Journal of Phonetics 74 (2019) 96–113

Contents lists available at ScienceDirect

Journal of Phonetics

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

Special Issue: Plasticity of Native Phonetic and Phonological Domains, eds. de Leeuw & Celata

Language change and linguistic inquiry in a world of multicompetence: Sustained phonetic drift and its implications for behavioral linguistic research

Charles B. Chang

Boston University, Department of Linguistics, 621 Commonwealth Avenue, Boston, MA 02215, USA

ARTICLE INFO

Article history:
Received 11 January 2018
Received in revised form 11 March 2019
Accepted 13 March 2019
Available online 11 April 2019

Keywords:
Ambient exposure
First language attrition
Plasticity
Crosslinguistic influence
Voice onset time
Fundamental frequency
Vowel formants

ABSTRACT

Linguistic studies focusing on monolinguals have often examined individuals with considerable experience using another language. Results of a methodological review suggest that conflating ostensibly 'multicompetent' individuals with monolinguals is still common practice. A year-long longitudinal study of speech production demonstrates why this practice is problematic. Adult native English speakers recently arrived in Korea showed significant changes in their production of English stops and vowels (in terms of voice onset time, fundamental frequency, and formant frequencies) during Korean classes and continued to show altered English production a year later, months after their last Korean class. Consistent with an Incidental Processing Hypothesis (IPH) concerning the processing of ambient linguistic input, some changes persisted even in speakers who reported limited active use of Korean in their daily life. These patterns thus suggest that the linguistic experience obtained in a foreign language environment induces and then prolongs restructuring of the native language, making the multicompetent native speaker in a foreign language environment unrepresentative of a monolingual in a native language environment. Such restructuring supports the view that one's native language continues to evolve in adulthood, highlighting the need for researchers to be explicit about a population under study and to accordingly control (and describe) language background in a study sample.

1. Introduction

A fundamental goal of linguistic inquiry has been to characterize the complex and largely unconscious knowledge that permits human beings to speak their first language like a native speaker, with little apparent effort. To make inroads on this task, researchers have often abstracted away from a native speaker's knowledge of other languages, with the result that, especially since the advent of the generative grammar movement in the 1950s (Chomsky, 1957), linguistic competence has been analyzed largely in accordance with a monolingual model of the native speaker. This approach is supported by some findings suggesting that bilinguals develop distinct systems for their two languages (Genesee, 1989; Kim, Relkin, Lee, & Hirsch, 1997; Paradis, 2001; Freedman & Barlow, 2012). However, the monolingual model is problematic for two reasons. First, it is inconsistent with the fact that the majority of language users across the world are not actually monolingual (Tucker, 1999). Second, there is abundant evidence that a bilingual's language systems are not completely separate, but rather shared to some degree (Schwanenflugel & Rey, 1986; Fox, 1996; Marian, Spivey, & Hirsch, 2003; Flege, 2007). Since a bilingual cannot be considered the sum total of two monolinguals (Grosjean, 1985; Grosjean, 1989), the investigation of bilinguals' competence in just one language without regard for their competence in the other language amounts to a questionable enterprise. Nevertheless, this remains common practice within the field of linguistics, and the ramifications of this practice for the study of language is the topic of this article.

This paper has three main objectives. The first is to show that the practice of conflating bilingual and monolingual individuals is indeed common in the linguistic literature and is, therefore, an issue that needs to be addressed. The second is to show why this issue does not pertain to fluent, 'balanced' bilinguals specifically, but rather to bilinguals more generally, including marginally bilingual individuals—those who use a second language (L2) much less proficiently and/or frequently than their native or first language (L1). The final objective is to discuss methods of addressing the matter of language

E-mail address: cc@bu.edu
URL: http://cbchang.com

https://doi.org/10.1016/j.wocn.2019.03.001

background in behavioral linguistic research so as to increase the rigor, transparency, and generalizability of empirical findings.

The rest of the paper is divided into five sections. In §2, I establish the premise that bilinguals differ from monolinguals by synthesizing the research on effects of L2 knowledge on the L1, with special attention to phonetic and phonological effects of a late-acquired L2. The argument that, in spite of these effects, ostensibly bilingual individuals are often conflated with monolinguals is developed in §3, which presents a review of population sampling methodology in behavioral linguistic research. The scope of monolingual-bilingual differences—in particular, whether they extend to bilinguals who show limited active use of the L2—is examined in §4, which reports a longitudinal study of L1 production in an L2 environment demonstrating the phonetic plasticity of the L1 in adulthood. In §5, the findings are contextualized within the broader study of lifespan linguistic development, with recommendations for the treatment of language background. Finally, §6 summarizes the main conclusions.

