L2 Influence on L1 Speech in the Production of VOT

Tetsuo Harada
University of Oregon, Eugene, OR 97403-1248, USA
E-mail: tharada@oregon.uoregon.edu

ABSTRACT

This research investigates the effect of the acquisition of English voice onset time (VOT) by early Japanese English bilinguals on their production of the Japanese counterparts in L1. The data were collected from 6 monolingual English speakers, 6 monolingual Japanese speakers, and 6 early Japanese English bilinguals. Results show that the early bilinguals, who speak Japanese at home and English elsewhere, make a distinction between Japanese and English VOT regardless of the place of articulation. This means that the bilinguals have successfully established two different phonetic categories for Japanese and English; however, their Japanese VOT values have ended up being longer than those of the Japanese monolinguals, while their English categories are not statistically different from those of the English monolingual group. This may imply that in order to maintain phonetic contrast in a common phonological space not L2 sounds but L1 sounds can be deviated from L1 phonetic categories.

1. INTRODUCTION

One of the challenging issues in second language phonetics and phonology is how bilinguals produce an L2 sound. Recent studies show that even proficient bilinguals may not produce an L2 sound exactly as monolinguals do [e.g., 1]. This prediction corresponds to Grosjean's view that the "mixing" of the two languages occurs inevitably because their two language systems are both engaged [2]. This view implies that interference is bi-directional [1, 3, 4].

The effect of an L1 sound on the production of an L2 sound has been well documented [e.g., 5, 6, 7, 8, 9]. However, it is still uncertain how the experience with L2 affects the production of sounds in L1. This research investigates the effect of learning of English voice onset time (VOT) by early Japanese English bilinguals on their production of the Japanese counterparts in L1.

Both Japanese and English have voiced and voiceless stops. Homma reports that the mean Japanese VOT values of initial /p, t, k/ were 24 ms, 32 ms, and 45 ms, respectively while those of medial /p, t, k/ were 7 ms, 16 ms, and 24 ms, respectively [10]. More recently, Han reports word-medial VOT values, based on 12 tokens for each of ten subjects: /p/ = 7.7 ms, /t/ = 12.0 ms, and /k/ = 17.9 ms [11]. As for English, Lisker and Abramson report means of VOT values for initial voiceless stops both in isolation (/p/ = 58 ms, /t/ = 70 ms, /k/ = 80 ms) and in sentences (/p/ = 28 ms, /t/ = 39 ms, /k/ = 43 ms) [12]. Judging from these available English data, it is quite predictable that the VOT values of Japanese /p, t, k/ produced by native English speakers (NE) are longer than those of native Japanese speakers (NJ). Although her data are limited to medial single stops, Han claims that NE's VOT values for the Japanese voiceless stops were found to be longer than those of NJ (NE's Japanese VOT /p/ = 15.6 ms, /t/ = 19.6 ms, /k/ = 30.8 ms; NJ's Japanese VOT /p/ = 7.7 ms, /t/ = 12.0 ms, /k/ = 17.9 ms) [11]. But there are no available reports of Japanese English early bilinguals' VOT values for Japanese voiceless stops that are relevant to this research study.

The authenticity of VOT is related to the age of learning [e.g., 13, 14]. Flege has developed the speech learning model (SLM) to explain the age-related limit on the ability to produce authentic L2 sounds [1, 5]. This model assumes that a perceptual phonetic category for an L2 sound initiates accurate production of the L2 sound and "many L2 production errors have a perceptual motivation" [1]. In early stages of L2 acquisition, learners may substitute an L2 sound with the closest sound in L1, but as they get more exposed to L2, they may discern phonetic differences between the L2 sound and the L1 sound. This perceptual awareness may lead them to establish a new phonetic category and to correctly produce the target sound. The hypothesis in this model is that the likelihood that L2 learners will discern the phonetic differences depends on both the phonetic dissimilarity between the L1 and the target L2 sounds as well as the age of learning a second language. In other words, the greater the phonetic difference between L1 and L2 sounds is and the earlier L2 learners are exposed to a second language, the more likely they are to notice the phonetic difference, which will lead to authentic production of L2 sounds.

