EFFECTS OF PHONETIC SIMILARITY AND L2 EXPERIENCE: PRODUCTION OF ENGLISH /s/-/ʃ/ BY ADULT KOREAN ESL LEARNERS

Sang Yee Cheon

Dept. of East Asian Languages and Literatures, University of Hawai'i at Mānoa, USA
scheon@hawaii.edu

ABSTRACT

This study examined the effects of both phonetic similarity between L1 and L2 sounds and L2 experience on the production improvement of L2 English /s/-/ʃ/ by adult Korean ESL learners. Two production judgment methods were employed. The result shows that a dissimilar L2 sound /ʃ/ was better produced than a similar L2 sound /s/ in terms of listener judgment. Adult L2 learners did not produce similar L2 sounds accurately, even with extensive L2 experience. However, in terms of acoustic-based judgment, ESL learners produced L2 sounds accurately regardless of their proficiency levels, resulting in a disparity in production performance between two production judgment methods. In assessing the production ability of L2 learners, more than one phonetic-based judgment in addition to listener judgment needs to be taken into consideration.

keywords: phonetic similarity, English sibilants, Korean, L2 experience, production.

1. INTRODUCTION

A great deal of experimental research has been conducted on various sounds in second language (L2) speech, especially on vowels, stops, and liquids. However, not much work has been done on sibilants, in particular, in terms of L2 acquisition. The L2 EN sibilants chosen for this study are an alveolar /s/ and a palato-alveolar /ʃ/. According to previous studies, these voiceless sibilants, /s/ and /ʃ/, each have a relatively strong acoustic intensity, and therefore register as darker patterns in a spectrogram when compared to non-sibilants. The overall spectral shape of the two voiceless sibilants is determined by the size and shape of the oral cavity in front of the constriction. /s/ is produced with a shorter anterior cavity than is /ʃ/, and thus shows a spectral peak at higher frequencies. This study dealt with acoustic measurement and with native listener judgment as two ways of assessing the acceptability of L2 English sibilants as produced by Korean ESL advanced and beginning learners.

2. PRODUCTION OF L2 SOUNDS

The main question addressed in Flege's Speech Learning Model (SLM, Flege [5]) is whether new sounds are more difficult or are easier for adult L2 learners to produce accurately than similar sounds. It has been traditionally accepted that similar sounds can be easier for adult L2 learners to produce than new sounds. According to the SLM, however, L2 experience influenced the perception of L2 sounds and this perceptual difference has an influence on the production of L2 sounds. Thus, the greater the perceived distance of an L2 sound from the closest L1 sound, the greater the likelihood that a new category will be established for the L2 sound. Flege and his colleagues have conducted many experimental phonetic studies on the role of similarities between L1 and L2 sounds and L2 experience. For example, Bohn & Flege [2], in their experimental study of perception and production of a new category by adult L2 learners showed that unlike inexperienced German speakers of English, experienced learners of English produce a new English vowel category accurately. Major [7] claimed that advanced learners have more difficulty in producing similar sounds, while beginning learners show difficulty in producing new sounds.

2.1. Cross-language phonetic similarity

In classifying sounds as either similar or dissimilar, there are no agreed definitions or a metric that would determine the degree of similarity or dissimilarity between L1 and L2 sounds (Major [7]). However, Cheon [3] conducted two experiments to gauge the degree of phonetic and perceived similarities between L1 and L2 sounds and between members of L1 or L2 contrasts. In a study of AX discrimination tasks of EN and KO sibilants by KO speakers, Cheon [3] showed that the degree of perceived similarity between EN /sa/ and KO /sa/ is different from that between EN /sa/ and KO /s*a/ (6% vs. 88%, respectively); on the other hand, EN /ʃi/ vs. KO /si/ as well as EN /ʃi/ vs. KO /s*i/ are found to be perceptually dissimilar to each other (15% vs. 0%, respectively); acoustically and perceptually, EN

/si/ is found to be more similar to KO /s*i/ (84%) than it is to KO /si/ (1%). Also, English (EN) voiceless sibilants are different from Korean (KO) voiceless sibilants in terms of articulation and acoustics. A previous study (Anderson et al. [1]) found that KO sibilants are not produced with the same place of articulation in the oral cavity as EN voiceless sibilants. For example, EN /s/ is an alveolar and EN /ʃ/ is a palato-alveolar, while KO lax and tense sibilants, /s/ and /s*/, are pre-alveolars. In addition, unlike KO [ʃ], which is an allophone of KO /s/ when it precedes the vowel /i/ or the glide /j/, EN /ʃ/ is produced with some lip rounding, which enhances the low frequencies in the frication noise.

