Journal of Phonetics 40 (2012) 249–268

Contents lists available at SciVerse ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

Rapid and multifaceted effects of second-language learning on first-language speech production

Charles B. Chang

University of Maryland, College Park, Center for Advanced Study of Language, 7005 52nd Avenue, College Park, MD 20742, USA

A R T I C L E   I N F O

Article history:
Received 26 April 2011
Received in revised form
17 October 2011
Accepted 27 October 2011
Available online 31 December 2011

A B S T R A C T

Despite abundant evidence of malleability in speech production, previous studies of the effects of late second-language learning on first-language speech production have been limited to advanced learners. This study examined these effects in novice learners, adult native English speakers enrolled in elementary Korean classes. In two acoustic studies, learners' production of English was found to be influenced by even brief experience with Korean. The effect was consistently one of assimilation to phonetic properties of Korean; moreover, it occurred at segmental, subsegmental, and global levels, often simultaneously. Taken together, the results suggest that cross-language linkages are established from the onset of second-language learning at multiple levels of phonological structure, allowing for pervasive influence of second-language experience on first-language representations. The findings are discussed with respect to current notions of cross-linguistic similarity, language development, and historical sound change.

1. Introduction

The dynamics of interaction between two phonological systems have long been the subject of linguistic inquiry. Early observations of cross-language interaction between phonologies focused on the influence of an early-acquired first language (L1) on a late-acquired second language (L2) and the phenomenon of foreign accent in the L2 as a result of two related assumptions: the existence of a so-called "critical period" for language acquisition and, by implication, a unidirectional kind of cross-language influence. The classic view of the critical period (Lenneberg, 1967; Penfield & Roberts, 1959; Scovel, 1969) held that biological changes in brain development were responsible for the fact that adults are generally less successful than children at acquiring language. Children are still within a critical period of neural maturation, whereas adults, now past this period, have become neurally rigid. From this view it follows that while the L1 may cause some interference in adult acquisition of an L2, the L1 itself should not be affected.

The notion of L1 invariance during late L2 acquisition, however, is at odds with findings from work in phonetics and second language acquisition, which has produced mounting evidence that the L1 can, in fact, be affected by L2 learning. Recognition of L2 influence on L1 speech goes back as early as Selishchev (1925), but is first discussed extensively in the work of Flege (1987, 1995, 2002, 2007). These studies, along with others in the literature on L2 and bilingual speech, have repeatedly reported cases of subtle phonological restructuring in the L1 as a consequence of L2 experience, a phenomenon referred to here as PHONETIC DRIFT.[1] The occurrence of phonetic drift in mature adult L1 speakers converges with an abundance of other evidence suggesting that the L1 system does not become static and invariable, but instead remains dynamic and ever-changing. Significant, and often rapid, adjustments to L1 speech in response to environmental factors have, for example, been reported in studies of the "Lombard effect" in hearing people (Lane & Tranel, 1971; Lombard, 1911), degradation and recalibration of production in cochlear implant users (Lane, Wozniak, Matthies, Svirsky, & Perkell, 1995; Matthies, Svirsky, Perkell, & Lane, 1996; Perkell, Lane, Svirsky, & Webster, 1992; Svirsky, Lane, Perkell, & Wozniak, 1992; Vick et al., 2001), perturbations caused by delayed auditory feedback (Lee, 1950; Yates, 1963), sensorimotor adaptation to altered auditory feedback (Houde & Jordan, 1998; Jones & Munhall, 2000, 2005; Pile, Dajani, Purcell, & Munhall, 2007), phonetic convergence and divergence in interactive conversation (Bourhis & Giles, 1977; Giles, 1973; Gregory, 1990; Natale, 1975; Pardo, 2006, 2010), and spontaneous imitation of model speech

E-mail addresses: cbchang@post.harvard.edu, cbchang@umd.edu
URL: http://terpconnect.umd.edu/~cbchang/

[1] L2-influenced phonological change in the L1 has also been discussed in the literature under several other names, including "cross-language influence", "cross-language interaction", "reverse interference", "convergence", and "attrition". Here it is referred to as phonetic drift in order to distinguish it from changes in the L1 at other levels of linguistic structure and time scales and to highlight the connection to diachronic sound change (Section 5.3).

0095-4470/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2011.10.007

(Goldinger, 1998; Nielsen, 2011). These findings are relevant to the study of L2 influence on L1 speech because they show a strong tendency for L1 speech production to be adjusted in the face of a mismatch between incoming auditory information and a talker's internal targets for production, and such a mismatch is the norm in the case of L2 learning (where there are typically substantial differences between L2 auditory input and L1 production targets).

The present study examined the extent of L1 change during L2 learning with the goal of illuminating the structural basis and temporal development of cross-language phonetic influence—that is, the nature of the phonological representations that are related cross-linguistically to give rise to this change, and the progress of the change over time. Specifically, the study focused on the first weeks of native English speakers' acquisition of Korean, thereby broadening the scope of previous research on phonetic drift in two ways: investigating novice L2 learners with no prior experience in the L2, and eliminating the confound of typological/orthographic relationship present in much of the work in this area. Given that L1 phonological categories can be influenced by the phonetic characteristics of similar L2 categories (Section 2.2), this study investigated whether they would be affected early in L2 acquisition. The generally accepted null hypothesis that the L1 remains unchanged in the short term has never been demonstrated empirically. Thus, the objective of the present study was to test the alternative hypothesis that, similar to the case of phonetic convergence in L1 conversation, phonetic drift during L2 learning would occur rapidly, beginning in the very first stages of L2 learning. This hypothesis was tested in two studies focusing on categories previously examined in the literature on phonetic change: stop consonant categories differing in terms of articulatory coordination between glottal and oral gestures (i.e., temporal parameters) and vowel categories differing in terms of articulatory targets (i.e., spectral parameters).

This report on the findings is organized as follows. Section 2 reviews models of L2 speech acquisition accounting for cross-linguistic phonetic phenomena, compares the phonetics of the English and Korean categories under study, and motivates a set of specific predictions for a longitudinal investigation of the L1 English of novice L2 learners of Korean. Section 3 describes the design of the longitudinal study, as well as the methods used to analyze learners' speech production over time. Section 4 presents results on phonetic drift in English stop consonants and vowels. Section 5 discusses the findings with respect to the conceptualization of cross-linguistic similarity, phonological development, and historical change in contact situations. Finally, Section 6 summarizes the main findings and suggests avenues for further research.

2. Background

At first glance, the phenomenon of between-language convergence and divergence between L1 and L2 sounds may seem parallel to the phenomenon of within-language convergence and divergence between L1 talkers, which has been extensively described in the literature on L1 speech accommodation (for a recent review, see Pardo, 2010). However, there are two good reasons to distinguish between these situations: neither the strong sociolinguistic motivation to accommodate nor the lexical overlap supporting within-language convergence are present in the case of between-language convergence. While there are clear social reasons why L1 talkers might come to speak more or less like other members of the same speech community (Chartrand & Bargh, 1999; Giles, Coupland, & Coupland, 1991), it is difficult to imagine similar social motivations for L2 learners to modify their L1 representations with respect to an L2. L1 accommodation to an L2 is unlikely to accomplish a modification of the social distance between L2 learners and native speakers of the L2 (given that the L1 is not necessarily a shared language) or to change the social distance between L2 learners and native speakers of the shared L1 in any intended way (given that the accommodation has nothing to do with L1 input).[2] Moreover, in the case of L1 speech accommodation there are several ways in which L1 tokens may be connected to each other that are not available in the case of L1–L2 phonological interaction in most models of bilingual speech processing and production (e.g., de Bot, 1992; Paradis, 2001). When one L1 English talker hears another L1 English talker utter the word *pot* [pʰat], for example, that token can link up to the first talker's previously experienced exemplars of *pot* on multiple levels of linguistic representation (semantic, syntactic, and phonological), which may allow for easier access to any socioindexical information stored with these exemplars (see, e.g., Johnson, 1997, Chap. 8; Johnson, 2006). In contrast, when an L1 English talker learning Korean hears a Korean speaker utter the unfamiliar Korean word 팥 [pʰat˺] "red bean", that token is not likely to link up to previously experienced exemplars of English words (e.g., *pot*) on any level but the phonological.

Consequently, the current study approached the question of L1 phonological change during L2 learning from an essentially psycholinguistic, not sociolinguistic, perspective. It would be imprudent to deny that the extent of L1 accommodation to the L2 may be modulated by socially relevant aspects of an individual's personality (e.g., affinity for the L2 culture). However, the view adopted here is that the basis of this phenomenon is not interactional, but internal—namely, perceived phonological similarity between L1 and L2. The following section provides the theoretical background from which this view follows.

2.1. Modeling the acquisition of L2 speech

The study of cross-language phonetic influence has been informed by several models of L2 speech acquisition (for a recent review, see Escudero, 2007). The two that have been most widely tested are the Speech Learning Model (SLM) developed by Flege (1995, 1996) and the Perceptual Assimilation Model-L2 (PAM-L2) developed by Best and Tyler (2007). While the SLM provides a unified account of cross-linguistic interaction in L2 speech perception and production, the PAM-L2, building upon an earlier Perceptual Assimilation Model (PAM) of non-native speech perception (Best, 1994, 1995), focuses on L2 perception and the various ways in which it may be influenced by L1 phonology.

Flege's (1995, p. 233) SLM is based upon the idea that "phonetic systems reorganize in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones". Contrary to the notion of a critical period, the SLM postulates that learning mechanisms used in L1 acquisition are available throughout life, such that an L1 phonetic category encoding language-specific features of an L1 sound continues to develop in adulthood under the influence of all sounds identified with that category. L1 and L2 sounds are posited to exist in a shared system, where there is a general pressure to keep them distinct, and to be related to each other on an allophonic, rather than phonemic, basis. Given sufficient dissimilarity from the closest L1 sound, a novel sound encountered in the L2 precipitates the formation of a new phonetic category, which, if specified with the same information as a native speaker's, will allow the L2 sound to be produced accurately.

[2] If anything, L1 accommodation to an L2 might—unintentionally—increase social distance with respect to an L1 interlocutor, as L2-influenced L1 speech is likely to be perceived by other L1 talkers as accented (see, e.g., Sancier & Fowler, 1997).

However, at older ages of learning and with L2 sounds similar to L1 sounds, formation of a new category is blocked by the perceptual mechanism of equivalence classification, which results in the perceptual linkage of similar L1 and L2 sounds to the same category. This cross-language linkage causes the perceptually equated sounds to eventually approximate each other in production, such that both sounds come to be produced differently from monolingual norms. At the same time, even non-linked L1 and L2 sounds associated with unique phonetic categories may come to be produced differently from monolingual norms if they dissimilate from each other in order to maintain cross-linguistic phonetic contrast.

The insight behind equivalence classification in the SLM is that it provides an explanatory account of apparent age effects in L2 speech acquisition. Flege (1987, p. 50) observes that "if humans rely increasingly less on sensory information in making categorical decisions as they mature, and if, at the same time, they become capable of identifying an increasingly wider range of phones as belonging to a phonetic category, it may become increasingly difficult for L2 learners to note the phonetic (but not auditory) difference between 'similar' phones in L1 and L2." To put it another way, it is the consequences of linguistic experience, not neurological developments *per se*, that result in the longitudinal decline in ability to acquire L2 sounds like a native speaker. A so-called "new" L2 sound, by virtue of being unlike any previously experienced L1 sound, is not analogized to an L1 sound and causes the formation of a new phonetic category. In contrast, a "similar" L2 sound, by virtue of being close to a previously experienced L1 sound, is analogized to this L1 sound, and, given the experiential basis of equivalence classification, this occurs with increasing probability as age of L2 learning increases. Thus, by claiming that speech learning continues throughout life with much the same learning mechanisms, the SLM attributes differences between L1 and L2 speech learning in large part to one's prior linguistic experience. Age, as a proxy for experience, is therefore expected to have a significant effect on learning outcomes.

Best and Tyler's PAM-L2 differs from the SLM in developing an elaborated account of L2 perception within a gestural framework. While the SLM generally discusses L2 perception as being guided by a notion of cross-linguistic similarity based on acoustic perceptual proximity (e.g., Flege, 1996), the PAM-L2 presents a typology of specific ways in which unfamiliar L2 speech contrasts may be interpreted relative to the gestural constellations of L1 phonological categories (so-called perceptual assimilations). The type of perceptual assimilation that occurs with members of an L2 contrast predicts the degree of difficulty that L2 learners will have with discriminating that contrast. If the L2 sounds are assimilated to different L1 categories, the contrast will be discriminated accurately; if not, the contrast will be discriminated less accurately, to a degree depending upon how equally well the L2 sounds are assimilated to the same L1 category. Thus, novel L2 contrasts, rather than being uniformly difficult for L2 learners to perceive, are predicted to differ in ease of discriminability according to how they assimilate to L1 categories.

The principal difference between the PAM-L2 and the earlier PAM (which accounts for non-native speech perception by naive listeners, not L2 learners) is that the PAM-L2 incorporates the influence of an L2 learner's developing phonetic and phonological knowledge of the L2, thus allowing for perceptual assimilation at the gestural, phonetic, and phonological levels. The use of gestural, phonetic, and/or phonological information in L2 perception is related to the L2 learner's level of analysis of L2 speech, which may vary according to the stage in acquisition (e.g., practicing individual sounds vs. acquiring new words). Although any of the three levels may play a larger role than the other two at a given point in L2 acquisition, the PAM-L2 adopts a broad view of the way they may interact, stating that "L1–L2 differences at a gestural, phonetic, or phonological level may each influence the L2 learner's discrimination abilities, separately or together, depending on the context or the perceiver's goals" (Best & Tyler, 2007, p. 25). The model focuses in particular on cross-linguistic similarity at the phonetic and phonological levels. In fact, the PAM-L2 claims that the phonological level is central to L2 speech perception, and in this respect the model departs from the SLM, which adheres to similarity relations between the L1 and the L2 at the phonetic level only.

A major way in which the SLM differs from the PAM-L2—and the main reason the SLM is more relevant to the present study—is that the SLM overtly addresses the connection between L2 perception and L2 production. Because "production of a sound eventually corresponds to the properties represented in its phonetic category representation" (Flege, 1995, p. 239), an L2 sound linked to a new, unique phonetic category resembling a native speaker's is predicted to be produced accurately, while L1 and L2 sounds perceptually linked to the same category are predicted over time to approximate each other in production. In this regard, it bears repeating that although the SLM's notion of perceptual linkage of similar L1 and L2 sounds has often been used to account for how the L1 is transferred to L2 production, it provides at the same time a theoretical formulation of how the L2 can be transferred to L1 production. In fact, the prediction of phonetic drift resulting from L2 learning is made explicit in the model:

> ...cross-language phonetic interference is bidirectional in nature. The SLM predicts two different effects of L2 learning on the production of sounds in an L1, depending upon whether or not a new category has been established for an L2 sound in the same portion of phonological space as an L1 sound (Flege, 1995, p. 241).

Specifically, if a new category is not established for an L2 sound (i.e., the L2 sound is linked to an L1 category), then L2-to-L1 interference is predicted to be convergent in nature. On the other hand, if a new category is established for the L2 sound, then L2-to-L1 interference (if it occurs; see Flege, 2002) is predicted to be divergent in nature, motivated by the pressure to maintain cross-linguistic contrast between L1 and L2 sounds in a shared phonological space. These predictions are borne out in the bilingual speech literature, which reports cases of both kinds of phonetic drift.

2.2. Phonetic drift of the first language

The L1 speech of bilinguals has often been found to differ from that of monolinguals. Studies of overall accent, for example, have documented non-monolingual-like L1 production in early Korean–English bilinguals (Yeni-Komshian, Flege, & Liu, 2000), early Mandarin–English bilinguals (Jiang, 2010), and late German–English and German–Dutch bilinguals (de Leeuw, Schmid, & Mennen, 2010). Other studies have provided acoustic evidence of phonetic drift in bilinguals, such as Flege's (1987) study of late French–English and English–French bilinguals, adult L2 learners who were highly experienced in the L2 after having lived in an L2 environment for years. This study compared the L1 speech of these individuals to monolingual phonetic norms, finding deviance for both groups in both languages. Voice onset time (VOT) of French /t/ was longer than native for both groups, while VOT of English /t/ was shorter than native for both groups. In a similar manner, the second-formant frequency (F_2) of French /u/ was higher than native for both groups, while the F_2 of English /u/

was lower than native for the L1 French speakers (but not for the L1 English speakers). In contrast, French /y/ was produced in a native-like fashion by both groups. These results supported the hypothesized classification of /t/ and /u/ as "similar" L2 sounds with L1 counterparts and the classification of /y/ (by L1 English speakers) as a "new" L2 sound with no L1 counterpart.