However, this model suggests that "even when categories are established for an L2 sound, the L2 sound might not be produced exactly as it is produced by native speakers." Traditionally, the term "interference" has referred only to the influence of the L1 on the production of an L2, but Flege claims that "cross-language phonetic interference is bi-directional in nature" [1]. He found that bilinguals produced stops in their L1 with VOT values resembling those of stops in L2. This example contradicts the traditional view that the dominant language may remain resistant to an influence by the non-dominant language, at least for certain phonetic dimensions [e.g., 15].

The primary goal of the present study is to test the

 ISBN 1-876346-48-5

hypothesis in Flege's SLM that cross-language phonetic interference is bi-directional in nature, specifically, among very proficient Japanese English bilinguals, who speak Japanese at home and English elsewhere. The study addresses the following research questions: a) is there any bi-directional interference in the production of Japanese and English VOT by early bilinguals; and b) is there any variation of bi-directional interference in the bilinguals' production of Japanese and English VOT across places of articulation.

2. METHOD

2.1 Subjects

The data for the study were collected from 6 Japanese monolinguals, 6 early Japanese English bilinguals, and 6 English monolinguals. The Japanese monolingual speakers participating in this study had never lived in an English speaking country and had never used English in their daily life. They were not always purely monolingual speakers because almost every student in Japan is required to study English as a foreign language in junior and senior high schools, but the subjects in this study did not have a functional command of English. The English monolingual speakers were undergraduate students at a university on the West Coast of the United States. They had never lived in a foreign country for more than six months and had never spoken other languages than English. The Japanese English bilinguals were from Japanese-speaking families in the United States and spoke Japanese at home and English elsewhere. Most of these subjects had been exposed to Japanese at a heritage language school or at a school in Japan for a certain period of time.

2.2 Procedures

First the English data for the early bilinguals were collected and then the Japanese data from the same bilinguals were elicited to compare their English and Japanese VOT. Each session consisted of a 20-minute face-to-face pronunciation elicitation test administered in a quiet room. During each session the informant was shown pictures of objects which had been designed to elicit words beginning with the target voiceless stop consonants. The data from the monolingual speakers were collected, using the same method.

2.3 Materials

The words were selected taking into consideration the following criteria: (1) the following vowel quality ([a] for Japanese words or [æ] for English words), (2) disyllabic words, and (3) the same accent or stress pattern (HL for Japanese VOT data, and stress on the first syllable for English VOT data).

Following a picture cue, the subjects 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 corpus three times. The corpus was as follows:

Japanese VOT corpus

/p/	/t/	/k/
papa (papa)	tako (octopus)	kame (turtle)
pari (Paris)	tane (seed)	kata (shoulder)
	tate (length)	kasa (umbrella)

English VOT Corpus

/p/	/t/	/k/
panda	tablet	carrot
parrot	tadpole	camel
package	taxi	candy

2.4 Data Measurement

Each of the experimental words occurred three times on each list except for *papa* in the Japanese VOT corpus. This means that the total number of tokens for the Japanese VOT corpus was 324 tokens (7 words x 3 repetitions x 12 subjects + 1 word x 6 repetitions x 12 subjects). For the English corpus there were 324 tokens (9 words x 3 repetitions x 12 subjects). Average VOT values for each place of articulation for each subject were based on 9 observations.

Macquirer, a speech analysis program, was used for data measurement. 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, the waveform was used as secondary information. VOT was identified in the waveforms from the beginning of the release burst to the first positive peak in the periodic portion of the following vowel.

3. RESULTS

Figure 1 shows the mean VOT values for Japanese voiceless stops by the Japanese monolinguals (JM), and the early Japanese English bilinguals (JB), and the mean VOT values for English counterparts by the same bilinguals (EB) and the English monolinguals (EM).

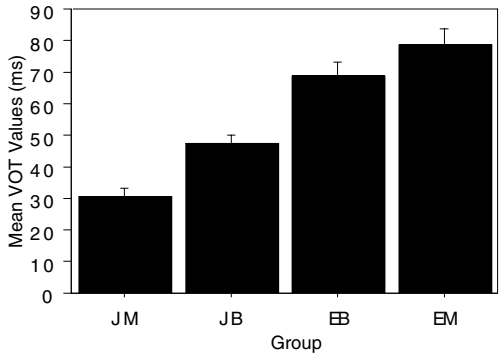

Figure 1 The mean VOT values for Japanese and English voiceless stops. The error bars enclose +/- one standard error.

