Age Effects in L2 Perception and Production by Learners of Korean

Jayeon Lim
(Korea University)


This study examines perception and production by second language learners of Korean. The particular focus is word-initial aspirated and lax stops and affricates in Korean. The results of the learners with differing age of onset suggest the following: Early L2 exposure leads to better performance, which correlates strongly to the age of onset of L2 exposure. The perception and production of the early learners approximate that of the native speakers in this study. There was no interaction between segment types and their performance. On the other hand, the performance of the late learners exhibited characteristics different from that of the early learners. Their accuracy was lower than that of the early learners in perception and production of lax consonants. Segment type was a factor in their performance, which suggests L1 interference.

Key words: second (L2) language learners, perception, production, Korean, lax and aspirated consonants

1. Introduction

Much of the discussion on second language (L2) acquisition has been centered on the question of whether there is an age-related effect on L2 acquisition. This trend is based on the fact that outcomes of adult L2 acquisition and child first language (L1) acquisition contrast distinctively in that the majority of adults often fail to reach the same level of mastery as that achieved by normal children. The goal of this study is to investigate the perception and production of Korean consonants by L2 learners with differing age of onset (AO). A basic question to be asked is whether age-related factors exist in L2 perception and production.
1.1. Previous Studies

A great number of studies that explain the difference between L1 acquisition and L2 learning of phonology support the Sensitive Period Hypothesis (SPH). The SPH states that there is a limited developmental period during which it is possible to acquire a language, whether it is L1 or L2, to normal, native-like levels (Birdsong, 1999). Once this window of opportunity has passed, however, the ability to learn language declines. Lenneberg (1967) stated that maturational processes constrain abilities for normal language acquisition. Reduction of neural plasticity after puberty due to the completion of cerebral lateralization prevents complete attainment of a first or second language.

Other than clinical evidence from aphasics, evidence from the study of Genie supports the notion of SPH (Curtiss, 1977). Genie was deprived of linguistic input and interaction during the presumed critical period (i.e., until age 13) and showed morphological and syntactic deficits. Congenitally deaf people who were first exposed to American Sign Language, an L1 for them, at differing ages also support the SPH (Johnson & Newport, 1989; 1991; Newport, 1988). These studies have shown that the AO of language exposure is the most predictive variable for ultimate attainment and performance.

1.1.1. Infant Speech Perception

Although the specific age for the SPH is yet to be discovered, results from several studies on L1 phonology provide some insights (Werker & Tees, 1984a; Werker & Tees, 1984b; Kuhl et al., 1992). The studies showed that 6- to 8-month-old infants in English-speaking environments had the ability to discriminate speech sounds while infants of 10 to 12 months of age showed perceptual abilities altered by language experience (Werker & Tees, 1984a, 1984b; Werker & Lalonde, 1988).

These studies suggested that perceptual ability of infants included language-universal categories until about the age of 6- to 8-month old, but the ability to discriminate shifts to the language-specific mode after that age. For example, Werker and Tees (1984a) found in both cross-sectional and longitudinal studies that 6- to 8-month-old English infants were able to discriminate Hindi and Thompson phonetic distinctions not phonemically distinguished in English, while English infants of ages 10 to 12 months were not. Those infants exposed to Hindi or Thompson could discriminate their native phonemic contrasts.
Although the studies mentioned above do not suggest that discriminatory capacities of sounds imply an innate knowledge of the phonetic categories, the ability to detect differences between the language-specific phonetic units may decline as maturation proceeds. The results of these studies on L1 demonstrated how innate abilities interacted with infants' early experiences to produce language-specific patterns of perception.

1.1.2. L2 Acquisition of Phonology

One interesting question in the area of L2 learning is whether learners' ability to learn additional categories of speech perception and production changes, when they already have their established L1 phonetic system. A related question to be asked is as follows: Is there an effect of age in L2 acquisition as well as in L1 acquisition? If so, children should be better L2 learners than adults are. Much research suggests that AO is an important predictor of ultimate attainment in L2 acquisition.

Much of the early research on the effect of age on L2 acquisition was concentrated in the area of phonology, which includes mostly production, specifically pronunciation abilities of nonnatives. The results suggested that the AO is the best predictor for successful acquisition. Starting from Asher and Garcia (1969), many studies have shown that learners are less likely to be judged as near-native or native speakers as the AO lowered (cf. Seliger, Krashen & Ladefoged, 1975; Oyama, 1976; 1978; Scovel, 1988, Patkowski, 1990, among others).

These claims have not gone unchallenged. Bongaerts, van Summersen, Planken, and Schils (1997) provided recent counterevidence to the SPH. They compared the highly advanced late learners of English (their L1 was Dutch) to native English speakers and found that pronunciation of five out of eleven late learners was rated to fall within the range of the native English speakers. Bongaerts et al. claimed that the native-like proficiency of some of their learners failed to confirm the SPH.

Doubt remains about the study of Bongaerts et al (1997) as well. Sentences used to elicit the subjects' production were limited to six in number and were sampled under carefully monitored conditions. There is controversy over whether proper production of six sentences can demonstrate that L2 learners have achieved phonological competence (Morris, 1998). In summary, the review of the studies that claim to provide the counter-evidence to the SPH leaves the issue of age-related effects still undetermined.

