Perception and Production in the Acquisition of L2 Phonemic Contrasts

Fred R. Eckman[1], Gregory K. Iverson[1], Robert Allen Fox[3], Ewa Jacewicz[3], Soyoung Lee[2]

[1]University of Maryland Center for Advanced Study of Language,
[2]University of Wisconsin-Milwaukee, [3]The Ohio State University
feckman@casl.umd.edu, iverson@uwm.edu

1. Introduction

This paper reports data from a pilot project that is part of an ongoing study investigating the second-language (L2) acquisition of phonemic contrasts. Our focus will be on the acquisition of the English /s/ - /š/ contrast by native speakers of Japanese and Korean. These native language backgrounds were chosen because Japanese and Korean employ the sounds in question as allophones of the same phoneme, with [š] occurring before high front vowels, [s] elsewhere. The task of the L2 learner in this case is to acquire the contrast by suppressing the application of the native language (NL) allophonic rule in pronunciation of the target language (TL) words.

Universal principles of phonology predict that learners will be able to suppress the application of the allophonic rule across a morpheme boundary only if they are also able to inhibit the rule from applying within morphemes. In other words, a learner will acquire the contrast in environments across morphemes only if that learner has acquired the contrast within morphemes. The purpose of this paper is to test this prediction in both the production and perception of the /s/ - /š/ contrast by Japanese and Korean learners of English.

The paper is structured as follows. The Background section sets the context for the study by briefly reviewing some of the literature on the relationship between perception and production of contrasts in second language acquisition (SLA), by illustrating the necessary phonological principles and constructs, and by stating the hypotheses underlying this pilot study. The Methodology section lays out the procedure used to elicit, record, transcribe and score the data. The next section describes the results of the study as they pertain to the hypotheses, which, in turn, is followed by a discussion of these results. The final section concludes the paper.

2. Background

One of the overarching questions in the L2 acquisition of phonemic distinctions is whether learners have to perceive contrasts before they can successfully produce them. Although it would seem to be intuitively clear that a learner must perceive any given contrast before being able to implement it in production, the literature in fact attests all four logical possibilities.

Two of these possibilities are straightforward and require little or no elaboration. There are numerous documented examples of learners who can neither perceive nor produce a TL contrast, as well as instances in which L2 learners can both perceive and produce TL contrasts (e.g., Bion et al., 2006, and others). The other two logical possibilities are not as straightforward, and therefore require discussion.

Flege's (1995) Speech Learning Model (SLM) is predicated on the notion of "equivalence classification" in which an L2 learner sets up categories for TL phonemes on the basis of that learner's perception of the segments in question. Examining correlations between perception and production in results from selected studies on L2 consonants and vowels (Flege, Bohn & Jang, 1997; Flege, MacKay & Meador, 1999). Flege (1999) found that L2

segmental production and perception are generally correlated, although factors such as age limits or methodological differences among the variables investigated by different researchers may weaken the correlation. Similarly, speaking rate and clarity may have differed across studies so that "there is no guarantee that the perception and production tests are matched in terms of rate and clarity" (p. 1275).

The results from several vowel studies, specifically, show that, in each case, there is a relation between their perception and production. Experience with the L2 and the learner's native language background are the two other factors that influence the way L2 vowels are perceived and produced.

Ingram and Park (1997) examined the influence of L1 vowel quality on the acoustic and perceptual outcome of the acquisition of L2 vowels. Perception and production of Australian English monophthongal vowels /i, ɪ, e, ɛ, a/ by Korean speakers were investigated in relation to cross-generational differences within the Korean speech community. The results indicate that representation of vowel contrasts in the L1 is a strong guiding principle in formation of contrasts in the L2.

Flege et al. (1997) assessed the relation between vowel perception and production as a function of experience with an L2 (English). Ninety subjects participated who were native speakers of German, Spanish, Mandarin, and Korean, and who differed in their length of residence in the United States. The major finding of the study was that the experienced L2 speakers (mean 25 years of residence in the U.S.) produced and perceived the English vowels /i, ɪ/ and /ɛ, æ/ more accurately than the inexperienced ones.

Flege, MacKay and Meador (1999) investigated the relation between production and perception of ten English vowels by native speakers of Italian. The subjects were selected on the basis of age of arrival in Canada, which ranged from 7 to 19 years. The obtained data showed that accuracy in L2 vowel production was related to how accurately the vowels were perceived in the L2. The later the native speakers of Italian arrived in Canada, the less accurately they perceived and produced English vowels. Not surprisingly, the degree of accuracy in L2 production was related more closely to perception of L2 vowels than to perception of L1 vowels, indicating a direct relationship between perception and production of vowels in the L2, at least in the case of highly experienced L2 speakers.

