Native and Non-Native English speakers' VOT Productions of Stops*

Mi-Ryoung Kim
(Korea Cyber University)

Kim, Mi-Ryoung. 2011. Native and Non-Native English speakers' VOT Productions of Stops. *The Linguistic Association of Korea Journal. 19*(1). 97-116. This study examined how native and non-native English (NE and NNE) speakers produced (L2) English stops in utterance-initial position in terms of Voice Onset Time (VOT). VOT was measured in English words spoken by NE (English monolinguals) and NNE speakers (Standard Chinese, French, Hindi, Japanese, and Korean). The results showed that NNE speakers produced remarkably different VOTs and their differences could be accounted for in terms of L1 influence and proficiency in L2. When the speakers' proficiency was taken into account, L1 influences on L2 speech held only for the low proficiency group. The results suggest that there is a close relationship between L1 influence and proficiency in L2 and that L1 influences on L2 speech diminish as L2 proficiency grows. Given that only a small number of participants for each language were used, it is hard to claim that this can be true for all L2 speakers. Further research is necessary.

Key Words: Voice Onset Time (VOT), L2 English stops, L1 influences, proficiency in L2 English, non-native or L2 speakers

1. Introduction

English has become a global language or world English (Canagarajah 2010)

---

* This paper was presented at a conference of the Linguistic Association of Korea on October 2010 in Gwangju, Korea. I am very grateful to the three anonymous reviewers for their insightful comments and helpful suggestions on an earlier version of this paper. All remaining errors and inadequacies are mine.

and hence it is often called "Globish (i. e. Global + English)." According to Crystal (1997), one out of four people in the world can communicate in English. Non-native speakers of English (hereafter, NNE speakers) have outnumbered native speakers of English (hereafter, NE speakers) by three to one. English can be frequently used as a means of communication even among NNE speakers whose native languages are different. When they produce English, it is easy to see that NNE speakers carry foreign accents because of the influence of their native language. The effect of L1 on L2 in pronunciation has been well documented[1)] (Flege and Eefting 1986, Flege 1992, Flege *et al.* 1995, Patkowski 1994). However, little research on how L2 in pronunciation can differ according to proficiency in L2 has been done.

Voice Onset Time (hereafter, VOT) refers to the time interval between the release of the stop and the onset of glottal vibration, or voicing. Since Lisker and Abramson (1964), VOT has become an important acoustic parameter to distinguish stops in the languages of the world (Ladefoged and Maddieson 1996, Shimizu 1996). It has also been used as a criterion to measure the effect of L1 on L2 in pronunciation (Flege and Eefting 1986, Flege 1992, Flege *et al.* 1995, Patkowski 1994, Fowler *et al.* 2008). In this study, VOT is measured to examine the effect of L1 on L2 in pronunciation.

English contrasts two stop categories: voiceless and voiced. Voiced stops are pronounced as either fully voiced or partially voiceless. These categories are reliably differentiated by VOT. In Figure 1, which shows wide-band spectrogram with displays of waveform, we have stops and vowel syllables illustrating three common conditions of VOT (Lisker and Abramson 1964: 389-390). In the first figure, voicing begins before the release of the stop; in the second, just after the release; in the third, voicing onset lags considerably behind the release. Measurements of VOT before the release are stated as negative numbers called *voicing lead*, while measurements of VOT after the release are stated as positive numbers called *voicing lag*. In Figure 1, we can see either voicing lead (left) or short voicing lag (middle) for English voiced stops. Voiceless stops in English carry a long voicing lag (right).

1) L2 (i.e., English in this study) is generally used as a cover term to refer to a target, foreign, or second language. L1 refers to a native language, mother tongue, or first language.

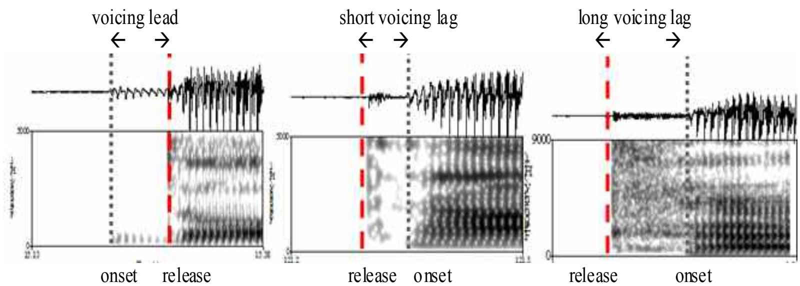


Figure 1. Three conditions of VOT (ms) for English /d/ and /t/ followed by the mid vowel: voicing lead for /d/, short voicing lag for /d/, and long voicing lag for /t/.