The issue of the age-related effect or the SPH also remains unsolved in
cases of perception. If it were determined that non-native accent is due to a loss of flexibility in the motor system, one would not encounter difficulties in perception. Nevertheless, findings of age related phenomena are also prevalent in perception. It is widely accepted that late learners are less sensitive than early learners in language-specific phonetic properties. However, the specific timing of the decline in perception abilities of L2 learners has not yet been established.

As we have seen in the reviews in the above sections, adult L2 learners often perceive L2 sounds in ways that differ from the native speakers of the language. The SPH predicts that AO is the most important factor for this difference. Other models such as the perceptual assimilation model (PAM) and the speech learning model (SLM) also address this difference. These models are based on the learners’ perception of the distance between L1 and L2 phones. They begin with a premise that L2 phones that are similar to a native-language category are particularly difficult for learners to perceive as different from the native language sound category.

According to the PAM (Best, 1995), sounds in a foreign language are perceived according to their similarities and dissimilarities to L1 phonemes that are closest articulatorily. L2 learners show differences in ability to discriminate based on the perceived distance between the unknown foreign sounds and the closest L1 sound. Thus, L1 influence on perception of L2 contrasts depends on the relation between phonetic details and phonemic categories.

The SLM developed by Flege (1995) is also based on how L2 learners perceive non-native sounds. According to Flege, the perceived phonetic distance of L1 and L2 sounds is one of the important factors that determines whether an L2 learner will be aware of the phonetic differences between an L2 sound and the closest sound in the L1. The awareness of a cross-language difference encourages the formation of a new L2 phonetic category. Another important factor in learning L2 speech sounds is the AO. Flege (1995) posits that the L2 learners are less likely to establish new categories for L2 sounds as the AO to an L2 setting increases. In this sense, the SLM does not differ from the claim of the SPH.1)

---

1) Although these recent theories such as the PAM and the SLM have derived from speech perception models, Flege does not always distinguish between perception and production in the application of his model. He posits that once L2 learners become capable of perception, production would follow (Flege, 1995).
The point where the SLM and the SPH differ is that the SLM predicts that the underlying capacity for speech learning, as well as the ability to establish new phonetic categories remains intact throughout life. In this sense, Flege's model differs significantly from the SPH.

The claim that learning abilities for phonetic categories remain intact across the lifetime is not supported by some studies. Research on Catalan-Spanish bilinguals is one example. Pallier, Bosch and Sebastian-Galles (1997) studied the perception of Catalan /e/ and /ɛ/ when a single vowel /e/ is available in Spanish. Significant differences were observed between the early bilinguals and native speakers of Catalan, suggesting that AO is a significant factor in the acquisition of these phonemic categories. Other studies that measuring the perceptual abilities of the early Catalan-Spanish bilinguals have also suggested that they exhibit significant differences from native speakers (e.g., Sebastian-Galles & Soto-Faraco, 1999; Bosch, Costa & Sebastian-Galles, 2000).

In summary, it still remains controversial as to whether the perception and production abilities of L2 learners are comparable to those of L1 learners. It was suggested that even though some studies (e.g., Flege, 1995; Lively, Logan & Pisoni, 1993) have shown that perception of non-native speech sounds can be learned through experience, the situation is not always so clear (e.g., Pallier et al., 1997).

To sum up, both the SLM and the SPH share the same assumption that the learning difficulties of most L2 learners originate from L1 interference. That is, older learners should face more difficulties than younger learners. However, the SLM and the SPH differ in terms of learnability issues. While the PAM does not make explicit predictions, the SLM claims that both early and late learners of L2 will have the same capacity to learn phonetic categories in L2. While the SLM model acknowledges that early learners will have advantages over late learners, problems can be conquered over the course of learning with naturalistic exposure.

The purpose of the present study can be stated in the following research question: Does AO affect L2 acquisition so that early L2 learners have an advantage over late L2 learners in terms of eventual L2 perception and production of phonetic categories? In the process of addressing this question, a subsequent question can be asked: If early learners are found to outperform late L2 learners, how does their process of learning differ?

Two experimental groups with differing AO but with comparable levels
of proficiency were compared to investigate the above research questions. The experimental groups of the present study differ from most of the previous studies on the age-related issue. The members of one group are the late L2 learners of Korean, who began learning as adults. The members of the other group are the early learners of Korean, who were immersed in Korean-speaking environment as infants, but were extracted from that environment as they entered school and lost their proficiency in Korean. The members of the two groups are currently at the same proficiency level in their L2. Subjects' abilities in perception and production of Korean stops, affricates, and fricatives were measured.

