LONG-TERM PHONOLOGICAL BENEFITS OF CHILDHOOD L2 EXPERIENCE IN A JAPANESE IMMERSION PROGRAM

Tetsuo Harada

Waseda University
tharada@waseda.jp

ABSTRACT

This study examines to what extent English-speaking adults who have attended a Japanese immersion program in childhood, in which many content subjects are taught in Japanese, can retain their L2 pronunciation ability even if L2 input dramatically decreases after they exit the program. The results show that the immersion graduates still retained their ability to control segmental timing (i.e., voice onset time (VOT), contrast between single and geminate stops in Japanese), although their L2 sounds deviated from those of monolingual speakers.

Keywords: immersion education, retention of L2 pronunciation, VOT, closure duration, Japanese.

1. INTRODUCTION

Age of acquisition effects have been well documented in the field of second language (L2) speech learning (e.g., [1, 8]). Most studies focused on the acquisition of L2 sounds in a naturalistic setting and claim that the earlier one gets exposed to an L2, the more likely s/he is to attain authentic pronunciation. Harada [4, 5] examined whether or not the findings in the naturalistic setting would apply to an instructional setting and found that English-speaking children in a Japanese immersion program, in which many content subjects were taught in Japanese, were successful in establishing phonetic categories for VOT and single/geminate stops in Japanese, but that their L2 sounds deviated from those of Japanese monolinguals and even their immersion teachers. On the other hand, adults who had overheard Spanish during childhood were found to have better Spanish pronunciation than adult learners of Spanish who had had no childhood experience with Spanish [7], which shows long-term benefits of childhood L2 experience in a naturalistic setting. However, it is unknown to what extent children in an instructional setting, that is, an immersion program, can retain their L2 pronunciation after they exit the program. This study addresses the following question: Do English-speaking immersion graduates, who were exposed to a substantial amount of Japanese at an early age, still retain their ability to distinguish between English and Japanese VOT and to contrast singletons with geminates in Japanese, even if the quality and quantity of input drop dramatically after childhood?

2. BACKGROUNDS

2.1. Production of VOT by L1 and L2 speakers

Japanese has substantially shorter VOT values than English (Japanese /p/ = 24 ms, /t/ = 32 ms, /k/ = 45 ms in [6]; English /p/ = 57 ms, /t/ = 78 ms, /k/ = 77 ms in [2]), although there is little agreement about whether or not Japanese is a short lag language [9].

With the data here in mind, it can be predicted that English-speaking learners of Japanese are likely to produce Japanese VOT with longer VOT than L1 Japanese speakers. Harada [4] found that American English-speaking children in an early total immersion program produced Japanese voiceless stops with significantly longer VOT values (/p/ = 61 ms, /t/ = 64 ms, /k/ = 73 ms) than monolingual Japanese children (/p/ = 19 ms, /t/ = 22 ms, /k/ = 37 ms). However, it is worth noting that the immersion children distinguished phonetically between Japanese and English VOT regardless of places of articulation (/p/ = 81 ms, /t/ = 87 ms, /k/ = 97 ms).

2.2. Production of single and geminate stops in Japanese by L1 and L2 speakers

In Japanese, single and geminate stops are contrastive as in *ita* "existed" and *itta* "said," and phonetically, the geminates are realized by holding the closure stage of the first stop for a certain period of time. Researchers do not agree on the closure duration ratio of geminate to single stops. Harada [5] measured closure duration of singletons

and geminates produced by monolingual Japanese children in grade 5 and found that the ratio of geminates with singletons was 2.44.

English has no phonemic distinction between single and geminate consonants; although there are double consonants across morpheme boundaries (e.g., catfish vs. cattail), these are only 1.8 to 2 times as long as single consonants [3]. Harada [5] analyzed the production of singletons and geminates by English-speaking children in an early total Japanese immersion program and found that their ratios of geminates with singletons were smaller, ranging from 1.51 to 1.99. However, statistical analysis indicated that the immersion children significantly differentiated between the single and geminate stops. This means that although their contrast between the two was not as clear, the Japanese immersion program allowed the children to make a phonological distinction between the two types of sounds.