English voiced and voiceless stops can be reliably accounted for under the three conditions of voicing because they are phonetically fully voiced (i.e., voicing lead), voiceless unaspirated (i.e., short voicing lag), and voiceless aspirated (i.e., long voicing lag). For example, Lisker and Abramson (1964) found voicing lag for English voiceless stops (i.e., positive VOT values; its range between 58ms and 80ms). Both voicing lead and voicing lag were found for voiced stops (i.e., positive and negative VOT values; its range between -102 and 21ms) (see also Klatt 1975).

VOT has been applied to word-initial stops and has been found to be effective as a means of separating phonemic categories when some languages are utilized in conjunction with voicing to yield two, three or four categories. However, in some languages, categories are distinguished solely by differences in aspiration. One such case, where VOT alone is not enough, is Korean. Since stops in Korean are phonetically voiceless in initial position, they are hard to be defined under the three conditions of VOT alone. In addition to VOT, another condition (i.e., tone) has been adopted to define Korean stops appropriately (see Kim 2000, Kim *et al.*, 2002, Kim and Duanmu 2004 for consonant-tone interaction). Since the main concern of this study is VOT, more specifically native and non-native VOT productions of English stops, the discussion of other dimensions in languages is omitted. The research questions to be answered in this study are:

1. How do NE and NNE speakers produce English stops in terms of VOT?

2. Can NNE speakers' VOT productions be accounted for in terms of proficiency in L2 English?

2. Methods

2.1. Participants

The data were collected from nine participants, two NE and seven NNE speakers (i.e., 2 American English monolinguals vs. 2 Japanese speakers, 1 Standard Chinese or Mandarin, 1 Hindi, 1 French, and 2 Korean speakers). None of the participants reported being diagnosed with a language or reading disorder. Their mean age was 22 (range 18-25). Participants were all students of the East Tennessee State University at the time of the recording. Two NE speakers were monolinguals and used Standard English. They were born and raised in the mid-part of the USA. All NNE speakers were in the United States as students. Their residencies in the United States, except for one, were less than a month. Only the Korean male participant NK1 had attended high school in the United States for three years. The French participant was tested and found to be fluent in L2 English. She reported that she had lived in France in her life except for a few months. She had stayed in Boston, Mass. for six months when she was a high school student.

NNE speakers were divided into two groups according to their proficiency scale based on the self-report questionnaire: High-Proficiency (hereafter, HP group, scale 4-5) and Low-Proficiency (hereafter, LP group, scale 1-3) (see Appendix for the questionnaire items). The LP group consisted of two Japanese speakers (NJ1 and NJ2), one Korean speaker (NK2), one Hindi speaker (NI), and one Standard or Mandarin Chinese speaker (NC). The Hindi and Mandarin speakers were slightly more fluent than the other three. Although English is an official language in India, he himself reported that his English was not very good (Scale 3). He reported that his native language is Hindi Gujarati. The HP group consisted of one Korean student (NK1), and one French student (NF).

2.2. Speech Materials and Procedures

The target words, given in Table 1, were balanced for voicing (voiceless and voiced), place of articulation (labial, alveolar, and velar), and vowel context (/a/ and /aɪ/). All words were real words.

Table 1. English stops for VOT measurements

voiceless		voiced	
/pat/ "pot"	/paɪ/ "pie"	/bat/ "bot"	/baɪ/ "bye"
/tat/ "tot"	/taɪ/ "tie"	/dat/ "dot"	/daɪ/ "die"
/kat/ "cot"	/kaɪ/ "kye"	/gat/ "got"	/gaɪ/ "guy"

Each of the participants was recorded using a portable Panasonic recorder in a sound-attenuated room. In order to record the target words in the same phonetic environment, each word was presented in a carrier sentence. The carrier sentence was "___ *is the word*" where the target words were located in sentence-initial position. Thus, the target words were fully stressed and emphasized. The participants read sentences on flashcards three times in randomized blocks. They were instructed, in English, to read the cards at a comfortable speaking rate and loudness level and to repeat any utterances when they were unsatisfied with their production. Before the recording session began, participants were given a chance to rehearse the sentences. All speakers reported familiarity with the target words.

All utterances were digitally recorded at a sampling rate of 44,000 Hz and digitally transferred to a personal computer as wave files with the sampling rate using Praat 4.6.40. The VOT measurements were performed using Praat 4.6.40; a speech analysis program.

2.3. VOT Measurements

The VOT of the initial stop in each target word was measured from the beginning of the stop-burst release to the onset of the periodic portion using waveform and spectrogram. The onset of the vowel in the waveform was determined by the onset of the first full glottal pulse of the vowel as well as the pitch of the spectrogram. The onset of the voicing energy in the second formant shown in a time-locked spectrogram was used to help determine voicing onset in conjunction with the waveform. In the few productions with pre-voicing, the VOT was measured as a negative number as the time interval between the onset of periodic pulsing during the closure up to the stop release (Lisker and Abramson 1964). English voiced stops were occasionally prevoiced by some NE and NNE speakers.