2. Background

2.1. L2 influence on the L1 at multiple levels

Over the previous decades, a growing body of evidence has suggested that the L1 system can be influenced by L2 experience. Clearly, knowledge and use of an L2 is associated with metalinguistic consequences (Yelland, Pollard, & Mercuri, 1993; Jessner, 1999; Bournot-Trites & Tellowitz, 2002) as well as domain-general effects (Cook, 1997; Bialystok & Craik, 2010; Bassetti & Cook, 2011; Kroll, Bobb, & Hoshino, 2014), but the main concern here is with linguistic effects, which often arise from late L2 learning in a variety of linguistic domains (Pavlenko, 2000). For example, L2 influence is observed at the conceptual and cognitive linguistic levels, where it is evident that neither advanced L2 proficiency nor L2 immersion is required for L2 knowledge to result in L1 modifications (Brown, 2008; Brown & Gullberg, 2011; Brown & Gullberg, 2012; for recent reviews, see Jarvis & Pavlenko, 2008; Higby & Obler, 2014). L2 influence is amply documented at the morphosyntactic level as well (Kecskes, 1998; Pavlenko & Jarvis, 2002; Jarvis, 2003; Tsimpli, Sorace, Heycock, & Filiaci, 2004), in certain cases only when the L2 experience is extensive (Dussias & Sagarra, 2007) or early-acquired (Kim, Montrul, & Yoon, 2010) but in other cases just with ambient L2 exposure (Laufer & Baladzhaeva, 2015). Additionally, L2 influence is reported in lexical semantics as well as lexical access and processing (Pavlenko & Jarvis, 2002; van Hell & Dijkstra, 2002; Pavlenko, 2003; Schmid & Köpke, 2009), even after a relatively short period of L2 immersion (Linck, Kroll, & Sunderman, 2009).

Studies of L2 users have led to specific claims about temporal and linguistic constraints on L2-to-L1 influence within the bilingual mind. In regard to temporal constraints, it has been stated, for example, that "L2 users who have been exposed to the L2 for 3 years or longer through intensive interaction in the target language context may start exhibiting bidirectional transfer effects in their two languages" (Pavlenko & Jarvis, 2002, p. 209), while "a L2 that is hardly mastered should not have much influence on L1" (Major, 1992, p. 201); such statements suggest that L2 influence is a phenomenon specific to advanced L2 users. As for linguistic constraints, it has been hypothesized that "changes in L1 syntax will be restricted to the interface with the conceptual/intentional cognitive systems" (Tsimpli et al., 2004, p. 257), reflecting the larger idea that L2 knowledge affects 'fuzzy' aspects of the L1 such as meaning rather than core structural properties such as syntactic parameters. As discussed above, however, while these hypothesized constraints on L2 influence are consistent with some findings, they are not fully supported by this literature, which also includes cases of L2 influence in non-advanced L2 users as well as in structural aspects of the L1 such as inflectional morphology (Jarvis, 2003) and phonemic contrast (de Leeuw, Tusha, & Schmid, 2018).

Positing globally restrictive constraints on L2 influence is especially difficult in light of the abundant—and, in many cases, rapid—L2 effects documented at the level of the sound (Kartushina, Frauenfelder, & Golestani, 2016; Celata, in press; Chang, in press; de Leeuw, in press). With respect to phonological rules and contrasts, relatively extensive L2 experience (in English) is found to alter the production and/or perception of final devoicing in L1 Russian (Dmitrieva, Jongman, & Sereno, 2010), /h/-merger in L1 Korean (Joh, Ko, Lim, & Lee, 2010; Cho & Lee, 2016), and the light-dark lateral contrast in L1 Albanian (de Leeuw et al., 2018). At the phonetic level, L2 experience influences various properties of L1 speech, such as voice onset time (VOT), fundamental frequency (f_0), and the first, second, and third formants (F_1, F_2, F_3). For example, late-onset L2 immersion in English (where voiceless stops have long-lag VOT) leads to lengthened VOTs in the short-lag voiceless stops of L1 French, and vice versa (Flege, 1987); in fact, this VOT shift has been reported even in functional monolinguals with only ambient L2 exposure (Caramazza & Yeni-Komshian, 1974; cf. Fowler et al., 2008). With respect to f_0, L2 experience in Greek influences peak f_0 alignment in L1 Dutch (Mennen, 2004), while L2 experience in English is correlated with higher onset f_0 values following lenis stops in L1 Korean (Yoon, 2015). As for vowel formants, early-onset L2 experience in Spanish is linked to lower F_1 values in L1 Quichua vowels (Guion, 2003), while late-onset L2 immersion in English is linked to higher F_1 values in most L1 Dutch vowels (Mayr, Price, & Mennen, 2012). Late-onset English immersion can also affect the production of laterals and rhotics in L1 German, as indexed by F_1, F_2, and F_3 (de Leeuw, Mennen, & Scobbie, 2013; Ulbrich & Ordin, 2014; Bergmann, Nota, Sprenger, & Schmid, 2016).

