Acoustics and Perception of Flaps in American English and Korean

Eun-kyung Sung
(Seoul National University)

Sung, Eun-kyung. (2003). Acoustics and perception of flaps in American English and Korean. *Language Research* 39(4), 933-958.

The goals of this paper are two-fold. The first goal is to explore the relationship between phonetic characteristics and phonological representation, and the second goal is to investigate the role of phonological functions of sounds in perceiving the sounds. The comparison of acoustic properties of flaps between American English and Korean shows that phonetic characteristics of sounds are not able to account for phonological functions of the sounds. The results of discrimination experiments reveal that although the L1 phonological system strongly affects perception, perceptual patterns cannot be explained by the phonological or phonetic status of sounds in L1. The subjects are sensitive to phonetic cues to contrastive features in their L1. That is, the American English speakers are sensitive to a VOT cue, whereas the Korean speakers are sensitive to both VOT and duration cues.

Key words: phonetic, phonology, acoustics, perception, discrimination, VOT, closure duration

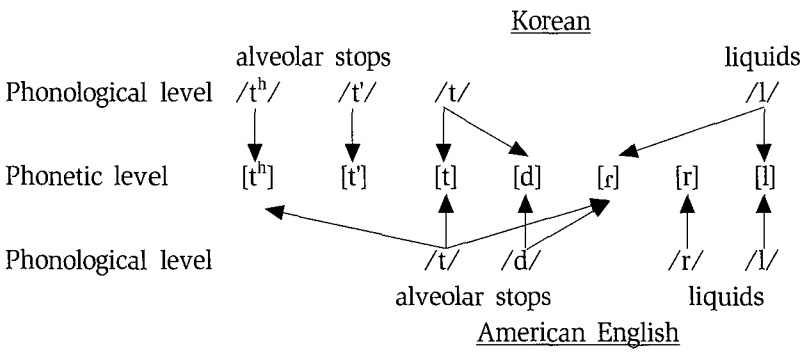
1. Introduction

This paper investigates the relationship between phonetic manifestations and underlying representations of sounds, and the effect of L1 phonological system in perception. These issues are addressed on the basis of in-depth production and perception studies of flaps in American English (AE) and Korean.

Flaps in AE and Korean provide an appropriate case in order to explore these issues 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 /l/ found when /l/ comes between vowels, e.g., *puli* [pu.ɾi] 'beak'.

It contrasts with the alveolar lax stop /t/, as we can see in a minimal pair, e.g., *muli-ta* [mu.ɾi.da] 'unreasonable' vs. *muti-ta* [mu.di.da] 'blunt, dull'. It should be noted that Korean /t/ further contrasts with /t'/ and /t^h/, alveolar tense and aspirated stops, respectively, e.g., *sato* [sa.do] 'an apostle' vs. *sat'o* [sa.t'o] 'mayor (archaic)' vs. *sat^{h}o* [sa.t^{h}o] 'sandy soil'. By contrast, the flap in AE is an allophone of the alveolar stops /d/ and /t/, e.g., *latter* and *ladder*, both pronounced [læɾɚ]. It contrasts with the liquids /l/ and /r/, e.g., *Betty* vs. *berry* vs. *belly*. Figure 1. shows the relationship of flaps and the main variants of liquid and alveolar stop sounds in Korean and AE.[1]

Figure 1. Partial Phonological Status of Flaps, Alveolar Stops and Liquids in Korean and AE[1]



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.[2] The effects of the phonological function of flaps in perception are investigated through two discrimination experiments involving phonetically relevant sounds in Korean and AE.

1) Coda neutralization in the Korena phonology is not considered here.

2) In this study, the acoustic properties of AE and Korean flaps are compared using three acoustic measures, closure duration, existence of voicing, and occurrence of burst. These three measures are used in many previous studies (Zue & Laferriere, 1979; Price, 1981; Rimac & Smith, 1984; Spiegel, 1987; de Jong, 1998) in order to examine acoustic characteristics of flaps and stops.

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., CVɾV. 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 vocalic environments 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) (see Appendix 1). In total, 240 items for each group (12 tokens×2 repetitions×10 subjects) were analyzed.[3)]

The AE speakers were given the words in English orthography, and the Korean speakers in Korean orthography when they read the test items. The test items were presented in random order. The productions were recorded onto a TEAC RW-800 CD recorder through a lavalier microphone attached to a subject's collar in a sound-attenuated room at UD.