3. Results and Discussion

3.1. Native VOT Productions of English Stops

As described in previous studies (Lisker and Abramson 1964, Klatt 1975, Keating 1984), English stops produced by NE speakers in this study showed three conditions of voicing: voicing lead and short voicing lag for voiced stops and long voicing lag for voiceless stops. Table 2 represents the mean VOT values of voiceless stops in English provided by two NE speakers.

Table 2. Voice Onset Time in Milliseconds (ms): Native English

	NE1			NE2		
	Mean	Range	SD.	Mean	Range	SD.
/p/	81	69~101	12	82	69~87	7
/t/	98	85~109	8	116	101~123	13
/k/	97	66~134	28	93	73~105	15
/p,t,k/	**92**	**66~134**	**19**	**97**	**69~123**	**19**
/b/	-26	-184~14	78	-92	-138~9	55
/d/	-28	-127~14	64	-100	-129~14	56
/g/	-68	-145~24	73	-72	-127~19	70
/b,d,g/	**-41**	**-184~24**	**71**	**-88**	**-138~19**	**59**

As seen in Table 2, the mean VOT values of voiceless stops for NE1 and NE2 were 92 ms and 97 ms, respectively. Both NE speakers produced very long VOT values (mean 95 ms across the two speakers) for voiceless stops. and they also produced a relatively wider range of VOT values between 66 ms and 134 ms, compared with previous studies (Lisker and Abramson 1964, Klatt 1975). For the place of articulations in voiceless stops, alveolars were the longest (mean 98 ms and 116 ms respectively), velars were longer (mean 97 ms and 93 ms respectively), and labials were the shortest (mean 81 and 82 ms respectively) for both speakers. Considering that the mean VOT values of English voiceless stops in previous studies (Lisker and Abramson 1964, Fowler *et al.* 2008) were about 72 ms, these two speakers' VOT values were relatively longer. This VOT lengthening might be due to the fact that the target words were fully stressed in utterance-initial position. It can be accounted for by a domain-initial strengthening effect (Keating *et al.* 2003). In this study, English stops were heavily aspirated similar to the aspirated stops of other languages such as Korean and Mandarin. Note that the aspirated stops of Korean and Mandarin have very long voicing lags (90 ms and 102 ms, respectively, in Shimizu 1996).

In Table 2, the mean VOT values of voiced stops in English for NE1 and NE2 were -41 ms and -88 ms, respectively. As expected from previous studies (Lisker and Abramson 1964), voiced stops were produced in two ways with either voicing lead (i.e., negative VOT value) or short lag voicing (i.e., positive VOT value). There was little speaker variation for voiceless stops but a noticeable speaker variation for voiced stops. Speaker NE2 was responsible for 95% of all the stops produced with a voicing lead whereas the other speaker NE1 produced 45% of such stops. Even for the same word, both voicing lead and short voicing lag were observed from two different tokens. The examples are shown in Figure 2.

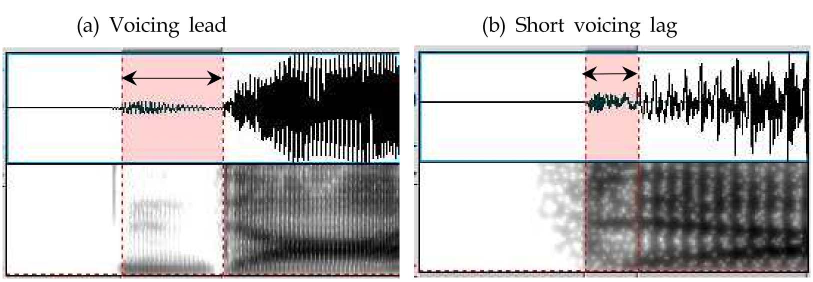


Figure 2. Voicing lead (a) vs short voicing lag (b) for the different tokens of the same word 'dot' produced by speaker NE1. The VOT values for (a) and (b) are −125ms and 16ms, respectively.

As can be seen in Figure 2(a), voicing lead consists of voicing onset, glottal voicing during the stop closure, a short silent duration without glottal voicing, and the release of the stop. It was comparable to Lisker and Abramson's voicing lead in Figure 1 where glottal voicing was entirely shown during the closure duration of the stop (Lisker and Abramson 1964: 390). In Figure 2(b), a short voicing lag consists of the release of the stop, some aspiration, and the onset of voicing. Note that the same word 'dot' was repeatedly differently in the tokens by the same speaker. Both NE speakers often produced voiced stops with either voicing lead or voicing lag (Keating 1984).