As expected, the monolingual Japanese speakers' voiceless stops had substantially shorter VOT values than those of the early bilinguals' Japanese (30 ms vs 47 ms). Also, all the bilinguals produced Japanese voiceless stops with shorter VOT values than their English voiceless stops (47 ms vs 69 ms), while their English voiceless stops had shorter VOT values than those of the English monolinguals (69 ms vs 79 ms).

Figure 2 will demonstrate the differences in VOT values across places of articulation for all of the subjects.

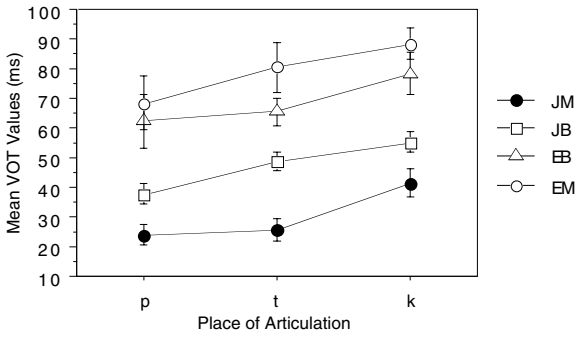

Figure 2 The differences in VOT values across places of articulation for all of the subjects. The error bars enclose +/- one standard error.

The mean VOT values obtained for each of the subjects were submitted to a (4) Group and (3) Place of Articulation ANOVA, which yielded a significant group main effect [Group, F (3, 60) = 40.702, $p < .0001$; Place, F (2, 60) = 9.171, $p = .0003$]. The duration pattern of VOT relative to the place of articulation is similar across the four groups: VOT is longest for /k/ ($p < .05$) and there is no significant difference between /p/ and /t/ ($p = .2555$). In addition, since there was no interaction between Group and Place [Group * Place, F (6, 60) = .252, $p = .9566$], the mean Japanese VOT values produced by the bilingual subjects were greater than those of the monolingual Japanese subjects regardless of the place of articulation. However, though statistically not significant, there is a trend that the bilinguals' mean VOT values for /t/ in both Japanese and English are more deviant from those of the monolinguals. Since there was no significant difference across places of articulation, the mean VOT values obtained for each of the subjects were submitted to a (4) Group one-way ANOVA. The analysis yielded a significant group main effect [F (3, 68) = 34.660, $p < .0001$]. Scheffe post hoc tests revealed that the bilinguals produced Japanese voiceless stops with significantly longer VOT values than the monolingual Japanese speakers, but they produced them with significantly shorter VOT values than their English VOT ($p < .0018$). This suggests that the bilinguals are making a phonetic distinction in VOT values between Japanese and English. However, the mean VOT values for Japanese voiceless stops produced by the bilinguals are significantly different from those of the Japanese monolingual group, while those of English voiceless stops produced by them are not significantly different from those of the English monolingual group ($p = .2830$).

4. DISUCSSION

This study clearly suggests that the early bilinguals, who speak Japanese at home and English elsewhere, make a distinction between the Japanese and English VOT values regardless of the place of articulation. This means that the bilinguals have successfully established two different phonetic categories for Japanese and English VOT; however, their Japanese VOT categories are not the same as those of the Japanese monolingual speakers. But their English categories are not statistically different from those of the English monolingual group.

The success in establishing a new category for English VOT shows that the bilinguals were able to notice the slight phonetic difference in VOT between English and Japanese. This finding supports Flege's hypothesis that bilinguals can establish a new phonetic category for an L2 sound only when they discern the phonetic difference between the L2 sound and the closest L1 sound [1]. However, it is worth pointing out that the bilinguals' Japanese VOT values are affected by those of English, their second language, and have ended up being longer than those of the Japanese monolinguals. Therefore, we can argue that there is clearly L2 interference in the production of Japanese VOT. But the finding that the bilinguals' English categories do not differ from those of the English monolinguals can imply that there may be no L1 interference in the production of English VOT.