3. METHOD

3.1. Participants

Three groups of paid participants consisted of 6 Japanese monolinguals, 6 American English monolinguals, and 6 English-speaking Japanese immersion graduates, whose L2 input dropped after they exited a Japanese immersion program in an elementary school. The type of the immersion program from which they graduated was early partial immersion, in which about 50% of the content courses were taught in Japanese from kindergarten or grade 1 to grade 5. In middle school (grades 6 to 8) 25 to 30% of the courses were instructed in Japanese, and in high school (grades 9 to 12) the Japanese input dramatically decreased to about 15% of the instruction. In high school, some students were in a regular Japanese course in a traditional curriculum. All the immersion graduates were native speakers of English, and were exposed to Japanese only in class. They were taking a 3rd year Japanese course at a university at the time of data collection. Their immersion teachers were either native speakers of Japanese or Japanese-English bilinguals who had been born in a Japanese-speaking family in the US.

3.2. Procedures

Each recording session consisted of a 20-minute face-to-face pronunciation elicitation test administered in a quiet room. Each participant was audiotaped while s/he interacted with the researcher in a quiet classroom. During each session the informant was shown pictures of objects which had been designed to elicit words including the target voiceless stop consonants.

3.3. Materials

The following criteria were taken into consideration in word selection: (a) the following vowel quality, (b) disyllabic words, and (c) the same accent pattern. Following a picture cue, the participants were asked to say a word, inserting it in the Japanese carrier phrase *sore wa* _____ *desu* "That is _____" or in the English carrier phrase *I see a* _____ *in the picture*. The subjects were asked to repeat each word in the VOT and geminate corpus three times, whereas they repeated the words in the singleton consonant corpus four times to balance out the number of the tokens in each category. Some words used in this experiment are *papa* 'papa,' *tako* 'octopus,' *kame* 'turtle' for Japanese VOT, *panda, taxi, candy* for English VOT, *happa* 'leaves,' *batta* 'grasshopper,' *mikka* 'the third day' for Japanese geminates, and *hata* 'flag,' *taka* 'hawk,' *saka* 'slope' for Japanese singletons.

3.4. Data Measurement

The total number of tokens measured for Japanese and English VOT was 486 tokens each. The VOT of initial stops was measured to the nearest millisecond from the beginning of the release burst to the onset of voicing energy in F2 formants. Also, when the onset of F2 formants was not clear, VOT was measured from the beginning of the release burst to the first positive peak in the periodic portion of the waveform. A total of 6 mean VOT values for initial /p, t, k/ in English and Japanese, which were based on 9 observations each, were calculated for each bilingual. For the monolinguals, a total of 3 mean VOT values were obtained.

For measurement of closure duration, the total number of tokens analyzed was 867 tokens: 408 for single stops and 459 for geminate stops. The voiceless stops—[p], [pp], [t], [tt], [k], and [kk]—were identified by a gap on the spectrogram showing the stop closure. The beginning of the closure was defined by a cessation of the F2 of the preceding vowel. The end of the closure was defined by the burst of the following single or geminate consonant. When the burst was not

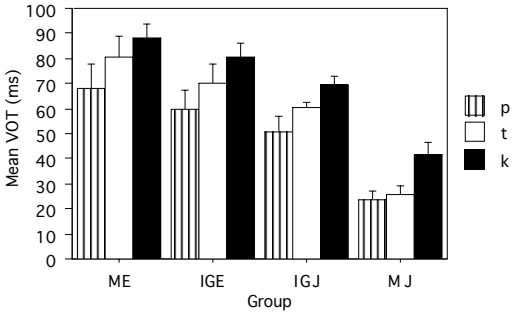
visible, it was measured up to the beginning of frication. The waveform was used as secondary information. Mean closure duration values for /pp, tt, kk/ and /p, t, k/ for each participant were calculated based on nine and eight observations, respectively.