3.2 Non-native VOT Productions of English Stops

In this section, the results of NNE speakers are compared with those of NE speakers discussed above. The Japanese, Korean, Hindi, Mandarin, and French data are discussed below in turn. The mean VOT values of L2 English stops for the two Japanese participants' VOT productions are presented in Table 3.

Table 3. VOT in Milliseconds (ms): Non-Native English (Japanese)

	NJ1			NJ2		
	Mean	Range	SD.	Mean	Range	SD.
/p/	53	33~68	13	38	23~67	15
/t/	51	35~75	14	44	33~57	10
/k/	70	59~80	12	68	50~83	13
/p,t,k/	**58**	**33~80**	**15**	**50**	**23~83**	**18**
/b/	-41	-48~-31	8	14	9~20	4
/d/	-100	-66~11	56	-25	-134~22	69
/g/	-24	-48~20	33	-23	-75~26	52
/b,d,g/	**-37**	**-66~20**	**26**	**-12**	**-134~26**	**50**

Consider the results of voiceless stops first. As seen in Table 3, the mean VOT values of L2 voiceless stops for NJ1 and NJ2 are 58 ms and 50 ms, respectively. Their ranges are between 23 ms and 83 ms. Note that the values of Japanese speakers were about 40 ms shorter than NE speakers' values. These speakers were able to produce long VOT values over 80 ms. However, comparing Table 2 with Table 3, they overall produced shorter VOT values of voiceless stops than NE speakers. For voiceless stops, there was a little speaker variation between the two Japanese speakers. Note that both belonged to the LP group. However, there was some speaker variation for the results of voiced stops, as seen in Table 3. The mean VOT values of L2 voiced stops are -37 ms for NJ1 and -12 ms for NJ2. Both speakers produced English voiced stops with either a voicing lead or voicing lag similar to NE speakers. However, speaker NJ1 produced more voicing leads than speaker NJ2 (90% vs. 25%). Overall, we can say that the mean VOT duration and its range between Japanese and English speakers were remarkably different in that, for voiceless stops, Japanese speakers produced shorter voicing lag than NE speakers (mean 54 ms vs. 95 ms) and for voiced stops, Japanese speakers produced relatively shorter voicing lead than NE speakers (mean -24 ms vs. -68 ms). How the differences between NJ and NE speakers' VOT productions be explained? Can Japanese speakers' VOT productions in L2 English be predictable from their L1 speech?

In Japanese, there is a two stop category, voiceless and voiced. Japanese voiceless stops have shorter voicing lag than English voiceless stop. It has been reported that the mean VOT values of initial voiceless stops in Japanese were about 40 ms across the three places of articulation (Shimizu 1996, Harda 2003). This roughly corresponded to the present results of L2 voiceless stops produced by the Japanese speakers in this study. Japanese voiced stops have either voicing lead or short voicing lag similar to English voiced stops (Harda 2003). This also corresponded to the present results. It was likely that the two speakers in the present study often produced L2 voiced stops with either a voicing lead or voicing lag. The overall results of Japanese speakers suggested that there were some L1 influence on L2 speech. The influence held well for both speakers who belonged to the LP group. The results imply that, for a similar category, L2 stops can be produced similar to L1 stops. As a result, their phones in L1 and L2 may be phonetically similar. One question arises. Do HP Japanese speakers also carry the L1 influence on L2 speech similar to LP speakers? Since HP Japanese speakers were not taken into consideration in the present study, further study is necessary to claim the implication.

The VOT productions of L2 English stops for Korean speakers' NK1 and

NK2 whose proficiency levels in L2 English were remarkably different (NK1 = HP and NK2 = LP) are also examined. Table 4 presents the mean VOT of voiceless and voiced stops in L2 English produced by the two Korean speakers NK1 and NK2. Their mean VOT values are 79 ms and 109 ms, respectively. Their ranges were between 45 ms and 122 ms for NK1 but between 76 ms and 145 ms for NK2. Those of L2 voiced stops were -30 ms and 26 ms each. Its ranges were very different for the two speakers (-127 ms and 34 ms for NK1 vs. 5 ms and 48 ms for NK2).

