「언어연구」, 제21권(2004), 75～96

The Effect of L1 Phonological System in Perception

Eun-Kyung Sung
(Cyber University of Foreign Studies)

Sung, Eun-Kyung. 2004. The Effect of L1 Phonological System in Perception. *Linguistic Research* 21, 75-96. The goal of this paper is to investigate the role of phonological functions of sounds in perceiving the sounds. The results of discrimination and categorization experiments reveal that the L1 phonological system strongly affects perception. In addition, the phonological processing is more involved in the categorization tasks than in the discrimination tasks. The discrimination experiments also show that 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. **(Cyber University of Foreign Studies)**

Keywords L1 phonological system, perception, categorization, discrimination, vot, duration

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. A number of cross-language studies have revealed language-specific patterns of perception by examining American English (AE) /r/ and /l/ (Goto, 1971; Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura, 1975; MacKain, Best & Strange, 1981; Polka & Strange, 1985), voice onset time (VOT) of stop consonants (Lisker & Abramson, 1964; Williams, 1977; Flege & Eefting, 1988; Bohn & Flege, 1993), and vowels (Beddor & Strange, 1982; Gottfried, 1984; Flege, Munro & Fox, 1994).

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 Korean and American English (AE) provide an appropriate case since flaps are differently categorized in the two languages. This paper investigates the ways in which different phonological functions of the flaps in the two languages affect the perceptual patterns of AE and Korean speakers.

Flaps in Korean and AE are phonetically very similar. In a pilot study, I found that when the flaps from the two languages were put into the same vowel contexts, neither American nor Korean listeners could discriminate between them. In fact, since the flaps in the two languages are so similar, a frequently used technique in teaching AE flaps to Korean students is to focus attention on the physical similarity between the flaps in the two languages. However, 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^ho* [sa.t^ho] '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 [læɾər]. 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.

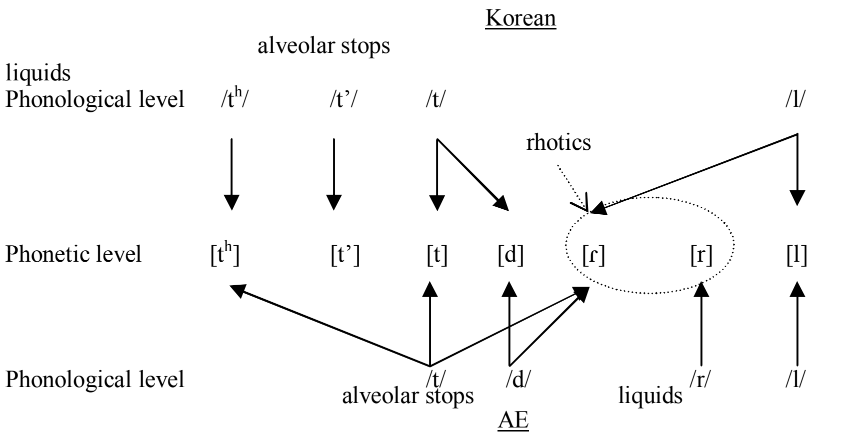

Figure 1. Partial phonological status of flaps, alveolar stops and liquids in Korean and AE

2. Discrimination Experiments

2.1 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

2.2 Experiment 1: Discrimination of American English Flaps, Alveolar Stops and Liquids

2.2.1 Methodology

2.2.1.1 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.

2.2.1.2 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) x 4 repetitions x 2 speakers], and 40 different pairs [5 contrasts (ɾ-d, ɾ-t^h, ɾ-l, ɾ-r, r-l) 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. 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.

Table 1. 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.

2.2.2 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.

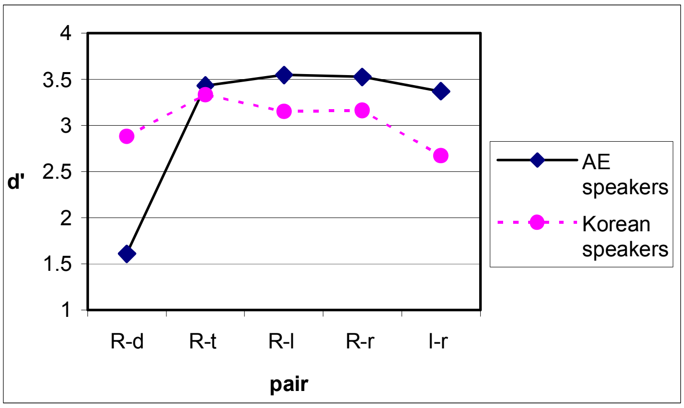

Figure 2. Distribution of mean sensitivity (d') functions for the AE stimuli (R = ɾ)

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 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.