The finding of phonetic drift in VOT has been reproduced by other researchers working with diverse bilingual populations. Major (1992, 1996) found similar evidence of VOT drift in his studies of English–Portuguese bilinguals, although he concluded that "the influence of L2 is most prevalent in casual styles of L1 and may or may not be present in formal varieties" (Major, 1992, p. 204). L2 influence on L1 VOT was also found in early Japanese–English bilinguals (Harada, 2003), late Korean–English bilinguals (Kang & Guion, 2006), and late English–Spanish bilinguals (Lord, 2008). Sometimes the VOT drift is dissimilatory, rather than assimilatory, although this effect is generally found in individuals with early L2 exposure (e.g., Mack, 1990; Yusa et al., 2010). In the latter study, the occurrence of dissimilatory VOT drift in Japanese children with little English experience was taken as evidence that "L2 affects the phonetic production of L1 even when users are not proficient" (Yusa et al., 2010, p. 583). Dissimilatory VOT drift was also found in Spanish–English and Dutch–English bilinguals, who produced their short-lag L1 voiceless stops with shortened VOT in comparison to the VOT norms of age-matched monolingual controls (Flege and Eefting, 1987a,b). What these findings of dissimilatory drift have in common is an explanation in terms of "polarization" (Keating, 1984; Laeufer, 1996): L1 sounds shift to maximize their perceptual distance from other categories in the system of contrasts. Thus, in the case of Spanish and Dutch, the VOT of L1 voiceless stops shortened to dissimilate from the long-lag VOT of L2 voiceless stops.

While the above studies generally examined cross-sectional samples at one time point, Sancier and Fowler (1997) conducted a longitudinal study of VOT in a late Portuguese–English bilingual as she traveled between the U.S. and her native Brazil. They found that she produced shorter VOTs in both Portuguese and English voiceless stops following months of immersion in Portuguese and, conversely, longer VOTs in both languages following months of immersion in English. Although the difference between the two conditions was small (on the order of 5 ms), it was statistically significant and, moreover, perceptible to native Brazilian listeners (though not to U.S. listeners). The phonetic drift of both languages in the direction of the ambient language was accounted for in terms of imitation, phonological correspondence, and recency effects. First, the effect of the ambient language on speech production was due to the tendency of humans to imitate what they hear. Second, the effect of the ambient language on production of a different language was due to a connection between phonologically corresponding categories across languages. Thus, hearing Portuguese /t/ affected the production of English /t/ because on the phonological level they are the same thing: voiceless coronal plosives. Finally, the effect of late-acquired L2 categories on early-acquired L1 categories was due to the heavy weighting of recently experienced exemplars in memory. In this way, recent L2 experience was able to influence L1 representations in spite of the greater cumulative experience with L1.

Phonetic drift occurs not only in temporal, but also spectral aspects of consonant production (Peng, 1993), in intonational properties (Mennen, 2004), and in vowel production (Baker & Trofimovich, 2005; Flege, 1987; Guion, 2003). In Guion's (2003) study, L1 vowels were found to undergo phonetic drift at the level of the vowel system. Guion tested Quichua–Spanish bilinguals on their production of both the Quichua vowels (/ɪ, a, ʊ/) and the Spanish vowels (/i, e, a, o, u/) and discovered that bilinguals who had accurately acquired the L2 vowel system produced the L1 vowels as raised relative to (near-)monolingual norms—that is, with lower first-formant frequency (F_1). A number of explanations for this result were considered. Assimilation to nearby L2 vowels was discounted because the raising of Quichua /a/ diverged from Spanish /a/, the closest L2 vowel. Global maximization of vowel dispersion was also discounted since Spanish /a/ was not produced in a manner maximizing its distance from Quichua /a/. Instead, the systemic raising of the L1 vowels was attributed to enhancement of perceptual distance between the L1 high vowels and the L2 non-high vowels, consistent with predictions of adaptive dispersion theory (Liljencrants & Lindblom, 1972; Lindblom, 1986), which posits that phonetic systems organize to maximize the perceptual distinctiveness of each of its members while minimizing production costs. This account explains why the L1 vowel system shifted in just one direction: raising of the Quichua vowels was sufficient to accommodate the non-high Spanish vowels at distinct positions in the L1–L2 vowel space, with no need for additional lowering or fronting/backing. However, it does not explain the upward direction of the shift (which was presumably due to the starting position of Quichua /a/ above Spanish /a/).

The findings of Flege (1987) and Guion (2003) are similar in that they both evince phonetic drift of vowels, but they differ in that Flege examined individual vowels, while Guion examined vowel systems. Investigation of the vowel system in its entirety allows for a level of analysis that is not possible when only pairs of close vowels are considered. Thus, whereas the forward drift of French /u/ seen in Flege (1987) was attributed to influence from L2 English /u/ specifically, the upward drift of Quichua /ɪ, a, ʊ/ seen in Guion (2003) was analyzed as a system-wide development motivated by pressures toward vowel dispersion. In this respect, the drift in Quichua vowels and the drift in French /u/ had opposite motivations: the French drift was assimilatory, bringing French /u/ closer to English /u/, while the Quichua drift was dissimilatory, shifting the Quichua vowels away from the Spanish non-high vowels. The crucial detail of the Quichua case is that dissimilation acted not at a segmental level, but at a systemic level to ensure sufficient dispersion between vowels with minimal movement.[3]

In short, an abundance of evidence from bilingual studies has shown that L1 production can be significantly affected by L2 experience. A recurrent finding of this literature is that L1 sounds tend to drift toward the closest L2 sounds; however, they may also drift away from L2 sounds in order to maximize contrast within a shared phonological system. In the current study, phonetic drift was examined in the two category types most widely discussed in the bilingual speech literature—stop consonants and vowels—with a focus on adult L1 English speakers learning Korean. The following section reviews the relevant phonetic differences between Korean and English.

2.3. Differences between Korean and English

2.3.1. Stop types

The study of L1 consonant production examined two distinguishing properties of stop types: VOT and fundamental frequency (f_0)

[3] It should be noted that the Quichua data are also amenable to an account in terms of three segment-level dissimilations. Quichua /ɪ/ and /ʊ/ could have raised to dissimilate from Spanish /e/ and /o/, respectively, while the low vowels could have shifted in a pull-and-push chain of events—Spanish /a/ raising in approximation to the higher Quichua /a/, which would then have raised to dissimilate from Spanish /a/ (cf. Mack, 1990). This segmental account differs from the systemic account in essentially positing three separate shifts that happened to go in the same direction; the systemic account, by contrast, views these shifts as the joint result of pressure toward sufficient vowel dispersion.

onset in the following vowel. These two acoustic parameters are identified with the highest degree of consensus in the literature as significant cues to the three-way laryngeal contrast among lenis, fortis, and aspirated stops in Korean. In utterance-initial stops, VOT increases going from fortis to lenis to aspirated, while f_0 onset increases going from lenis to fortis to aspirated (Cho, Jun, & Ladefoged, 2002; Kang & Guion, 2008; Kim, 2004; Silva, 2006a,b). However, with each of these parameters there is considerable overlap between categories, such that both are necessary cues for making a full three-way contrast. As shown in Chang (2010b), the stop production of the native Korean participants in the current study is consistent with the phonetic space of Korean stop types reported in Kim (2004). Consequently, in the Korean speech to which L2 learner participants had the most exposure, initial fortis stops can be assumed to have had short-lag VOT and relatively high f_0 onset; initial lenis stops, medium- to long-lag VOT and relatively low f_0 onset; and initial aspirated stops, the longest VOT and the highest f_0 onset.

In contrast to the necessary use of VOT and f_0 in making a full three-way contrast among Korean stop types, VOT alone largely suffices to make the two-way contrast between English stop types: voiceless stops are characterized by consistently longer VOT than voiced stops (Lisker & Abramson, 1964). The English stops, however, differ in terms of closeness in VOT to Korean stops, and the cross-linguistic differences are mediated to some extent by gender differences in VOT that have been reported in the literature for both languages (Oh, 2011; Whiteside & Irving, 1998). As shown in Table 1, English voiced stops are so close to Korean fortis stops that the difference between the two falls well below the within-category just-noticeable difference (JND) for VOT in quiet conditions of 23 ms (Hazan, Messaoud-Golusi, Rosen, Nouwens, & Shakespeare, 2009), both on average and at each place of articulation. Similarly, English voiceless stops are close to Korean lenis stops, with the differences between the two consistently falling under the JND for VOT. In contrast, the VOT of Korean aspirated stops is substantially longer than that of English voiceless stops at every place of articulation, and the differences exceed the JND for VOT in every case except coronals in females. Thus, on the basis of VOT alone, Korean fortis stops are not distinguishable from English voiced stops, nor Korean lenis stops from English voiceless stops. Korean aspirated stops, however, are generally distinguishable from English voiceless stops, the exception being females' coronal stops.

While English voiced and voiceless stops are distinguishable in terms of VOT, these categories also differ with respect to f_0 onset:

Table 1
Native VOT norms (in ms) with standard deviations (in parentheses) for plosives in Korean and English. Korean figures are for monosyllabic items with /a/ read in isolation (see Table 3) and are averaged over the seven female and two male native speakers of Korean from the current study (see Section 3.1). English figures are for monosyllabic items with /ɑ/ read in isolation and are averaged over the 40 female and 40 male native speakers of General/Southern American English in Morris et al. (2008, p. 312).

Gender	Place	Short-lag		Long-lag		
		Korean fortis	English voiced	Korean lenis	Korean aspirated	English voiceless
Female	Labial	7 (5)	12 (5)	57 (18)	82 (30)	54 (15)
	Coronal	9 (3)	16 (6)	61 (14)	81 (21)	69 (16)
	Dorsal	18 (5)	22 (8)	74 (21)	108 (20)	63 (14)
	AVERAGE	11 (4)	17 (6)	64 (18)	90 (24)	62 (15)
Male	Labial	11 (4)	13 (6)	54 (35)	90 (32)	53 (27)
	Coronal	14 (7)	16 (5)	41 (20)	85 (14)	58 (15)
	Dorsal	26 (8)	24 (10)	71 (29)	115 (26)	62 (15)
	AVERAGE	17 (6)	18 (7)	55 (28)	97 (24)	58 (19)

on average, f_0 starts off lower following voiced stops than voiceless stops (Haggard, Ambler, & Callow, 1970; Hombert, 1978). The f_0 difference between English stop types has consistently been shown to be relatively subtle in comparison to the f_0 differences between Korean stop types, however. In comparison to an f_0 difference between English voiced and voiceless stops that may approach 15–20 Hz (Hombert, 1978; House & Fairbanks, 1953; Lehiste & Peterson, 1961), the f_0 onset difference between Korean lenis and aspirated stops, for example, averages 57 Hz over the native Korean participants in the current study (Chang, 2010a), a difference that is so pronounced it has sometimes been analyzed as tonal contrast (e.g., Silva, 2006a). The differences in vowel midpoint f_0 between Korean lenis and aspirated stops are also larger than the f_0 differences between English stop types: 44 Hz (3.3 semitones) for females and 27 Hz (3.5 semitones) for males (Oh, 2011). Thus, although direct cross-linguistic comparisons of phonetic norms for f_0 are not possible on the basis of what has been reported in the literature (due to disparities in gender balance, measurement methods, and vowel contexts), comparisons of f_0 differences suggest that the elevated f_0 onset typical of the two laryngeally marked Korean stop types (fortis and aspirated) is substantially higher than the f_0 onset typical of both English stop types. Furthermore, the cross-linguistic differences in f_0 are likely to be noticeable in light of JNDs that have been reported for frequencies in the range of f_0, which do not exceed 4–5 Hz (Harrison, 1996; Roederer, 1973).

In perceptual judgments of cross-linguistic similarity between the Korean and English stop types, VOT and f_0 both seem to play a role. If VOT were the only consideration, the cross-language linkages established by L2 learners would be predicted to be between English voiced stops and Korean fortis stops and between English voiceless stops and Korean lenis stops, since these are the pairs that are most similar to each other in VOT. However, cross-linguistic perceptual data, while consistent with the former linkage, are inconsistent with the latter one. Schmidt (2007, Chap. 11) found that L1 English speakers with no prior knowledge of Korean overwhelmingly labeled Korean lenis and aspirated stops as English voiceless stops and Korean fortis stops as English voiced stops, but with different degrees of cross-linguistic similarity: Korean aspirated stops were rated as more similar to English stops than Korean lenis or fortis stops were. Thus, despite the closer proximity of Korean lenis stops to English voiceless stops in VOT, Korean aspirated stops were perceived as more similar, presumably due to closer proximity in f_0. These results suggest that English-speaking learners of Korean are most likely to perceptually link English voiced and voiceless stops to Korean fortis and aspirated stops, respectively.

2.3.2. Vowels

The study of L1 vowel production examined two distinguishing properties of vowel quality, F_1 and F_2, in Korean and English monophthongs: Korean /i, ɛ, u, ɨ, o, ʌ, a/[4] (Ingram & Park, 1997; Ko, 2009; Lee, 1993; Silva, 2010) and American English /i, ɪ, e, ɛ, æ, u, ʊ, o, ɑ, ɔ, ʌ/. The American English vowels have been observed to be the locus of much dialectal variation, which is relevant insofar as it results in differences in cross-linguistic proximity across dialects. Phonetic norms for American English vowels have not been published for all dialect regions,[5] but four studies have

[4] The high back unrounded vowel, often transcribed as /ɯ/ in phonological descriptions of Korean, is transcribed here as /ɨ/ to better represent its central quality in contemporary Seoul Korean.

[5] Labov, Ash, and Boberg (2006, pp. 77–116) provide a thorough overview of English vowel variation in the dialects of North America, but in mostly comparative terms (with ranges for F_1 and F_2, but no measures of central tendency such as means).

provided data on talkers from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal): Peterson and Barney (1952, p. 183), Hillenbrand, Getty, Clark, and Wheeler (1995, p. 3103), Yang (1996, p. 250), and Hagiwara (1997, p. 656). Spectral norms reported in these studies differ in two main ways. First, NMidW talkers produce /æ/ with a relatively low F_1 and high F_2 and /a/ with a relatively high F_2, features of the Northern Cities Shift characteristic of the NMidW area (Labov, 1994). Second, SoCal talkers produce /u, ʊ/, along with /ʌ/, with a relatively high F_2 consistent with the California Vowel Shift (Hinton et al., 1987). S&SW talkers produce /u, ʊ/ with a relatively high F_2 as well. Consequently, in contrast to the familiar trapezoidal vowel space of MidA talkers, the English vowel space has a triangular configuration for NMidW talkers, while it is shaped like a parallelogram for S&SW talkers and SoCal talkers, as shown in Fig. 1.[6]

Despite dialectal variation in acoustic proximity between English and Korean vowels, several cross-language vowel pairs emerge as consistently close across dialects. If cross-language perceptual similarity is based on acoustics (in this case, F_1 and F_2),[7] the potential influence of Korean vowels on English /i, ɪ, e, ɛ, ʊ, ʌ/ is similar across dialects, as in each of these cases the closest Korean vowel to the English vowel—and, thus, the most likely L2 attractor—is, with few exceptions, the same across dialects and in a similar position relative to the English vowel in $F_1 \times F_2$ space. However, Korean /a/, while consistently the closest Korean vowel to English /ɑ/, is positioned differently across dialects (Fig. 1). There is also a salient disparity in the proximity of Korean vowels to English /o, u/, due to the discrepancy in F_2 values between fronting dialects and non-fronting dialects. Therefore, if Korean-to-English influence were to occur vowel to vowel based on acoustic proximity in F_1 x F_2 space, it would sometimes vary depending on English dialect. English /ɑ/, for instance, might be expected to drift backwards for NMidW talkers, but forwards for MidA, S&SW, and SoCal talkers.