Table 4. VOT in Milliseconds (ms): Non-Native English (Korean)

		NK1	HP		NK2	LP
	Mean	Range	SD.	Mean	Range	SD.
/p/	59	45~79	13	102	80~124	18
/t/	82	65~110	16	105	76~136	21
/k/	96	49~122	27	111	75~145	29
/p,t,k/	**79**	**45~122**	**24**	**106**	**76~145**	**22**
/b/	-27	-87~20	70	14	5~21	6
/d/	-27	-70~34	82	29	16~28	18
/g/	-35	-127~30	67	34	26~48	8
/b,d,g/	**-30**	**-127~34**	**69**	**26**	**5~48**	**14**

Unlike Japanese and NE speakers, there was remarkable speaker variation between these two speakers for both stop categories. For voiceless stops, NK1 produced shorter VOT than NK2. For voiced stops, NK1 was able to produce L2 voiced stops with either a voicing lead or a voicing lag whereas speaker NK2 was not. How can we interpret these speaker differences? Can they be accounted for in terms of proficiency in L2 English as well as L1 influence on L2 speech?

In Korean, there are three stop types, so-called tense (or fortis), lax (or lenis), and aspirated (or voiceless unaspirated, voiced, voiceless aspirated in Kim and Duanmu 2004). All the three stops are phonetically voiceless in utterance-initial position. Therefore, they are expected to have voicing lag only (i.e., positive VOT). According to Lisker and Abramson (1964), the mean VOT values of voiceless stops in Korean were 12 ms for the tense series, 30 ms for the lax series, and 103 ms for the aspirated series, respectively (see also Shimizu 1996). Since English voiceless stops and Korean aspirated stops are very similar in terms of the amount of aspiration, it is expected that, if there is a strong L1 influence on L2 speech, Korean L2 speakers produce L2 voiceless stops similar to L1 aspirated stops. In addition, since all three Korean stops are phonetically voiceless in utterance-initial position, it is expected that, if there is an L1 influence on L2 speech, Korean L2 learners were able to produce L2 voiced stops as voiceless stops without any voicing lead. These two aspects held well for NK2 but not for NK1. Note that NK2 belonged to the LP group but NK1 belonged to the HP group. The results suggested that the two Korean speakers' VOT variations can be accounted for in terms of L1 influence on L2 speech as well as their proficiency in L2 English.

Next, consider the results of the Hindi, Mandarin and French speakers together in Table 5 where the standard deviations are omitted.

Table 5. VOT in Milliseconds (ms): Non-Native English (Hindi, Mandarin, and French)

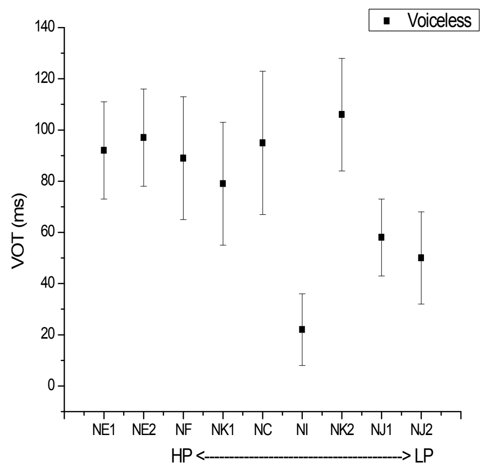
	NI		NC		NF	
	Mean	Range	Mean	Range	Mean	Range
/p/	18	15~21	82	45~104	99	80~106
/t/	11	9~12	103	52~148	85	61~123
/k/	38	20~57	100	76~106	82	51~99
/p,t,k/	**22**	**9~57**	**95**	**45~148**	**88**	**51~123**
/b/	-96	-171~-50	11	4~21	-8	-87~34
/d/	-82	-105~-28	15	12~22	0	-95~42
/g/	-116	-163~-80	23	16~27	9	-84~35
/b,d,g/	**-98**	**-171~-28**	**17**	**4~27**	**10**	**-95~42**

Remarkable speaker variations among the three speakers for the two stop categories can be observed in Table 5. For L2 voiceless stops, the results of the Hindi speaker NI with those of NC and NF speakers are compared. The mean VOT values of L2 voiceless stops were 22 ms for NI, 95 ms for NC, and 88 ms for NF. The VOT productions of NC and NF speakers corresponded to those of NE speakers. Unlike speakers NC and NF, speaker NI almost always produced L2 voiceless stops with a very short voicing lag ranging between 9 ms and 57 ms. Consider the results of L2 voiced stops. The mean VOT values were -98 ms for NI, 17 ms for NC, and 10 ms for NF. The Mandarin speaker NC consistently produced them with a short voicing lag (100%) whereas the Hindi speaker NI consistently produced them with a voicing lead (100%). Speaker NF alone produced voiced stops with either voicing lead or short voicing lag similar to the NE speakers. For this speaker, voicing lag showed higher percentage than voicing lead (80% vs 20%). How can we explain these speakers' differences where the same L2 stops were differently produced in terms of VOT? Can their differences be accounted for in terms of L1 influence and proficiency in L2 discussed above?