2.3. Results

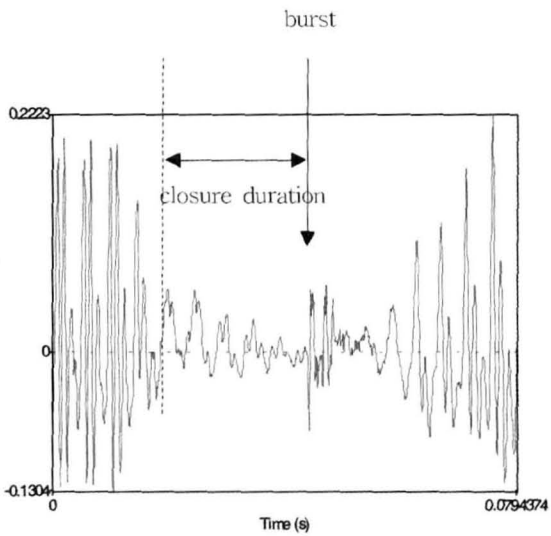
2.3.1. Closure Duration

Closure duration was determined by measuring the time in milliseconds (ms) from a waveform with the assistance of a spectrogram. The duration

3) The measures discussed in this study were obtained from waveforms and broad-band (200Hz) spectrographic displays generated by Praat 3.9.3, speech analysis software on a PC.

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. Figure 2. shows an example of an AE flap token with a burst and closure.

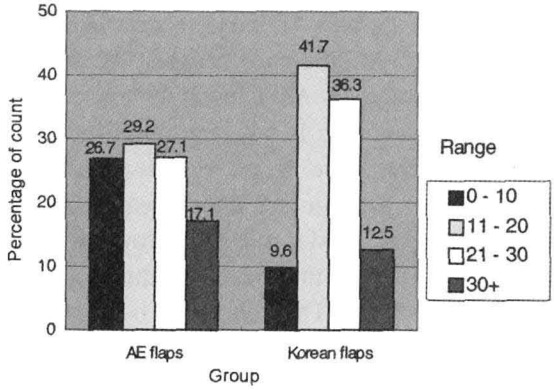
Figure 2. A Waveform of a Flap in AE Speaker(AF2)'s Production of *motor*[4)]



The following histogram shows the distribution of flap closure duration in AE and Korean. The *x*-axis shows ranges of closure duration, pooled into 10 ms ranges. The *y*-axis indicates percentage of the count in each range. Flaps whose duration was more than 30 ms were merged into one range (30+ ms) since the percentage of the count of those flaps was not high in either AE or Korean.

4) AM = American English male speaker, AF = American English female speaker, KM = Korean male speaker, KF = Korean female speaker

Figure 3. Distribution of Flap Closure Duration in AE and Korean



The distribution patterns for flap closure duration show that in both AE and Korean, the closure duration of more than 80% of the flap tokens is in the 0 - 30 ms range (83% of AE flaps and 88% of Korean flaps). The distribution of the AE flaps, however, shows a flatter shape than that of the Korean flaps. That is, the AE flaps are almost equally distributed in the three ranges (0-10, 11-20, and 21-30 ms), while the Korean flaps have a peak in the range of 11-20 ms. The AE flaps have more tokens that fall into the 0-10 ms range than the Korean flaps. The Korean flaps, instead, have more tokens whose closure duration falls between 11 to 30 ms than the AE flaps. There is not any difference between the AE and Korean flaps in the range of more than 30 ms closure duration.[5)]

The count of the number of flaps in each range for each group was submitted to a chi-square test to determine whether or not two variables (group and range) were statistically independent or related to each other. The chi-square test of independent variables shows the existence of a significant relationship between the variables of group and range [$x(3) = 29.5$, $p < .001$]. Thus, the distribution patterns of closure duration are statistically different between AE and Korean flaps. Individual data, however, indicate that two subjects (AM1 & AM4) produced very short flaps whose closure duration is less than 10 ms for more than 50% of the time (54% & 59% respectively), and the flap production of AM1, in particular, did not

5) The shortest flaps observed in this experiment do not have any closure duration (0 ms), and the longest flap observed shows 58 ms closure duration.

show any closure duration for 54% of the time. All other AE speakers except these two speakers produced short flaps (0-10 ms) for less than 30% of the time. It seems that the two AE subjects' data (AM1 & AM4) contributed the most to the observed difference between AE and Korean flaps.