Direct evidence of L1 change due to L2 influence has also been provided by a number of longitudinal studies. For instance, in a one-year longitudinal study, Oh et al. (2011) show that L2 immersion in English results in increased F_2 values for some L1 Japanese vowels, although only in children and not in adults. Additional longitudinal data come from the L2 training literature (e.g., Kartushina, 2015; Kartushina, Hervais-Adelman, Frauenfelder, & Golestani, 2016), as well as a case study of an L1 Portuguese late learner of English (Sancier & Fowler, 1997; for related research on Spanish–English bilinguals, see Tobin, Nam, & Fowler, 2017). In the latter study, VOT in both L1 and L2 voiceless stops is found to be influenced by the VOT norms of the most recently experienced ambient language; thus, short-lag Portuguese stops are produced with significantly longer VOTs following a few months of immersion in English, an effect that is perceptible to native Portuguese listeners. The fact that this speaker's L1 production is detectably affected by recent L2 immersion despite her greater total experience in the L1 is attributed to three factors: crosslinguistic phonological similarity (which leads to L1 sounds becoming perceptually linked to, and thus influenced by, similar L2 sounds), a tendency toward imitation (even of L2 exposure; see, e.g., Ward, Sundara, Conboy, & Kuhl, 2009), and the recency effect on memory.

Recent L2 experience, however, has an effect that is modulated by learners' prior familiarity with the L2, as shown in longitudinal work on L1 English learners of Korean (Chang, 2012; Chang, 2013; for related research on L1 Mandarin learners of Korean, see Holliday, 2015). Korean is a language that, unlike English, has a three-way stop laryngeal contrast distinguished in terms of VOT and f_0 (Yoon, 2015; Bang, Sonderegger, Kang, Clayards, & Yoon, 2018). In Chang's results, both of these properties in the L1 show evidence of change due to recent L2 experience (PHONETIC DRIFT) within the first five weeks of L2 instruction in an immersion environment. Drift is found in the VOT of English voiceless stops (which lengthens in approximation to the longer VOT of the perceptually similar Korean aspirated stops), onset f_0 (which increases due to the f_0 elevation associated with Korean fortis and aspirated stops), and mean F_1 of the vowel system (which decreases due to the lower mean F_1 of the Korean vowel system). Notably, regardless of the acoustic property examined, the magnitude of drift is found to be larger in inexperienced learners (true beginners) than in experienced learners with prior exposure to Korean, suggesting that phonetic drift due to L2 learning decreases over the course of L2 development. Although there are individual differences in phonetic drift (see, e.g., Huffman & Schuhmann, 2015), the same general phenomenon is found in foreign language (i.e., non-immersion) contexts (Herd, Walden, Knight, & Alexander, 2015; Schuhmann & Huffman, 2015). Furthermore, recent work on phonetic drift in perception has demonstrated that, like L1 production, L1 perception can undergo rapid shifts during elementary L2 learning as well (Tice & Woodley, 2012; see also Namjoshi et al., 2015).

Thus, while there is some evidence that L2 learners can pattern like monolinguals in their L1, the bulk of the literature suggests that L2 experience tends to influence L1 performance, regardless of when the L2 was learned. At the level of syntax and semantics, there is some indication that L2 influence may be strongest with an early onset and/or high level of L2 experience. At the phonetic level, by contrast, effects of recent L2 experience are commonly found in late L2 learners and may be strongest at low levels of prior L2 experience. The occurrence of crosslinguistic phonetic interaction in late L2 learners is, in fact, predicted by three core principles of the Speech Learning Model (SLM; Flege, 1995; Flege, 1996; Flege, 2002): (1) that sound categories continue to develop over the lifespan; (2) that the sounds of an L1 and L2 exist in a shared mental phonetic space; and (3) that 'similar' (as opposed to 'new') L2 sounds tend to undergo a perceptually-based, automatic 'equivalence classification' with L1 sounds, resulting in a merging of their phonetic properties. According to the SLM, equivalence classification of L2 sounds with L1 sounds becomes more likely as L1 categories evolve over the course of normal L1 development (Flege, 1995), so the probability of perceptual linkage between L1 and L2 sounds increases with a late onset of L2 learning. Adult L2 learners are thus particularly subject to L2-to-L1 influence, because they have the greatest tendency to link L2 sounds to L1 sounds rather than creating distinct L2 categories.