In the questionnaire, the Hindi speaker NI responded that his native language was Hindi Gujarati. In Hindi Gujarati, there are four series of stops including voiced stops, breathy voiced stops, voiceless unaspirated, and voiceless aspirated stops. According to Shimizu's study (1996), the mean VOT value of voiceless aspirated stops in Hindi is 91.3 ms. The value is similar to that of English voiceless stops where their mean VOT was 95 ms in the present study. Since voiceless aspirated stops between Hindi Gujarati and English are very similar in terms VOT, it is highly expected that the Hindi speaker produced L2 voiceless stops similar to his L1 aspirated stops. Very interestingly, however, he produced L2 English voiceless stops similar to L1 voiceless unaspirated stops where its mean VOT duration is 19 ms (Shimizu 1996: 130). In addition, when he produced L2 voiced stops, he consistently produced them with long voicing lead. According to Shimizu's results, Hindi voiced stops are produced with voicing lead alone. Overall, the Hindi speaker NI's results suggested that his VOT productions of L2 English stops were influenced by his L1 speech.

In Mandarin, there are two stop types, voiceless aspirated and voiceless unaspirated. According to Shimizu's study (1964), VOT values in Standard Chinese voiceless stops are relatively longer than those in English ones (102 ms vs. 72 ms). It was likely that speaker NC's mean VOT values corresponded to the NE speakers' (95 ms vs. 95 ms). However, her VOT range was much wider and longer (45~148 ms) than the NE speakers (66~134 ms). Considering that Standard Chinese stops are all voiceless in the initial position, it is expected that, if there is a strong L1 influence on L2 speech, Standard Chinese speakers may produce L2 voiced stops with a voicing lag only. As expected, speaker NC who belonged to the LP group produced voiced stops with a voicing lag only. The overall results of the Mandarin speaker suggested that her L2 speech was strongly influenced by her native language.

In French, there are two stop types with a voicing contrast similar to English as well as Japanese. Different from English and Japanese stops, French voiceless stops have a short voicing lag and voiced stops have a voicing lead. For example, [p] in French has VOT similar to a [b] in English when English /b/ is produced voiceless unaspirated. If there is an influence on L2 speech, it is expected that the French speaker produced L2 voiceless stops with a short voicing lag. However, the French speaker NF produced L2 voiceless stops with a long voicing lag instead of a short one. This indicated that there was little L1 influence on her L2 speech. Comparing with the NE speech, her VOT values were slightly shorter (95 ms vs. 88 ms) but very close to NE speakers' results. Unlike her native voiced stops where they were produced with voicing lead alone, she produced voiced stops with either voicing lead or voicing lag similar to the NE speakers'. As a result, her VOT productions for both voiceless and voiced stops were very close to the NE productions'. The results of the French speaker NF, who belonged to the HP group, suggested that she carried little L1 influence on L2 speech. Considering the fact that her exposure period toward the target language was very short (i.e, under 7 months), it was very surprising to see that the French speaker produced near native-like pronunciation in L2 English. She reported that she had learned English in schools in France. However, she was very fluent in English. The results of this speaker were comparable to those of Fowler *et al.* (2008) where even simultaneous bilingual (English-French) speakers carried cross language phonetic influences between the two languages.



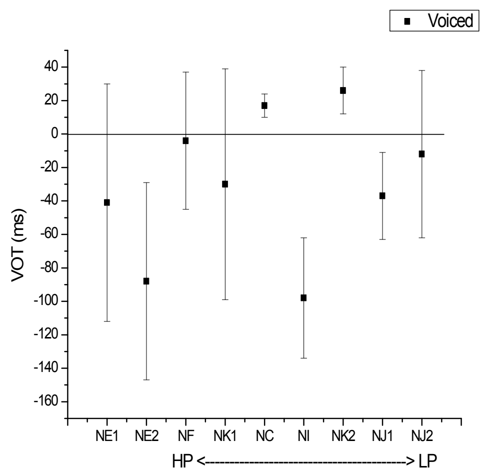


Figure 3. The mean VOT values of English voiceless stops (top figure) and voiced stops (bottom figure) across native and non-native speakers. (NE=Native English, NF=Native French, NK=Native Korean, NC=Native Standard Chinese, NI=Native Hindi, NJ=Native Japanese; HP=High Proficiency, LP=Low Proficiency)

3.3 The Relationship between L1 Influence and Proficiency in L2

Overall, NNE speakers produced L2 English stops very differently. Their productions were different from NE speakers' productions. The differences could be interpreted in terms of the two aspects: L1 influence on L2 speech and proficiency in L2. There was a close relationship between the two aspects. When NNE speakers were not fluent (i.e., LP), their productions of L2 stops were highly predictable from their L1 counterparts. When NNE speakers were fluent (i.e., HP), their productions of L2 stops were close to those of NE speakers. The relationship between L1 influence and proficiency in L2 can be clearly seen in Figure 3.