Several other studies have focused on consonants. Aoyama et al. (2003) investigated Japanese-speaking learners of English on the perception of English /r/ and /l/. English /r/ is perceptually less similar to Japanese /r/ than is English /l/. According to the SLM, English /r/ should be acquired by Japanese learners faster than English /l/, and this is what the results supported. More recently, Kluge et al. (2007) studied the production and perception of English /m/ and /n/ in coda position by native speakers of Brazilian Portuguese. The results showed a positive correlation between the perception and production tests.

It has also been reported that L2 learners' production of certain contrasts can exceed their ability to produce that contrast. Sheldon and Strange (1982) argued that L2 learners can produce a TL contrast better than they can perceive it. The authors tested native speakers of Japanese learning English on their ability to perceive and produce the distinction between /r/ and /l/, a contrast lacking in Japanese. Specifically, it was found that native speakers of English, when listening to recordings of the subjects' productions of minimal pairs containing /r/ and /l/, could successfully distinguish /r/ from /l/ better than the subjects could distinguish this contrast in their own productions.

Within the context of the above-referenced work on the relationship between the production and perception of phonemic distinctions by L2 learners, this paper reports the results of a pilot study investigating the production and perception of a TL contrast under two conditions. The first is where the contrasting TL segments are allophones of the same phoneme in the NL, and the second is where the segments are in two phonological contexts designated as BASIC (or NON-DERIVED), on the one hand, and DERIVED, on the other. These

terms refer to the morphological structure of representations that would trigger the application of a rule, in this case, the NL allophonic rule. In general, representations to which the rule can apply are basic, or non-derived, if the form is mono-morphemic. For example: *she* and *sea* are both single morphemes, and any rule defined on the phonological content of these forms would be applying in a basic, or non-derived, environment. A representation in which the rule could apply over the span of more than one morpheme comprises a derived environment. For example, the words *messy* and *passing* represent derived environments with respect to application of a rule that affects the fricative /s/ before the vowel /i/.

Along these lines, we are interested in the ability of Japanese and Korean L2 learners of English both to produce and perceive the contrast between /s/ and /š/ in basic and derived environments.

This contrast was chosen because the phonemic status of the two segments is different in these two languages than it is in English. In Japanese, there is a surface contrast between the two sounds when they occur before the vowels [a, o, u]; before the vowel [e] only [s] (and not [š]) occurs natively, and before [i] only /š/ (and not [s]) appears. In Korean, on the other hand, [s] and [š] are in complementary distribution, with [š] occurring only before high front vowels and [s] elsewhere. Thus, a rule changing /s/ to [š] before [i] is motivated for the grammar of both of these languages. For Japanese and Korean learners of English to acquire a contrast between these segments, the learners must suppress the application of their NL rule. If this rule is transferred into their interlanguage (IL) grammar, then words such as *seep* and *sheep* would be pronounced homophonously.

Given this background, there are two hypotheses underlying the study, stated in (1) and (2) below.

(1) Hypothesis 1
In the production data, a subject's IL grammar will evidence a contrast between /s/ and /š/ in derived environments only if the contrast is found also in basic environments.

(2) Hypothesis 2
In the perception data, an IL grammar may show the contrast between /s/ and /š/ in derived environments without also having it in basic environments.

The rationale for Hypothesis 2 stems from the fact that virtually all (if not truly all) generalizations about L2 grammars have been made on the basis of data from production, not perception. Hypothesis 1 has its roots in two general principles of phonology, Structure Preservation and the Derived Environment Constraint, shown in (3) and (4), respectively.

(3) Structure Preservation
Representations within the lexicon may be composed only of elements drawn from the phonemic inventory.

(4) Derived Environment Constraint
Structure-preserving rule applications are restricted to derived environments (i.e., rule applications that involve phonemes of the language apply only across morpheme boundaries).

As outlined in Eckman et al. (2003), Structure Preservation and the Derived Environment Constraint have important implications for learnability, in general, and for the acquisition of the English contrast between /s/ and /š/, in particular. It follows from these principles that the acquisition of a TL phonemic distinction whose contrasting segments correspond to allophones of the same NL will take place in stages. At the beginning, when

the IL grammar lacks the contrast, the transferred NL rule will apply "across the board" in both basic and derived environments. As the learner begins to acquire the contrast in question, the two segments become part of the IL lexicon, and the Derived Environment Constraint permits the rule to apply only in derived contexts, that is, only across morpheme boundaries. The last stage would be one in which the contrast has become acquired to the point where the rule is suppressed altogether.