In addition to cross-linguistic differences between individual vowels, there are also differences between the aggregate vowel systems of Korean and English. Due to the greater number of basic vowels in English, the English vowel space is, regardless of dialect, characterized by closer average inter-vowel spacing than the Korean vowel space (917, 811, 753, and 733 Hz, respectively, for female MidA, NMidW, S&SW, and SoCal English vs. 1035 Hz for female Standard Korean; 693, 651, 575, and 488 Hz, respectively, for male MidA, NMidW, S&SW, and SoCal English vs. 767 Hz for male Standard Korean). In addition, vowels have a different distribution within the English vowel space compared to the Korean vowel space: the English vowel space is more crowded in both the front and lower (i.e., non-high) regions. In the front region, the English vowel space contains five front vowels, compared to two in Korean, while in the lower region, it includes /e/, /æ/, and, for many dialects, /ɔ/—vowels absent from the Korean inventory. Moreover, several corresponding vowels are located at different points in the vowel space, with the Korean vowel being realized as higher (and, often, as more back). Korean /o/, /ʌ/, and /ɨ/ are each higher than English /o/, /ʌ/, and /ʊ/, respectively; /o/ and /ʌ/, moreover, are further back in Korean than in English. These comparisons suggest that the Korean vowel system is higher and more back than the English vowel system overall, and this conclusion is consistent with figures for mean F_1 and F_2 of the two languages' basic vowel inventories, which are consistently lower for Korean than English (Table 2).[8]

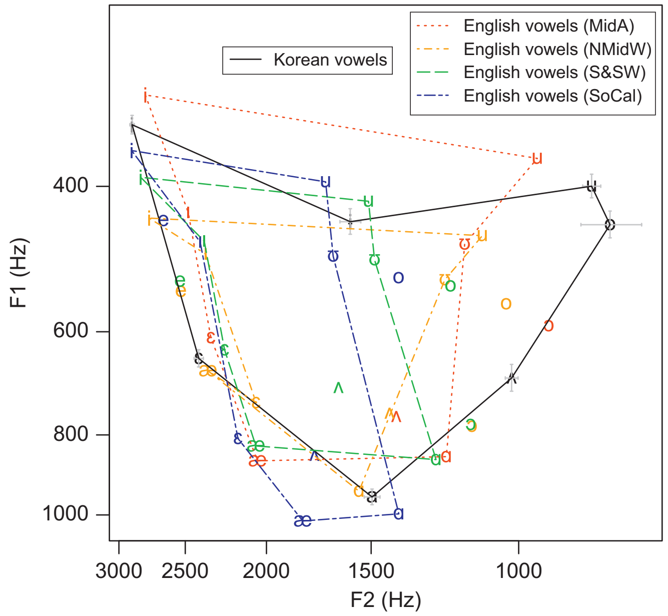

Fig. 1. Logarithmic plot of Korean and English vowels as produced by female talkers. Korean values are from native speakers of Korean in the present study (see Section 3.1). English values are from native speakers of American English from the Mid-Atlantic (MidA), northern Midwest (NMidW), South and Southwest (S&SW), and southern California (SoCal) in the studies of Peterson and Barney (1952), Hillenbrand et al. (1995), Yang (1996), and Hagiwara (1997). Error bars indicate ±1 mean standard error.

Table 2
Mean F_1 and F_2 of the Korean and English vowel inventories. Figures (in Hz) are means of norms for Korean monophthongs as produced by native speakers of Korean in the current study (see Section 3.1) and for English monophthongs as produced by native speakers of American English from the Mid-Atlantic (MidA), the Northern Midwest (NMid), the South and Southwest (S&SW), and Southern California (SoCal) in the studies of Peterson and Barney (1952), Hillenbrand et al. (1995), Yang (1996), and Hagiwara (1997).

Gender	Formant	Korean	American English			
		Standard	MidA	NMid	S&SW	SoCal
Female	F_1	558	583	624	600	634
	F_2	1572	1700	1776	1843	1982
Male	F_1	435	503	527	510	481
	F_2	1314	1428	1523	1580	1608

2.4. Research questions and predictions

This study had one main goal: to arrive at a better understanding of *how* and *why* phonetic drift occurs. The research reviewed in Section 2.2 suggested that while a decline in L1 use

[6] Note that the vowel spaces plotted in Fig. 1 are based on values from different studies that are not normalized, but this is probably not problematic as there are no systematic between-language differences for the extreme vowels /i/ and /ɑ/ ~ /a/.

[7] Some findings suggest that cross-language perceptual similarity does not follow directly from acoustic proximity (e.g., Strange, Levy, & Lehnholf, 2004), although it must be kept in mind that comparisons of acoustic similarity and perceptual similarity are dependent on a choice of acoustic parameters and, therefore, can provide evidence only against a relationship between a particular set of acoustic similarity metrics and perceptual similarity, not against the general premise of a relationship between acoustic and perceptual similarity.

[8] Norms for English /e, o/ are not reported by Peterson and Barney (1952); however, the absence of these non-low vowels in the calculation of a grand mean for F_1 in MidA English is likely to slightly lower, rather than raise, the average, thus strengthening the case for lower mean formants in Korean. Norms for the English vowel /ɔ/ are not reported by Hagiwara (1997) since /ɔ/ is generally merged with /ɑ/ in SoCal English.

may contribute to phonetic drift, this is not the main cause. Rather, L2 experience appears to be the primary factor driving changes in L1 production. The amount of L2 experience required for phonetic drift to occur remains unclear, however, since previous studies have documented this phenomenon almost entirely in individuals who were highly experienced in L2—either L2-dominant speakers in attrition situations (e.g., Campbell & Muntzel, 1989; Chang, 2009; Chang, Yao, Haynes, & Rhodes, 2011; Oh, Jun, Knightly, & Au, 2003) or fluent bilinguals. This bias in the literature reflects the view that a large amount of L2 experience (as indicated by advanced L2 proficiency) is needed to influence L1 in a substantial way:

> ...a L2 that is hardly mastered should not have much influence on L1, while a L2 which is mastered to a high degree should exert more influence (Major, 1992, p. 201).

Although Yusa et al. (2010) showed that children with relatively little L2 experience may also show signs of phonetic drift, these findings are ambiguous since children also have relatively little L1 experience; hence, phonetic drift here could be attributed to underdeveloped L1 representations that are still maturing. Phonetic changes in the L1 of child bilinguals can, moreover, often be explained in terms of normal L1 development, rather than cross-linguistic influence from the L2 (Khattab, 2000). Therefore, the present study focused on adults—individuals with fully developed L1 representations—from the onset of L2 exposure, examining the conditions least likely to produce evidence of phonetic drift: acquisition of an L2 that is both typologically and orthographically unrelated to the L1, and production of formal (i.e., citation-style) speech.

The first question posed in this study was whether adult L2 learners would show phonetic drift early in L2 acquisition, before gaining considerable proficiency in L2. Although one might not expect so given the empirical focus of previous studies on advanced learners, the opposite prediction followed from principles of the SLM. Recall that cross-language influence in the SLM is based on equivalence classification, a basic cognitive process that aids in perceptual categorization. As a mechanism used in normalization and categorization, equivalence classification can be assumed to apply automatically. If, consequently, novel L2 sounds are perceptually linked to close L1 sounds automatically, it follows that the potential for cross-language influence will be immediate. Thus, the first hypothesis of this study was that adult L2 learners would manifest phonetic drift during the first weeks of L2 acquisition, since the cross-language linkages supporting phonetic drift would be established at the onset of L2 experience, allowing accruing L2 phonetic input to promptly begin influencing L1 representations.

The second question was whether phonetic drift would be assimilatory or dissimilatory. Both patterns of phonetic drift were found in the literature; however, the principles of the SLM implied that phonetic drift in this case would be assimilatory. Recall that according to the SLM, dissimilation between L1 and L2 sounds only occurs when the L2 sound was perceived as different enough from the closest L1 sound to have formed its own phonetic category. The SLM also states that new phonetic categories for L2 sounds are established less often as age of L2 learning increases, suggesting that dissimilation is less likely to happen in adult learners. Therefore, the second hypothesis of this study was that phonetic drift in adult L2 learners would generally be assimilatory to properties of the L2.

The third question was whether phonetic drift would occur exclusively at a segmental level or also at a more general level (e.g., a natural class of sounds). The SLM and the PAM-L2 are agnostic on this point, discussing L1-L2 relations only between segment-sized units. The corresponding bias in the literature toward examining cross-linguistic influence between segments might thus lead one to expect phonetic drift to be largely limited to segment-to-segment cross-linguistic connections. However, previous findings—in particular, those of Nielsen (2011) and Guion (2003)—suggested otherwise. The results of Nielsen (2011) for VOT imitation showed talkers being influenced by exposure to model speech at a level generalizing to all stops with the same laryngeal specification, suggesting that, in the domain of VOT, L1 representations change at a subphonemic level. The results of Guion (2003) for vocalic drift showed all vowels in the L1 inventory moving in concert, suggesting that L1 representations for vowels change at a systemic level. Thus, the third hypothesis in this study was that phonetic drift would occur via cross-language linkages at multiple levels of phonological structure—at a segmental level (the level of the individual speech sound), at a subsegmental level (the level of the natural class), and at a global level (the level of the system)—resulting in phonetic drift of varying degrees of generality.

Following from these three hypotheses, the assumption that cross-language differences must be perceptible to cause phonetic drift, and the phonetic norms and JNDs described in Section 2.3, it was predicted that novice L2 learners of Korean would manifest early phonetic drift in English VOT, f_0, and vowel formants. With regard to VOT, the VOT of English voiced stops was predicted not to change significantly, since the disparity between English voiced stops and Korean fortis stops is likely too small to be perceived. In contrast, the VOT of English voiceless stops was predicted to lengthen due to the longer VOT norms of Korean aspirated stops, and this lengthening was expected to occur at a subsegmental level generalizing to the coronal stops (the only place where the English-Korean disparity was probably too small to be perceived for most learners).

With regard to f_0, the f_0 onset following English voiced and voiceless stops was predicted to rise due to the elevated f_0 onset following Korean fortis and aspirated stops. This f_0 increase could occur in one of two ways: at the level of the natural class (i.e., via linkages between English voiced stops and Korean fortis stops, and between English voiceless stops and Korean aspirated stops), or at a global level. In the former case, vowel onsets following English voiced and voiceless stops would both drift upwards in f_0, but the f_0 increase would be limited to just these environments, while in the latter case, the f_0 increase would extend to all English words including non-stop-initial ones. Details about the f_0 properties of English and Korean specifically did not adjudicate between these two possibilities; however, the speech adaptation literature mentioned in Section 1 provided evidence that f_0 is modulated at least in part by a control mechanism separate from the internal model of segmental control, suggesting that changes in f_0 production, rather than being tied to properties of specific stop categories, would occur more generally. Thus, it was predicted that f_0 in English would drift upwards at a global level, affecting both stop-initial and non-stop-initial (in this case, vowel-initial) words.

Finally, with regard to vowel formants, phonetic drift of English vowels was predicted to occur at the level of the vowel system, resulting in a global shift in formants affecting all the vowels in a similar manner. Specifically, it was predicted that the English vowels would generally shift upward and backward (i.e., F_1 and F_2 values would decrease) due to the lower overall F_1 and F_2 levels of Korean. Crucially, this shift was expected to take place via global linkages to overall F_1 and F_2 levels in Korean, such that individual vowel shifts would not be explicable in terms of movement toward the closest individual Korean vowels. The overall pattern of vocalic drift, moreover, was not expected to differ across English dialects, since it would have arisen due to cross-language differences in overall F_1 and F_2 levels that are a function of basic differences between the English and Korean vowel inventories.[9]

3. Methods

3.1. Participants

A sample of 36 adult native speakers of American English learning Korean participated in the study to completion. From this sample a group of ‘‘functionally monolingual'' (in the sense of Best & Tyler, 2007, p. 16) learners were selected for analysis on the basis of responses in a background questionnaire. These learners came from various dialect regions of the U.S., including Eastern New England, the Mid-Atlantic, the Inland North, the North Midland, the South, the Southwest, and the Pacific Northwest (Labov et al., 2006). Although all participants had formally studied foreign languages (most often Spanish or French, for a period ranging from 4 months to 15 years), the 19 learners examined here (16 females, 3 males; mean age of 22.1 years, range of 21–26) were consistent in reporting English to be their native language, their best language, and the language used at home and no regular communicative use of another language besides English. Furthermore, they reported no significant prior exposure to Korean, prior study of Korean, or prior travel to Korea lasting longer than one week. None, moreover, reported any history of hearing or speech impairments in childhood. All were paid for their participation.

At the time of data collection, learners were living on the campus of a South Korean university and starting a six-week course of intensive Korean language instruction (roughly equivalent in content to one semester of college-level Korean). Every weekday during this period, learners had class with two female instructors for 4 h on average, which, according to exit questionnaires, constituted the majority of the time they heard and spoke Korean during the time period of the study. Classes were conducted in Korean, but outside of class learners functioned mostly in English, thus making the type of language learning situation in which they found themselves a cross between typical second language acquisition (SLA), in which learners are immersed in an L2 environment and acquire the L2 largely ‘‘in the wild'', and typical foreign language acquisition (FLA), in which learners study the L2 formally in an L1 environment. In the current study, learners were living in Korea, but receiving most of their L2 exposure via structured formal instruction in a classroom setting where all their fellow learners shared the same L1 background.

In addition to the group of L2 learners, a control group of nine native Korean speakers participated. All were paid for their participation. These nine Korean speakers (seven females, two males; mean age of 27.8 years, range of 22–34) were the L2 learners' Korean instructors and resident assistants and, thus, provided most of the learners' L2 input. Having been educated in South Korea, where formal English instruction is compulsory from as early as primary school, these Korean speakers had all received some degree of schooling in English, but they were functionally monolingual in Korean at the time of this study and can be considered representative of young Korean speakers in contemporary South Korea. All the teachers spoke Standard Korean, having been trained in Seoul; five of the seven, moreover, had grown up in Seoul or the surrounding Gyeonggi province. The two resident assistants had grown up in the adjacent Gangwon province and also spoke a relatively standard variety.

3.2. Procedure

Learners participated in two production experiments during the study: Experiment 1K elicited production of Korean, while Experiment 2E elicited production of English. Both experiments were run by the same experimenter (the author), who always provided instructions in English. In each experiment, production of monosyllables in isolation was elicited via a reading task in which participants were shown the native spelling of the items they were to pronounce. The experiments were longitudinal in nature and run a total of five times, each time in the space of 48 h between one week of classes and the following week of classes so as to keep the amount of Korean instruction received equal across participants.

Experiments 1K and 2E were almost always completed in one session, with a break between experiments. Since Experiment 1K was preceded by another experiment involving a task in Korean, Experiment 1K was completed before Experiment 2E in order to require only one switch between languages (see, e.g., Grosjean, 2001). This order was kept constant across time points and participants in order to control for effects of task order and allow for both cross-sectional and longitudinal comparisons.[10] The experiments were generally run in a quiet room in the learners' dormitory. In each experiment, stimuli were presented and responses recorded in DMDX (Forster, 2008) on a Sony Vaio PCG-TR5L laptop computer. The stimuli were randomized and presented in four blocks, such that four tokens were collected of each item. Audio was recorded at 44.1 kHz and 16 bps using an AKG C420 or C520 head-mounted condenser microphone, which was connected either to the computer via an M-AUDIO USB preamp or to a Marantz PMD660 solid-state recording device.

3.3. Materials

The speech materials for Experiments 1K and 2E consisted of, respectively, 22 Korean and 23 English monosyllabic items representing most of the phonemic contrasts in the two languages. Members of a subgroup of items were maximally similar to each other in segmental makeup (e.g., Korean 후 /hu/ vs. English *who'd* /hud/) for the purposes of cross-language acoustic comparisons. English items were generally of the form CVC to allow for lax vowels, while Korean items were generally of the form CV to make them easier for novice learners to read. The materials were the same in every week of the study and are presented in Table 3.[11]

[9] An anonymous reviewer wondered what the implications of this prediction were for non-systemic types of similarity. In short, the prediction of phonetic drift at a systemic level does not discount the grammatical relevance of cross-linguistic similarity at other levels. Rather, it suggests that effects of similarity at a systemic level may override effects of similarity at, e.g., a segmental (in this case, vowel-to-vowel) level.