A look at the top figure for the results of voiceless stops in Figure 3 shows that, among NNE speakers' VOT productions, the HP speakers' productions (NF and NK1) are very close to the NE speakers' wherein their VOT values are completely or partially overlapped with NE speakers'. In contrast, the LP speakers' productions, except for NC and NK2, differ from NE speakers' productions. It was obvious that speakers NC and NK2 produced L2 stops as the VOT values of their L1 voiceless aspirated stops. Their L2 productions corresponded to those of L1 stops as can be seen in diagram (1).

(1) Mean VOT values for voiceless stops across languages (adapted from Shimizu 1996, Lisker and Abramson 1964, Fowler *et al.* 2008)

shorter voicing lag ← → longer voicing lag

Release burst (reference point)

0ms 50ms 100ms

French /p,t,k/ (15~40 ms)

Hindi /p,t,k/ (12~34 ms) and /p^h,t^h,k^h/ (75~119 ms)

Japanese /p,t,k/ (30~66 ms)

English /p,t,k/ (61~80 ms)

Korean /p^h,t^h,k^h/ (85~100 ms)

Standard Chinese /p^h,t^h,k^h/ (96~112 ms)

The diagram (1) shows that French has the shortest voicing lag (15~40 ms), Japanese the shorter voicing lag (30~66 ms), and English, Chinese, and Korean have a long voicing lag (60~120 ms) (adopted from Shimizu 1996 for the Standard Chinese, Japanese, and Korean data; Fowler *et al.* 2008 for the French and English data). From the diagram (1), it seemed clear that LP speakers, whose native languages were different, produced L2 voiceless stops differently because of their L1 influence. All the LP speakers' VOT results (NJ1, NJ2, NK2, NI, and NC) in the present study corresponded to the VOT pattern in (1) whereas the HP speakers' VOT patterns (NF and NK1) did not. Unlike the LP speakers, HP speakers showed a different pattern from (1). The NF speaker showed a remarkably long voicing lag which was very different from the L1 pattern in (1) indicating that the L1 influence on L2 speech is very weak. The Korean speakers showed different VOT patterns in terms of their proficiency in L2. The LP speaker NK2 showed quite longer voicing lag as in (1) than the HP speaker NK1, suggesting that the L1 influence on L2 speech is robust for the LP speaker but not for the HP speakers.

The relationship between L1 influence on L2 speech and proficiency in L2 English is much clearer in the results of voiced stops as in the bottom figure in Figure 3. When there is a voicing contrast in L1, L2 voiced stops were produced similar to L1 voiced stops. This held well for the LP speakers, NJ1, NJ2, and NI. Unlike the LP speakers, the HP speakers produced L2 stops similar to native-like stops with either voicing lead or short voicing lag. When there is no voicing contrast in L1, L2 voiced stops were produced similar to their counterparts in L1. This also held well for the LP speakers. For example, speakers NC and NK2 produced voiced stops with a short voicing lag only. They both could not produce voicing lead because their L1 stops were phonetically voiceless in the initial position. It would be very interesting to see whether L2 speakers, whose languages do not carry voicing in initial position, can acquire prevoicing according to proficiency in L2. The results of voiced stops were also well accounted for in terms of L1 influence on L2 speech and proficiency in L2. It was robust that L1 influences on L2 speech were clear for the LP speaker but not for the HP speaker. It was also notable that when NNE speakers' proficiency grew their VOT productions approached those of the NE speakers', indicating that the L1 influences on L2 speech were diminished.

One question arises. Can L2 adult learner acquire native-like pronunciations as their proficiency in L2 grows? Fowler *et al.* (2008) reported that even simultaneous bilingual do show cross language phonetic influences. Considering that it is hard to remove a foreign accent for adult L2 learners, it would be interesting to see whether the relationship between L1 influence and proficiency in L2 can be applied also to other late L2 learners. There can be counter examples for the relation between L1 influence on L2 speech and proficiency in L2. For example, although some speakers are highly fluent in L2 English, their pronunciations in L2 might be strongly influenced by L1. Given that only a small number of speakers were examined for each language, it is hard to generalize. Further research is necessary.

Recently, Kim (2008) reported that the VOT values between aspirated and lax stops have been neutralized in that the VOT values of aspirated stops are reduced whereas those of lax stops are increased (see also Silva 2006). Kim claimed that the two phones are in the process of undergoing some changes. The sound change, especially the VOT shortening of voiceless aspirated stops, may occur due to the L2 influence on L1 speech. Since few studies on the L2 influence on L1 speech have been done, further research is necessary to see whether the VOT change in L1 is due to L2 influence.