This study was designed 1) to evaluate the effect of phonetic similarity on the production of L2 contrast that is allophonic in L1 and 2) to examine the effect of L2 experience on the production improvement of the L2 contrast. EN /s/ is perceptually and acoustically more similar to KO /s*/ in the context of /a/; EN /ʃ/ is perceptually dissimilar to KO [ʃ] in the context of /i/. It was predicted that KO ESL learners would be good at producing EN /s/-/ʃ/ because the two L2 sibilants would be assimilated to different sounds by KO speakers, even though they are not phonemic in their L1. Experiment 1 examined the spectral peak frequency of EN /s/-/ʃ/ as spoken by KO ESL learners. Experiment 2 examined Koreans' production ability of the EN sibilants in terms of listener judgment.

2.2. Experiment 1

2.2.1. Method
Subjects. Fifteen paid volunteers participated: five American English (AE) speakers and ten KO ESL learners from Seoul. All subjects were male students at the University of Hawai'i at Mānoa. Five KO learners who had studied English in the U.S. for less than one year belonged to the beginners' experimental group (on average, 0.8 years). Five KO learners who had studied English in the U.S. for more than five years belonged to the advanced learners' experimental group (on average, 6.2 years) at the time of recording. A control group of five native speakers of AE was included. None of the subjects had a history of speech or hearing disorders.

Stimuli and Procedures. EN stimuli consisted of one of the sibilants /s/ and /ʃ/ followed by one of the vowels /i, a, u/ (i.e., /si/ *see* or /ʃi/ *she*). The target word was located in a carrier phrase, *I'm saying ___ now*. Speakers were asked to repeat the

randomized list five times in a natural conversational style. A total of 30 tokens were produced by each subject. In total, 450 sentences (2 sibilants x 3 vowels x 5 repetitions x 15 speakers) were analyzed for the acoustic measurement of EN sibilants. In previous research, EN sibilants /s/ and /ʃ/ statistically do not show any significant difference in amplitude and in duration. However, EN /s/ and EN /ʃ/ are distinguished by the shape of their noise spectra (Jongman et al. [6]). The spectral characteristics of the noise tend to be relatively consistent across time, though there are changes near frication-aspiration or frication-vowel boundaries. In Experiment 1, the highest frequency at the durational midpoint of sibilants was measured by using Fast Fourier Transform (FFT) analysis.

2.2.2. Results
A three-way repeated measures ANOVA that tested the effects of Group, Sibilant contrast, and Vowel context on spectral peak frequency of EN /s/-/ʃ/ yielded two main effects (Sibilant contrast and Vowel context) [i] and two significant interactions (Group x Sibilant contrast and Sibilant contrast x Vowel context). Surprisingly, there was no main effect of Group on spectral peak frequency, implying that KO learners of AE were able to differentiate EN /s/ and EN /ʃ/ in terms of spectral peak frequency regardless of their proficiency level. A main effect of *Sibilant contrast* indicated that there was a significant difference between /s/ and /ʃ/ (F [1, 12] = 181.46, p<.0001). The mean spectral peak of /s/ was always higher than that of /ʃ/. A Group x Sibilant contrast interaction indicated that the mean spectral peak frequency of /s/ did not show a statistically significant difference between AE speakers and KO ESL advanced learners, but these two groups were significantly different from L2 beginning learners (F [2, 12] = 68.21, p<.0001). In the production of EN /ʃ/, there were no statistical significant differences between groups. A dissimilar sound EN /ʃ/ is better produced than a similar sound EN /s/ by KO ESL beginning learners.
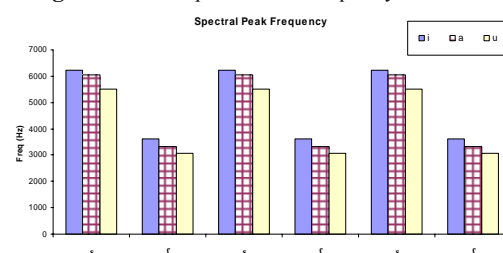
Figure 1: Mean Spectral Peak Frequency

Figure 1 shows the mean spectral peak frequency of EN /s/ and EN /ʃ/ in the three vowel contexts as produced by the three groups.