[10] Although both experiments were monolingual and thus involved no code-switching, an anonymous reviewer was concerned that completing a Korean task and then an English task shortly afterward might resemble a code switch and elicit temporary cross-linguistic convergence, as is well-documented in the study of code-switching (e.g., Bullock & Toribio, 2009, Chap. 8). First, it should be noted that by virtue of maintaining an identical task order any possible effect of a ‘‘code switch'' between experiments was, by design, parceled out of the longitudinal results. Second, there appeared to be no effect of a language switch on production in Experiment 2E anyway. The findings reported in Section 4 were constant over the course of the experiment, with participants' English not significantly differing between the beginning of Experiment 2E (minutes after they had completed Experiment 1K) and the end of Experiment 2E (approximately 8 min later), suggesting that at no time point did patterns of English production result from short-term effects of a language switch between experiments.

[11] IPA transcriptions use the extended IPA symbols for weaker and stronger articulations (International Phonetic Association, 1999, p. 189) to transcribe lenis obstruents (e.g., /t͉/) and fortis obstruents (e.g., /t͈/).

Table 3
Korean and English items used in the production experiments.

Korean		English		Korean		English	
바	/p͉a/	*'bot*	/bɑt/	히	/hi/	*heed*	/hid/
빠	/p͈a/					*hid*	/hɪd/
파	/pʰa/	*pot*	/pɑt/	해	/hɛ/	*hate*	/het/
다	/t̪͉a/	*dot*	/dɑt/			*head*	/hɛd/
따	/t̪͈a/					*had*	/hæd/
타	/t̪ʰa/	*tot*	/tɑt/	호	/ho/	*hoed*	/hod/
가	/k͉a/	*got*	/gɑt/	후	/hu/	*who'd*	/hud/
까	/k͈a/			흐	/hɨ/	*hood*	/hʊd/
카	/kʰa/	*cot*	/kɑt/	허	/hʌ/	*hut*	/hʌt/
						hawk	/hɔk/
사	/s̪͉ʰa/	*sod*	/sɑd/	외	/wɛ/	*wait*	/wet/
싸	/s̪͈a/	*shot*	/ʃɑt/			*wet*	/wɛt/
시	/s̪͉ʰi/	*seed*	/sid/	위	/wi/	*wee*	/wi/
씨	/s̪͈i/	*sheet*	/ʃit/	알	/aɭ/	*all*	/ɔl/

Each subset of items listed in Table 3 was meant to test one or more of the hypotheses formulated in Section 2.4. The items beginning with stop consonants were meant to test for phonetic drift in VOT (specifically, the hypothesis that drift in VOT would be subsegmental and occur at the level of the natural class). Thus, stop-initial items were included to elicit productions of stop consonants that could be measured for VOT, the primary distinguishing characteristic of different stop voicing (laryngeal) categories, as well as for f_0 onset in the following vowel. The items beginning with /h/ were meant to test for phonetic drift in vowels (specifically, the hypothesis that phonetic drift in vowel production would be systemic and occur over the whole vowel system). Thus, items beginning with onsets having no oral place of articulation were included so as to elicit productions of vowels that could be measured for formant frequencies, the primary distinguishing characteristic of different vowel qualities, with minimal coarticulatory influence from onset consonants. The remaining items comprised control and filler items, such that in each study approximately one third of the speech materials were critical items subjected to analysis.

3.4. Acoustic analysis

Acoustic data from recordings comprised measurements of VOT in word-initial plosives, f_0 at the onset of the following vowels, and F_1 and F_2 at vowel midpoint. All acoustic measurements were taken manually in Praat (Boersma & Weenink, 2008) on a wide-band Fourier spectrogram with a Gaussian window shape (window length of 5 ms, dynamic range of 50 dB, pre-emphasis of 6.0 dB/oct) or the corresponding waveform.

The plosive study examined VOT and f_0 onset. Critical items subjected to analysis were the 15 items beginning with plosives, which were always followed by a low vowel. VOT was defined as the time at voicing onset minus the time at plosive release (the beginning of the burst interval). The point of voicing onset was taken to be the first point at which a voicing bar with clear glottal striations appeared in the spectrogram. The f_0 at voicing onset was measured manually by taking the combined wavelength of the first three regular glottal periods in the vowel and converting into a frequency value. The three-period interval was marked off on the waveform, with an initial period generally being skipped if it was more than 33% longer or shorter than the following period. However, if the earliest interval of three regular periods occurred more than five periods into the vowel (i.e., the vowel began with an extended interval of creaky or irregular phonation), the token was discarded.[12]

The vowel study examined F_1 and F_2, measured automatically from spectrograms annotated manually for vowel onset and offset. Critical items subjected to analysis comprised the 16 items with a glottal fricative onset and the two items with an aspirated bilabial stop onset (i.e., Korean /pʰa/, English [pʰɑt]). Vowel onset was marked at the first glottal striation, while vowel offset was marked either at the final glottal striation (in words with a following coda consonant) or at the point where a clear F_1 and F_2 were no longer visible (in words with no following consonant). Mean values of F_1 and F_2 were measured over the middle 50 ms of the vowel interval demarcated in this way.[13] The analysis method was linear predictive coding, using the Burg algorithm (Childers, 1978) in Praat. The frequency range and number of formants entered into the formant analysis were obtained by visually inspecting a few spectrograms from the given participant and adjusting the default parameters until tracking of F_1 and F_2 was smooth and accurately followed the formants visible in the spectrogram.

Once the formant measurements were extracted, they were inspected for outliers by vowel and formant, and potential errors were flagged. Spectrograms of all tokens were then individually inspected to check the accuracy of the formant tracking. When the formant tracking was irregular or inaccurate, the analysis parameters were adjusted until tracking was smooth, and new measurements were extracted. If the formant tracking could not be made satisfactory via adjustment of the analysis parameters, then measurements were taken manually on an average spectrum of the middle 50 ms of the vowel.[14] Tokens were discarded if they contained a pronunciation anomaly or a speech error.[15]

Tests of intra-rater reliability indicated that the measurements collected were highly reliable. Six months after the original measurements were taken, approximately 20% of the analyzed tokens were randomly selected and reanalyzed. Pearson's correlation coefficients showed the two rounds of measurements to be closely correlated for all measures: VOT [$r = 0.91$, $p < 0.0001$], f_0 [$r = 0.98$, $p < 0.0001$], F_1 [$r = 0.97$, $p < 0.0001$], and F_2 [$r = 0.94$, $p < 0.0001$]. The average difference between paired VOT measurements was 3 ms; between paired f_0 measurements, 4 Hz; between paired F1 measurements, 7 Hz; and between paired F2 measurements, 15 Hz.

3.5. Statistical analysis

In order to achieve a fair comparison of the English and Korean laryngeal categories in the plosive study, tokens were divided into bins representing three universal phonetic categories of stop voicing (Keating, 1984) in accordance with VOT boundaries estimated from the literature (Lisker & Abramson, 1964; Lisker et al., 1977; Keating, 1984): voicing that begins prior to release (prevoiced stops; VOT < 0 ms), voicing that begins with or shortly

[12] The percentage of tokens discarded by this criterion or for other pronunciation anomalies (e.g., mid-utterance coughing) was small: 1.9–5.0%, depending on the language.

[13] Midpoint measurement is not without its shortcomings, given that so-called monophthongs in American English are diphthongized to different degrees (rendering the vowel midpoint more representative of the overall quality of the vowel for some vowels than for others). However, since vowel duration did not differ substantially across time points in this study, midpoint measurements still allowed it to be apparent when there was a change in overall vowel quality over time.

[14] The need to resort to manual measurement was rare, occurring in 0.3% of L2 learners' English tokens and 0.2% of native Korean speakers' Korean tokens.

[15] These tokens were few in number, amounting to 1.8% of L2 learners' English tokens and 0.4% of native Korean speakers' Korean tokens.

after release (short-lag stops; VOT of 0–30 ms), and voicing that begins long after release (long-lag stops; VOT >30 ms). All three phonetic voicing types occur in American English generally, and this tripartite nature of voicing production was reflected in the data of the current study, which showed a trimodal distribution of VOT in learners' productions of both English stops and Korean stops. However, the distribution of tokens across the three phonetic voicing types was not always equal for similar English and Korean laryngeal categories (e.g., prevoiced stop tokens were produced unevenly between the two languages). Consequently, cross-linguistic comparisons matched tokens of laryngeal categories by phonetic voicing type (the canonical type for the category, which was the same for paired categories). Therefore, English voiced stops and Korean fortis stops were compared with respect to short-lag productions, while English voiceless stops and Korean aspirated stops were compared with respect to long-lag productions.

To account for inter-speaker differences in speech production, the acoustic data in both studies were analyzed with mixed-effects linear regression (see, e.g., Johnson, 2008, pp. 237–247) using the lme() function in R (Development Core Team, 2010). All regression models had one random effect—namely, Participant—and fixed effects that included Gender (of the participant) and Time (point in the language program). Models of VOT and f_0 also included Place (of articulation) as a fixed effect, while models of F_1 and F_2 also included Vowel; the two- and three-way interactions among predictors were considered as well. Place and Vowel were included as fixed effects in the models because they have been shown to have a significant influence on the given dependent variables (e.g., Ladefoged, 2005; Nearey & Rochet, 1994). In a departure from previous studies of bilingual speech production, which have generally not accounted for gender differences explicitly, Gender was also included as a fixed effect because it, too, has been shown to have a significant influence on the dependent variables (e.g., Whiteside, 1998; Whiteside & Irving, 1998; Simpson, 2009). Each regression model was built incrementally in order to avoid overfitting it to the data, and its robustness was measured via cross-validation (Johnson, 2008, pp. 238–240): the same type of model was built 1000 times based upon a random subset of 85% of the data, and its predictions were then tested against the actual values for the dependent variable in the remaining 15% of the data. Given the focus of the study on the effect of time, models were built in three stages: first with only one fixed effect, then with fixed effects excluding Time, and finally with fixed effects including Time.

4. Results

4.1. Phonetic drift in stop production

Over time, learners' English voiced stops changed little in VOT, but increased significantly in f_0 onset. Regression models fit to the VOT data with one fixed effect indicated that of Gender and Place, Place was by far the better predictor of VOT. The 95% confidence interval for the percentage of variance accounted for (r^2) in a model with only Place as a fixed effect was 0.306–0.561, compared to 0.031–0.225 for a model with only Gender as a fixed effect. Adding Gender did not increase the mean r^2 of a Place-only model [$t(1994) = -1.01$, *n.s.*], though adding the Gender × Place interaction did [$t(1998) = 2.14$, $p < 0.05$], as males showed a greater increase in the VOT of velars over the other places of articulation than did females.[16] Adding Time, however, did not further improve

[16] Note, however, that the gender effects reported here as well as in Section 4.2 should be regarded with caution due to the small number of male participants.

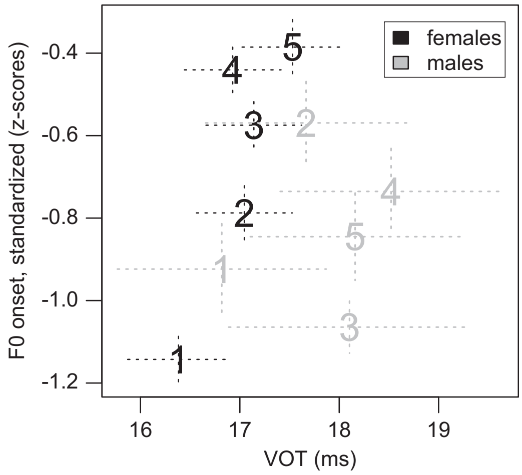

Fig. 2. Mean VOT vs. mean f_0 onset (standardized) over time for English voiced plosives. Numerical symbols plot means of the respective weeks for female learners in black and for male learners in gray. Error bars indicate ± 1 standard error about the mean.

the model [$t(1993) = 0.88$, *n.s.*], nor did adding any other predictors. Regression models fit to the f_0 data with one fixed effect indicated that Gender was a more informative predictor of f_0 onset than Place, and adding Place did not improve upon a Gender-only model of f_0 onset [$t(1993) = -0.35$, *n.s.*]. Including Time, however, significantly improved the model, increasing mean r^2 from 0.840 to 0.857 [$t(1957) = 13.23$, $p < 0.0001$]. Including the Gender × Time interaction improved the model further, increasing mean r^2 to 0.860 [$t(1995) = 2.93$, $p < 0.01$], as males did not show the increase in f_0 onset that females did. No other predictors emerged as significant. Post-hoc comparisons of adjacent time points using Tukey's HSD test showed that significant changes in f_0 onset occurred between Weeks 1 and 2 [$p < 0.001$] and Weeks 3 and 4 [$p < 0.05$]. In short, English voiced stops were only found to increase in f_0 onset—an effect that was consistent across the three places of articulation and continued beyond the first week of the language program, but was isolated to female learners.[17] This gender difference can be seen in Fig. 2, which compares means for males and females in terms of standard f_0 (standardized by learner with respect to the learner's variance in f_0 over the entire study).

In contrast to the English voiced stops, learners' English voiceless stops increased significantly in both VOT and f_0 onset. Regression models fit to the VOT data with one fixed effect indicated that Place was a better predictor of VOT than Gender, and adding Gender again failed to improve significantly upon a Place-only model of VOT [$t(1998) = -0.52$, *n.s.*]. Adding Time, however, resulted in a significant increase of mean r^2 from 0.333 to 0.383 [$t(1991) = 17.25$, $p < 0.0001$]. No other predictors improved the model further. Regression models fit to the f_0 data with one fixed effect indicated that, as with f_0 of the voiced stops, Gender was a more informative predictor of f_0 onset than Place, and adding Place did not improve upon a Gender-only model of f_0 onset [$t(1995) = -1.23$, *n.s.*]. On the other hand, adding Time significantly improved the model, increasing mean r^2 from 0.860 to 0.873 [$t(1979) = 12.73$, $p < 0.0001$]. Including the Gender × Time interaction improved the model further, increasing mean r^2 to 0.876 [$t(1989) = 2.30$, $p < 0.05$], since males again did not show the increase in f_0 onset that females did. No other predictors emerged as significant. Post-hoc comparisons of adjacent time

[17] These patterns of change in the production of English voiced stops held not only of short-lag tokens, but of prevoiced tokens as well. Like short-lag tokens, prevoiced tokens showed little change in VOT over time, and they also increased in f_0 onset, a trend that was marginally significant [$F(4, 10) = 2.61$, $p = 0.09$].

points using Tukey's HSD test showed that significant changes in VOT occurred between Weeks 1 and 2 [$p<0.01$] and Weeks 4 and 5 [$p<0.05$], and that significant changes in f_0 onset occurred between Weeks 1 and 2 [$p<0.001$] and Weeks 3 and 4 [$p<0.05$]. In short, English voiceless stops were found to increase in both VOT and f_0 onset, and these effects were neither limited to the first week of the language program nor specific to one particular place of articulation. However, as was the case with the English voiced stops, the increase in f_0 onset was specific to female learners. The disparity between the genders in this regard is illustrated in Fig. 3.

By virtue of the fact that the Place × Time interaction did not emerge as a significant predictor, the regression model of VOT in English voiceless stops implied that VOT drifted over time at all three places of articulation. However, it did not indicate whether stops at different places of articulation drifted by the same amount. For this reason, post-hoc analyses of VOT differences between Weeks 1 and 5 were done using Tukey's HSD test on English voiceless stop tokens divided by place of articulation. As illustrated in Fig. 4, these analyses showed that the VOT of English /p/ lengthened by 22 ms from Week 1 to Week 5, an increase that was statistically significant [$p<0.0001$]. The VOT of English /k/ lengthened by a similar amount (20 ms), an increase that was also statistically significant [$p<0.001$]. However, the VOT of English /t/ lengthened by a smaller amount (14 ms), an increase that was only marginally significant [$p=0.06$]. These results suggest that while the VOT of English voiceless stops at all places of articulation lengthened in approximation to the longer VOT of Korean aspirated stops, the VOT of English /t/ lengthened to a lesser degree than the VOTs of English /p/ and /k/, in accordance with the differences in cross-linguistic similarity discussed in Section 2.3. Thus, the drift in VOT of English /t/ evinced both subsegmental and segmental influence from Korean: subsegmental influence from the significantly longer VOT norms of Korean aspirated stops (which served to lengthen VOT), and segmental influence from the similar VOT norms of Korean /tʰ/ (which served to moderate the amount of VOT lengthening).