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 [$t(18) = -.818$, $p > .05$]. 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.

Figure 4. Boxplots of Mean Closure Duration in AE and Korean Flaps

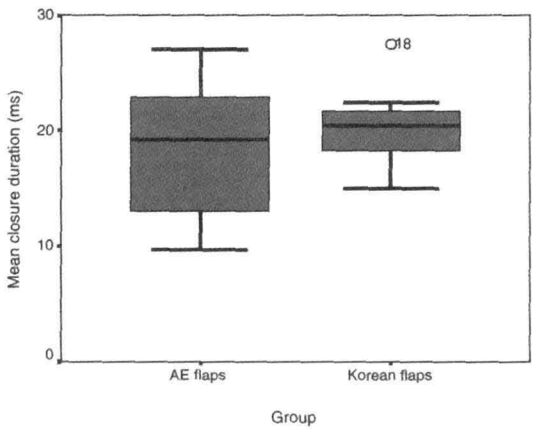


Figure 4. shows that the Korean flaps have a more compact shape than the AE flaps, and the AE flaps have more variability than the Korean flaps. The large variability in the AE flaps came from the subgroup of AE speakers who often produced very short flaps.

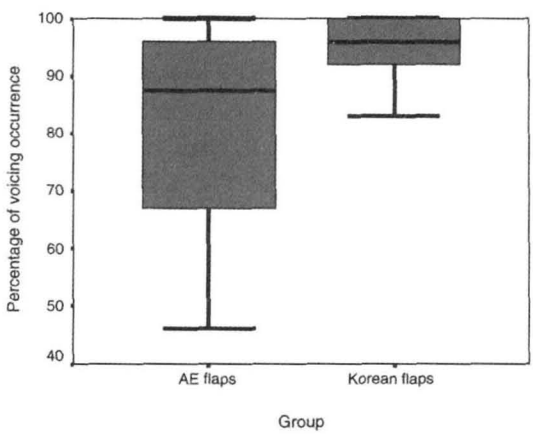
2.3.2. Voicing

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 as 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 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 voicing in AE flaps (86%) given by Fox and Terbeek (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%). The other two speakers also produced flaps with voicing more than 80% of the time. 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 [$z = -1.468$, $p > .05$]. The following boxplots show the percentage of voicing occurrence for the two groups.

Figure 5. 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

In cases where the waveforms or spectrograms manifested a closure interval, they showed or did not show a release burst. The release burst is an aperiodic signal portion preceding the first glottal pulse of the following vowel. The burst is created as air pressure built up by the closure of the oral cavity is suddenly released. 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. It seems that these four AE speakers' production makes the AE flaps different from the Korean flaps in terms of bursts. 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. In addition, the release burst is very rare in their production. 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 [$z = -2.164$, $p = < .05$]. The Korean flaps are more often accompanied by release bursts than the AE flaps. The following boxplots show the percentage of burst occurrence.

Figure 6. also shows 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.

Figure 6. Boxplots of Percentage of Burst Occurrence in AE and Korean Flaps

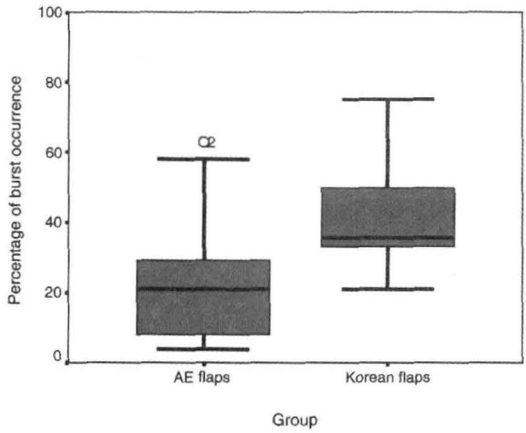


Table 1. Mean Values of Flaps in AE and Korean (Standard Deviation in parentheses)

	AE	Korean
Closure Duration	19 ms (13)	20 ms (9)
Voicing	82% (19)	95% (6)
Burst	25% (6)	42% (4)

2.4. Discussion

The current production experiment investigates whether or not there are significant differences between AE and Korean flaps in terms of closure duration, voicing and burst. Although in general AE flaps are acoustically similar to Korean flaps, some differences were found. The differences found between the AE and Korean flaps include the following: (1) the distribution of flap closure duration shows that the AE flaps are quite equally distributed across three ranges (0-10, 11-20 and 21-30 ms), while more Korean flaps fall into two ranges (11-20 and 21-30 ms) than into the 0 -10 ms range; (2) the Korean flaps are more often accompanied by a release burst than the AE flaps.

