In
addition, it is demonstrated that listeners perceive the flaps differently
according to their L1 backgrounds. That is, they perceive the sounds through
the filter of their L1 phonological system. The phonological or phonetic
status of sounds, however, does not predict the perceptual patterns of the
sounds. Listeners are generally sensitive to phonetic cues to features that are
contrastively used in their L1. While the Korean speakers show sensitivity to
both VOT and closure duration, the AE speakers are only sensitive to the
difference of VOT.

References

Brown C 2000 The interrelation between speech perception and phonological
acquisition from infant to adult In J Archibald (ed.), Second language
acquisition & linguistic theory:14-63 Oxford: Blackwell

Carney, A E, G P Widin, & N F Viezeister 1977 Non-categorical perception of
stop consonants differing in VOT, Journal of the Acoustic Society of America:
62:961-970

Fox, R A & D Terboek 1977 Dental flaps, vowel duration and rule ordering in

Green, D M, & J A Swets 1965 Signal detection theory and psychophysics New
York: Wiley

Han, M-S & R S Weitzman 1970 Acoustic features of Korean /p, t, k/, /p, t, k/
and /p', t', k'/. Phonetics 22:112-128

Ingram, J & S Park 1998 Language, context, and speaker effects in the identification
and discrimination of English /n' and /n/ by Japanese and Korean listeners
Journal of the Acoustic Society of America 103:1161-1174
Multiple Wh-interpretations

Yoshihisa Kitagawa, Dorian Roehrs, and Satoshi Iomioka
Indiana University and University of Delaware

1. Observations and Generalizations

1.1. Illusory Wh-Effects in Japanese

It probably is no exaggeration to say that there is growing concern in the field of Japanese syntax that many important and influential works on island effects in the past might not have been based upon precise empirical observations. For instance, Deguchi and Kitagawa (2002) and Ishihara (2002) challenged the claim that Wh-in-situ in Japanese exhibits Subjacency effects Deguchi and Kitagawa pointed out: (i) that Wh-questions in Japanese are accompanied by Empathic Prosody (henceforth EPD), which consists of an emphatic accent on the focused Wh-phrase followed by deaccenting of all lexical accents up to some COMP, and (ii) that the [+WH] CP at which EPD ends coincides with the scope domain of a Wh-phrase. They reexamined many relevant examples, assigning two patterns of EPD, as in (1).2

---

1. This correlation may not necessarily hold when a certain type of presupposition is involved, which we will not discuss in this work

2. In our examples, an emphatic accent is indicated by X's deaccented strings by ++X++X, and retention of a lexical accent by X's # placed on a sentence means that it is unacceptable with the indicated prosody