4. RESULTS

4.1. VOT

Figure 1 shows that the Japanese immersion graduates (IGJ) produced Japanese voiceless stops with longer mean VOT values (60 ms) than the Japanese monolinguals (MJ: 30 ms), while they tended to produce them with shorter VOT values than their own English counterparts (IGE: 70 ms). It is also worth noting that their English VOT values were shorter than those of the English monolinguals (ME: 79 ms). The mean VOT values obtained for each participant were submitted to a (4) Group and (3) Place of Articulation two-way ANOVA, which yielded significant main effects but no interaction between them; group: $F(3, 60) = 35.979$, $p = .000$; place: $F(2, 60) = 10.099$, $p = .000$; Group x Place: $F(6, 60) = .162$, $p = .986$. Pair-wise comparisons revealed that although they were significantly different from those of the Japanese monolinguals ($p = .000$), the immersion graduates' VOT values in Japanese were marginally different from their VOT values in English ($p = .052$) and significantly different from the English monolinguals' ($p = .000$) regardless of the place of articulation. The results show that the Japanese immersion graduates retained their ability to produce VOT in Japanese after they exited the program and their Japanese input dramatically decreased.

Figure 1 The mean VOT values across places of articulation for English voiceless stops by the English monolinguals (ME) and the immersion graduates (IGE), and Japanese voiceless stops by the immersion graduates (IGJ) and the Japanese monolinguals (MJ). The error bars enclose ± one standard error.



4.2. Closure Duration

As shown in Figure 2, the Japanese monolinguals (MJ) produced their singletons with smaller closure duration values than the geminates (72 ms vs. 168 ms). Their geminates were more than two times the length of the single consonants. The immersion graduates (IGJ) produced singletons with longer duration values and geminates with shorter duration values (94 ms vs. 147 ms) than the monolinguals. To control for the speech rate, the mean duration values obtained for each participant were submitted to a Group (2) and Type (6) two-way ANCOVA, in which the duration of the initial word *orewa* in the carrier phrase was a covariate. The procedure yielded a significant main effect due to type only and a significant interaction between group and type; group: $F(1, 59) = .016$, $p = .899$; type: $F(5, 59) = 45.331$, $p = .000$; Group x Type: $F(5, 59) = 4.369$, $p = .002$. This means that the immersion graduates did not significantly differ from the monolinguals. In order to determine whether the immersion graduates can clearly differentiate singletons from geminates at each place of articulation, the data only for the immersion graduates were analyzed using a Type (6) one-way ANCOVA, which revealed a main effect due to type; $F(5, 29) = 9.158$, $p = .000$. Pair-wise comparisons showed that they distinguished singletons from geminates at each place of articulation ($p = .001$ for /p/ vs. /pp/, $p = .000$ for /t/ vs. /tt/, $p = .004$ for /k/ vs. /kk/).

Figure 2 The mean closure duration of single and geminate stops across the three places of articulation for the monolingual Japanese speakers (MJ) and the immersion graduates (IGJ). The error bars enclose ± one standard error.

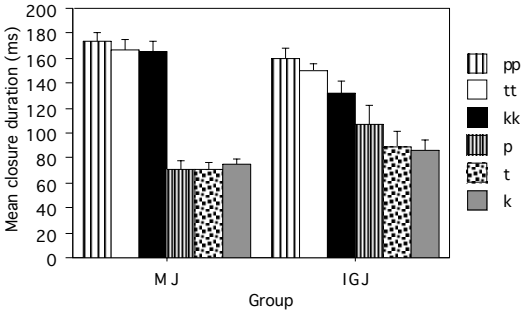
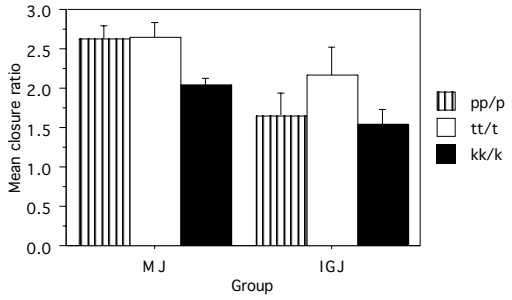