AE flaps, however, do not differ from Korean flaps in terms of mean closure duration and percentage of voicing occurrence. It should also be noted that a subgroup of the AE speakers (three speakers, AM1, AM4 &

AF1, in particular) are different from the other AE speakers since they often produced very short flaps whose mean duration is less than 15 ms. Their flaps are voiced more than 95% of the time, and are not accompanied by a burst more than 85% of the time. There is no such variability in the Korean flaps.

Further, the fine acoustic differences between the AE and Korean flaps found in the present study are not able to explain the different categorization of flaps in the phonological systems of the two languages. For example, although there is a slight difference in the patterns of closure duration, more than 80% of both the AE and Korean flaps have a short closure duration that falls in the range of 0-30 ms. The question is whether or not the 10 ms or 20 ms difference of closure duration is perceptually important. Listeners may not perceive that small difference such as 10 or 20 ms since 30 ms itself is very short. Thus, it would be difficult to argue that flaps in the two languages are categorized as two contrastive classes because of that small difference.

Another difference found between the AE and Korean flaps is that the Korean flaps are more often accompanied by a burst than the AE flaps. Since the burst is one of the main phonetic characteristics of obstruents, stops in particular, the difference in the burst occurrence is not relevant to the different categorization of flaps, obstruents in AE and sonorants in Korean. Thus, the raw phonetic characteristics of flaps are not able to explain the functioning of the sounds in the abstract phonological systems of the two languages.

Furthermore, a unified module which does not distinguish between phonetics and phonology, cannot account for the results of the present study. Flemming (1995, 2001). Kirchner (1997), and Steriade (2000) argue that phonetics and phonology are integrated into a single grammar instead of separate components by including more phonetic details into phonological representations. In the present study, however, flaps are categorized as two distinctive phonemes in AE and Korean despite the strikingly similar phonetic properties. It is assumed that there is a distinction between phonetic and phonological representations although there is a considerable amount of interface between phonetics and phonology. The following section shows how the abstract functions of the flaps in the phonological system affect perception of the sounds.

3. Discrimination Experiments

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. The two phones may also be categorized as the same phoneme 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. Experiment 2: Discrimination of AE Flaps, Alveolar Stops and Liquids

3.2.1. Methodology

3.2.1.1. 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). The AE subjects were undergraduate students who were in an introductory linguistics course at the University of Delaware (UD). The age-range of both the AE and Korean speakers was from 19 to 35 years old. All subjects reported having normal hearing. AE speakers who had

significant experience in Spanish (more than 4 years) were excluded from the subjects.[6]

3.2.1.2. 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 speakers. Each pair consisted of two very short pieces of words produced by the two speakers with standard pronunciation. Each token included an intervocalic flap [ɾ], an alveolar stop [d] or [t^h], or a liquid [r] or [l] between the preceding vowel [u] and the following vowel [i], e.g., [uɾi], [udi]. In the stimuli including [ɾ], [r] and [l], the preceding vowel was always stressed. In the stimuli including [d] and [t^h], however, the following vowel was always stressed, otherwise the sounds in question would be pronounced as a flap. In AE, voiceless alveolar stops are aspirated when followed by a stressed vowel. In each token, 40 ms of each vowel was included, and the other part of the vowel was cut off.

The reasons for using the truncated stimuli are 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 (ɾ, d, t^h, r, l)×4 repetitions×2 speakers], and 40 different pairs [5 contrasts (ɾ-d, ɾ-t^h, ɾ-l, ɾ-r, r-l)×4 repetitions×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. Since every pair consisted of two items produced by two speakers, no item was paired with itself. There was a 1500 ms interval between the two members of a pair, and a 3000 ms interval between the pairs.

The pairs in the present experiment were chosen on the basis of the phonological or phonetic status of the phones in each pair. The following table shows the phonemic relations of pairs of segments in AE stimuli.

---

6) Spanish has a flap which is acoustically very similar to AE flaps, and the Spanish flap is spelled as "r". Thus, the subjects with significant Spanish experience may be influenced by the Spanish flap in the perception experiment.