Developments in learners' Korean stops concurrent with the changes in their English stops provided additional support for the conclusion that phonetic drift of learners' English occurred in approximation to Korean. While the VOTs of English voiced stops and Korean fortis stops both stayed relatively steady (Fig. 5a), the VOTs of English voiceless stops and Korean aspirated stops both lengthened significantly (Fig. 5b). Over five weeks, the VOT of Korean aspirated stops lengthened by approximately 25 ms over its initial level in Week 1, while the VOT of English voiceless stops lengthened by 19 ms. The initial increase in the VOT of Korean aspirated stops was expected, as it was consistent with learners' approximation of the relatively long VOT norms for Korean aspirated stops. However, the continued increase in the VOT of Korean aspirated stops was unexpected, since by Week 2 learners had already reached native-like VOT levels for this stop series. By continuing to lengthen VOT beyond this point, learners were thus over-aspirating the Korean aspirated stops. In contrast to the Korean aspirated stops, the increase in VOT of the English voiceless stops cannot be explained in terms of phonetic norm approximation, since—at a VOT of 86 ms in Week 1—they started off well above the appropriate VOT norm for English voiceless stops, which is estimated to be approximately 60 ms at the beginning of isolated syllables with the vowel /ɑ/ when averaging over talker gender and place of articulation (Morris, McCrea, & Herring, 2008, p. 312). Instead, the manner in which Korean aspirated stops and English voiceless stops changed in lockstep (such that they did not differ significantly in VOT at any time point) suggests that their similar increases in VOT were not coincidental, but coordinated: as the Korean aspirated stops came to be produced with more native-like and then exaggerated VOT, the English voiceless stops "went along for the ride" and lengthened in VOT as well, even though from the outset this change resulted in the English stops becoming less native-like vis-à-vis the phonetic norms of American English.[18]

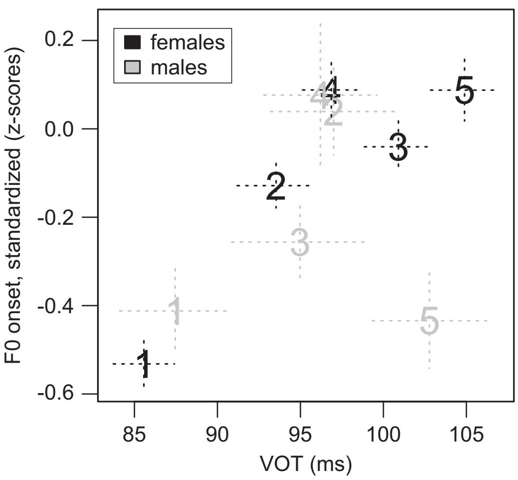

Fig. 3. Mean VOT vs. mean f_0 onset (standardized) over time for English voiceless plosives. Numerical symbols plot means of the respective weeks for female learners in black and for male learners in gray. Error bars indicate ±1 standard error about the mean.

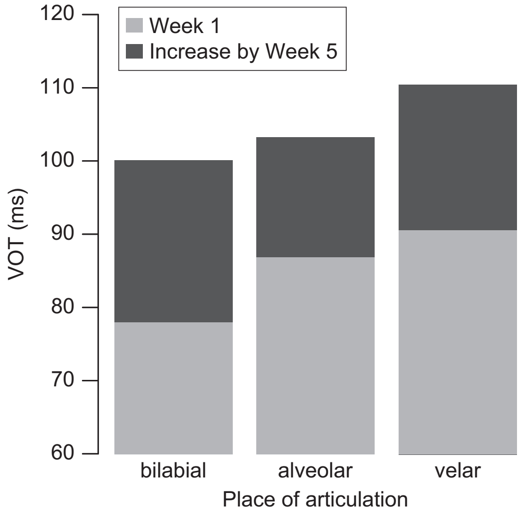

Fig. 4. Mean VOT of English voiceless plosives in Week 1 (light gray), compared to the increase in VOT by Week 5 (dark gray), by place of articulation.

[18] An anonymous reviewer wondered whether the VOT lengthening was the product of hyperarticulation arising from a task familiarity effect (increased familiarity with the materials may have allowed participants to give more confident productions, thereby resulting in hyperarticulation of the long-lag stops). There are three main reasons why this alternative account of the VOT lengthening is implausible. First, the observed pattern is contrary to the expected effect of familiarity, which typically results in articulatory reduction even from a very early age (see, e.g., Schwartz, 1995). Second, if, for the sake of argument, increased familiarity were assumed to lead to increased hyperarticulation, there would be no reason not to expect this effect to appear across the board, yet this is hardly the case. English voiced stops, for example, did not shorten in VOT (or, for that matter, lower in f_0 onset). Third, effects of increases in familiarity should be more pronounced for the Korean stops, since the Korean materials would have started out as much less familiar to the participants than the English ones. However, as discussed below, changes in f_0 were actually less pronounced for the Korean stop types than the English stop types. Together, these facts suggest that the observed patterns of phonetic drift did not arise as an artifact of increases in task familiarity.

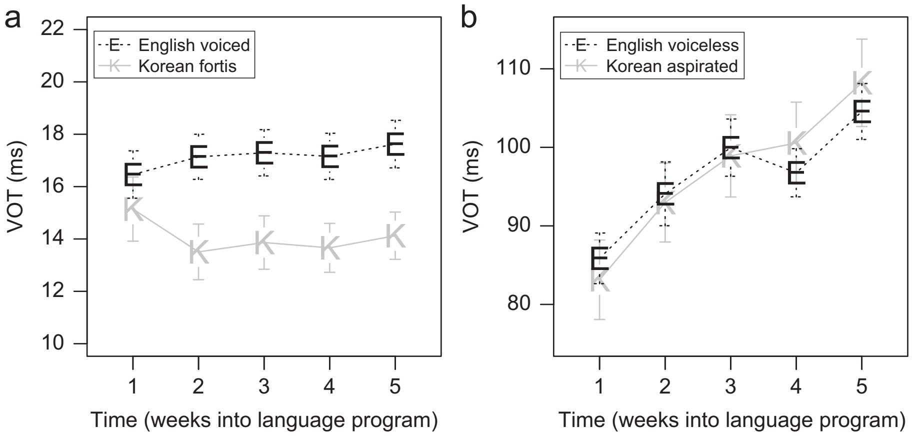

Fig. 5. Mean VOT of English plosives ('E') compared to mean VOT of the perceptually linked Korean plosives ('K') over time: (a) English voiced vs. Korean fortis, and (b) English voiceless vs. Korean aspirated. Error bars indicate 95% confidence intervals.

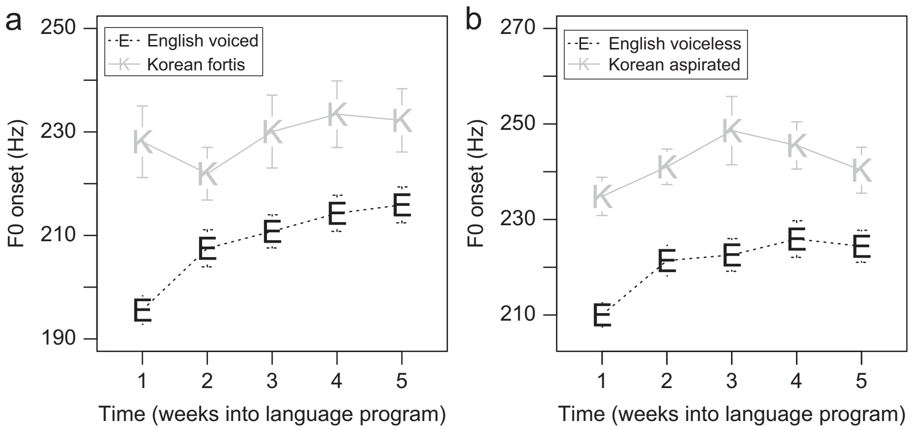

Fig. 6. Mean f_0 onset of female learners' English plosives ('E') compared to mean f_0 onset of the perceptually linked Korean plosives ('K') over time: (a) English voiced vs. Korean fortis, and (b) English voiceless vs. Korean aspirated. Error bars indicate 95% confidence intervals.

Developments in f_0 onset for female learners (i.e., the learners who showed a change in their English f_0) were consistent with the assimilatory pattern of changes in VOT. With the exception of a dip in Week 2, the f_0 following Korean fortis stops stayed steady over time; however, the f_0 following English voiced stops drifted upward toward the f_0 associated with the Korean fortis stops, such that the f_0 distance between English voiced stops and Korean fortis stops shrank from 33 Hz in Week 1 to 16 Hz in Week 5 (Fig. 6a). Similarly, while the f_0 following Korean aspirated stops did not show a net change over five weeks, the f_0 following English voiceless stops also drifted upward, resulting in the f_0 distance between English voiceless stops and Korean aspirated stops decreasing from 25 Hz in Week 1 to 16 Hz in Week 5 (Fig. 6b). Thus, both the voiced and voiceless stops of English became more similar to the fortis and aspirated stops of Korean in terms of f_0 onset. However, it is not possible to conclude that these f_0 increases arose via cross-language linkages between perceptually similar stop types specifically, since the Korean stop types do not differ in terms of their potential effect in this respect. The Korean stop types are both higher in f_0 than the corresponding English stop types; therefore, they may have triggered upward drift in f_0 following English stops via linkages between parallel stop types or simply via a global link in overall f_0 level across languages.

What is required to conclude that the observed f_0 increases following English stops resulted solely from cross-language linkages at the level of the laryngeal natural class is evidence that f_0 did not similarly increase in English words that would have been unaffected by the f_0 properties of Korean stop onsets—namely, onsetless (i.e., vowel-initial) words. On the other hand, if f_0 were found to increase in vowel-initial words as well, this would constitute evidence that the observed f_0 increases resulted at least in part from a global linkage of overall f_0 level across languages. Thus, in addition to stop-initial English words, f_0 onset was also measured in the vowel-initial English word *all*. Contrary to the hypothesis of f_0 drift via class-level linkage exclusively, the f_0 onset of English vowel-initial tokens was also found to increase over time, and a mixed-model analysis of variance showed that this effect was statistically significant [$F(4,60)=4.51$, $p<0.01$]. Nevertheless, the magnitude of f_0 increase in English stop-initials was larger than that in English vowel-initials (Fig. 7). While the mean f_0 onset of English vowel-initials rose by 10 Hz from Week 1 to Week 5, that of English voiced and voiceless stops rose by 20 Hz and 14 Hz, respectively. Therefore, these data suggest that the upward drift in the f_0 of English voiced and voiceless stops actually arose via cross-language linkages at two levels: a global linkage to the higher overall f_0 level of Korean (which resulted in f_0 drift for English vowel-initials, too), as well as class-level linkages to Korean fortis and aspirated stops (which resulted in greater f_0 drift for English stop-initials than for English vowel-initials).

To summarize, patterns of phonetic drift in the VOT and f_0 onset of English stop consonants followed the predictions presented in Section 2.4: phonetic drift in adult L2 learners occurred

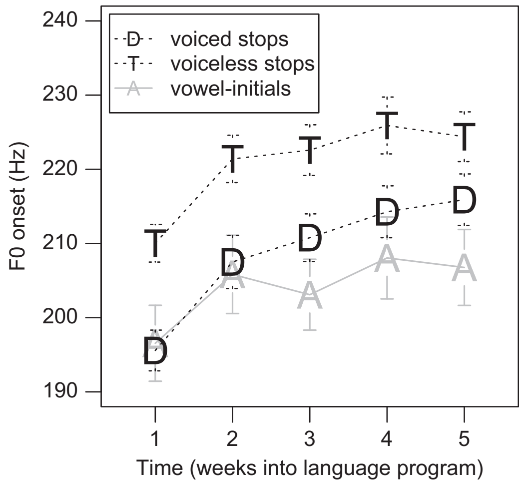

Fig. 7. Mean f_0 onset in female learners' English voiced plosives ('D'), English voiceless plosives ('T'), and English vowel-initials ('A') over time. Error bars indicate 95% confidence intervals.

rapidly, approximated perceptually linked L2 structures, and showed both generality and specificity. When cross-linguistic differences in a given dimension were, by virtue of their size, unlikely to be perceived (as in the case of English voiced stop VOT vis-à-vis Korean fortis stop VOT), phonetic drift did not take place. In contrast, when cross-linguistic differences were likely to be perceptible, phonetic drift occurred after only weeks of L2 learning; however, the drift patterns showed effects of cross-language linkage at more than one level. The drift in VOT of English /t/ showed influence from the VOT norms of both the corresponding Korean stop type generally and the corresponding Korean segment specifically, while the drift in f_0 onset of English voiced and voiceless stops showed influence from both general f_0 level in Korean and specific f_0 properties of the corresponding Korean stop types.

4.2. Phonetic drift in vowel production

Over time, the English vowels drifted for both female and male learners, although effects for the few male learners did not generally reach statistical significance. The progression of phonetic drift from Week 1 to Week 5 is shown in Figs. 8 and 9 for female and male learners, respectively.[19] For female learners, the English vowels generally drifted upward (i.e., F_1 decreased over time), while for male learners, they drifted downward and, to some extent, forward in the vowel space (i.e., F_1 and F_2 increased over time). Regression models fit to the F_1 data with one fixed effect indicated that of Vowel, Gender, and Time, Vowel had, unsurprisingly, the strongest effect on F_1. The 95% confidence interval for r^2 was 0.914–0.934 in models with only Vowel as a fixed effect, compared to 0.034–0.090 in models with only Gender or Time as a fixed effect. Although Vowel accounted by itself for over 90% of the variance in F_1, inclusion of Gender as well as Time significantly improved upon Vowel-only models of F_1. When Gender and the Vowel × Gender interaction were added to Vowel

[19] Weeks 2 and 4 are omitted from vowel plots for clarity of presentation, though they were included in all the statistical analyses. Error bars in the vowel plots represent the average of individual participants' standard errors for a given vowel, rather than standard errors calculated over the distribution for the entire group. This calculation prevents formant disparities due to physiological differences between male and female participants from inflating the displayed error, allowing for a more accurate representation of the range of variation in the average participant's production of each vowel.

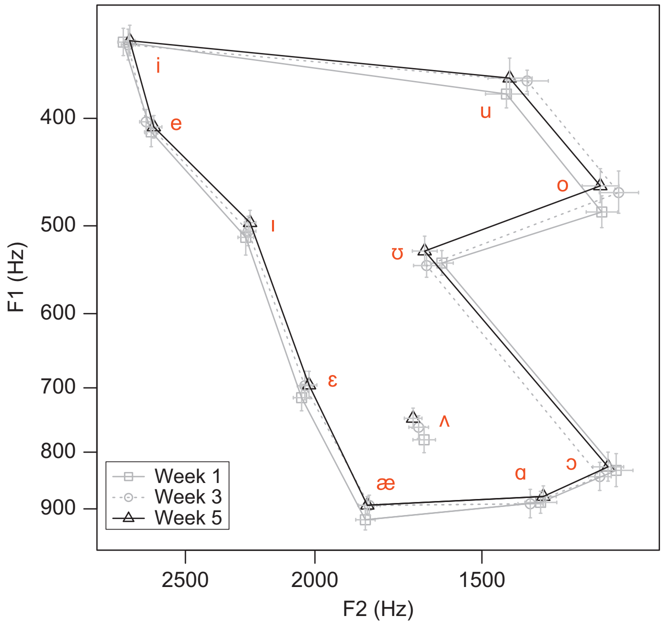

Fig. 8. Mean F_1 by mean F_2 of English vowels over time as produced by female L2 learners of Korean. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 mean standard error.