2.2. Multicompetence and the notion of 'native speaker'

The broad susceptibility of the L1 to L2 influence in adulthood is consistent with a view of linguistic knowledge as fluid and holistic, encapsulated in Cook's notion of MULTICOMPETENCE (Cook, 1991; Cook, 1992; Cook, 2003; see also Kecskes & Papp, 2000). In a multicompetence framework, the acquisition of additional languages is conceptualized not as mere accrual, but instead as restructuring of knowledge, a process that changes the language user fundamentally. Thus, compared to unicompetent (monolingual) language users, multicompetent users have not only *more* knowledge (of L2, L3, etc.), but also *different* knowledge overall, including of the L1. As such, it is unexceptional—and actually expected—for an L1 to be perceived or produced differently by L2 users compared to monolinguals, who represent a mental state prior to 'initiation' into L2 knowledge.

In contradicting the idea of language separation within the multilingual mind, the multicompetence framework also problematizes the term NATIVE SPEAKER, an ambiguous descriptor for a linguistic profile that may or may not correspond to 'monolingual' (see, e.g., Beinhoff, 2008). As a model for L2 learners, native speakers are commonly thought of as individuals who have attained 'full' command of the target variety, the type of command that is often the object of description in linguistic research. In practice, however, the native (*qua* the most proficient) speakers of a language are rarely identified as such by proficiency measures, but rather by proxy measures (e.g., age/onset of learning) or by self-report, which may be based on the same proxy measures (e.g., being exposed to the language from birth).

Building a sample of native speakers via proxy measures, while expedient, may not pick out the intended population of native users—those who have "special control" and "insider knowledge" of the language, who "control its maintenance and shape its direction" (Davies, 2003, p. 1)—because most proxy measures do not account for the fact that language knowledge is dynamic and, consequently, subject to change as well as loss (de Bot, Gommans, & Rossing, 1991; Seliger & Vago, 1991; Stolberg & Münch, 2010; Schmid, 2013). That is to say, native-like command at one point in time does not necessarily imply native-like command at the time of study. For this reason, the target population in linguistic research is often *monolingual* native speakers, since monolinguals should be exempt from the transformative influences of L2 knowledge, and this monolingual model is the one typically assumed in research on one language (see, e.g., Chomsky, 1986; Piller, 2002).

2.3. Monolingual studies recruiting monolinguals?

Despite the commonness of the monolingual model in linguistic research, the prevalence of multicompetence raises two important questions for this model. The first question is whether this model makes sense for the language under study. If, for example, the language is not generally spoken by monolinguals, but by multilinguals, is it reasonable to examine knowledge of that one language without considering the other languages within the user's linguistic repertoire (see, e.g., Lüpke & Storch, 2013)? This paper does not address this question, although it should be noted that the answer to this question might very well be no, depending on whether the empirical simplification imposed by the monolingual model stands to produce misleading results (e.g., where a study sample is composed of users whose profile of multicompetence is not typical for users of that language).

Assuming that the monolingual model does make sense for the language under study (because it is in fact spoken largely by monolinguals), the second question is whether the monolingual model is being applied appropriately. That is, does a study of 'native' (*qua* monolingual) knowledge of a given language actually examine monolingual users? Discrepancies in the linguistic literature suggest that the answer to this question may often be No. For example, the phonology of Swedish has been described inconsistently as contrasting either voiced and voiceless aspirated stops (Helgason & Ringen, 2008) or voiceless unaspirated and voiceless aspirated stops (Keating, Linker, & Huffman, 1983), which is attributable to the latter study's examination of speakers immersed in L2 English.[1] In other words, the two studies describe Swedish phonology differently because although they target the same population (namely, 'native' Swedish speakers), one examines functional monolinguals in a Swedish-speaking environment, whereas the other examines multicompetent speakers in an English-speaking environment.

This type of methodological disparity reflects the assumption of an unchanging L1—that is, the idea that L1 users maintain the same L1 competence regardless of variation in language background and environment. Given that this assumption is questionable (see §2.1), it should generally be rejected; however, if it is rejected while the monolingual model is maintained, then recruiting the target demographic for a behavioral study of one language (i.e., monolinguals) requires understanding language users not just in terms of their self-identified L1, but in terms of their broader language background and environment. This is because these latter variables, which both affect the linguistic behavior on which the study's conclusions are based, cannot be presumed to be the same (in particular, monolingual) across so-called 'native speakers'.