4. Conclusion

The present study examined NE and NNE VOT productions of English stops. The findings showed that there were remarkable VOT differences of both voiceless and voiced stops among NNE (or L2) speakers. Their VOT differences were accounted for by the influence of L1 and proficiency in L2. When proficiency in L2 was low, NNE speakers consistently produced L2 stops similar to L1 stops (see Figure 3). When proficiency in L2 was high, however, they produced L2 stops similar to native-like stops. Korean participants' VOT differences were also accounted for in terms of their proficiency levels in L2. For voiced stops, the LP speaker was not able to produce voicing lead whereas the HP speaker could. For voiceless stops, the LP speaker produced L2 stops similar to L1 stops whereas the HP speaker produced results similar to those of the NE speakers'. The results implied that, as proficiency in L2 grows, the L1 influence on L2 speech is decreased. Given that only a small number of participants for each language were used, it is hard to claim that this can be true for all L2 speakers. Further research is necessary.

References

Canagarajah, S. (2010). New definitions of English proficiency and changing pedagogical priorities. *Proceedings at the Korea Association of Teachers of English 2010 International Conference*, Seoul National University.

Crystal, D. (1997). *English as a global language*. Cambridge: CUP.

Flege, J.E., Munro, M.J., and Mackay, I.R.A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*,1-26.

Flege, J.E. (1992). *Speech learning in a second language*. Phonological Development: Models, Research, Implications. C.A.Ferguson, L. Mennand C. Stoel-Gammon, Eds. pp.565-604. Timonium MD: York Press.

Flege, J.E., and Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica 43*, 155-171.

Fowler, C.A., Sramko, V., Ostry, D. J., Rowland, S.A., and Halle, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics 36*, 649-663.

Harda, T. (2003). L2 Influence on L1 Speech in the Production of VOT. *Proceedings of 15th ICPhS Barcelona*.

Keating, P.A. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language 60*, (2), 286-319.

Keating, P.A, Cho, T., Fougeron, C., and Hsu, C. (2003). Domain-initial strengthening in four languages. *Laboratory Phonology 6*, 145-163.

Kim, M.-R. (2000). *Segmental and Tonal Interactions in English and Korean: A Phonetic and Phonological Study*. Unpublished doctoral dissertation. The University of Michigan. Ann Arbor, MI.

Kim, M.-R. (2008). "Lax" stops in Korean revisited: VOT neutralization. *Studies in Phonetics, Phonology and Morphology 14*, (2), 3-20.

Kim, M.-R., Beddor, P., and Horrocks, J. (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics, 30,* 77-100.

Kim, M.-R and Duanmu, S. (2004). "Tense" and "lax" stops in Korean. *Journal of East Asian Linguistics 13*, 59-104.

Klatt, D.H. (1975). Voice onset time, frication and aspiration in word-initial consonant clusters. *Journal of Acoustical Society of America, 18*, 686-705.

Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Chapter 3. Blackwell Publishers.

Lisker, L. and Abramson, A. S. (1964). A cross-language study of voicing initial stops: Acoustic measurements. *Word, 21*,384-422.

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

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology 23*, 287-308.

Shimizu, K. (1996). *A Cross-Language Study of Voicing Contrasts of Stop Consonants in Asian Languages.* Seibido Publishing Co. Tokyo, Japan.

APPENDIX. QUESTIONNAIRE

Part A. Biographical information

1. Gender: Male Female
2. What is your age? ____________________
3. What is your nationality? ____________________________
4. What is your native language? _______________________
   (if you grew up with more than one language, please specify)
5. What is your status here in college?
   Undergraduate Exchange student Please specify _______

Part B. English Background information

1. Rate yourself according to the categories	VP	P	So-So	G	VG
Your overall comprehension ability in English	1	2	3	4	5
Your overall speaking ability in English:	1	2	3	4	5
Your overall reading ability in English:	1	2	3	4	5
Your overall writing ability in English:	1	2	3	4	5
Your overall native-like accent ability:	1	2	3	4	5

2. How long have you lived in the United States? ______________
3. When did you come to the United States? Write years of school you attended.
   ____

(where VP=Very Poor, P=Poor, So-So, G=Good, VG=Very Good)

Mi−Ryoung Kim
Department of Practical English, Korea Cyber University
#307 Jongno Biz-well, 34, Ikseon-dong,
Jongno-gu, Seoul, Korea, 110-340
Phone: 82-2-708-7845
E-mail: kmrg@mail.kcu.ac

Received: 18 January, 2011
Revised: 06 March, 2011
Accepted: 14 March, 2011