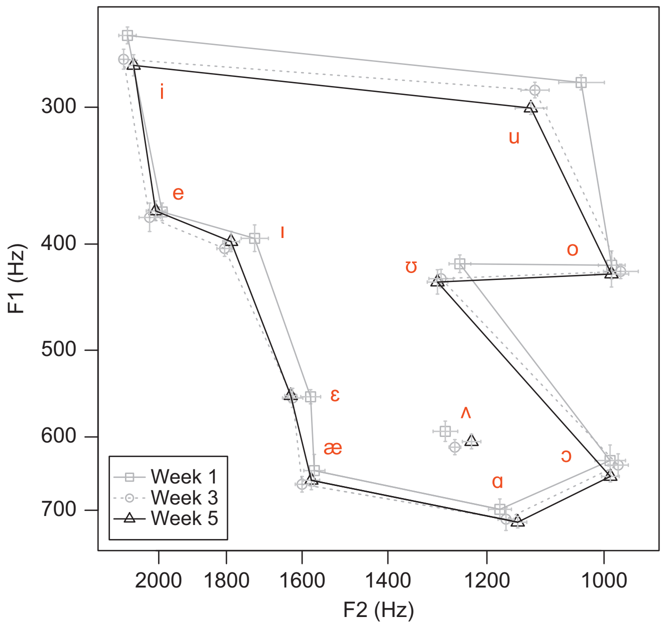

Fig. 9. Mean F_1 by mean F_2 of English vowels over time as produced by male L2 learners of Korean. The scale of both axes is logarithmic. Week 1 means are represented with squares connected by solid gray lines; Week 3 means, with circles connected by dotted gray lines; and Week 5 means, with triangles connected by solid black lines. Error bars indicate ±1 mean standard error.

as fixed effects, mean r^2 increased by approximately 0.013 from 0.924 to 0.937 [$t(1949) = 63.35$, $p < 0.0001$], and it further increased to 0.938 [$t(1993) = 2.52$, $p < 0.01$] when Time and the Gender × Time interaction were added to the model. However, neither the Vowel × Time interaction nor the Vowel × Gender × Time interaction emerged as a significant predictor [$Fs(40, 918) < 0.45$, *n.s.*], suggesting that the overall trends were not isolated to just a few vowels. Similar to models of F_1,

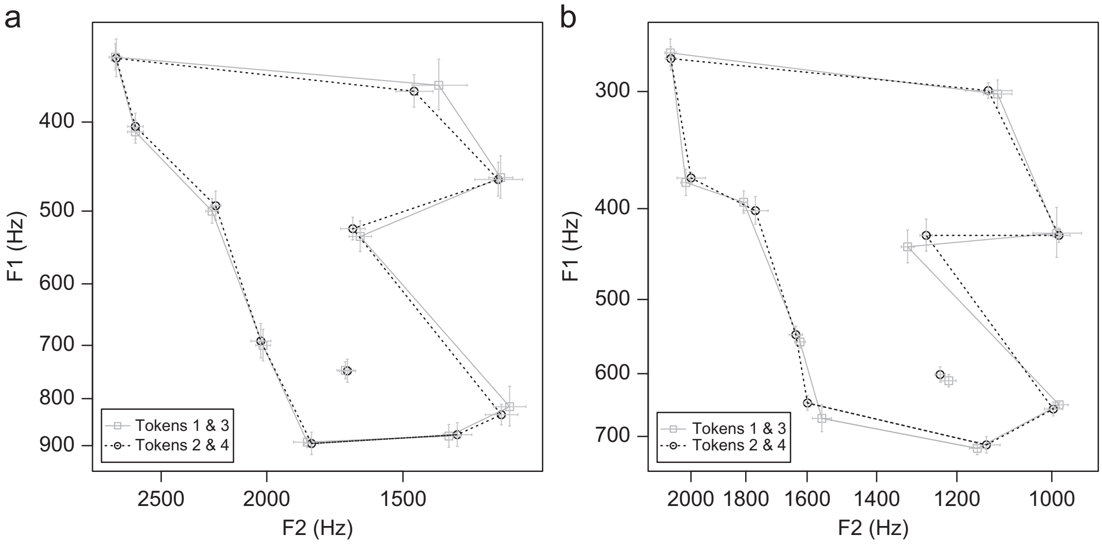

Fig. 10. Variability of English vowel production in Week 5, by talker gender: (a) female L2 learners, and (b) male L2 learners. In each plot, mean F_1 is plotted by mean F_2 on a logarithmic scale. Means over the first half of productions (tokens 1 and 3) are represented with squares connected by solid gray lines; means over the second half of productions (tokens 2 and 4), with circles connected by dotted black lines. Error bars indicate ± 1 mean standard error.

regression models fit to the F_2 data indicated that Vowel had the strongest effect on F_2, but that including Gender significantly improved upon Vowel-only models. When Gender and the Vowel × Gender interaction were included in the model, mean r^2 increased by approximately 0.012 from 0.929 to 0.941 [$t(1949)=63.35$, $p<0.0001$]. However, adding Time failed to improve the model's performance [$t(1998)=-1.33$, *n.s.*], and there were no significant effects of any interactions with Time, either [Gender × Time: $F(4,918)=0.23$, *n.s.*; Vowel × Time, Vowel × Gender × Time: $Fs(40,918)<0.33$, *n.s.*].

In order to examine the generality of the F_1 decrease among the more numerous female learners, mean F_1 of female learners was further examined by vowel. Post-hoc comparisons of F_1 between Week 1 and Week 5 via paired one-tailed *t*-tests showed a statistically significant or marginally significant decline in F_1 for six of the eleven vowels examined: /ɪ/ [$t(15)=1.66$, $p=0.06$], /ɛ/ [$t(15)=1.56$, $p=0.07$], /æ/ [$t(15)=1.99$, $p<0.05$], /ʊ/ [$t(15)=1.68$, $p=0.06$], /o/ [$t(15)=2.07$, $p<0.05$], and /ʌ/ [$t(15)=1.89$, $p<0.05$]. The differences for the remaining five vowels (/i, e, u, ɑ, ɔ/) did not reach significance; however, they, too, showed a decrease in F_1 that was consistent with the pattern for the other vowels. These results support the conclusion that the decrease in F_1 for female learners was a general phenomenon characterizing the English vowel system as a whole, rather than just a few vowels in particular.

Thus, after accounting for the effects of vowel, gender, and participant, time was still found to have an effect on F_1 (but not F_2) of the English vowels, though males and females differed in terms of the effect that both vowel and time had on their produced F_1. The gender difference with respect to the effect of vowel can be seen in comparing Fig. 8 with Fig. 9: although the male and female vowel spaces were similar in overall shape, they differed in terms of the precise organization of the vowels with respect to each other. For example, /æ/ was basically even with /ɑ/ for female learners, but higher than /ɑ/ for male learners; conversely, /o/ was even with /ʊ/ for male learners, but higher than /ʊ/ for female learners. The gender difference with respect to the effect of time was in the directionality of phonetic drift in F_1, which generally decreased for female learners, but increased for male learners.

Although the English vowels were found to drift over time, the effect was quite subtle. In order to check whether the observed shifts were in fact different from normal speech variability one would expect to find at a given point in time, the consistency of learners' English vowel production was examined at the last measurement point (Week 5 of the language program) by splitting the data into two halves of non-consecutive tokens (i.e., tokens 1 and 3, and tokens 2 and 4). Separate means were calculated for each half of the data, and the two halves were then compared with each other. This comparison showed that, for both male and female learners, there was some variability in the production of certain vowels at this time point—in particular, /u/ for females (Fig. 10a), and /æ/ and /ʊ/ for males (Fig. 10b). Nevertheless, the overall picture that emerged from this analysis was that, at one point in time, learners were highly consistent in their English vowel production, suggesting that the directional drift observed in their English vowels over time was not simply an artifact of normal speech variability.

With regard to vowel dispersion, in contrast to the phonetic drift of Quichua vowels in Guion (2003), which resulted in increased cross-language dispersion of the Quichua and Spanish vowels, here phonetic drift of the English vowels resulted in decreased cross-language dispersion of the English and Korean vowels. When the mean acoustic distance between English and Korean vowels (i.e., the average of the distances in $F_1 \times F_2$ space between every possible English–Korean vowel pair) was calculated for each time point, it was found that cross-linguistic vowel spacing did not increase over time for either female or male learners (Fig. 11). On the contrary, there was a tendency for the mean spacing between vowels to decrease over time, a trend that was marginally significant for female learners [$F(4,60)=2.24$, $p=0.07$].

Given that cross-linguistic vowel spacing tended to decrease, it follows that phonetic drift of English vowels should have occurred in approximation to Korean vowels, yet the drift patterns of individual English vowels are inconsistent with respect to whether they converged with or diverged from nearby Korean vowels. In the case of female learners, for example, the raising of English /ʊ/ brought it closer to the nearby Korean /ɨ/, but the raising of English /u/ took it farther away from both Korean /ɨ/ and /u/, the two closest L2 vowels (Fig. 12). Similarly, the raising of English /e/ was convergent with the closest Korean vowel, /i/, while the raising of English /ɪ/ was divergent from the closest Korean vowel, /ɛ/. The existence of many such contradictory patterns in the phonetic drift of individual vowels indicates that the observed vowel shifts cannot be accounted for coherently in

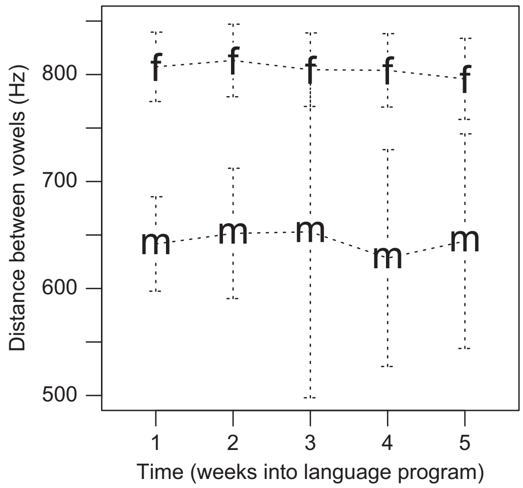

Fig. 11. Mean spacing between model (native female) Korean vowels and English vowels over time, by talker gender. Error bars indicate 95% confidence intervals.

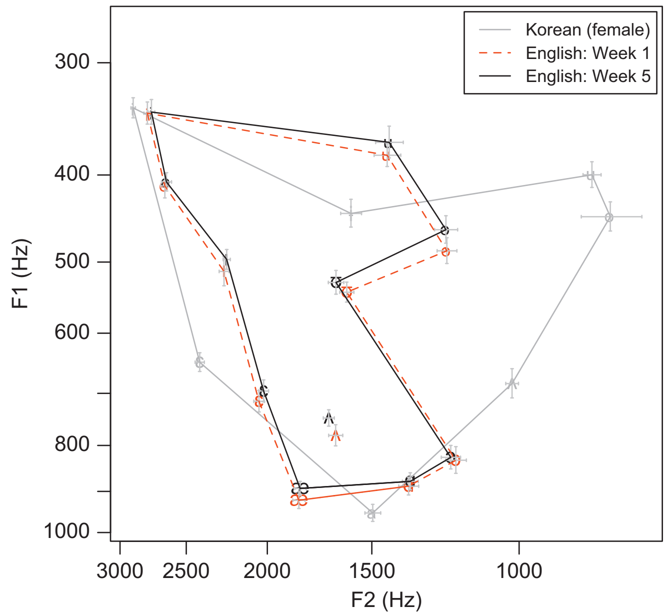

Fig. 12. Phonetic drift of the English vowels, female L2 learners of Korean. The scale of both axes is logarithmic. Plot symbols are the standard IPA transcriptions of the vowels. Means for model female Korean vowels are connected with gray solid lines; for English vowels in Week 5, with black solid lines; and for English vowels in Week 1, with dotted lines. Error bars indicate ±1 mean standard error.

terms of cross-linguistic influence at the level of individual vocalic segments.

On the other hand, comparisons of mean formant values of the English vowel system and the Korean vowel system suggest that the observed phonetic drift was indeed assimilatory, but systemic in nature. Over time, mean F_1 of female learners' English vowels decreased toward the lower mean F_1 of female Korean vowels (Fig. 13a), while mean F_2 of female learners' English vowels stayed steady (Fig. 13b). A post-hoc examination of differences between weeks using Tukey's HSD test indicated that while the F_1 increases between Weeks 1 and 2 and Weeks 3 and 4 were not statistically significant, the decreases between Weeks 2 and 3 and Weeks 4 and 5 were [$p < 0.001$], resulting in a net F_1 decrease of approximately 17 Hz from Week 1 to Week 5. Conversely, mean F_1 and F_2 of male learners' English vowels increased, respectively, by 8 Hz and 26 Hz toward the higher formant norms of female Korean vowels (Fig. 13a and b), although, with only three male learners, neither of these trends reached significance.

To summarize, phonetic drift in vowel formants—like phonetic drift in VOT and f_0—provided evidence in support of the hypotheses presented in Section 2.4. Again, significant changes in L1 production were evident after mere weeks of L2 learning, and the changes were convergent with the L2. In the case of vowel production, however, phonetic drift was systemic, occurring uniformly over the entire vowel system rather than disparately for different individual vowels. Female learners showed a decrease in mean F_1 of the English vowel system approximating the mean F_1 of the model Korean vowel system, but little change in mean F_2. On the other hand, male learners showed slight increases in mean F_1 and F_2 of the English vowel system, changes that also appeared to approximate the model female Korean vowel system.

5. Discussion

This study documented phonetic drift in L1 as a consequence of elementary experience in an L2. As predicted, modifications to L1 production occurred early in L2 acquisition, in assimilation to the L2, and on multiple structural levels. Here we consider the implications of these findings for the conceptualization of cross-linguistic perceptual relations, L1 development, and diachronic sound change.

5.1. Cross-linguistic similarity and perceptual linkage

A recurring theme of this study has been the dual nature of phonetic drift as both general and specific. Phonetic drift in English vowel formants was found to occur generally over the vowel system, rather than diversely for different individual vowels. Phonetic drift in English VOT also showed generalization: VOT of English voiceless stops lengthened at every place of articulation including the alveolars, which for females should not have drifted on the basis of their distance from Korean aspirated alveolar stops. However, at the same time the amount of VOT drift differed across places of articulation: more drift was found for bilabials and velars than for alveolars, in accordance with differences in VOT norms between English and Korean stops at each place of articulation. In a similar way, phonetic drift in English f_0 showed generalization to vowel-initials, while still showing effects of the higher f_0 norms associated with Korean fortis and aspirated stops specifically: f_0 increased to a greater degree for the English stop-initials perceptually linked to the Korean fortis and aspirated stops than for English vowel-initials. Thus, although there was generalization of phonetic drift, L1 modifications nevertheless remained sensitive to differences between L1 and L2 phonetic norms at lower levels of phonological structure.

The picture that emerges from these results is of a multi-layered network of cross-linguistic linkages in bilingualism. To the extent that the phonology of an individual language is organized into units of various size, it can be perceptually linked to the phonology of a different language in various ways, and cross-linguistic linkages at different levels of structure seem to jointly influence the realization of phonetic drift. Phonetic drift in English VOT was influenced by a segmental linkage between English /t/ and Korean /tʰ/ and a subsegmental linkage between English voiceless stops and Korean aspirated stops, while phonetic drift in English f_0 was influenced by subsegmental linkages between English voiced stops and Korean fortis stops and between English voiceless stops and Korean aspirated stops, as well as a global linkage between overall f_0 levels in English and

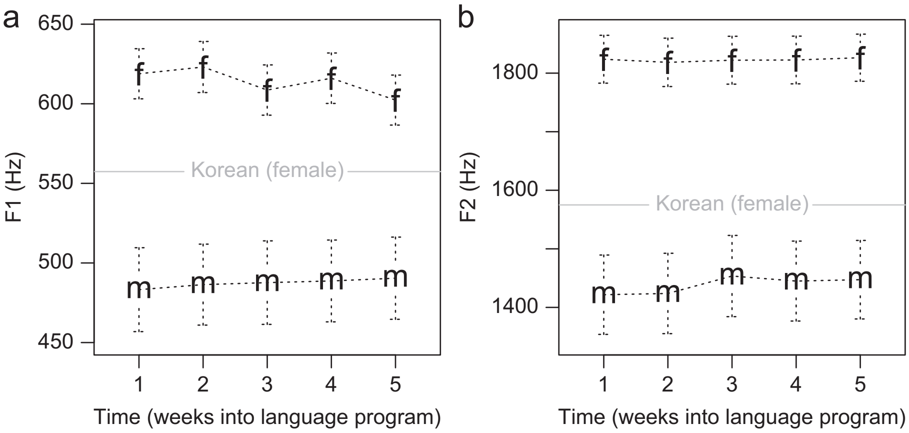

Fig. 13. Mean F_1 (a) and mean F_2 (b) of the English vowel system over time in comparison to the model Korean vowel system, by talker gender. Error bars indicate 95% confidence intervals.