To my knowledge, the extent to which the field of linguistics, including the subfield of phonetics, has addressed the potentially problematic confluence of the monolingual model and the assumption of an unchanging L1 has not been examined systematically, which leads to the following question: to what degree does contemporary linguistic research adopting the monolingual model in fact reject the assumption of an unchanging L1? That is, do monolingual studies reliably distinguish between monolingual and multicompetent language users? To address this question, which has implications for both the interpretation and the replication of empirical findings on language, a methodological review was conducted of recent behavioral linguistic studies focusing on monolingual populations.

3. Population sampling in behavioral linguistic research: a review

3.1. Methods

The basis for the methodological review was a corpus of linguistic studies constructed from recent publications in high-impact journals. Given that the primary concern was with behavioral research meant to generalize to populations of mature, functionally monolingual native speakers, the journals included were the top two linguistics journals according to Google Scholar's 2014 *h*5-index ('the largest number *h* such that *h* articles published in [the preceding five years] have at least *h* citations each') that primarily publish behavioral studies (i.e., studies that require participants) directly related to language and whose focus is neither on children nor on topics related to multilingualism (e.g., L2 learners, bilingualism, language contact). These journals were the *Journal of Phonetics* (*JPhon*) and *Language and Cognitive Processes* (*LgCog*; as of 2014, *Language, Cognition and Neuroscience*). To further limit the scope, the review focused on a target time period consisting of the first three years of the current decade (i.e., 2011–2013).

The final corpus comprised all, and only, studies that were ostensibly meant to generalize to the population of adult monolingual native speakers of the subject language.[2] All 363 articles published in *JPhon* and *LgCog* during 2011–2013 were reviewed to determine whether the article met any of three exclusion criteria: (1) reporting no novel adult data (e.g., child-focused or computational studies), (2) focusing on multilingualism, broadly construed, and including no monolingual control group, or (3) stating explicitly that the findings might not generalize to monolinguals. Every article which did not meet an exclusion criterion was included in the corpus, for a total of 286 studies (127 from *JPhon*, 159 from *LgCog*).

Data for each study were collected by consulting the abstract, introduction, methods, and discussion sections of the paper, as well as the institutional affiliations in the byline. To obtain a realistic picture of the potential for misinterpretation and/or overgeneralization of findings, these data were compiled with two simplifying assumptions. First, it was assumed, unless information reported on the participants suggested otherwise, that participants recruited from the study location had knowledge of the language(s) dominant in that region, where they were presumed to have been living for a significant amount of time. For example, Spanish speakers recruited from the US were assumed to have knowledge of American English

[1] Since 'the recordings for the experiment of Keating et al. (1983) were made in the US (Keating, p.c.)' (Helgason & Ringen, 2008, p. 620), most likely at UCLA, the speakers' place of residence was presumably somewhere in southern California, although the paper does not specify where the speakers were living at the time of the study. However, given the phonetic implementation of 'voiced' stops in American English as voiceless unaspirated, it follows that Swedish speakers influenced by American English would produce Swedish voiced stops as voiceless unaspirated. Along the same lines, the growing population of proficient L2 speakers of English in Sweden might introduce further variation into results obtained on 'native Swedish'.

[2] To clarify what is meant by 'language' in this context, this term is being used broadly to refer to varieties that may be called 'languages' or 'dialects' by different researchers, meaning that studies focusing on so-called 'dialects' were eligible for inclusion in the corpus as long as the target population was not bidialectal speakers.

unless it was specified that they were monolingual. Second, if it was not stated explicitly where the study took place, it was assumed that the study location corresponded to the institutional affiliation in the byline. In the case of multiple affiliations, the study location was taken to be the one with the highest number of speakers of the subject language according to Ethnologue (Lewis, Simons, & Fennig, 2015).

3.2. Results

The corpus (publicly accessible at https://osf.io/u7864/) was diverse in terms of topic areas, methodologies, and subject languages. Given the target journals, the corpus included research mainly in phonetics and psycholinguistics; however, the psycholinguistic studies addressed questions related to virtually all areas of linguistics (e.g., lexicon, phonology, syntax, semantics, pragmatics, reading, gesture). The vast majority (94%) of studies involved laboratory-based experimental work, but there were also examples of archival, ethnographic, corpus-based, and web-based studies. As for subject languages, the corpus was skewed toward research on varieties of English (52% of studies), but a wide range of other languages was represented as well (e.g., Berber, Central Arrernte, Hmong, Sign Language of the Netherlands).