2.3. Experiment 2

2.3.1. Method

Subjects. 30 male subjects participated, including the same 15 subjects from Experiment 1. After acoustic measurements were taken in Experiment 1, data from 15 more subjects were collected for Experiment 2. The approximate average age and the average length of residence in the U.S. for the 10 beginning subjects were 30.3 years and 1.0 years, respectively; the approximate average age and the average length of residence in the U.S. for the 10 advanced subjects were 36.1 years and 5.6 years, respectively. A control group of 10 AE speakers was included. The approximate average for the control group was 30.

Stimuli and Procedures. The EN stimuli in Experiment 2 were the same as those in Experiment 1. In order to evaluate production performance relative to acoustic performance on EN sibilants, listener judgments were carried out separately on the three groups. Production accuracy in the L2 speech was judged by three native speakers of AE as correct or incorrect, along with the aid of acoustic analyses: the first rater was a phonetically trained listener; the second rater was a naïve listener; the third rater was a bilingual speaker of KO and EN. In total, 900 sentences were analyzed (2 sibilants x 3 vowels x 5 repetitions x 30 speakers). Each subject's scores for the five repetitions were averaged prior to statistical analysis.

2.3.2. Results

A three-way repeated measures ANOVA (3 x 2 x 3 design) yielded two significant main effects of Group and Vowel context and one interaction effect of Sibilant contrast x Vowel context. A main effect of Group and a subsequent Tukey post-hoc test indicated that mean production scores were significantly higher for native AE speakers than they were for KO ESL learners. KO ESL advanced learners were significantly higher in their mean production scores than were KO ESL beginning learners (F [2, 27]= 38.124, $p<.0001$). L2 experience made a difference in the production of EN sibilants by Korean ESL learners. There was a Sibilant contrast x Vowel context interaction (F [2, 54] = 7.923, $p<.0010$). Subsequent post-hoc tests revealed that there was a statistically significant difference between /su/ and /ʃu/, but there were no

statistically significant differences between /s/ and /ʃ/ before /a/ and /i/.

Table 1 shows the mean correct production scores of EN /s/ and /ʃ/ in the three vowel contexts as produced by the three groups.

Table 1: Mean Correct Production Scores[ii]

	/s/			/ʃ/		
	a	**i**	u	**a**	**i**	u
Cont	**1**	**1**	1	**1**	**1**	1
L2 Adv	**0.68**	**1**	0.68	**0.76**	**0.7**	1
L2 Beg	**0.18**	**0.5**	0.34	**0.36**	**0.46**	0.66

Overall, EN /s/ in the context of /i/ was better produced by KO ESL beginning and advanced learners than it was in the context of /a/ and /u/. EN /si/ and /ʃu/ were authentically produced but EN /sa/ was poorly produced by KO ESL advanced learners. It may be due to negative transfer from L1 Korean because EN /sa/ is a nonsense word unlike EN /si/ *see* and EN /ʃu/ *shoe*. According to native speakers of AE, most of the EN /sa/ produced by ESL beginning learners was not acceptable. A KO and EN bilingual speaker judged the L2 EN /sa/ as L1 Korean-transferred sound.

3. DISCUSSION AND CONCLUSION

The effects of Group (American English vs. KO ESL advanced learners vs. KO ESL beginning learners), of Sibilant contrast (/s/ vs. /ʃ/) and of Vowel context (/i/ vs. /u/ vs. /a/) both on spectral peak frequency in Experiment 1 and on correct production scores in Experiment 2 were examined. In Experiment 1, both groups of KO ESL learners produced EN /s/ with higher spectral frequency than EN /ʃ/, just as in the production of American English. That is, KO ESL learners did differentiate EN /s/ and /ʃ/ in terms of spectral peak frequency, regardless of their L2 experience (Cont ≡ L2 Adv ≡ L2 Beg).