Table 2. Phonemic Status of the Pairs of Segments in the AE Stimuli

AE pairs	Korean	AE
[ɾ - d]	**phonemic** difference (different phonemes)	**allophonic** difference (same phoneme)
[ɾ - t^h]	**phonemic** difference (different phonemes)	**allophonic** difference (same phoneme)
[ɾ - l]	**partially phonemic** difference (same phoneme, but phonemic difference intervocalically in [ɾ-ll])	**phonemic** difference (different phonemes)
[ɾ - r]	**non-native** difference ([r]; non-native phone)	**phonemic** difference (different phonemes)
[l - r]	**non-native** difference ([r]; non-native phone)	**phonemic** difference (different phonemes)

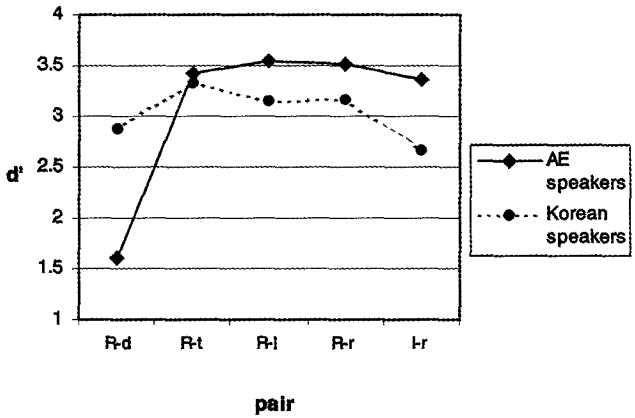
All the subjects were tested individually in a sound-attenuated chamber. They listened to 160 pairs of short pieces of naturally produced AE words through the Psyscope program on a Macintosh computer. An AX paradigm was employed with truncated stimuli to assess the subjects' ability to acoustically discriminate two items in each pair. The stimuli were presented through headsets in the following patterns: AA, BB, AB, and BA. The subjects indicated whether the two items were repetitions of the same word or pronunciations of two different words by pressing the *f* or the *j* button, respectively.

3.2.2. 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 figure displays the mean discrimination sensitivity for each pair by the AE and Korean speakers.

Whereas the AE speakers show variation among the pairs, between the [ɾ-d] pair and all other pairs in particular, the Korean speakers performance of perception is more or less consistent across the pairs. In general, the AE speakers performed well on all the pairs except the [ɾ-d] pair, and

Figure 7. Distribution of Mean Sensitivity (d') Functions for the AE Stimuli (R = ɾ)



individual data show that all the AE speakers uniformly performed poorly on the [ɾ-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 [ɾ-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: [ɾ-d], [ɾ-t^h], [ɾ-l], [ɾ-r], [l-r]) as the within-subjects factor, the results confirmed that there was a significant effect of pair [$F(4, 152) = 31.04$, $p < .001$]. Group was not significant [$F(1, 38) = .154$, $p > .05$], but there was a significant pair x group interaction [$F(4, 152) = 19.92$, $p < .001$]. Group as a whole did not show a significant effect since although the AE speakers performed much worse than the Korean speakers on the [ɾ-d] pair, the AE speakers performed better than Korean speakers on all other pairs. Thus, the group effect on one pair, [ɾ-d], was obscured by the opposite group effect on the other pairs, [ɾ-t^h], [ɾ-l], [ɾ-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 [$F(4, 76) = 51.11$, $p < .001$]. Bonferroni post hoc tests showed that the AE speakers' mean d' scores of the [ɾ-d] pair was significantly lower than those of all other pairs ($p < .001$ 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 < .01]. Bonferroni post hoc tests showed a marginal difference between [ɾ-r] and [l-r] (p = .40 < .05). That is, Korean speakers' performance on [l-r] was a little worse than that on [ɾ-r]. 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, [ɾ-d] and [l-r] [F(1, 38) = 24.274, p < .001; F(1, 38) = 9.614, p = .004 < .005, respectively]. The performance of the AE speakers on the [ɾ-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 [ɾ-t^h] pair.

3.3. Experiment 3: Discrimination of Korean Flaps, Alveolar Stops and Liquids

3.3.1. Methodology

3.3.1.1. Subjects

The same subjects who performed Experiment 2 participated in Experiment.