2. Focus of the Present Study

2.1. Korean Stops, Affricates, and Fricatives

For people learning Korean as a foreign language, the notoriously difficult phonetic/phonological inventories are the stop, affricate, and fricative consonants. The stop and affricate consonants in Korean can be classified into three categories according to the place of articulation, i.e. labials, dentals and palatal-velars. In the same place of articulation, there are three different types of consonants with respect to the manner of articulation. The three types are referred to in the literature as: (a) “forced”, “strong”, “glottalized” or “tensed and unaspirated”, (b) “lax”, “weak” or “slightly aspirated”, and (c) “aspirated” or “tense and heavily aspirated” (Kim, 1975). In this paper, (a) will be referred to as the “tense”, (b) as the “lax” and (c) as the “aspirated.” The same manner classification also applies to affricates, while fricatives are classified as two types, the tense and the aspirated. The pertinent consonant classification may be summarized as in Table 1.

<table>
<thead>
<tr>
<th>Table 1. Consonant Classification of Korean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tense</strong></td>
</tr>
<tr>
<td>Stops</td>
</tr>
<tr>
<td>Affricates</td>
</tr>
<tr>
<td>Fricatives</td>
</tr>
</tbody>
</table>

Both the tense series and the aspirated series are always voiceless in
any environment. The lax series is voiced in word-medial position when both the preceding and the following sounds are voiced, but voiceless in word-initial position. Hence, the only distinction between the lax and aspirated stops, affricates and fricatives is the degree of aspiration in word-initial position.

Several researchers have studied the acoustic characteristics of the Korean stops, affricates, and fricatives. Lisker and Abramson (1964) classified stop consonants in various languages based on VOT measured on sound-spectrograms. VOT is defined as the interval between the release burst of an initial stop and the onset of laryngeal pulsing. Speech samples of the Lisker and Abramson (1964) study were the initial stops in meaningful isolated words in sentences produced by a native speaker of the Seoul dialect. Their results show that there is not much difference between the two lower-valued categories—the tense and the lax types in VOT while the third category ‘the aspirated type’ shows greater average values compared to those found in any of the other languages which were analyzed.

Table 2, adapted from Hardcastle (1973), shows the average value and range of VOT associated with each of the stop types measured. The table shows the clear difference in VOT for each type of stop in any given place of articulation. The VOT values associated with the lax stops are usually about three to five times longer than those for the tense stops. The aspirated stops average about two to five times longer than the lax ones. These values agree in general with those found by other investigators (Han & Weitzman, 1970). From these studies, the acoustic difference among the three categories of consonants in Korean is quite clear with regard to VOT.

<table>
<thead>
<tr>
<th></th>
<th>p^h</th>
<th>p</th>
<th>p^h</th>
<th>t^h</th>
<th>t</th>
<th>t^h</th>
<th>k^h</th>
<th>k</th>
<th>k^h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>4.5</td>
<td>24.0</td>
<td>65.5</td>
<td>3.0</td>
<td>13.0</td>
<td>55.0</td>
<td>9.7</td>
<td>43.0</td>
<td>85.5</td>
</tr>
<tr>
<td>Maximum</td>
<td>12.0</td>
<td>61.0</td>
<td>111.0</td>
<td>8.7</td>
<td>49.5</td>
<td>98.0</td>
<td>21.2</td>
<td>70.5</td>
<td>130.0</td>
</tr>
<tr>
<td>Mean</td>
<td>7.6</td>
<td>37.0</td>
<td>83.0</td>
<td>5.4</td>
<td>22.7</td>
<td>80.1</td>
<td>15.9</td>
<td>51.5</td>
<td>103.0</td>
</tr>
</tbody>
</table>

2.2. The Present Study

The present study differs from previously reported studies in the following ways. Even though the Korean consonants in question were of
interest to phoneticians, there has been little research done with regard to second language learning by learners with either same or differing AO. This study investigates whether different groups of L2 learners (with the same overall proficiency level but different AO) demonstrate proficiency in phonological perception and production that differs from each other and also from that of L1 speakers. Specifically considered were perception and production of Korean stops and affricates in aspirated and lax categories by second language learners. Tense stops and affricates were excluded from this study, since it was proposed above that additional features other than VOT were involved in distinguishing tense consonants from lax and aspirated ones (Kang, 2000). The assumption is that earlier exposure to the language would lead to greater accuracy in perception and production of Korean stops and affricates.

Given this, the following hypotheses were tested:

Hypothesis 1: Adult learners of Korean with no early exposure to the language (Non-Korean American or the NKA group) will show less proficiency in their perception of stops and affricates than learners with early exposure (Korean American or the KA group), as many studies in L1 acquisition have suggested that language-specific attunement to sounds becomes evident during infancy (Werker & Tees, 1984a; Werker & Tees, 1984b; Kuhl, 1993). Results of L2 studies also support the hypothesis in that the acquisition of new speech sounds can be seriously compromised (Bosch et al., 2000; Pallier et al., 1997; Sebastian-Galles, 1999, among others).

Hypothesis 2: Adult learners of Korean with no early exposure to the language (NKA group) will show less proficiency in their production of stops and affricates than learners with early exposure (KA group), as is claimed by studies of L2 acquisition of speech (e.g., Asher and Garcia (1969), Oyama (1976), among others).