Analysis of this corpus revealed recurring information gaps regarding the language background of participants who were ostensibly monolingual native users (speakers or signers) of a particular language. It was much more common for participants' language background to be left unclear (80% of studies) than described unambiguously as monolingual (10% of studies). The many cases of ambiguity stemmed from the fact that when researchers did not describe participants as monolingual, they included no information regarding knowledge of other languages. Surprisingly, it was also not uncommon for there to be no clear description of participants as L1 users. For example, in 15% of studies, participants were never described as 'native' or their native/first language was left unspecified. This is noteworthy because nearly all of these studies involved linguistic tasks.

Although it was most common for studies to be ambiguous about the language background of participants, a considerable number expressly examined participants who were multicompetent. Approximately 13% of studies included a native group that regularly used or had significant knowledge of one or more additional languages. Since none of these studies offered a reason for examining multicompetent rather than monolingual native users, the language situation in each of the relevant speech communities was investigated to determine whether the use of a multicompetent sample was consistent with the characteristics of the contemporary population of language users. This investigation revealed that the multicompetence of nine user samples (Blackfoot, Catalan, Central Arrernte, Dutch, Gujarati, Hindi, Oneida, Q'eqch'i, Trique) could be considered representative of the norm in the speech community. However, these nine groups accounted for only a minority of the aforementioned 13% of studies; discounting these cases still left 10% of studies which examined multicompetent native users for no apparent reason. These results thus suggest that about 1 in 10 studies targeting monolinguals actually examines multicompetent users, although note that by collecting and reporting information about language background (see §5.3) these studies are transparent about this, thus allowing the results to be properly interpreted.

For the majority of studies, data collection sites were appropriate given the study's aims and subject language(s); nevertheless, the practice of collecting data from individuals living in a foreign language environment was found in 15% of studies. These studies rarely addressed the implications of the language environment: in only two cases was the foreign language environment acknowledged as a limitation, and in only one was a reason provided for collecting data in this environment (logistical constraints). If the latter study is excluded along with one other study that was presumably constrained by the availability of necessary equipment (an electromagnetic articulometer), this leads to the conclusion that about 1 in 7 studies of monolingual language performance is conducted in a nonnative environment (namely, the researchers' location), meaning that the so-called 'monolinguals' are probably exposed to, and possibly proficient in, another language that has potentially influenced the target language examined in the study.

3.3. Discussion of methodological review

In sum, a review of recent publications on adult monolingual language performance revealed two problematic aspects of the literature in this area: (1) insufficient description of study samples, and (2) disparities between study samples and target populations. Across a range of topic areas and subject languages, researchers were found to neglect describing participants' language background—sometimes even failing to identify their L1—making it difficult to tell precisely what kind of individuals made up a study sample. Moreover, when information about language background was provided, the given study sample often did not match the target population (i.e., monolingual users in the native language environment) due to multicompetence and/or residence in a foreign language environment. These characteristics of the sample were never required by the research questions and rarely discussed in terms of their ramifications for the results.

Overall, these results present a troubling view of the interpretability (and replicability) of published behavioral linguistic studies, but it should be noted that there were exceptions to the patterns described above. First, some studies examined samples that clearly matched the populations they were meant to represent (e.g., French monolinguals in France: Abdelli-Beruh, 2012; Greek monolinguals in Greece and Australian English monolinguals in Australia: Antoniou, Tyler, & Best, 2012), even if these studies were in the minority. Second, although they did not provide a holistic picture of language background that would make replication straightforward, some studies did provide details beyond native language, which were generally framed in terms of limitations on knowledge of other languages. A few commented on aspects relevant to the materials used (e.g., no experience with a crucial L2; no experience with lexical tone; no experience with vowel/consonant harmony), while others highlighted general restrictions on experience with additional languages, particularly with the ambient language in a nonnative language environment (e.g., low proficiency; late onset of learning).

In regard to studies of speakers in a nonnative language environment, one pattern that stood out was an emphasis on their short length of residence (LoR) in the L2 environment. This pattern was puzzling because different studies assumed different notional thresholds (ranging from three months to two years) for the maximum LoR participants could report and still be considered monolingual-like in their L1. None of the thresholds, however, were justified explicitly, suggesting that they were either arbitrary or conventional. If the most frequently observed LoR threshold (namely, two years) is in fact convention, this raises the question of whether this is the right convention for recruiting monolingual-like participants in an L2 environment. Studies of L1 attrition have generally examined long-term migrants, those who have been residing in an L2 environment for many years (de Bot et al., 1991; de Leeuw, Schmid, & Mennen, 2010; Schmid, 2013); there are no known studies providing evidence for two years as a valid LoR threshold.[3]