Taking the case of this study, we predict, following Hypothesis 1, the following stages of acquisition.

(5) Acquisition sequence according to Hypothesis 1

Stage I, No Contrast: not able to make the relevant target language contrast, applying the native language rule in both derived and basic environments (e.g., a Korean learner says the pairs *sea–she* and *messing–meshing* homophonously, as [ši] and [mɛšɪŋ]);

Stage II, Partial Contrast: able to make the contrast in some words, applying the native rule only in derived environments (a Korean learner says *sea–she* correctly but errs by producing *messing–meshing* homophonously)

Stage III, Contrast: able to make the contrast in all words, applying the native rule in neither derived nor basic environments (a Korean learner says the pairs *sea–she* and *messing–meshing* correctly);

Also according to the hypothesis, the following stage of acquisition would be excluded.

(6) Excluded stage: to make the contrast in some words, applying the native rule only in basic contexts (a Korean learner says the pair *sea–she* homophonously, but says *messing–meshing* correctly).

As embodied in Hypothesis 2, however, we do not predict that the stages of acquisition in (5) will necessarily hold for perception of the contrast.

Given this background along with the two hypotheses, we turn now to a description of our data gathering.

3. Method

For the purposes of this pilot study, data focusing on the ability to produce and perceive the contrast between /s/ and /š/ were gathered from four native speakers of Japanese and six native speakers of Korean. The subjects were between the ages of 19 and 35. The length of time that the subjects had studied English in their own country ranged from one to ten years (mean 7.2 years) and their stay in the U.S. ranged from one month to two and a half years (mean fifteen months). All four of the Japanese subjects, and four of the six Korean subjects were enrolled in the Intensive English Program at the University of Wisconsin-Milwaukee. Two of the Korean subjects were not students.

3.1 Production task

For the production task, a set of 60 target words and 30 fillers were elicited from each subject and recorded. (The list of words is shown in appendix A.) Excluding the fillers, each word was constructed to contain /s/ or /š/ in either a basic or derived context. The contexts for the application of the NL rule that specifies [š] rather than [s] before a high front vowel were word-initial for the basic environment (e.g., *sip/sick*), and medial before the suffix *-ing* or *-y* for the derived environment (e.g., *passing/messy*). For purposes of balance, we elicited words containing /s/ and /š/ medially and finally; however, these segments never occurred in an environment subject to the rule in question, because final segments are, by definition, not before a vowel, and because English monomorphemic words with intervocalic /s/ or /š/ occur chiefly before a reduced vowel (*license*, *lesson*, *message*) rather than the high front vowel.

The productions were elicited and recorded under computer control through a program written in Matlab, displaying on a computer screen a set of pictures and clues designed to evoke the word in question. Also on the screen were a set of commands such as "Wait" or "Speak" to guide the subject and the experimenter through the elicitations. If a word could not be elicited through the picture display or written clues that followed, a verbal model was played, and the subject repeated the word. The pictures were presented in a random order, and were recorded directly into a hard disc drive at the sampling rate of 44.1 kHz. The subjects spoke into a head-mounted microphone at a distance of one inch from the lips.

In order to implement a "blind" transcription, the words were elicited and recorded at the University of Wisconsin-Milwaukee and then sent to The Ohio State University, where they were transcribed by a native speaker of American English. The transcription process was fully automated using a program written in Matlab that displayed the following choices: "good /š/," "distorted /š/," "good /s/," "distorted /s/," "another fricative," and "comments on segment." After hearing the token, the transcriber selected an appropriate box on the computer screen.

3.2 Perception task

Stimuli for the perceptual testing consisted of single-word minimal pairs with either /s/ or /š/ in a basic or derived phonological environment. As was the case with the production task, the basic environment was word-initial before a high front vowel and the derived environment was intervocalic before the suffix *-ing* or *-y*. The words were recorded by a male, native speaker of American English, and were different from those used in the production task. (The list of words is shown in Appendix B.) The contrast in the basic medial position was not tested due to lack of a sufficient number of minimal pairs with /s/ and /š/ in intervocalic position. The stimulus set consisted of 72 items (4 stimuli x 2 contrastive segments x 3 positions in a word x 3 repetitions). The stimuli were presented diotically via Sennheiser HD600 headphones at a comfortable listening level (~70 dB HL). A single-interval, two-alternative, forced-choice (2AFC) identification task was used with the two response choices of /s/ and /š/ displayed on the computer monitor. After hearing the word, the subject indicated with a button mouse whether the word contained /s/ or /š/. The word tokens were presented in a random order. To make the perception task more demanding, the stimuli were presented in masking white noise at three different levels of sound-to-noise (S/N): 0 dB, -4 dB, and -8 dB.