Hypothesis 3: Following Flege (1995), interference from L1 is expected in learning of L2 speech. The NKA subjects, who began learning the target language as adults, are expected to show interference from their L1. They are predicted to map Korean aspirated consonants onto voiceless stops in their L1. More difficulty is expected with the lax category than the aspirated ones since Korean lax consonants have VOT values that fall between voiced and voiceless consonants in English.
3. Method

3.1. Subjects

The subjects for this study were students enrolled in first and second year Korean language classes at the University of Oregon (N=24) and a control group of native Korean speakers (N=10), also enrolled at the same university. Background information on each of the subjects was collected. Their mean age was 22 years, ranging from 19 to 26. As mentioned above, the subjects are divided into a control group and two experimental groups:

1) Native Speaker Group: NS (N=10)
   The control group, that of native speakers of Korean, are students enrolled at the University of Oregon, with varying years of residence in the English speaking country during their adulthood (mean age 24 years, age range 20-28). They are all from Seoul, and speak standardized dialects. They are expected to show good perception and production, serving as a native-speaker baseline.

2) The Korean-American Group: KA (N=13)
   Since the characteristics of this KA group differ from most of the learners discussed in the SLA literature, as well as other groups in this study, it deserves a more in depth explanation. According to Fillmore (1991), it was Lambert (1981) who first discussed terminology of 'subtractive bilingualism', in relation to French-Canadian and Canadian immigrant children whose acquisition of English in school resulted not in bilingualism, but in the erosion or loss of their primary languages. The phenomenon is an often-observed one in the United States, where few American-born children of immigrant parents are fully proficient in the ethnic language, even if it was the only language they spoke when they first entered school. Once these children learn English, they tend not to maintain the language spoken at home, even if it is the only one their parents know. At the age of three or four, the children are in a language-learning mode. They learn whatever language or languages they hear, as long as the conditions for language learning are present. What happens when young children enter the American school and learn English at the age of five, is that they often drop their first language.
The background information obtained from the subjects shows the exact same situation for the KA group.

Korean-American subjects in the experiment were either born in the United States or moved here from Korea at the pre-school age. They all have Korean parents with whom they rarely communicate in Korean, but there are individual variations regarding this issue. This information was collected prior to the experiment in the format of background information. The summary of their background information is provided in Table 3.

Table 3. Background Information of the Subjects

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Arrival Age in US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>20.46</td>
<td>2.08</td>
</tr>
<tr>
<td>SD</td>
<td>1.13</td>
<td>2.56</td>
</tr>
</tbody>
</table>

3) The Non-Korean American Group: NKA (N=11)

The members of the second group are either native speakers of English or Japanese. Japanese and English have the same phonemic consonant categories (/p/ and /b/ in both word-initial and word-medial positions as well as /t/ and /d/, /k/ and /g/, and /c/ and /j/). However, their phonetic forms are quite different. Their mean age was 21.46 and the age range was from 20 to 26.

3.2. Materials

The variables dealt with in this study are stops and affricates of two types-lax and aspirated. Fricatives were excluded here because they lack the lax type. The tense type of all stops and affricates is also excluded. One reason was that the minimal pairs among these three types are not easy to find. Also, since L2 learners of Korean did not show much difficulty in learning to perceive these tense consonants, the distinction of tense consonants was not dealt with in this study.

Materials consisted of a list of 148 randomized words, with 78 tokens of experimental words and 70 distracters. The experimental words were composed of either the CV, CVC, CVCV, or CVCVCV minimal pair words that began with any of the eight consonants (see Table 4). Distracter items were also composed of either one of the CV, CVC, CVCV, or CVCVCV minimal pairs that began with consonants other than ones used
in the experimental words (for example, /nan/-orchid- vs. /san/-mountain-).

Table 4. Number of Stimulus Words

<table>
<thead>
<tr>
<th>Consonants</th>
<th>pʰ</th>
<th>p</th>
<th>tʰ</th>
<th>t</th>
<th>kʰ</th>
<th>k</th>
<th>cʰ</th>
<th>c</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>11</td>
<td>9</td>
<td>78</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>19</td>
<td>19</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the perception task, a test tape was constructed from the 148 randomized words. A native speaker who had been living in the United States for 2 months and was from the city of Seoul produced the material. For the production task, identical words that were used for the perception task were provided in a different order.

3.3. Procedures

The experiments were performed in the following order of schedule: (1) proficiency test (day 1), (2) training session (day 2), (3) perception task (day 3), (4) production task (day 6, 7, or 8), and (5) orthography test (day 9). The members of the control group participated only in (3) the perception task and (4) the production task.

To measure whether the subjects of the experimental groups have equal proficiency in the target language, they were given written proficiency test. The test was administered during the first week of the second term of three-term sequence. The tests consisted of fifty questions consisting of multiple-choice and short answer questions. The questions measured overall proficiency in the target language, including grammar and lexical abilities. The level of the test was 1st year college foreign language class.