Korean, which may arise from a shared control mechanism for f_0 modulation. Phonetic drift in English vowel formants, too, was influenced by global linkages between overall F_1 and F_2 levels in English and Korean.[20] These findings serve as a reminder that, despite the segmental focus of the L2 speech literature, cross-linguistic similarity is complex and multifaceted. Languages resemble each other more or less at several levels of granularity, and these different levels of analysis all seem to shape patterns of cross-linguistic phonetic influence.

In short, the present findings are consistent with principles of the SLM and the PAM-L2, but only partly predicted by these models, since these models do not address cross-linguistic perceptual relations beyond the segmental level. A complete model of cross-linguistic phonetic influence must also account for cross-language developments that occur at a non-segmental level, which constitute a large part of cross-language phonetic effects, as shown in the current study. Such a model will need to acknowledge the multiple sources of information and influence in speech production, including general mechanisms of speech motor control and somatosensory feedback (see, e.g., Larson, Altman, Liu, & Hain, 2008; Perkell et al., 1997; Tremblay, Shiller, & Ostry, 2003).

5.2. Continuity of language development

One of the main empirical contributions of this study is the finding that experience in another language rapidly alters production of the native language. For the late L2 learners examined here, a few weeks of learning Korean promptly affected their English, in a manner that both generalized across segmental categories and approximated the phonetic properties of the L2. These results obtained in spite of the study's focus on adult learners, formal speech, and an L1–L2 pair with relatively little in common—all factors that were predicted to reduce the likelihood of finding early L2 influence on the L1. The fact that phonetic drift occurred even in these ostensibly adverse conditions suggests that there is nothing out of the ordinary about phonetic drift, and that thinking of this phenomenon in terms of attrition (see, e.g., Schmid, 2010; Seliger & Vago, 1991) misses the bigger picture. The current findings are remarkable precisely because they cannot be attributed to L1 attrition stemming from lack of L1 use, as many previous findings of phonetic drift could be. As mentioned in Section 3.1, outside of actual class time learners in the current study functioned almost entirely in English—an expected pattern, since even by the end of their language program they had obtained only a rudimentary command of Korean.[21] However, in spite of their continued high L1 use and low L2 attainment, learners still manifested phonetic drift, even after the first week of classes when they knew very few lexical items of the L2. This suggests that, contrary to what has been assumed in the L2 speech literature, a high level of L2 proficiency is not necessary for L2 input to influence L1 representations.

Instead of being symptomatic of attrition, the occurrence of phonetic drift seems to be indicative of a fluid, multifaceted quality to language development over the lifespan, wherein production of a particular language can be shaped not only by experience in that language, but by experience in all languages within an individual's linguistic system. This system is constantly changing over time, such that—far from fossilizing—L1 representations developed in childhood continue to be updated in adulthood, consistent with the SLM's notion of continuity in language development as well as with the conceptual framework of dynamic systems theory (de Bot, Lowie, & Verspoor, 2007). Although phonetic drift is not necessarily a mark of attrition, there is a natural link between the two, and it may be most accurate to think of phonetic drift as one step in a continuum of cross-linguistic effects in bilinguals dependent on relative use of the L1 vs. the L2. As they gain proficiency and confidence in the L2 over time, L2 learners may increase their use of the L2 at the expense of the L1, raising the likelihood of the L2 exerting an influence on L1 representations. A significant decline in L1 use, especially in situations of L2 immersion, has been shown to affect

[20] As discussed in Section 4.2, female and male learners differed with respect to the direction of drift in the English vowel system, but both directions of drift were assimilatory to the L2 in light of the fact that the L2 model for all learners was female Korean. Compared to female learners' English vowels, male learners' English vowels started on the opposite side of the L2 Korean vowels (i.e., lower in overall F_1 and F_2), and so males' approximation of the L2 naturally occurred in the opposite direction. Nevertheless, it is somewhat surprising that male learners showed evidence of approximation to female Korean vowels, especially since they did not show such approximation to female Korean f_0. The reason for these differences in patterning between vowel formants and f_0 is unclear, although it is possible that they are related to sociolinguistic differences between the genders in f_0 modulation (see, e.g., Daly & Warren, 2001).

[21] This is not unexpected, since Korean is often cited as one of the most difficult languages for native English speakers to acquire as an L2 (see, e.g., the National Security Agency's report on comparative difficulty of foreign language learning). As a Category IV language (Association of the United States Army, 2010; Quigley, 2008), Korean is estimated to require 64 weeks of intensive instruction for L2 learners to reach the high elementary speaking level of 1+ on the ILR scale, more than twice as long as it takes to reach that level in a Category I language such as Spanish or French. Thus, even without formal assessment data it is reasonable to assume that after only six weeks of Korean instruction participants were not very proficient.

aspects of L1 performance such as lexical access relatively quickly, suggesting that inhibitory processes facilitating L2 acquisition play an important role in early L1 attrition (Linck, Kroll, & Sunderman, 2009). Thus, while the phenomenon of cross-linguistic influence instantiated in phonetic drift may provide the basic mechanism of phonological attrition, it does not fully determine the occurrence of attrition. Rather, there seem to be other contributing factors, complicating the question of when phonetic drift gives way to phonological attrition.

Although the current findings suggest that the starting point of phonetic drift coincides with the beginning of L2 experience, they do not have anything to say about the endpoint of phonetic drift. One wonders whether after extensive experience in another language, an individual can ever return to producing his L1 like a monolingual again. Studies on speech adaptation have shown that production adjustments made in response to altered auditory feedback persevere for a short while, but eventually dissipate after the feedback alteration is removed. Thus, it is possible that phonetic drift in response to L2 exposure might also subside if L2 exposure were cut off. There are no controlled studies on the effects of distant, rather than recent or current, L2 experience in adulthood on L1 production, but it is probable that the degree to which L1 production drifts back to monolingual-like values in the continued absence of L2 exposure will depend greatly on the total amount of L2 experience accrued. For example, over the same period of time with only L1 exposure, less "return" drift is predicted for a speaker with twelve years of L2 experience than for one with twelve months of L2 experience. However, the fact remains that virtually nothing is known about the temporal progression of such "return" drift and its relation to amount of accrued L2 experience. Can an eighty-year-old immigrant with sixty years of L2 experience return to pronouncing her L1 like a monolingual upon moving back to her (monolingual) native country? Such questions highlight the need for empirical studies of the effects on the L1 of distant L2 experience gained in adulthood. Individuals with this sort of linguistic experience, typically excluded from research because they qualify as neither monolingual nor bilingual, should be studied in their own right—-much as heritage speakers are starting to be—in order to better our understanding of linguistic experience: how it is acquired, stored, and forgotten (or not forgotten).

5.3. From individual to community-wide drift

In the study of historical language change, a distinction has been drawn between change at the lexical level (i.e., borrowing of words) and change at a structural level (e.g., phonology, morphology, syntax). Cross-linguistic influence at the lexical level is thought to be relatively common, whereas cross-linguistic influence at a structural level is thought to be less common, requiring a high degree of L2 experience among a large segment of the speech community:

> Although lexical borrowing frequently takes place without widespread bilingualism, extensive structural borrowing, as has often been pointed out, apparently requires extensive (though **not** universal) bilingualism among borrowing-language speakers over a considerable period of time (Thomason & Kaufman, 1988, p. 37).

The implication is that diachronic sound change arising from language contact (Boretzky, 1991; Campbell, 1976), the kind of change that occurs in a language over generations of speakers, arises in conditions of prolonged contact between an L1 and an L2. Sapir (1921, pp. 193–214) wrote of the progression of such change as "phonetic drift."

In the current study, the term "phonetic drift" has been used to refer to phonetic change at a micro level (the idiolect of an individual speaker), a terminological conflation that serves to emphasize the contiguity with phonetic change at a macro level (the language of a speech community). Simply put, phonetic drift at the micro level may be thought of as planting the seeds of contact-induced historical sound change, consistent with the idea that "the locus of language contact is the bilingual speaker" (Sankoff, 2002, Chap. 25, p. 643). Such sound change is likely to occur in areas of high within- and cross-language phonetic variability, in accordance with the claim that "bilingual phonologies may become particularly permeable to inter-linguistic influence precisely where they are acoustically and perceptually unstable, and where they are already congruent to some degree" (Bullock & Gerfen, 2004, p. 103). What is important to note is that the current study suggests, contrary to the discourse of historical linguistics, that structural change in a domain like phonetics can be significantly accelerated by a relatively brief period of L2 experience.

The propagation of phonetic drift throughout a population might occur in much the same manner as sound changes due to "hypocorrection", one of two complementary mechanisms of change in the theory of sound change proposed by Ohala (1993, Chap. 9). Ohala accounted for sound change in terms of non-veridical perception by listeners. Change could result from a perceptual "miss", where a listener failed to correct for features of the acoustic signal resulting from speaker-centric articulatory modifications ("hypocorrection") or from a perceptual "false alarm", where a listener incorrectly interpreted the signal in terms of articulatory modifications that had not in fact occurred ("hypercorrection"). Hypocorrection may provide the link between individual-level phonetic drift and community-level sound change, in that child acquirers of the L1 receiving L2-influenced L1 input from a late L2 learner (i.e., a native L1 talker with L2 experience) probably do not "correct" for the L2 influence in the L1 speech to which they are exposed. Unless they happen to be acquainted with the talker's linguistic experience (and this plays a role in how the talker's speech is stored), adult interlocutors are also unlikely to normalize for L2 influence in the L1 speech of an ostensibly native L1 talker. Thus, following interactions with L1 talkers manifesting phonetic drift, both child acquirers of the L1 and adult monolingual speakers of the L1, having failed to correct for the L2 influence, stand to retain L2 influence. In this way, phonetic drift within a few bilingual speakers in one generation may be passed on to later generations of speakers, both bilingual and monolingual.

Note that this scenario is not quite like either of the externally motivated routes to sound change that have previously been described in the historical linguistics literature—namely, borrowing and substratum influence.[22] As summarized by Sankoff (2002, Chap. 25, pp. 644–649), borrowing refers to the incorporation of elements from an L2 into the L1 (e.g., incorporation of a Romance stop voicing system into Dutch; see Simon, 2011), while substratum influence refers to the effect that non-native transformations of a language by L2 learners can have on L1 speakers of the language (e.g., influences of non-native features in the speech of

[22] Although this discussion emphasizes the effects of external factors on the development of sound change, it is important to recognize that such external factors typically combine with internal factors in determining the outcome of language contact (Sankoff, 2002, Chap. 25, p. 641). We cannot forget that individuals may show longitudinal changes in their speech production that have nothing to do with influence from another language (e.g., Harrington, Palethorpe, & Watson, 2000; Sankoff & Blondeau, 2007). As such, these internal developments must be considered in conjunction with external influences in evaluating the phonological consequences of language contact.

first-generation immigrants on the speech of the second generation; see, e.g., Fought, 1999). Naturally, sociohistorical variables (e.g., relative status of the languages, relative numbers of speakers, duration of contact) have played a large part in the explanation of when externally motivated sound change occurs via borrowing, substratum influence, or some combination of the two. The contribution of the present study is to point out another possible mechanism of externally motivated sound change: phonetic drift, which is coreferent neither with borrowing nor with substratum influence. Phonetic drift differs from borrowing in that it does not (necessarily) result in categorical change. The drift in VOT of English stops seen in Section 4.1, for example, did not alter the basic nature of the English voicing contrast. Phonetic drift also differs from substratum influence in that the locus of change is the native speaker, not the L2 learner. Thus, if the right sociolinguistic conditions were to obtain in this case of contact between English and Korean, a diachronic lengthening of VOT in English voiceless stops, for example, would be predicted to occur primarily by way of Korean-influenced VOT in native English speakers, not by way of Korean-influenced VOT in native Korean learners of English.[23]

6. Conclusion

The theme of this article has been the influential role of language experience in shaping the linguistic representations drawn upon in speech production. The present study demonstrated that learning an L2 influences production of the L1 in the short term, resulting in assimilatory modifications to the L1 on a number of structural levels. Future work on cross-linguistic production has the potential to contribute to a holistic picture of how L1 speech production can be affected by a wide range of linguistic experiences, such as ambient L2 exposure (e.g., Caramazza & Yeni-Komshian, 1974; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008) and interactions with non-native speakers (e.g., Ferguson, 1975; Kim, Horton, & Bradlow, 2011; Long, 1983).

In addition, the present findings speak to the need to be ever more careful about defining the subject population for a linguistic study. To be specific, conflating even marginally bilingual participants with monolingual participants is likely to obscure our understanding of the object of study. The empirical problem with using L1 speakers who have significant L2 experience as representative of monolingual L1 speakers is particularly evident in studies that make claims about phonetic norms for the L1. Disparities result when one study examines native speakers in the native speech community, while another study examines expatriates (e.g., Helgason & Ringen, 2008; Keating, Linker, & Huffman, 1983). The contribution of the current study in this respect is demonstrating that L2 experience must be especially rigorously controlled in studies of speech production, since even brief periods of L2 learning can trigger phonetic drift away from L1 norms.

The point to take away, however, is not that study participants should always be monolingual. Rather, the experiential characteristics of the study sample should accurately represent the population which the study means to investigate. Hence, a phonetic study of Vietnamese meant to generalize to monolingual Vietnamese speakers, for instance, should examine Vietnamese speakers without significant—in particular, recent—L2 experience. In effect, this means that such a study should be conducted in Vietnam, where the ambient language is Vietnamese, instead of in a country like the U.S., where the ambient language is largely English. More generally, this means that researchers studying languages that are not native to the surrounding community should expect to work in collaboration with colleagues abroad, since an accurate depiction of these language varieties as they are spoken in their native language environments might not be possible otherwise.

In closing, it bears repeating that the sorts of L1 phonetic developments documented in this study appear to be entirely normal and are probably much more common than is reflected in the literature on L2 speech. Anecdotally, the L2 learners under study did not show any detectable loss of L1 fluency while manifesting phonetic drift, nor did their casual speech in regular conversations sound noticeably accented. For this reason, use of the term "attrition" to describe the phenomenon of phonetic drift has been deliberately avoided, as it is a misnomer. Individuals undergoing attrition experience a deterioration in their L1 production as communication is accomplished increasingly in an L2, while individuals undergoing phonetic drift experience a change, but not necessarily a deterioration, in their L1 production due to the accumulation of L2 experience. In short, phonetic drift happens as a matter of course in L2 learners and should, therefore, be taken into account in any study of speech production.

[23] One wonders whether phonetic drift might not be partly responsible for the converse situation that has obtained in Korean—namely, diachronic shortening of VOT in Korean aspirated stops (Silva, 2006b)—as this development has been essentially contemporaneous with the rapid increase in English language instruction and programming in South Korea over the last half a century.

Acknowledgments

This study was conducted with the support of a National Science Foundation Graduate Research Fellowship and funding from the Center for Korean Studies and the Department of Linguistics at the University of California, Berkeley. The paper benefited immensely from the comments of Ocke-Schwen Bohn, Ann Bradlow, Susanne Gahl, Carla Hudson Kam, Sharon Inkelas, Keith Johnson, John Ohala, several anonymous reviewers, and audiences at the University of California, Berkeley, Yonsei University, and meetings of the Linguistic Society of America and the Acoustical Society of America. The author is also grateful to Daiana Chang and Kevin Sitek for research assistance and to the Fulbright Korean-American Educational Commission for logistical help.

References

Association of the United States Army (2010). *DLI's language guidelines*. AUSA News: Special Report, August.