2.3 Experiment 2: Discrimination of Korean Flaps, Alveolar Stops and Liquids

2.3.1 Methodology

2.3.1.1 Subjects

The same subjects who performed Experiment 1 participated in Experiment 2.

2.3.1.2 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) x 2 speakers x 4 repetition] and 40 different items [5 contrasts ([ɾ-d], [ɾ-t'], [ɾ-t^h], [ɾ-ll], [d-t']) x 2 speakers x 4 repetition]. 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 2. 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)

2.3.2 Results

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

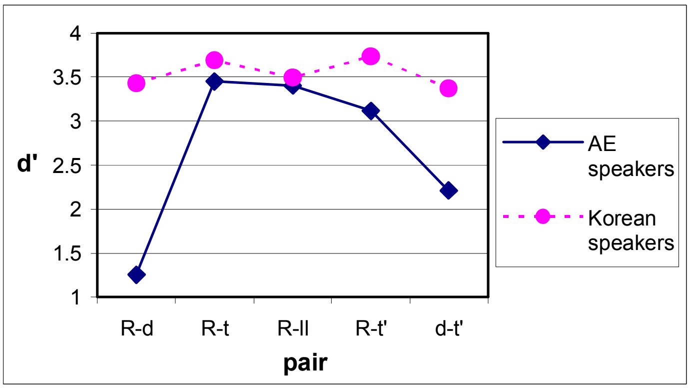

Figure 3. Distribution of mean sensitivity (d') functions for the Korean stimuli(R = ɾ)

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 [$F(4, 152) = 44.69$, $p < .001$], 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-t^h], 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-t^h], 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-t^h], and $p = .023 < .05$ for [ɾ-l]). Their sensitivity to these pairs in the Korean stimuli was higher than in the AE stimuli.

3. Categorization experiments

3.1 Experiment 3: Categorizing American English Flaps, Alveolar Stops and Liquids

3.1.1 Predictions

The AE speakers would categorize the AE flaps as alveolar stops, whereas the Korean speakers would categorize the flaps as liquids. In addition, the influence of the L1 phonological system would be greater in the categorization tasks than in the discrimination tasks.

3.1.2 Methodology

3.1.2.1 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. All of them participated in the present experiments after they completed the previous discrimination tasks. 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 the Spanish spelling. This tendency was, in fact, observed in a pilot study. Another six AE speakers were recruited and they participated in the discrimination and categorization experiments.

3.1.2.2 Stimuli

The stimuli in the present study consisted of 40 test words and 40 fillers that were all nonce words in AE. The stimuli were produced by a female AE speaker with standard pronunciation. Nonce words were used to prevent the subjects' perceptions from being influenced by their familiarity with particular lexical items. The 40 test words included AE flap [ɾ], alveolar stops [d] and [t^h], and liquids [r] and [l] between various vowels, e.g., [niɾo], [fot^hu]. The preceding vowels were chosen among /i, a, o, u/, and the following vowels among /i, o, u/. The vowel [a] was excluded from the following vowels since words which end with [a] are very rare in AE. Approximately the same number of each vowel was used in both preceding and following vowels. 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 [t^h] were produced between a preceding unstressed vowel and a following stressed vowel. The AE liquids [r] 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. The following are the stimuli used in this experiment.

All the AE stimuli were recorded onto the TEAC RW-800 CD recorder, and the recordings were digitized at 44.1 kHz sampling frequency and 16 bit resolution using the EAC software program. Each nonce word was saved as a separate AIFF file using the Praat program, version 3.9.3. The sound files were later converted using the SoundApp program and transferred to the PsyScope program, version 2.5.1. All the items were automatically randomized for each subject.

3.1.2.3 Procedure

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. They were also told that they were listening to the AE stimuli and they were instructed to write down each word with normal English spelling on the response sheet. All the words were randomized for each subject. Each target word was repeated after a pause of 1000 ms. The subjects heard the next item by clicking the mouse, and there was no fixed intertrial interval. After the experiments ended, the subjects had a brief informal interview with the experimenter about whether the task was difficult and whether they were influenced by foreign language spellings.