Since one of the purposes of the study is to investigate perception in the second language, it is important that their perceptual ability in Korean is ensured. Lobo and Yoshida (1982) reported that the discriminability of discrete sounds by Japanese learners of English was more accurate when those sounds were in English words they were familiar with, than when they were in words they did not know. Based on this, prior instruction on unfamiliar lexical items was given to the experimental groups. This training session took place one day prior to the perception task for a period of 50 minutes. Instruction on the vocabulary was given for the words to be tested as well as distracter words. The
purpose was to ensure that the subjects had a considerable amount of exposure to the Korean words used in the experimental tasks. Throughout the experiment, including the training session, subjects were told they were being tested on their semantic ability, for the purpose of avoiding subjects' deliberate attention to certain consonants.

Two main parts of this study are: (1) testing perception and production differences among groups and (2) exploring the bases for the production results by measuring VOT across the groups.

3.5. Perception Task

During the perception task, subjects listened to the tape via headphones in a language-listening laboratory. They were told to select the correct word in Korean among the two minimal pairs (i.e. /pul/-fire- or /phul/-glue-) upon hearing the test items. The subject saw these particular minimal pairs as a list on a sheet of paper, written in Korean alphabet, and heard one of the pairs (e.g., /pul/) and had to decide to which of these words it belonged.

They were also instructed to write down the meaning of the chosen word in English, since the subjects were told they were being tested for their semantic ability in order to avoid any deliberate attention to the target phonetic items. All three groups were in the same room at the same time. To give the subjects equal and appropriate response time, the experimenter presented the recorded words to all subjects at the same time.

The perception task took place 3-5 days prior to the production task. The reason was to clear away any immediate auditory residue the subjects might have received from the process of the perception experiment. Any interval longer than this period on the other hand might allow for possibilities of acquiring phonetic or phonological distinctions.

3.6. Production Task

For the production task of the study, each subject was asked to read randomized word lists written in Hangul. The lists were identical to the ones used in the perception task, but differently ordered. The production task differed from the perception task in that the subjects were only presented with one of the minimal pairs balanced for number of types. There were 78 experimental words. As in the perception task, 70
distracter words were included. There were 148 words altogether. The subjects were tested individually, unlike in the perception task. They were told to read the words presented to them.

3.7. Orthography Test

The consonants that differ only in their aspiration not only resemble each other acoustically, but are also similar in their orthography. To ensure that learners had mastered the orthography, a brief test was given. This ensured that their lack of abilities in perception and production was not due to their poor ability in orthography.

3.8. Measurements

For the perception task, accuracy scores were calculated. For the Production Task, the subjects' word readings were recorded individually in a sound-attenuated booth using a DAT (digital) recorder. The performance of the Production Task was measured by two different methods. First, three native speakers judged the recorded production of the subjects. They used a scale of 0, 0.5, and 1 with 1 being native-like. The numbers were labeled on pieces of paper that were provided to the judges. The native speaking judges were all students at the American English Institute at the University of Oregon, and native speakers of Korean with a maximum of six months of residence in the United States. Second, the recorded production was analyzed and VOT values were measured by the commercially available program SoundEdit 16 and Signalyze on a Macintosh computer. When any problematic or ambiguous cases arose from factors such as unclear separation of the cue word from the previously presented ones, the words were not included for measurement. Only 32 cases or 1.7% were excluded for this reason.

4. Results

4.1. Preliminary Results

Subjects in the experimental groups exhibited a mean score of 76.8% in their proficiency test of Korean. By group, the KA and NKA subjects scored 77.9% and 75.7% respectively. The difference in the two groups was not statistically significant \((F = 1.4)\).
For subjects in the experimental group, an orthography test was given in order to see whether they could discriminate the orthography of voiced and voiceless consonants. The subjects' scores were nearly perfect, with a mean accuracy of 98%, and they were considered to have mastered the pertinent orthography. There was no significant difference between the two experimental groups ($F = 0.8$).

4.2. Perception Task Results

Next, the results of the perception task are presented. Mean accuracy and standard deviations are presented in Table 5. Specifically, the results are as follows: (i) For the perception task, the Non Korean-American (NKA) group performed less accurately than the Korean-American (KA) group and the native speaker (NS) group. The difference between the KA and the NS groups was not significant. (ii) Between the KA and NKA groups, the NKA group responded less accurately in their perception task.

<table>
<thead>
<tr>
<th></th>
<th>$p^h$</th>
<th>$p$</th>
<th>$t^h$</th>
<th>$t$</th>
<th>$k^h$</th>
<th>$k$</th>
<th>$c^h$</th>
<th>$c$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>Mean</td>
<td>9.7</td>
<td>9.7</td>
<td>8.9</td>
<td>9.6</td>
<td>9.8</td>
<td>8.5</td>
<td>9.7</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.48</td>
<td>0.48</td>
<td>0.32</td>
<td>0.7</td>
<td>0.42</td>
<td>0.71</td>
<td>1.7</td>
<td>1.27</td>
</tr>
<tr>
<td>KA</td>
<td>Mean</td>
<td>9.62</td>
<td>8.46</td>
<td>9.0</td>
<td>8.08</td>
<td>9.62</td>
<td>7.38</td>
<td>10.69</td>
<td>7.77</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.65</td>
<td>1.45</td>
<td>0.0</td>
<td>1.98</td>
<td>0.96</td>
<td>0.96</td>
<td>0.48</td>
<td>0.48</td>
</tr>
<tr>
<td>NKA</td>
<td>Mean</td>
<td>8.91</td>
<td>3.18</td>
<td>7.09</td>
<td>4.0</td>
<td>8.6</td>
<td>5.0</td>
<td>9.27</td>
<td>4.73</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>0.83</td>
<td>1.99</td>
<td>1.58</td>
<td>1.73</td>
<td>1.26</td>
<td>1.61</td>
<td>1.27</td>
<td>1.1</td>
</tr>
</tbody>
</table>