Figure 3 shows that the monolinguals (MJ) made the contrast between the two types of stops with ratios of more than two times (2.44). However, the ratio for the immersion graduates (IGJ) was smaller (1.78). A Group (2) and Type (3:

pp/p, tt/t, kk/k) two-way ANOVA revealed a main effect due to group and type but no interaction between group and type; group: $F(1, 30) = 12.418$, $p = .001$; type: $F(2, 30) = 3.774$, $p = .035$; Group x Type: $F(2, 30) = .761$, $p = .476$. This means that in terms of the ratio of geminates to singletons the immersion graduates significantly differed from the monolinguals. In sum, although their contrast between the two types of sounds was not as clear as that of the Japanese monolingual speakers, they retained the ability to make the contrast which they had acquired in their Japanese immersion program in childhood.

Figure 3 The mean closure duration ratio of geminates to singletons across the place of articulation for the monolingual Japanese speakers (MJ) and the immersion graduates (IGJ). The error bars enclose ± one standard error.



5. DISCUSSION AND CONCLUSION

The immersion graduates' success in both the phonetic distinction between Japanese and English VOT and the contrast between single and geminate stops in Japanese may be attributable to their long-term retention of the L2 phonetic categories established during their enrollment in an early immersion program. The long-term benefits of childhood L2 experience in a naturalistic setting reported in [7] may apply to L2 exposure in an immersion program, where a large amount of L2 input is guaranteed. These findings may hypothesize that once a phonetic category is established during childhood in a naturalistic or instructional setting, it may be retained till adulthood even if L2 input decreases after childhood.

However, we must note that the immersion graduates left their L2 phonetic categories to have deviated from L1 speakers' and could not fine-tune them even long after they exited the immersion program. The immersion children's longer VOT and smaller ratio of geminates to singletons in Japanese, which were also found in [4, 5], remained as they were in adulthood. In other words, L2 pronunciation competence acquired resulting from a large amount of L2 input may be invariable till adulthood without deteriorating to any large extent, even if the input decreases dramatically after childhood.

Finally, we need to keep in mind that to argue strongly for the findings in this study, the data from participants with no L2 input during childhood who are enrolled in a university-level intermediate or advanced Japanese course must be analyzed for comparison. In addition, a perception study is called for to investigate the retention of a phonetic category established in childhood.

6. REFERENCES

[1] Flege, J. E. 1995. Second-language speech learning: Theory, findings, and problems. In: Strange, W. (ed), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 233-277.

[2] Flege, J. E., Munro, M. J., MacKay, I. R. A. 1995. Effects of age of second-language learning on the production of English consonants. *Speech Communication* 16, 1-26.

[3] Han, M. S. 1992. The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica* 49, 102-127.

[4] Harada, T. 2002. The acquisition of voice onset time by children in a Japanese immersion program. In: James, A., Leather, J. (eds), *New Sounds 2000: Proc. of the Fourth International Symposium on the Acquisition of Second-Language Speech*. Klagenfurt, Austria: University of Klagenfurt, 158-165.

[5] Harada, T. 2006. The acquisition of single and geminate consonants by English-speaking children in a Japanese immersion program. *Studies in Second Language Acquisition* 28(4), 601-632.

[6] Homma, Y. 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9, 273-281.

[7] Knightly, L, M., Jun, S.-A., Oh, J. S., Au, T. K. 2003. Production benefits of childhood overhearing. *J. Acoust. Soc. of Am*. 114(1), 465-474.

[8] Long, M. 1990. Maturational constrains on language development. *Studies in Second Language Acquisition* 12, 251-285.

[9] Vance, T. J. 1987. *An introduction to Japanese phonology*. Albany: State University of New York Press.