Contrary to the view that it takes two years for an L1 speaker to show evidence of L2 influence in their L1, findings on phonetic drift suggest that L2-influenced modifications to the L1 can occur within weeks of L2 exposure (Chang, 2012; Chang, 2013). However, because these L1 modifications were observed during a period of formal L2 instruction, which may be special in terms of relative L2 activation and/or L1 inhibition, it is possible that rapid L2 influence might be limited to the situation of intense L2 engagement. When learners are removed from this situation, such that active L2 use goes down and active L1 use goes back up, does their L1 drift back to monolingual norms, or does it continue to diverge from these norms due to continued ambient exposure to a now-familiar L2? This question provided the motivation for a longer longitudinal study of L1 users living abroad which could address the persistence of L1 drift after the end of L2 instruction.

4. Persistence of phonetic drift: an acoustic study

In light of the wide variation in assumptions regarding temporal constraints on L1 restructuring, the present study investigated the time course of L1 phonetic restructuring due to recent L2 experience (PHONETIC DRIFT, or 'drift' for short). The point of departure was the phonetic drift observed in Chang (2012, 2013). On the one hand, because the drift in those studies was found during a period of intensive L2 instruction immersing learners in the L2 for more than 30 h/week, it may have been due specifically to high L2 engagement associated with *learning*, which would imply that the L1 should drift back to monolingual norms once active L2 learning ends; on the other hand, if the crucial ingredient for prolonging drift is L2 *exposure*, then the discontinuation of active L2 learning should fail to fully reverse drift as long as learners are living in an L2 environment.

In the present study, the latter outcome was predicted on the basis of an INCIDENTAL PROCESSING HYPOTHESIS (IPH), which addresses the degree to which ambient linguistic input (i.e., input that is in the environment but not directed at the listener) may be *incidentally processed*, rather than ignored. The IPH posits that ambient linguistic input becomes increasingly difficult to ignore as one's knowledge of that language increases, consistent with the finding that auditory stimuli are more distracting when they are informative as opposed to uninformative (Parmentier, Elsley, & Ljungberg, 2010). The logic underlying the IPH is that, whereas ambient input in an as-yet unknown (therefore, uninformative) L2 may be treated as 'noise', thus avoiding at least deep linguistic processing, acquiring a 'critical mass' of L2 knowledge (e.g., a sizable lexicon, phonological categories) leads to ambient L2 input becoming potentially informative, such that it tends to be processed as a linguistic stimulus activating linguistic representations. Crucially, this means that ambient input in a *known* L2 is relatively likely to undergo some degree of processing, even if it is not actively attended to.

Thus, even after the end of L2 instruction, the learners from Chang (2012) and Chang (2013) were predicted to continue showing drift while in the L2 environment because incidental processing of ambient L2 input would maintain a high activation level of the L2. This prediction was tested by analyzing the L1 (English) speech production of a subset of the participants in Chang (2012, 2013) one year after the initial period of L2 (Korean) instruction. The dependent measures were the same: VOT, onset f_0 in the vowel following a stop, F_1, and F_2. Given the previous findings, there were four specific predictions regarding sustained drift toward phonetic norms of Korean (summarized in Chang, 2012, pp. 253–4):

(1) Since the initial drift in VOT of English voiceless stops had been pronounced (on the order of 20 ms in novice learners), drift in VOT was predicted to persist, resulting in longer-than-baseline VOT for voiceless stops (cf. VOT norms for Korean aspirated stops, about 28–39 ms longer than for English voiceless stops).
(2) Since the initial drift in onset f_0 following English stops had also been pronounced, drift in onset f_0 was predicted to persist as well, resulting in higher-than-baseline f_0 for both voiced and voiceless stops (cf. f_0 norms following Korean fortis and aspirated stops, estimated to be at least 10–15 Hz higher than following English stops).
(3) Since initial drift in F_1 and F_2 of the English vowel system had been subtle (F_1) or not significant (F_2), drift in vowel formants was not predicted to persist.
(4) Since active use of the L2 might encourage the persistence of L2 influence, it was predicted that the remnants of phonetic drift a year later would be more obvious for L2 learners who continued to speak the L2 frequently compared to those who spoke less frequently.

To address prediction (4) in particular, learners were analyzed in two groups differing in frequency of active L2 use.