A two-way (group X segment types) repeated measures ANOVA was performed on the means of the accuracy scores on consonants to see whether the observed differences of (i) and (ii) above can be attributed to chance or whether there are true differences between groups. The results revealed a highly significant difference between the groups ($F (2, 31) = 180.91, p<0.0001$) and also between the segment types ($F (1, 31) = 331.99, p<0.0001$). An interaction between group and segment types ($F (2, 31) = 77.91, p<0.0001$) was also significant. As can be seen from the table, the three groups have different responses by segment types. The effect of segment types was more intense in the NKA group than in the other two groups. Separate analyses for the aspirated and lax consonants showed...
that the effect of segment types was significant for the NKA \((F (1, 86) = 138.105, p<0.001)\), but not for the KA \((F (1, 102) = 3.135, p<0.001)\) or NS \((F (1, 78) = 5.989, p<0.001)\). Compared to the other two groups, the NKA had much lower scores for lax consonants than for aspirated consonants. A Tukey's test \((p<0.01)\) further showed that significant differences existed between the NS and the NKA when perception scores were collapsed over segment type, as well as between the KA and the NKA. It was revealed that the difference between the NS and the KA is not significant (see Table 6).

Table 6. Tukey’s Test for Comparison of Three Groups’ Perception Task

<table>
<thead>
<tr>
<th>Group</th>
<th>Comparison</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS vs. KA</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td>NS vs. NKA</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>KA vs. NKA</td>
<td>***</td>
<td></td>
</tr>
</tbody>
</table>

(*** \(p < 0.01\))

Further, a correlation analysis was performed on the responses to the lax stops by the NKA, in order to see if their accuracy scores were correlated to the length of VOT of the consonants presented to them. The results were significant \((r=0.638, p<0.05)\). It showed that the subjects of the NKA were more accurate for the lax consonants with shorter VOT in that there was a significant correlation between the types of the consonants and the length of VOT.

4.3. Production Results: Native Speaker Judges

The results of the native speaker judges are displayed in Table 7. Results of a two-way (group X segment types) repeated-measures ANOVA yielded a significant difference among groups \((F (2, 31) = 59.630, p<0.001)\). Effect of segment types \((F (1, 31) = 3.128)\) and an interaction between segment types and group \((F (2, 31) = 0.683)\) were not significant, unlike in the perception task. A Tukey's test revealed that the differences were significant for the NS vs. the NKA and the KA vs. the NKA. The difference was marginally significant for the NS and the KA.
Table 7. Tukey's Test for Comparison of Three Groups' Production Task (Native Speaker Judges)

<table>
<thead>
<tr>
<th>Group</th>
<th>NS vs. KA</th>
<th>NS vs. NKA</th>
<th>KA vs. NKA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.054</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>

(*** $p < 0.01$)

4.4. Production Results: Acoustic Analysis of VOT Measurements

The mean VOT measurements (in msec.) are presented in Table 8.

Table 8. Mean VOT in msec. and Standard Deviations, by Group

<table>
<thead>
<tr>
<th></th>
<th>$p^h$</th>
<th>$p$</th>
<th>$t^h$</th>
<th>$t$</th>
<th>$k^h$</th>
<th>$k$</th>
<th>$c^h$</th>
<th>$c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS</td>
<td>131.53</td>
<td>73.87</td>
<td>129.23</td>
<td>71.83</td>
<td>142.67</td>
<td>35.33</td>
<td>161.94</td>
<td>83.58</td>
</tr>
<tr>
<td>SD</td>
<td>11.55</td>
<td>11.00</td>
<td>11.45</td>
<td>9.77</td>
<td>11.00</td>
<td>12.33</td>
<td>13.67</td>
<td>5.74</td>
</tr>
<tr>
<td>KA</td>
<td>107.20</td>
<td>41.20</td>
<td>109.95</td>
<td>58.70</td>
<td>122.82</td>
<td>36.48</td>
<td>134.72</td>
<td>53.90</td>
</tr>
<tr>
<td>SD</td>
<td>15.92</td>
<td>7.72</td>
<td>19.08</td>
<td>12.12</td>
<td>9.99</td>
<td>14.65</td>
<td>18.95</td>
<td>10.23</td>
</tr>
<tr>
<td>NKA</td>
<td>102.93</td>
<td>11.07</td>
<td>104.08</td>
<td>10.52</td>
<td>98.08</td>
<td>15.63</td>
<td>129.04</td>
<td>23.48</td>
</tr>
<tr>
<td>SD</td>
<td>28.70</td>
<td>12.94</td>
<td>33.42</td>
<td>21.98</td>
<td>17.48</td>
<td>25.59</td>
<td>41.72</td>
<td>27.76</td>
</tr>
</tbody>
</table>