4. Results

We report the results according to how they pertain to each of the hypotheses, beginning with the production data and Hypothesis 1.

In analyzing the production results, we counted as a successful production all transcriptions designated by the transcriber as a "good" instance or "distorted" instance of the intended consonant. Transcriptions that were categorized as "some other fricative" or "comment on segment" or which were designated as good or distorted instances of the segment that was not intended were counted as unsuccessful productions. For example, transcriptions of an intended /s/ that were designated by the transcriber as "good [s]" and "distorted [s]" were categorized as /s/; transcriptions of an intended /s/ that were designated as "good /š/" or "some other fricative" were scored as unsuccessful productions. A subject's performance was scored in both basic (word-initial before a high front vowel) and derived (intervocalic before -*ing* or -*y*) environments. For the production data, the basic environment was word-initial before a high front vowel, and the derived environment was intervocalic before the suffix -*ing* or -*y*.

We considered a contrast to be acquired if the subject's performance on each member of the contrasting pair of segments reached the 80% criterial threshold. The hypothesis is supported if a subject exhibits the contrast a) in both derived and basic environments; b) in neither derived nor basic environments; c) in basic, but not derived, environments. The hypothesis is not supported if the subject exhibits the contrast only in derived environments.

Turning first to the results for the six Korean subjects, we see in Table 1 that the performance of each of the Koreans supports Hypothesis 1: subjects K1, K2 and K4 support the hypothesis by virtue of having the /s/ - /š/ contrast in both basic and derived environments; subject K3 is also supportive by evincing the contrast in the basic environment but not in the derived environment; and K5 and K6 lack the contrast in both environments.[1] There are no subjects evidencing the excluded stage of having the contrast in derived environments but lacking it in basic contexts.

Table 1. Production (in %) of target words with the segment in initial position (basic environment) and in medial position (derived environment). Responses are shown for each individual subject.

	Acceptable /s/		Acceptable /š/	
Subject	Initial basic	Medial derived	Initial basic	Medial derived
J1	70	95	100	100
J2	90	95	80	95
J3	60	100	80	90
J4	50	100	100	100
K1	90	95	100	90
K2	100	100	100	100
K3	90	100	100	75
K4	100	95	100	95
K5	100	95	100	50
K6	30	15	100	100

[1] The percentage of correct productions of /s/ in the derived environment for K5, 95%, is higher than what one would expect for a Korean ESL learner who lacks the /s/ - /š/ contrast in this position. For the time being, we can only speculate on the basis for this score, but the unexpectedly high performance may lie in the fact that the transcriptions are in fact truly "blind". The 95% performance on intended /s/ by K5 is in fact a composite of 45% "good /s/" and 50% "distorted /s/". The Korean (as well as Japanese) /š/, unlike the English /š/, is unrounded, i.e., produced with spread lips. It is possible, then, that, because of these phonetic differences, the transcriber categorized the speaker's un-English-like production of a Korean /š/ as "distorted /s/".

The production results for the four Japanese subjects are different from those for the Koreans. Subjects J1, J3 and J4 show the contrast in the derived environment but not in the basic context. Only J2 is consistent with Hypothesis 1 by exhibiting the contrast in both environment types. We will take up this difference in the performance of the Japanese and Korea subjects below in the Discussion section.

The results from the perception data bear on Hypothesis 2 and are shown in Table 2. These data are supportive of Hypothesis 2 in that all four of the logically possible stages of acquisition are attested, including the "excluded" stage shown in (6) above. Looking first at the results in the 0 dB signal-to-noise ratio (S/N), we see that subjects J1, J2, J4, K1, K2 and K4 perceive the contrast in both basic and derived environments; subjects K3, K5 and K6, on the other hand, have the contrast in neither environment; subject J3 represents the excluded stage by showing the contrast in derived contexts, but not in the basic environment. None of the subjects has the contrast in basic environments but not in derived. The results are similar, but more robust, when the speech signal is further degraded. Thus, under -4 dB S/N, four subjects have the contrast in both basic and derived environments (J2, J4, K2, K4), four subjects lack the contrast in these two contexts (J1, J3, K5, K6), and two subjects exhibit the excluded stage (K1, K3). Under further degradation of the signal (-8 dB S/N), only one subject has the contrast in both environment types (J4), six lack the contrast in both contexts (J1, J3, K1, K3, K5, K6), and one subject shows the excluded stage (J2).