However, in terms of listener judgment in Experiment 2, KO ESL beginning learners showed poor production scores. KO ESL advanced learners were significantly higher in their mean production scores than were KO ESL beginning learners (Cont > L2 Adv > L2 Beg). L2 experience made a difference in the production of EN sibilants by KO ESL learners. Regarding degrees of perceived similarity between EN /s/ and EN /ʃ/, EN /s/ is a similar L2 sound (88% between EN /sa/ and KO /s*a/ vs. 84% between EN /si/ and KO /s*i/) and EN /ʃ/ is a dissimilar sound (15% between EN /ʃi/ and KO /si/ vs. 0% between EN

/ʃi/ and KO /s*i/) to Korean. In Experiment 2, a dissimilar L2 sound /ʃ/ is better produced than a similar L2 sound /s/. The result of this study partially supports Flege's hypotheses that with extensive L2 experience, adult L2 learners will not produce similar L2 sounds accurately. Surprisingly, EN /ʃu/ and EN /si/ produced by KO ESL advanced learners were correctly judged by AE speakers more than any other EN sibilant plus vowel sequences. Cheon [4] found in the production of KO sibilants by AE learners of KO that more similar sounds (77% perceived similarity) are produced more exactly than less similar sounds (56% perceived similarity) by AE advanced learners of KO, while less similar sounds are produced poorly by both L2 advanced and beginning learners. It is suggested here that L2 sounds are produced differently by L2 learners depending on the degree of phonetic similarity between L1 and L2 sounds.

It is interesting to note that the two experiments show differences in production performance between acoustic-based judgment and listener judgment. KO ESL learners can distinguish EN /s/ from EN /ʃ/ in terms of spectral peak frequency. Thus, KO ESL learners and AE speakers statistically exhibited no significant difference in the acoustics of EN sibilants. However, this does not mean that KO ESL learners reached attainment of native-like pronunciation of EN sibilants. Even in reading speech, the production of EN /s/ and /ʃ/ by KO ESL learners was not judged correctly by native AE speakers. Overall, unlike the result of acoustic measurement, KO ESL beginning learners and some ESL advanced learners showed poor production scores around "chance production accuracy," even though ESL advanced learners were statistically a little better in the production of EN /s/-/ʃ/ than ESL beginning learners. This disparity may be due to an inherent acoustic dissimilarity between EN /s/ and EN /ʃ/. In the case of KO /s/-KO /s*/ produced by both KO speakers and AE learners of KO, Cheon [4] suggested that more than one acoustic parameter plays a role in the distinction of the KO sibilant contrast /s-s*/ that is made by KO speakers. KO native speakers used both duration and amplitude, but AE KSL advanced learners did not rely on amplitude but rather on duration alone. AE learners of KO seem to rely heavily on a single acoustic parameter to differentiate KO /s/ and KO /s*/. That is, temporal acoustic parameters seem to be easier for L2 learners in general to access than any other acoustic parameters. It is assumed that one or more acoustic parameters are necessary for KO speakers to distinguish the EN sibilant contrast. It seems that one acoustic parameter such as spectral peak frequency may not be adequate enough to assess the production accuracy of EN /s/-/ʃ/ in L2 speech.

It is suggested that the production scores of EN sibilants in terms of listener judgment are more accurate in assessing the acceptability of L2 sounds as produced by L2 learners. In assessing the L2 production ability of L2 learners, we need to consider both acoustic measurements and native listener judgment.

ACKNOWLEDGMENT

REFERENCES

[1] Anderson, V., Ko I., O'Grady, W., Choo, M. 2004. A palatographic investigation of place of articulation in Korean coronal obstruents. *Journal of Korean Linguistics* 12, 1−24.

[2] Bohn, O., Flege, J. 1990. Perception and production of a new vowel category by adult second language learners. In *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second - Language Speech,* edited by J. Leather and A. James. Amsterdam: University of Amsterdam.

[3] Cheon, S. 2006. *Phonetic similarities between English and Korean sibilants: Implications for second language acquisition*. Presented at the 152nd ASA Honolulu meeting.

[4] Cheon, S. 2006. Production of Korean fricatives in second language acquisition: Acoustic characteristics. *Korean Linguistics* 13, 17-48.

[5] Flege, J. 1995. Second language speech learning: theory, finding, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-linguistic Research*, edited by W. Strange. MD: York Press.

[6] Jongman, A, Wayland, R. Wong, S. 2000. Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am.* 108, 1252-1263.

[7] Major, R. 2001. *Foreign Accent: the Ontogeny and Phylogeny of Second Language Acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

---

[i] In Experiment 2, listeners were also presented with the vowel, and their reactions to the sibilants might have been influenced by non-native vowel productions. I regret that the potential factor like a vowel environment was not taken into consideration for data analyses.

[ii] The production scores as a dependable variable yielded a value of either 0 or 1, where 0 represented incorrect responses and 1 correct or acceptable responses. However, as a reviewer suggested, a multiple scale like a 3, 5, or 7 points rather than a binary scale would be better.