Two-way (group X segment type) repeated measures ANOVA revealed significant differences among groups ($F(2, 31) = 142.58, p<0.01$) and also between the segment types ($F(1, 31) = 184.31, p<0.01$). An interaction between group and segment types ($F(2, 31)=57.91, p<0.01$) was also significant. The significant interaction was explored with a series of ANOVAs performed on each group. Again, it was revealed the interaction was due largely to the NKA group ($F(1, 86) = 162.62, p < 0.001$): The difference in the performance of two different types of segment in the NKA group was significant. Within the KA ($F = 5.116$), or the NS ($F = 6.514$), the difference between VOT of the two segment types was not significant. A close observation of the numbers reveals that the NKA display much shorter VOT for lax consonants. Results of the Tukey's test showed that significant differences existed among all three groups.
Table 9. Tukey’s Test for Comparison of Three Groups’ Production Task (VOT Measurements)

<table>
<thead>
<tr>
<th>Group</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NS vs. KA</td>
<td>***</td>
</tr>
<tr>
<td>NS vs. NKA</td>
<td>***</td>
</tr>
<tr>
<td>KA vs. NKA</td>
<td>***</td>
</tr>
</tbody>
</table>

(*** p < 0.01)

In addition, a correlation analysis was performed to explore the nature of the KA’s VOT of the lax consonants. Recall the background information of the KA subjects, where they had a differing age of arriving in the US, ranging from ages of 0-5.1 with a mean age of 2.08. An object of the correlation analysis was to see if the AO of the KA correlates with VOT of the lax consonants. The results yielded a significant correlation \( r = 0.631, p < 0.05 \), implying that the earlier they were immersed into their primary language (i.e., English) environment, more likely they were to behave like the NKA in terms of their VOT of the lax consonants (i.e., they produced the lax consonants with shorter VOT).

In summary, the results of the perception task revealed a significant difference between the groups’ discrimination of lax consonants. No significant difference was exhibited in the case of aspirated consonants. Further analysis showed that the NKA group was the one with the difference between aspirated and lax consonants. Results of the correlation analysis showed that within the NKA group, who had difficulty with the lax consonants, the subjects were more likely to be accurate for the lax consonants with shorter VOT. In the case of the production task, the results of the judges’ ratings revealed a significant difference among the KA and the NKA. In the case of VOT measurements, an interaction of group X type was present. All three groups differed significantly from each other. As for the lax consonants, the NKA group exhibited shorter VOT for the lax group than the KA group. Correlation analysis revealed that the earlier the subjects of the KA group were immersed in English-speaking environments, the more likelihood there was that their VOT of the lax consonants were shorter.
5. Discussion and Conclusion

Overall, the NKA were less accurate, in both perception and production, than the KA or the NS. Acquisition of L2 perception and production by the learners in this study was determined, to some extent, by the age of initial exposure to the target language. Hypothesis 1 predicted that the perception of the NKA group in this study would be less accurate than the KA. The significant difference between the perception of KA and the NKA groups partially confirms Hypothesis 1, in that the NKA group showed poorer performance in the perception of lax stops and affricates than the KA group. Hypothesis 2, which predicted that the production of the NKA group would be less accurate than the KA, is confirmed. The finding is indicated by the results by native speaking judges and through VOT analysis. Hypothesis 3 is also confirmed. Within the NKA group, the learners were found to exhibit more difficulties with the lax consonants as opposed to the aspirated consonants, as predicted by the Hypothesis 3. They showed interference from their L1. Each experimental group will be discussed below in detail.

The KA group's early exposure to, but discontinued use of, the target language measurably influenced their perception of the target consonants. Their identification ability did not differ significantly from the NS. This finding appears to reflect the establishment and maintenance of a native-like perceptual system in Korean.

The ability of the KA to produce the target consonants (as judged by the native speakers of Korean), as well as their perception ability, reflects a system similar to that of the NS. The results of VOT of the KA group also approximate that of the NS. Within the KA group, a correlation analysis showed a significant correlation between shorter VOT of the lax consonants and the AO of the learners. The earlier the learners were immersed in English-speaking environment (i.e., earlier displacement from the Korean-speaking environment), the closer their production abilities were to the NKA.

These results suggest that, with respect to certain types of tasks and segment types, L2 speakers of Korean who were exposed to the target language at an early age were found to make perceptual identifications that nearly matched the native Korean speakers. That is, the members of the KA group did not rely on English but rather resembled the NS. The
finding is consistent with the predictions of the SPH. Their early linguistic experience enabled them to behave differently from the NKA, as will be discussed below.

In contrast to the KA group, the identification of the lax consonants was inaccurate for the NKA group, who seriously missed the native phonemic target. Unlike the KA group, the NKA group demonstrates a greater difficulty with one type of segment: They have almost twice as much difficulty in the lax consonants as in the aspirated consonants, as shown by the statistical results. Compared with VOT of the KA and the NS, VOT of the NKA were significantly lower. In fact, mean VOT of the lax consonants of the NKA was 23% of VOT of the NS. On the other hand, mean VOT of the aspirated consonants was 76.8% of VOT of the NS.