3.1.3 Results

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

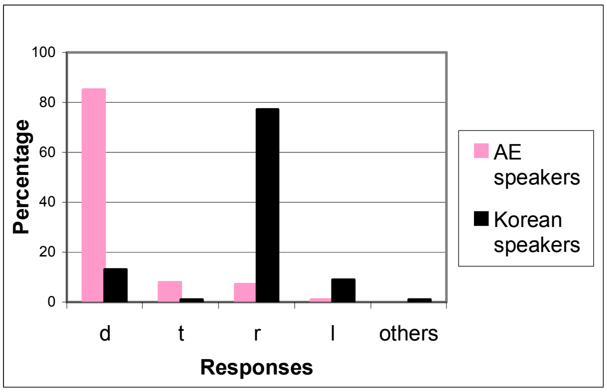

Figure 4. Percentage of Responses in Categorizing AE Flaps

The AE speakers show a strong tendency to categorize the AE flaps as an AE alveolar stop ("d" or "t") (93%), "d" in particular (85%), while the Korean speakers most often categorize the AE flaps as an AE liquid ("r" or "l") (86%), "r" in particular (77%). 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 are particularly interesting. The Korean speakers 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 [ɾ] and [r]. 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. Thus, the differences in phonetic cues and knowledge of the AE flapping rule do not prevent the Korean subjects from categorizing the AE flaps as “r”.

It seems that the two tasks, discrimination and categorization, demand different levels of perceptual processing, and phonological processing is employed more in the categorization task. The discrimination tasks, with the AX paradigm in particular, minimize memory load since the information from only two successive intervals needs to be stored. The categorization tasks, by contrast, involve matching a phone to internally stored abstract representations, and impose greater memory loads than discrimination tasks (Carney, Widin & Viemeister, 1977; Ingram & Park, 1998).

As expected, both the AE and Korean speakers had more "d" responses than "t" responses 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. Although in AE both "d" and "t" are pronounced as flaps, the acoustic properties of flaps made the subjects biased for the "d" responses.

3.2 Experiment 4: Categorizing Korean Flaps, Alveolar Stops and Liquids

3.2.1 Predictions

The AE speakers would categorize the Korean flaps as alveolar stops, while the Korean speakers would consistently categorize the Korean flaps as Korean liquids. In addition, There would be no significant difference in the categorization patterns between the AE and Korean flaps.

3.2.2 Methodology

3.2.2.1 Subjects

The subjects were the same as those in Experiment 3. All of them participated in Experiment 4 after they completed Experiment 3.

3.2.2.2 Stimuli

The stimuli consisted of 30 test words and 30 fillers which were all nonce words in Korean. The stimuli were produced by a female Korean speaker with standard pronunciation. The 30 test words included the Korean flap [ɾ], alveolar plain stop [d], alveolar tense stop [t'] and alveolar aspirated stop [t^h], and geminate lateral [ll] between various vowels, e.g., [niɾa], [sat^hi]. The vowels preceding and following the flaps, alveolar stops and laterals were /i, a, o, u/. The onset of each item was chosen among the Korean consonants with the exception of flaps, alveolar stops and liquids. Fillers included [p^h, b, k^h, g, s, n] between vowels. All the items were automatically randomized for each subject. No specific stress patterns were employed in the Korean stimuli, as Korean does not contrast lexical positions for stress. Recording procedures were the same as those used in the previous categorization experiment.

3.3.2.3 Procedure

The subjects heard 60 Korean nonce words through headphones presented in the Psyscope program on a Macintosh computer in a sound-attenuated chamber. They heard automatically randomized stimuli and were instructed to write down each word with normal English spelling (for the AE speakers), or with normal Korean spelling (for the Korean speakers) on the response sheets.

3.3.2 Results

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

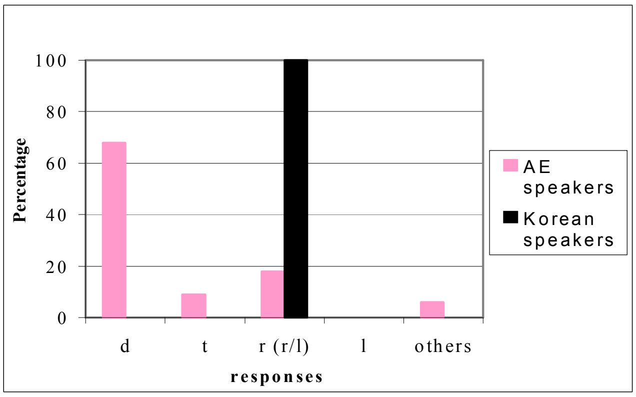

Figure 5. Percentage of Responses in Categorizing Korean Flaps
cf) (r/l) =ㄹ, others= ‘dr’ or ‘rd’