3.3.1.2. Stimuli and Procedure

The stimuli were constructed using edited natural speech which was derived from the productions of two female Korean speakers with standard pronunciation. Each pair consisted of two very short pieces of words produced by the two speakers. The short pieces of words included intervocalic flap [ɾ], alveolar lax stop [d], alveolar aspirated stop [t^h], alveolar tense stop [t'], and geminate lateral [ll] between the preceding vowel [u] and the following vowel [i], e.g., [uɾi], [udi]. The same vowel editing procedures as those used for the AE stimuli were employed for the Korean stimuli. Thus, only 40 ms of each vowel was included in each token.

The stimuli consisted of 80 test items and 80 fillers. The test items contained 40 same items [5 consonants (ɾ, d, t', t^h, ll)×2 speakers×4 repetitions] and 40 different items [5 contrasts ([ɾ-d], [ɾ-t'], [ɾ-t^h], [ɾ-ll], [d-t'])× 2 speakers×4 repetitions]. In the same pairs, no item was paired with itself since every pair consisted of two items produced by two speakers.

The fillers also consisted of 40 same items and 40 different items, and each item in a pair included [b], [p^h] [p], [s] or [s] between vowels. There was a 1500 ms interval between the two tokens in a pair, and a 3000 ms interval between the pairs.

The same procedures as those used for the previous perception experiment involving the AE stimuli were employed for this experiment. The following table shows the phonemic and phonetic status of the phones in each pair used in the Korean stimuli.

Table 3. Phonemic Status of the Pairs of Segments in the Korean Stimuli

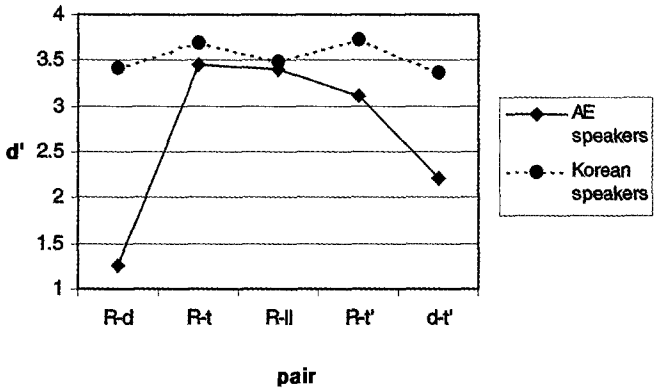
Korean pairs	Korean	AE
[ɾ - d]	**phonemic** difference (different phonemes)	**allophonic** difference (same phoneme)
[ɾ - t^h]	**phonemic** difference (different phonemes)	**allophonic** difference (same phoneme)
[ɾ - t']	**phonemic** difference (different phonemes)	**non-native** difference ([t']; non-native phone)
[ɾ - ll]	**partially phonemic** difference (same phoneme, but phonemic difference intervocalically)	**phonemic** difference (different phonemes)
[d - t']	**phonemic** difference (different phonemes)	**non-native** difference ([t']; non-native phone)

3.3.2. Results

The following figure displays the mean discrimination sensitivity for each pair by the AE and the Korean speakers. 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 [ɾ-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 [ɾ-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: [ɾ-d], [ɾ-t'], [ɾ-t^h], [ɾ-ll], [d-t']) as the within-subjects factor. The results confirmed that there was a significant effect of pair [$F(4, 152)$ =

Figure 8. Distribution of Mean Sensitivity (d') Functions for the Korean Stimuli (R = ɾ)



44.69, $p < .001$], group [$F(1, 38) = 41.95$, $p < .001$], and pair x group interaction [$F(4, 152) = 28.75$, $p < .001$]