Table 2. Identification responses (in %) for the /s/-/š/ contrast in initial (basic) and medial (derived) position presented at three levels of signal-to-noise ratio (S/N). Responses are shown for each individual subject.

	Overall /s/-/š/ correct						
Subject	S/N = 0 dB		S/N = -4 dB		S/N = -8 dB		
	Initial	Medial	Initial	Medial	Initial	Medial	Final
J1	71	100	54	75	54	75	
J2	79	100	79	88	71	88	
J3	75	92	71	71	71	71	
J4	100	100	100	96	83	96	
K1	92	88	75	100	63	71	
K2	100	100	96	100	75	96	
K3	75	79	63	100	58	58	
K4	100	92	88	100	75	96	
K5	54	67	67	54	54	71	
K6	63	46	54	46	54	58	

5. Discussion

We begin with the findings for Hypothesis 1. There are two points to observe with respect to Hypothesis 1. The first is the fact that the data from the Korean subjects supported the hypothesis and those from the Japanese subjects did not. The performance of the six Korean subjects attested the three stages predicted by Structure Preservation and the Derived Environment Constraint, specifically, that an IL grammar that exhibited the contrast between /s/ and /š/ in derived environments would necessarily exhibit the contrast in basic environments, but not vice versa. Thus, three of the subjects, K1, K2 and K4, had the contrast in both contexts; K3 had the contrast only in basic environments but not in derived positions, and K5 and K6 lacked the contrast in both environments. The production results from the four Japanese, on the other hand, were contrary to the predicted stages in that three of the subjects,

J1, J3 and J4, had the contrast in the derived environment but lacked it in the basic environment. Why do the two subject groups respond differently to the contrast?

As we have results from only ten subjects and therefore our claims in this arena must remain tentative, we nevertheless wish to suggest that the reason for this discrepancy revolves around the differing phonemic status of [s] and [š] in each language, as elucidated by Heo (2006).

In Korean, the sounds [s] and [š] are in complementary distribution, with [š] occurring always and only before the vowel [i] or the palatal glide [y] ([ši] 'poem', not *[si]) and [s] occurring elsewhere ([saram] 'person', not *[šaram]). Consequently, in Korean these sounds are allophones of the same phoneme, /s/, and their distribution motivates an allophonic rule whereby /s/ is realized as [š] before high front vowels. Whereas Japanese also employs the sounds [s] and [š], their distribution is different from that in Korean. In Japanese, /s/ and /š/ contrast word-initially before the vowels [a], [o] and [u] (e.g., [sakai] 'boundary' vs. [šakai] 'society'). Before [e] only [s] occurs natively (modulo recent borrowings such as [še:do] '(lamp) shade'), while before [i] only [š] appears ([šiki] 'ceremony', not *[siki]). A rule changing /s/ to [š] before high front vowels is motivated by the fact that stem-final [s] in Japanese alternates with [š] when a suffix beginning with [i] is added to the stem (/hanas+imasu/ → [hanašimasu] 'speak [polite]'). Our proposal is that it is the different distribution and status of these sounds in Korean and Japanese that accounts for the different performance of the Korean and Japanese learners on English /s/ and /š/.

When the Korean learners of English transfer the NL allophonic rule into the IL, the rule applies in both basic and derived environments and consequently the learners err on the /s/ - /š/ contrast in all words. As we have outlined above, when learners establish the /s/ - /š/ contrast in some words, /š/ enters the lexicon as a phoneme and the rule applies only in derived environments. The final stage is suppression of the rule altogether and successful production of the contrast in all contexts. These are the stages we see attested by the Korean subjects.