The AE speakers show a tendency to categorize the Korean flaps as the AE alveolar stops (77%), "d" in particular (68%). As expected, the Korean speakers always categorize the Korean flaps as the Korean liquid, “ㄹ”. The percentage of the AE speakers' responses of alveolar stops for the Korean flaps is lower than that for the AE flaps (77% vs. 93%). Thus, the AE speakers' responses for the Korean flaps are less consistent than those for the AE flaps although their responses are biased for the alveolar stops for both the AE and Korean flaps. The general patterns of the present results along with those in the previous experiments indicate that the L1 phonological system predominates over the acoustic properties of sounds in subjects' performance in categorizing both the AE and Korean stimuli. It seems, furthermore, that the L1 phonological effect is more apparent in the native stimuli than in the non-native stimuli.

In a brief informal interview after the experiment, the AE speakers explained that they were more influenced by the spellings of foreign languages when they heard the Korean stimuli than when they heard the AE stimuli. In other words, they knew that the stimuli were not English when they heard the Korean stimuli. They sometimes provided ambiguous responses like "rd" or "dr" for the Korean flaps, while they did not provide such responses at all for the AE flaps. Although some AE speakers fluent in Spanish were excluded from the experiments, many AE subjects still had some knowledge of foreign languages such as Spanish, German, or French. Thus, the difference in AE speakers' responses between the AE and Korean stimuli might not reflect the subtle acoustic difference between the AE and Korean flaps. Other factors may influence the AE speakers' responses for the Korean flaps as opposed to the AE flaps, including experience with foreign languages, difference in vowel quality between the AE and Korean stimuli, or difference between the native and non-native stimuli in general. These factors were apparently more engaged in the categorization experiments than in the discrimination experiments, since in the discrimination tasks, the AE speakers did not show any difference between the AE and Korean stimuli involving flaps. Since truncated stimuli were used in the discrimination tasks, the effects of the differences between the native and non-native stimuli were reduced.

4. Discussion

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 [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.

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 speculated that the Korean speakers' sensitivity to the durational difference facilitates the Korean speakers' performance on the [ɾ-r] pair.

The 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.

The overall results of the categorization tasks along with those of the discrimination tasks suggest that the L1 phonological system influence both tasks, although the categorization tasks are more influenced by the L1 phonological system than the discrimination tasks. In other words, the phonemic status of sounds takes precedence over acoustic discriminability in categorization tasks. For example, Korean speakers often categorized AE flap [ɾ] as AE retroflex approximant "r" although they could discriminate between AE [ɾ] and [r] in the discrimination test, and [ɾ] is acoustically more similar to [d] than to [r]. The greater effect of the L1 phonological system in the categorization tasks than in the discrimination tasks can be attributed to the difference of demand on memory load in the two tasks (Carney et al., 1977; Ingram & Park, 1998). Another possible explanation for the stronger effect of the L1 phonemic inventory in the categorization tasks than in the discrimination tasks would be that the vowel length in the stimuli is different between the two tasks. Whereas in the discrimination tasks, L1 phonetic information in vowels may be degraded due to the truncated stimuli, in the categorization tasks, no phonetic information involving vowels is missing.

In addition, the L1 phonological influence is stronger when the subjects hear the native stimuli than when they heard the non-native stimuli in the categorization tasks. For example, the AE speakers show more obstruent responses for the AE flaps than for the Korean flaps although their responses were strongly biased for obstruents for both the AE and Korean flaps. The AE speakers' performance is not different between the native and non-native stimuli in the discrimination tasks, but they differentiate between the two kinds of stimuli in the categorization tasks. That is, they are aware that the stimuli are not the native language when they hear Korean stimuli in the categorization tasks.

5. Conclusion and Implication

The present study shows that listeners perceive the flaps differently according to their L1 backgrounds. That is, they perceive sounds through their mental re-presentation rather than on the basis of acoustic properties of the sounds. In the discrimination tasks, listeners are also 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. The phonological processing, however, predominates over contrastive phonetic cues or acoustic properties of the sounds in the categorization tasks. Thus, the effect of the L1 phonological system has been found to be greater in the categorization tasks than in the discrimination tasks.