A question raised is this: Why did the NKA perform differently from the KA with regards to the lax consonants only? Clearly, the NKA group differs from the KA not only quantitatively in their perception abilities, but also qualitatively: Their abilities were not only lower in ranges, but also exhibited characteristics different from both the KA and the NS: the NKA group relied more on their respective L1s than the KA group as was evidenced by a significant correlation analysis between the lower VOT and their relatively more accurate perception. This implies that they were using their native perceptual categories.

English and Japanese both make a distinction within voicing categories. Bond and Fokes (1991) notes the following: In English and Japanese, the category of voiced stops can be unaspirated or produced with VOT values in the short lag range or the lead; these stops contrast with voiceless aspirated stops, which are produced with VOT values in the long lag range (Lisker & Abramson, 1964; Flege & Eefting, 1987). In word-initial positions of Korean, stops and affricates are always produced either as lax or aspirated. For this reason, the learners of the NKA group would have found their own way of making the distinctions possible: producing short-lag/lead vs. long-lag instead of lax vs. aspirated. Such use of the strategy is convincing in that aspirated consonants are in complementary distribution with lax consonants in Korean. Results from the production task should complement this view.

The production task showed results similar to the perception task in that the NKA group did worse than the KA group in both native speaker ratings and VOT. Again, the possibility can be raised that the learners use
cues available from their L1. Since Korean native speaking judges do not have access to the distinction between short-lag vs. long-lag word-initially, they would have judged the word-initial short-lag consonants as correct responses. Results from the VOT measurements need to be discussed to supplement this view.

The results have shown that clear differences existed in VOT among the NS, the KA and the NKA. All groups differ from each other. Again, we should investigate the lax consonants in more detail. The NKA are the ones with significantly lower VOT in lax consonants than the NS and the KA. In this vein, the NKA group subjects are found to rely more heavily on their L1 as opposed to the KA group.

The findings of the NKA group are in agreement with one premise of the SLM in that the established L1 phonetic categories interfere with the process of the L2 learning at later AO. According to the SLM, the distance of L1 and L2 sounds determines the L2 learner's awareness of the phonetic differences between an L2 sound and the closest sound in the L1. The aspirated consonants are approximate to the voiceless consonants in English and the mapping is not expected to cause a problem. With regards to the lax consonants, mapping should cause relatively more difficulty, since VOT of the lax consonants (for Korean /p/, /t/, /k/ they are 37, 22.7, 51.5 ms respectively) fall between the values of English voiceless (78, 89, 94 ms. for English [p], [t], [k]) and voiced consonants (-98/18, -85/15, -81/22 ms. for English [b], [d], [g]). As was mentioned above, the results of the NKA were consistent with the prediction of the SLM in that they had more difficulty with the lax consonants.

Overall, the NKA group showed poorer acquisition of perception and production than the KA. The behaviors examined were found to be different from that of the KA or the NS in that they had great difficulty with lax consonants in both perception and production. Although a detailed cross-linguistic data on phonetic distance is needed in order to provide precise information, the results of the NKA provide some insight into the L2 perception and production. In Guion, Flege, Yamada and Pratt (2000), both perceived cross-language phonetic distance and the discrimination of a wide range of L2 consonants were examined. The results indicated that the perceived phonetic distance of L1 and L2 sounds was found to predict learning L2 sounds at least in some cases. The results of the NKA in the present study support the findings of Guion et al. in that the distance of the L1 and the L2 influences late L2 learning.
In other words, learning of L2 is not completely independent of L1 and L1 plays a role in the learning of L2 speech sounds at least with adult learners.

A caveat remains in this study. Although the subjects in this study, especially the KAs were recruited carefully as to include only those who had the least amount of exposure to the target language and lost their ability in Korean since arriving in the US, one cannot be completely free from the question: Did the exposure to Korean in their home, whether explicit or implicit, influence the results of this study? The claim of Fillmore (1991) in the previous section provides a partial answer to this question. In order to fully escape from the question of continued exposure, a subsequent study with proficiency tests measuring all four aspect of language (speaking and listening as well as reading and writing) may be desirable.

To summarize, the purpose of this study was to see if L2 learners of Korean with different initial exposure to the target language showed differences in their perception and production. The findings of this study provide answers to this question. First of all, the NKA group with no early initial exposure to the language did show lower accuracy in perception of lax consonants and production of both lax and aspirated consonants. The NKA relied more on cues from their L1. Furthermore, among the KA, those who were exposed to English-only environments earlier produced the lax consonants with shorter VOT and resembled the results of the NKA. The KA group, with early initial exposure to the target language showed accuracy in perception and production that approximates to the NS. Segment type was not a factor in the performance of the KA and the NS, unlike the NKA.

References


Kuhl, P. K. (1993). Early linguistic experience and phonetic perception:
Implications for theories of developmental speech perception. *Journal of Phonetics* 21, 125-139.


Received: Mar. 9, 2003
Revised version accepted: Aug. 11, 2003
Accepted: Aug. 14, 2003