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 [$F(4, 76) = 47.49$, $p < .001$]. Bonferroni post hoc tests showed that the AE speakers' mean d' scores of the [ɾ-d] pair were significantly lower than those of all the other pairs ($p < .001$ for the comparisons of [ɾ-d] vs. [ɾ-t'], [ɾ-d] vs. [ɾ-t^h], [ɾ-d] vs. [ɾ-ll]; $p < .01$ for the comparison of [ɾ-d] vs. [d-t']). In addition, the mean d' scores of the [d-t'] pair was significantly lower than those of three other pairs, [ɾ-t'], [ɾ-t^h] and [ɾ-ll] ($p < .001$ for all comparisons), and significantly higher than those of the [ɾ-d] pair ($p < .01$). On the other hand, for the Korean speakers, the results of a one-way repeated measures ANOVA revealed that pair was marginally significant [$F(4, 76) = 3.85$, $p = .007 < .05$]. The follow-up Bonferroni post hoc test, however, did not reveal any significant difference between pairs ($p > .05$ 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, [ɾ-d], [ɾ-t'], and [d-t'] [$F(1, 38) = 68.17$, $F(1, 38) = 17.76$, $F(1, 38) = 24.62$, respectively, $p < .001$ for all three pairs). The AE speakers' performance on [ɾ-t^h] and [ɾ-ll] 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, [ɾ-d], [ɾ-tʰ], and [ɾ-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 ($p > .05$ for all three pairs). Their sensitivity to the [ɾ-d] pair was low in both the AE and Korean stimuli, and their sensitivity to the [ɾ-tʰ], and [ɾ-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 ($p = .003 < .005$ for [ɾ-d], $p = .024 < .05$ for [ɾ-tʰ], and $p = .023 < .05$ for [ɾ-l]). Their sensitivity to these pairs in the Korean stimuli was higher than in the AE stimuli.

4. 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 the mean values of closure duration and in 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 abstractphonological system, rather than the numerical values of the acousticparameters, 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 [ɾ-d] pair in both the AE and Korean stimuli, and the [d-t'] pair in the Korean stimuli. In addition, the Korean speakers performance of perception 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 hypotheses 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 [ɾ-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 [ɾ-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 [ɾ-d] and [ɾ-t^h] 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 [ɾ-d] pair at all, their performance on the [ɾ-t^h] 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 [ɾ-ll] 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 [ɾ-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. In

other words, the performance on the pairs involving an allophonic difference is not necessarily worse than that on the pairs involving a phonemic difference. Additionally, the performance on the pairs involving a non-native difference is not always worse than that on the pairs involving an allophonic difference. Further, there is no difference in the performance between the pairs involving a partially phonemic relation and the pairs involving a fully phonemic relation.

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 [ɾ] 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 [ɾ] 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 (see Appendix 2 for the acoustic properties of items used in the discrimination experiments).

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 [ɾ-t^h] 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 [ɾ] and [t^h] 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 [ɾ] and /d/ is different from

that between [ɾ] 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 [ɾ] 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 [ɾ-d] and [ɾ-t^h] pairs.

The AE speakers' high sensitivity to the Korean [ɾ-t'] pair can be explained in two ways. First, although the AE speakers are insensitive to the durational difference in general, they may perceive the large difference of the duration. In the present stimuli, the closure duration values of [t'] are almost four times those of [ɾ]. The question of how much length difference affects the AE speakers' perception, however, is left to future study. Second, the AE speakers may perceive phonetic characteristics of Korean [t'] involving the [constricted] gesture. Although, in general, it is assumed that the [constricted] gesture is not employed in AE phonology, the gesture is not completely absent. According to Zue and Laferriere (1979), in the nasal-release context (i.e., VCn), the phonetic realization of [t] in *sweeten* is different from that of [d] in *Sweden*. That is, [t] is realized as a glottal stop by forming the constriction at the glottis, whereas [d] is released through the nasal cavity. Presumably, the experience of glottalization as the cue to distinguish /t/ from /d/ in this context, may help the AE speakers perceive the [constricted] gesture in the Korean tense alveolar stop [t'], enhancing their perception of the [ɾ-t'] pair.

As noted above, the performance on the [ɾ-d] pair reveals the different perceptual sensitivity to the durational difference between the AE and Korean speakers. Further, it is assumed that the Korean speakers' sensitivity to the durational difference facilitates the Korean speakers' performance on the [ɾ-r] 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). That is, the presence or absence of a feature in the L1 can explain the different performance on L2 contrasts among different language groups. For example, Chinese speakers' ability to discriminate /l/ and /r/ can be attributed to the presence of the feature [coronal] in their L1, whereas Korean and Japanese speakers' inability to perceive this contrast can be explained by the absence of the feature in their L1s. Since all three languages lack the /l/-/r/ phonemic contrast, the results cannot be understood as a direct

consequence of L1 phoneme inventories. Brown argues that not phonemes, but features, guide the mapping process between the L2 input and the L1 grammar.

5. 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, (pp.* 14-63). Oxford: Blackwell.

Flemming, E. (1995). *Auditory Representations in Phonology.* Doctoral dissertation, UCLA.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18, 7-44.

Fox, R. A. and D. Terbeek. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics* 5, 27-34.