The results provide some implications for teaching foreign languages. The teachers should be aware of how strongly the L1 phonological system affects the foreign language learning. Although students are able to discriminate between two L2 sounds, that does not mean that the students have acquired the sounds. They may still categorize the two sounds as one L2 phoneme. For example, the Korean speakers do not have any difficulty in discriminating between AE [ɾ] and [r], but they still categorize these two phones as the same phoneme, "r". Lesson plans based on phonology, not phonetics, could be designed by the teachers.

References

Beddor, P.S.& W. Strange. 1982. Cross-language study of perception of the oral-nasal distinction. *Journal of Acoustic Society of America,* 71, 1551-1561.

Bohn, O. S. & J. Flege. 1993. Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments. *Journal of Phonetics,* 21, 267-290.

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. Viemeister. 1977. Non-categorical perception of stop consonants differing in VOT, *Journal of the Acoustic Society of America,* 62, 961-970.

De Jong, K. 1998. Stress-related variation in the articulation of coda alveolar stops: flapping revisited, *Journal of Phonetics,* 26, 283-310.

Flege, J. & W. Eefting. 1988. Imitation of a VOT continuum by native speakers of English a Evidence for phonetic category formation. *Journal of the Acoustic Society of America,* 83, 729-740.

Flege, J., M. Munro & R. Fox. 1994. Auditory and categorical affects on cross-language vowel perception. *Journal of the Acoustic Society of America,* 95, 3623-3641.

Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds “L” and “R”, *Neuropsychologia* 9, 317-323.

Gottfried, T. L. 1984. Effects of consonant context on the perception of French vowels. *Journal of Phonetics,* 12, 91-114.

Green, D.M., & J.A. Swets. 1966. *Signal detection theory and psychophysics*. New York: Wiley.

Ingram, J. & S. Park. 1998. Language, context, and speaker effects in the identification and discrimination of English /r/ an d/l/ by Japanese and Korean.

Iverson, G. & S-C. Ahn. 2003. *English voicing in dimensional theory.* Paper presented at the Conference of English Phonology. Université de Toulouse.

Iverson, G. & J. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology,* 12, 369-396.

Iverson, G. & H. Sohn. 1994. Liquid Representation. In Y. Kim-Renaud (ed.), *Theoretical Issues in Korean Linguistics,* 77-100. Center for the Study of Language and Information.

Kim, C-W. 1965. On the autonomy of the tensity feature in stop classification. *Word,* 21, 339-359.

Kim, C-W. & S-G. Park. 1995. Pronunciation problem of Australian students learning Korean: intervocalic liquid consonants. *Australian Review of Applied Linguistics,* supplement 12, 183-202.

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. 2002. *Signal Detection Theory.* In H. Pashler & J. Wixted (eds.), Stevens' handbook of experimental psychology, Vol 4., Methodology in experimental psychology , 43-90.

Macmillan, N.A., & Creelman, C.D. 1991. *Detection theory: A user's guide.* New York: Cambridge University Press.

McKain, K. S., C. T. Best, & W. Strange. 1981. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psychophysics,* 2, 369-390.

Miyawaki K., W. Strange, R. Verbrugge, A. M. Liberman, J. J. Jenkins & O. Fujimura. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 13, 253-260.

Monnet, A. & M. Feeman. 1972. A comparison of Spanish single-tap /r/ with American /t/ and /d/ in post-stress intervocalic position. In A. Valdman (ed.), *Papers in Linguistics to the memory of Pierre Delatte,* 409-416. Mouton: The Hague.

Polka, L. & W. Strange. 1985. Perceptual equivalence of acoustic cues that differentiate /r/ and /l/. *Journal of the Acoustic Society of America,* 78 (4), 1187-1197.

Price, P. 1981. *A cross-linguistic study of flaps in Japanese and in American English*. Unpublished doctoral dissertation. University of Pennsylvania, Philadelphia.

Werker, J. 1995. Exploring developmental changes in cross-linguistic speech perception. In L. Gleitman & M. Liberman (eds.), *An invitation to cognitive scienc: language* Vol 1, 87-106. MIT Press.

Werker, J. & J. Logan 1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics,* 37, 35-44.

Werker, J. & R. Tees. 1984b. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustic Society of America,* 75, 1865-1878.

Williams, L. 1977. The voicing contrast in Spanish. *Journal of Phonetics,* 5, 169-184.

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

Department of English
Cyber University of Foreign Studies
270 Imun-dong Dongdaemun-gu
Seoul, Korea 130-791
E-mail: eks@cufs.ac.kr

접수일자: 2004. 10. 09
게재일자: 2004. 12. 15