The finding that the Japanese categories of the early bilinguals are deviated from those of L1 Japanese speakers may be accounted for by the maintenance of phonetic contrast in a common phonological space [1, Susan Guion, personal communication]. VOT is represented as either negative VOT values standing for "voicing lead (onset of glottal vibration prior to articulatory release)" or positive VOT values meaning "voicing lag (onset of glottal vibration following release)" [14]. For example, in utterance-initial voiced stops in Japanese, the voicing usually begins -35-0 ms on average before the release of the stop closure [16], while in utterance-initial voiced stops in English, it more frequently begins shortly after the release of the closure. Utterance-initial voiceless stops in Japanese and English are identified in terms of VOT: in Japanese voiceless stops, the voicing starts shortly after the release of the closure, whereas in English it starts 60-80 ms after the release of the closure. In other words, the English voiced stops fall roughly in the same acoustic space as the Japanese voiceless stops. This can be illustrated in the following diagram:

ISBN 1-876346-48-5

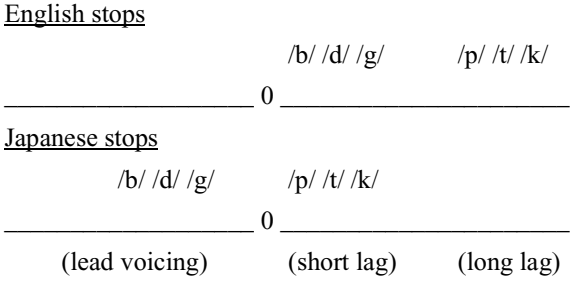

Figure 3 Voice onset time in English and Japanese

The conflict of English voiced stops with Japanese voiceless stops in the acoustic space may have caused the bilinguals' Japanese voiceless stops to be longer than those of the monolinguals so that they could maintain the phonetic contrast between the two stops. This may imply that in order to maintain phonetic contrast in a common phonological space not L2 sounds but L1 sounds can be deviated from L1 phonetic categories. Although this will require additional studies because VOT data for voiced stops were not collected in this study, due to the conflict of an L2 phonetic category with an L1 phonetic category a bilingual's L1 sound may not be produced in the same way it is produced by monolingual speakers.

ACKNOWLEDGMENTS

This research was supported by the Freeman Foundation Faculty Research Fellowship, and the Junior Professorship Development Award, both at the University of Oregon.

REFERENCES

[1] J. E. Flege, "Second-language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, W. Strange, Ed., pp. 233-277. Timonium MD: York Press, 1995.

[2] F. Grosjean, "Neurolinguists, beware! The bilingual is not two monolinguals in one person," *Brain and Language,* vol. 36, pp. 3-15, 1989.

[3] O. Bohn and J. E. Flege, "Perceptual switching in Spanish/English bilinguals," *Journal of Phonetics*, vol. 21, pp. 267-290, 1993.

[4] J. E. Flege, "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification," *Journal of Phonetics*, vol. 15, pp. 47-65, 1987.

[5] J. E. Flege, "Speech learning in a second language," in *Phonological Development: Models, Research, Implications*, C. A. Ferguson, L. Menn and C. Stoel-Gammon, Eds., pp. 565-604. Timonium MD: York Press, 1992.

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

[7] S. Oyama, "A sensitive period for the acquisition of a nonnative phonological system," *Journal of Psycholinguistic Research*, vol. 5, pp. 261-285, 1976.

[8] M. S. Patkowski, "The critical age hypothesis and interlanguage phonology," in *First and Second Language Phonology*, M. S. Yavas, Ed., pp. 205-221. San Diego CA: Singular Publishing Group, 1994.

[9] T. Scovel, *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech*, Cambridge MA: Newbury House, 1988.

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

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

[12] L. Lisker and A. Abramson, "A cross-language study of voicing in initial stops: Acoustical measurements," *Word*, vol. 20, pp. 384-422, 1964.

[13] J. E. Flege, "Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language," *Journal of the Acoustical Society of America*, vol. 89, pp. 395-411, 1991.

[14] L. Williams, "Phonetic variation as a function of second language learning," in *Child Phonology 2: Perception*, G. H. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson, Eds., pp. 185-215. New York: Academic Press, 1980.

[15] M. Mack, "Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals," *Perception and Psychophysics*, vol. 46, pp. 187-200, 1989.

[16] Y. Homma, "Voice onset time in Japanese stops," *Onsei Gakkai Kaiho*, vol. 163, 7-9, 1980.