Green, D. M., and J. A. Swets. (1966). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Kent, R. and C. Read. (1992). *The Acoustic Analysis of Speech.* San Diego:

Singular Publishing.

Kirchner, R. (1997). Contrastiveness and faithfulness. *Phonology* 14, 83-111.

Klatt, D. (1976). Linguistic uses of segmental duration in English; Acoustic and perceptual evidence. *Journal of the Acoustic Society of America* 59, 1208-1221.

Lisker, L. (1978). Rapid vs. Rabid: A catalogue of acoustic features that may cue the distinction. *Status Report on Speech Research* 54, 127-132.

Macmillan, N. A., and C. D. Creelman. (1991). *Detection Theory: A users' guide.* New York: Cambridge University Press.

Macmillan, N. A. (2002). Signal Detection Theory. In H. Pashler & J. Wixted eds., *Stevens Handbook of Experimental Psychology, Vol 4.,* Methodology in experimental psychology, 43-90.

Price, P. (1981). *A Cross-linguistic Study of Flaps in Japanese and in American English.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Steriade, D. (1993). Closure, release and nasal contours. In M. Huffman and R. Krakow eds., *Nasality,* (pp. 401-470). San Diego: Academic Press.

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. In J. Pierrehumbert and M. Broe eds., *Acquisition and the Lexicon: papers in laboratory phonology V,* (pp. 313-334.) Cambridge: Cambridge University Press.

Zue, V. (1976). *Acoustic Characteristics of Stop Consonants.* Indiana University Linguistics Club.

Zue, V. and M. Laferriere. (1979). Acoustic study of medial /t, d/ in American English. *Journal of the Acoustic Society of America 66*(4), 1039-1050.

Appendix 1:

Word Lists for AE Speakers (a) and Korean Speakers (b), and Carrier Phrases (c) in the Production Experiment

(a)	AE		(b)	Korean	
	meaty	[miɾi]		[miɾi]	'in advance'
	potty	[pʰaɾi]		[pʰaɾi]	'fly'
	moody	[muɾi]		[muɾi]	'herd'
	muddy	[m∧ɾi]		[məɾi]	'head'
	motive	[moɾɪv]		[moɾi]	'chasing'
	Peter	[pʰiɾə]		[piɾə]	'to beg'
	potter	[pʰaɾə]		[pʰaɾə]	'to dig'
	tutor	[tʰuɾə]		[tuɾə]	'to put'
	motor	[moɾə]		[poɾə]	'to see'
	butter	[b∧ɾə]		[pəɾə]	'to earn'
	motto	[maɾo]		[kaɾo]	'width'
	ghetto	[gɛɾo]		[sɛɾo]	'length'

(c) Carrier phrases

AE: Please say _____ several times.

Korean: cigum pute _____ seben malhaseyo.

'Now say _____ three times.'

Appendix 2:

Acoustic Properties of the Items Used in the Perception Experiments AE stimuli (a) and Korean stimuli (b)

(a) Acoustic Properties of the Items in the AE Stimuli (Flaps and Stops)

AE Item	Closure Duration (ms)	VOT (ms)
[uɾi]	16	2
[uɾi]	19	0
[uɾi]	24	0
[uɾi]	33	1
[udi]	82	19
[udi]	83	23
[udi]	79	17
[udi]	81	13
[ut^{h}i]	80	120
[ut^{h}i]	90	123
[ut^{h}i]	135	71
[ut^{h}i]	95	70

(b) Acoustic Properties of the Items in the Korean Stimuli (Flaps and Stops)

Korean Item	Closure Duration (ms)	VOT (ms)
[uɾi]	20	9
[uɾi]	31	2
[uɾi]	34	5
[uɾi]	28	8
[udi]	46	15
[udi]	46	18
[udi]	53	18
[udi]	35	24
[ut^{h}i]	139	60
[ut^{h}i]	151	59
[ut^{h}i]	227	58
[ut^{h}i]	243	53
[ut'i]	97	18
[ut'i]	94	22
[ut'i]	99	16
[ut'i]	93	10

Sung, Eun-kyung
Department of English Language Education
Seoul National University
San 56-1, Sillim-dong, Gwanak-gu,
Seoul 151-742, Korea
E-mail: eunkyungsung@hanmail.net

Received: Aug. 29, 2003
Revised version received: Nov. 1, 2003
Accepted: Nov. 7, 2003