Baker, W., & Trofimovich, P. (2005). Interaction of native and second language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In: J. C. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In: W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore, MD: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam, The Netherlands: John Benjamins Publishing.

Boersma, P., & Weenink, D. (2008). *Praat: Doing phonetics by computer, version 5.0.26*. 〈http://www.praat.org〉.

Boretzky, N. (1991). Contact-induced sound change. *Diachronica*, *8*(1), 1–15.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In: H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 119–136). London, UK: Academic Press.

Bullock, B. E., & Gerfen, C. (2004). Phonological convergence in a contracting language variety. *Bilingualism: Language and Cognition*, *7*(2), 95–104.

Bullock, B. E., & Toribio, A. J. (2009). Trying to hit a moving target: On the sociophonetics of code-switching. In: L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 189–206). Amsterdam, The Netherlands: John Benjamins Publishing.

Campbell, L. (1976). Language contact and sound change. In: W. M. Christie, Jr. (Ed.), *Current progress in historical linguistics* (pp. 181–194). Amsterdam, The Netherlands: North-Holland Publishing.

Campbell, L., & Muntzel, M. C. (1989). The structural consequences of language death. In: N. C. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death* (pp. 181–196). Cambridge, UK: Cambridge University Press.

Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, *2*(3), 239–245.

Chang, C. B. (2009). Convergence and divergence in language obsolescence. In: M. Pak (Ed.), *Current issues in unity and diversity of languages* (pp. 933–952). Seoul, South Korea: Linguistic Society of Korea.

Chang, C. B. (2010a). *First language phonetic drift during second language acquisition*. Ph.D. Thesis, Berkeley, CA: University of California, Berkeley.

Chang, C.B. (2010b). The implementation of laryngeal contrast in Korean as a second language. In: *Harvard studies in Korean linguistics* (pp. 91–104), vol. 13.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America*, *129*(6), 3964–3980.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Childers, D. G. (1978). *Modern spectrum analysis*. New York, NY: John Wiley & Sons.

Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193–228.

Daly, N., & Warren, P. (2001). Pitching it differently in New Zealand English: Speaker sex and intonation patterns. *Journal of Sociolinguistics*, *5*(1), 85–96.

de Bot, K. (1992). A bilingual processing model: Levelt's 'speaking' model adapted. *Applied Linguistics*, *13*(1), 1–24.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(1), 7–21.

de Leeuw, E., Schmid, M. S., & Mennen, I. (2010). The effects of contact on native language pronunciation in an L2 migrant setting. *Bilingualism: Language and Cognition*, *13*(1), 33–40.

Escudero, P. (2007). Second-language phonology: The role of perception. In: M. C. Pennington (Ed.), *Phonology in context* (pp. 109–134). New York, NY: Palgrave Macmillan.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, *17*(1), 1–14.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47–65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In: W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–272). Baltimore, MD: York Press.

Flege, J. E. (1996). English vowel productions by Dutch talkers: More evidence for the "similar" vs "new" distinction. In: A. James, & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 11–52). Berlin, Germany: Mouton de Gruyter.

Flege, J. E. (2002). Interactions between the native and second-language phonetic systems. In: P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217–244). Trier, Germany: Wissenschaftlicher Verlag.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In: J. Cole, & J. I. Hualde (Eds.), *Laboratory phonology*, Vol. 9 (pp. 353–382) Berlin, Germany: Walter de Gruyter.

Flege, J. E., & Eefting, W. (1987a). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, *6*(3), 185–202.

Flege, J. E., & Eefting, W. (1987b). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*(1), 67–83.

Forster, J. C. (2008). *DMDX, version 3.2.6.3*. 〈http://www.u.arizona.edu/∼ kforster/dmdx/dmdx.htm〉.

Fought, C. (1999). A majority sound change in a minority community: /u/-fronting in Chicano English. *Journal of Sociolinguistics*, *3*(1), 5–23.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., & Hallé, P. (2008). Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*, *36*(4), 649–663.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, *15*(2), 87–105.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In: H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge, UK: Cambridge University Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior*, *14*(4), 237–251.

Grosjean, F. (2001). The bilingual's language modes. In: J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Malden, MA: Blackwell Publishers.

Guion, S. G. (2003). The vowel systems of Quichua–Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica*, *60*(2), 98–128.

Haggard, M., Ambler, S., & Callow, M. (1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America*, *47*(2B), 613–617.

Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America*, *102*(1), 655–658.

Harada, T. (2003). L2 influence on L1 speech in the production of VOT. In: M.-J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th international congress of phonetic sciences* (pp. 1085–1088). Barcelona, Spain: Causal Productions.

Harrington, J., Palethorpe, S., & Watson, C. I. (2000). Does the Queen speak the Queen's English?. *Nature*, *408*(6815), 927–928.

Harrison, P. (1996). An experiment with tone. *UCL Working Papers in Linguistics*, *8*, 575–593.

Hazan, V., Messaoud-Golusi, S., Rosen, S., Nouwens, S., & Shakespeare, B. (2009). Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit?. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1510–1529.

Helgason, P., & Ringen, C. (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics*, *36*(4), 607–628.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*(5), 3099–3111.

Hinton, L., Moonwomon, B., Bremner, S., Luthin, H., van Clay, M., & Lerner, J., et al. (1987). It's not just the valley girls: A study of California English. In: J. Aske, N. Beery, L. Michaelis, & H. Filip (Eds.), *Proceedings of the 13th annual meeting of the Berkeley linguistics society: General session and parasession on grammar and cognition* (pp. 117–128). Berkeley, CA: Berkeley Linguistics Society.

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In: V. A. Fromkin (Ed.), *Tone: A linguistic survey* (pp. 77–111). New York, NY: Academic Press.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, *25*(1), 105–113.

Ingram, J. C. L., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, *25*(3), 343–370.

Interagency Language Roundtable (2011). *ILR speaking skill scale*. Available at 〈http://govtilr.org/Skills/ILRscale2.htm〉. Last retrieved October 1, 2011.

International Phonetic Association (1999). *Handbook of the international phonetic association*. Cambridge University Press, Cambridge, UK.

Jiang, H. (2010). Effect of L2 phonetic learning on the production of L1 vowels: A study of Mandarin–English bilinguals in Canada. In: K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *New sounds 2010: Proceedings of the 6th international symposium on the acquisition of second language speech* (pp. 227–232). Poznań, Poland: Adam Mickiewicz University.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In: K. Johnson, & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego, CA: Academic Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*(4), 485–499.

Johnson, K. (2008). *Quantitative methods in linguistics*. Malden, MA: Blackwell Publishing.

Jones, J. A., & Munhall, K. G. (2000). Perceptual calibration of *F*0 production: Evidence from feedback perturbation. *Journal of the Acoustical Society of America*, *108*(3), 1246–1251.

Jones, J. A., & Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Current Biology*, *15*(19), 1768–1772.

Kang, K.-H., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean–English bilinguals. *Journal of the Acoustical Society of America*, *119*(3), 1672–1683.

Kang, K.-H., & Guion, S. G. (2008). Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America*, *124*(6), 3909–3917.

Keating, P., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics*, *11*(3), 277–290.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–319.

Khattab, G. (2000). VOT production in English and Arabic bilingual and monolingual children. *Leeds Working Papers in Linguistics*, *8*, 95–122.

Kim, M. (2004). Correlation between VOT and F0 in the perception of Korean stops and affricates. In: *Proceedings of the 8th international conference on spoken language processing (INTERSPEECH-2004)* (pp. 49–52). Jeju Island, Korea: International Speech Communication Association.

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*(1), 125–156.

Ko, I. (2009). The merger of *ey* /e/ and *ay* /ɛ/ of Seoul Korean. *Eoneo Jeongbo [Language Information]*, *10*, 73–89.

Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors*. Malden, MA: Blackwell Publishers.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Berlin, Germany: Mouton de Gruyter.

Ladefoged, P. (2005). *Vowels and consonants* (2nd ed.). Malden, MA: Blackwell Publishing.

Laeufer, C. (1996). Towards a typology of bilingual phonological systems. In: A. James, & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 325–342). Berlin, Germany: Mouton de Gruyter.

Lane, H., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, *14*(4), 677–709.

Lane, H., Wozniak, J., Matthies, M., Svirsky, M., & Perkell, J. (1995). Phonemic resetting versus postural adjustments in the speech of cochlear implant users: An exploration of voice-onset time. *Journal of the Acoustical Society of America*, *98*(6), 3096–3106.

Larson, C. R., Altman, K. W., Liu, H., & Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice F_0 control. *Experimental Brain Research*, *187*(4), 613–621.

Lee, B. S. (1950). Some effects of side-tone delay. *Journal of the Acoustical Society of America*, *22*(5), 639–640.

Lee, H. B. (1993). Korean. *Journal of the International Phonetic Association*, *23*(1), 28–31.

Lehiste, I., & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America*, *33*(4), 419–425.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York, NY: John Wiley & Sons.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, *48*(4), 839–862.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*(12), 1507–1515.

Lindblom, B. (1986). Phonetic universals in vowel systems. In: J. J. Ohala, & J. J. Jaeger (Eds.), *Experimental phonology* (pp. 13–44). Orlando, FL: Academic Press.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384–422.

Lisker, L., Liberman, A. M., Erickson, D. M., Dechovitz, D., & Mandler, R. (1977). On pushing the voice-onset-time (VOT) boundary about. *Language and Speech*, *20*(3), 209–216.

Lombard, É. (1911). Le signe de l'élévation de la voix. *Annales des Maladies de l'Oreille, du Larynx, du Nez et du Pharynx*, *37*(2), 101–119.

Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, *5*(2), 177–193.

Lord, G. (2008). Second language acquisition and first language phonological modification. In: J. B. de Garavito, & E. Valenzuela (Eds.), *Selected proceedings of the 10th hispanic linguistics symposium* (pp. 184–193). Somerville, MA: Cascadilla Proceedings Project.

Mack, M. (1990). Phonetic transfer in a French–English bilingual child. In: P. H. Nelde (Ed.), *Language attitudes and language conflict* (pp. 107–124). Bonn, Germany: Dümmler.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, *76*(2), 190–208.

Major, R. C. (1996). L2 acquisition, L1 loss, and the critical period hypothesis. In: A. James, & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 147–159). Berlin, Germany: Mouton de Gruyter.

Matthies, M. L., Svirsky, M., Perkell, J., & Lane, H. (1996). Acoustic and articulatory measures of sibilant production with and without auditory feedback from a cochlear implant. *Journal of Speech and Hearing Research*, *39*(5), 936–946.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics*, *32*(4), 543–563.

Morris, R. J., McCrea, C. R., & Herring, K. D. (2008). Voice onset time differences between adult males and females: Isolated syllables. *Journal of Phonetics*, *36*(2), 308–317.

Natale, M. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, *32*(5), 790–804.

National Security Agency (2011). *Foreign language learning: A comparative analysis of relative difficulty*. Available at 〈http://www.nsa.gov/public_info/_files/cryptologic_spectrum/foreign_language.pdf〉. Last retrieved October 1, 2011.

Nearey, T. M., & Rochet, B. L. (1994). Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association*, *24*(1), 1–18.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142.

Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, *39*(1), 59–67.

Oh, J., Jun, S.-A., Knightly, L., & Au, T. (2003). Holding on to childhood language memory. *Cognition*, *86*(3), B53–B64.

Ohala, J. (1993). The phonetics of sound change. In: C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London, UK: Addison Wesley Publishing.

Paradis, M. (2001). An integrated neurolinguistic theory of bilingualism (1976–2000). *LACUS Forum*, *27*, 5–15.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, *119*(4), 2382–2393.

Pardo, J. S. (2010). Expressing oneself in conversational interaction. In: E. Morsella (Ed.), *Expressing oneself/expressing one's self: Communication, cognition, language, and identity* (pp. 183–196). London, UK: Psychology Press.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. New York, NY: Atheneum.

Peng, S. (1993). Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. *Phonetica*, *50*(4), 245–260.

Perkell, J., Lane, H., Svirsky, M., & Webster, J. (1992). Speech of cochlear implant patients: A longitudinal study of vowel production. *Journal of the Acoustical Society of America*, *91*(5), 2961–2978.

Perkell, J., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., & Wozniak, J., et al. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication*, *22*(2–3), 227–250.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*(1), 175–184.

Pile, E. J. S., Dajani, H. R., Purcell, D. W., & Munhall, K. G. (2007). Talking under conditions of altered auditory feedback: Does adaptation of one vowel generalize to other vowels? In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 645–648). Pirrot, Dudweiler, Germany.

Quigley, S. L. (2008). *Student linguists meet high standards at language institute*. Available at 〈http://www.defense.gov/news/newsarticle.aspx?id=51675〉. Last retrieved October 1, 2011.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roederer, J. G. (1973). *Introduction to the physics and psychophysics of music*. London, UK: English Universities Press.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *27*(4), 421–436.

Sankoff, G. (2002). Linguistic outcomes of language contact. In: J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 638–668). Malden, MA: Blackwell Publishing.

Sankoff, G., & Blondeau, H. (2007). Language change across the lifespan: /r/ in Montreal French. *Language*, *83*(3), 560–588.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York, NY: Harcourt, Brace and Co.

Schmid, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, *13*(1), 1–7.

Schmidt, A. M. (2007). Cross-language consonant identification: English and Korean. In: O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 185–200). Amsterdam, The Netherlands: John Benjamins Publishing.

Schwartz, R. G. (1995). Effect of familiarity on word duration in children's speech: A preliminary investigation. *Journal of Speech and Hearing Research*, *38*(1), 76–84.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, *19*(3–4), 245–253.

Seliger, H. W., & Vago, R. M. (1991). The study of first language attrition: An overview. In: H. W. Seliger, & R. M. Vago (Eds.), *First language attrition* (pp. 3–15). Cambridge, UK: Cambridge University Press.

Selishchev, A. M. (1925). Des traits linguistiques communes aux langues balkaniques: Un balkanisme ancien en bulgare. *Revue des Études Slaves*, *5*, 38–57.

Silva, D. J. (2006a). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(2), 287–308.

Silva, D. J. (2006b). Variation in voice onset time for Korean stops: A case for recent sound change. *Korean Linguistics*, *13*, 1–16.

Silva, D. J. (2010). Out of one, many: The emergence of world Korean(s). In: *Keynote address delivered at the 17th meeting of the international circle of Korean linguistics*. Ulaanbaatar, Mongolia: University of the Humanities.

Simon, E. (2011). Laryngeal stop systems in contact: Connecting present-day acquisition findings and historical contact hypotheses. *Diachronica*, *28*(2), 225–254.

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, *3*(2), 621–640.

Strange, W., Levy, E., & Lehnholf, R., Jr. (2004). Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. *Journal of the Acoustical Society of America*, *115*(5), 2606.

Svirsky, M. A., Lane, H., Perkell, J. S., & Wozniak, J. (1992). Effects of short-term auditory deprivation on speech production in adult cochlear implant users. *Journal of the Acoustical Society of America*, *92*(3), 1284–1300.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, *423*(6942), 866–869.

Vick, J. C., Lane, H., Perkell, J. S., Matthies, M. L., Gould, J., & Zandipour, M. (2001). Covariation of cochlear implant users' perception and production of vowel contrasts and their identification by listeners with normal hearing. *Journal of Speech, Language, and Hearing Research*, *44*(6), 1257–1267.

Whiteside, S. P. (1998). Identification of a speaker's sex: A study of vowels. *Perceptual and Motor Skills*, *86*(2), 579–584.

Whiteside, S. P., & Irving, C. J. (1998). Speakers' sex differences in voice onset time: A study of isolated word production. *Perceptual and Motor Skills*, *86*(2), 651–654.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, *24*(2), 245–261.

Yates, A. J. (1963). Delayed auditory feedback. *Psychological Bulletin*, *60*(3), 213–232.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and Cognition*, *3*(2), 131–149.

Yusa, N., Nasukawa, K., Koizumi, M., Kim, J., Kimura, N., & Emura, K. (2010). Unexpected effects of the second language on the first. In: K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *New Sounds 2010: Proceedings of the 6th international symposium on the acquisition of second language speech* (pp. 580–584). Poznań, Poland: Adam Mickiewicz University.