For the Japanese subjects, on the other hand, a contrast between /s/ and /š/ already exists in some NL words, and so this contrast can be transferred to the IL, along with the NL palatalization rule that changes /s/ to [š] before high front vowels. One other feature of the NL, the static absence of the /s/ - /š/ contrast within morphemes before high front vowels, can also be transferred into the IL grammar. An important aspect of the fact that [š] but not [s] occurs before [i] in Japanese is that this phenomenon is not related dynamically to the NL rule changing /s/ to [š]. There is no motivation, in other words, for the derivation of Japanese initial [š] before [i] via the application of a rule, as this is a non-alternating (albeit neutralized) morpheme-internal context, i.e., a Japanese word or morpheme such as [šiki] is phonemically /šiki/, not /siki/. Thus, the NL rule palatalizing /s/ to [š] before high front vowels in Japanese and the occurrence of /š/ to the exclusion of /s/ before word-initial high front vowels are derivationally independent of each other. If this idea is correct, then we would predict that the stages of acquisition outlined in (5) and the excluded stage in (6) would not come into play in the case of Japanese learners acquiring the English /s/ - /š/ contrast. And this seems to be what we have found: subjects J1, J3 and J4 were able to suppress the application of the NL palatalization rule, which applies only in derived environments, but still had difficulty overcoming the NL pattern of word-initial /š/ before high front vowels. J2, on the other hand, apparently was able to suppress both NL aspects, and evinces the contrast in both contexts.

We now turn to the results from the perception task, which bear on Hypothesis 2, and the claims that the stages outlined in (5) and (6) above may not hold for perception. Our data support this claim in that the subjects evinced all four of the logical possibilities for the combination of /s/ - /š/ contrasts in both basic and derived environments.

It is noteworthy, we believe, that universal generalizations such as the one asserting an implicational relationship between contrasts in basic and derived environments have been

formulated, exclusively to our knowledge, on the basis of production data. Thus, it is an open and empirical question whether these generalizations hold also for perception data. One basis on which one might expect the perception data not to adhere to the generalization is that the segments in question are contextually determined in the NL and thus more of a factor in production than in perception. Perhaps this is simply a restatement of the facts. In any event, this is the first test, to our knowledge, of whether a generalization postulated on the basis of production data is true also for perception.

6. Conclusion

This paper has reported the results of a pilot study on the acquisition of the English /s/ - /š/ contrast by native speakers of Japanese and Korean in terms of both production and perception. We tested two hypotheses concerning the acquisition of the contrast in basic and derived environments, one hypothesis pertaining to production and the other to perception. The first hypothesis was supported by the Korean subjects, whereas the second was confirmed by the subjects from both NL backgrounds. We suggested that differences in the phonemic status of /s/ and /š/ in the two native languages account for the difference in the production performance of the Korean and Japanese subjects.

Acknowledgements

This work was sponsored by a grant from the National Institutes of Health, grant number 5R01 HD046908-02, whose support is hereby gratefully acknowledged. The authors wish to thank Anne Hoffmann, John Olstad and Heather Povletich for their contributions.

References

Aoyama, K., Flege, J. E., Guion, S., Akahane-Yamada, R., & Tsuneo, Y. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 23,* 233-250.

Bion, R. A. H., Escudero, P., Rauber, A. S., & Baptista, B. O. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. *Proceedings of Interspeech 2006,* 1363-1366.

Eckman, F., Elreyes, A., & Iverson, G. (2003). Some principles of second language phonology. *Second Language Research*, *19*, 169-208.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech production and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, Md: York Press.

Flege, J. E. (1999). The relation between L2 production and perception. *ICPhS 99*, 1273-1276.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

Flege, J. E., McKay, I., & Meador, E. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America, 106*, 2937-2987.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia, 9*, 317-323.

Heo, Y. (2006). Postpalatal epenthesis in Korean and Japanese: The role of contrast. In W. O'Grady, et al. (Eds.), *Inquiries into Korean Linguistics II* (pp. 157-165). Seoul: Taehaksa.

Ingram, J., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25*, 343-370.

Kluge, C. D., Rauber, A. S., Reis, M. S., & Bion, R. A. H. (2007). The relationship between perception and production of English nasal codas by Brazilian learners of English. *Interspeech* 2007, 2297-2300.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech perception can precede speech perception. *Applied Psycholinguistics*, *3*, 243-261.

Appendix A: Production words

sink	fish
sick	wash
scissors	wish
seats	rush
see	crush
bicycle	brush
lesson	crash
license	splash
message	polish
motorcycle	busing
bus	crossing
cross	grassy
dress	kissing
grass	messy
mess	racing
race	dressy
pass	passing
price	pricey
kiss	pricing
she	brushing
sheet	bushy
ship	crashing
sheep	fishing
shoe	crushing
ocean	splashing
parachute	washing
patient	wishing
tissue	polishing
vacation	rushing
bush	

Appendix B: Words for the identification task

ship	clashing	clash
sip	classing	class
shock	leashing	leash
sock	leasing	lease
sheep	meshing	mesh
seep	messing	mess
shed	mashing	plush
said	massing	plus