4.1. Methods

4.1.1. Participants

A total of 36 L1 speakers of American English entered and completed the initial five-week study reported in Chang (2012, 2013). All were recent college graduates who had traveled to South Korea to teach English. They were invited to participate in an additional study session approximately one year after their arrival to Korea, and 17 elected to participate in this session. Two of these 17 participants reported problems with

[3] Although Pavlenko and Jarvis (2002) report L2 influence in participants with an LoR of at least three years, their results cannot be interpreted as evidence for an LoR threshold of three years because that is the shortest LoR included in that study (Pavlenko & Jarvis, 2002, pp. 193–4).

hearing and/or speech in early childhood, which they perceived as having been remedied with speech therapy; removing them from the dataset did not affect the results, so they are included in all results reported below.

Self-identified native speakers of English, the 17 participants analyzed here were raised primarily in the US and identified English as their strongest language and at least one of the languages used at home. Two participants also spoke a heritage language (Mandarin in one case, Russian in the other), which they coidentified with English as a native language. The other 15 participants identified only English as a native language and reported speaking only English at home. All participants had previously studied at least one foreign language (most often Spanish or French) for a period of 1–13 years; however, only one reported significant communicative use of a foreign language (Japanese), which was often the language of e-chat with friends. Thus, the majority (14/17) of participants were "functionally monolingual" L1 English speakers (in the sense of Best & Tyler, 2007, p. 16: "not actively learning or using an L2"), while the other three were bilingual L1 English speakers.

Based on data from a detailed questionnaire about their year in Korea (publicly accessible at https://osf.io/d5qzj/), participants were assigned to one of two groups according to whether they showed low active use (LU) or high active use (HU) of the L2 after the initial five-week study. This was done by ordering the sample by self-reported L2 speaking frequency and splitting it evenly into two groups. Nine participants (mean age 24.4 yr, *SD* 1.9; eight female) were thus assigned to the LU group, and eight (mean age 23.6 yr, *SD* 0.7; seven female) to the HU group. The groups did not differ significantly in terms of age [Welch-corrected two-sample $t(10.7) = 1.207, p = .254$] although the group division put all three bilinguals into the HU group. In addition to being closely matched on age and gender, the two groups were similar with respect to several variables related to use of English during the year: frequency of personal interactions with native English speakers in Korea, frequency of phone/e-chat interactions with native English speakers in the US, and time spent away from Korea in an English-speaking country, none of which differed significantly between groups [all $ps > .05$].

The principal difference between the LU and HU groups was in the nature of their experience using Korean over the year. Whereas HU participants described using Korean at home and/or work, LU participants described using mostly English both at home and at work. As a result, the LU group reported spending much less time speaking Korean (M_{LU} = 2.2 h/wk, M_{HU} = 13.8 h/wk; $t(7.5) = -5.679, p < .001$) whereas they heard Korean around them much of the time, just as the HU group did (M_{LU} = 34.1 h/wk, M_{HU} = 49.6 h/wk; $t(14.9) = -1.066, p = .303$). The LU group's limited active use of Korean was further reflected in lower self-ratings of Korean proficiency across a range of communicative tasks (M_{LU} = 2.1/6 ≈ 'poor', M_{HU} = 2.8/6 ≈ 'fair'; $t(8.8) = -3.247, p = .010$).

4.1.2. Learning context

In the initial five-week study, participants were enrolled in a Korean language program at a Korean university. Prior to beginning this program, most LU participants had received no significant exposure to Korean, so they were enrolled in an elementary-level class; the two exceptions had taken Korean in college and were enrolled in an intermediate-level class. Most HU participants were also enrolled in an elementary-level class, with two enrolled in an intermediate-level class as in the LU group. Despite the similarity in their enrollments, however, the HU group was the mirror image of the LU group in terms of experience: most HU participants had received significant prior exposure to Korean, by virtue of having been adopted from Korea ($n = 3$; mean age of adoption 0;11) and/or having studied Korean in college ($n = 4$; total class contact hours ranging from 60 to 600).

Both elementary- and intermediate-level classes in the language program followed the same intensive schedule over the six-week duration of the program. On most weekdays, there were four hours of instruction, for a total of more than 80 class contact hours by the end of the program (roughly equivalent in content to one semester of college-level Korean). Classes were conducted in Korean, and participants lived on campus during the program; however, they stayed in a dormitory with their fellow students, who were all native English speakers as well. Consequently, the type of L2 learning environment provided in this program might best be described as in between typical second language acquisition (in which learners acquire the L2 naturalistically in an L2 environment) and typical foreign language acquisition (in which learners study the L2 formally in an L1 environment).

Following the end of the language program, participants began working as English teachers in various host locations, where most (eight LU and six HU participants) reported receiving additional Korean instruction in the form of classes and/or one-on-one tutoring. The amount of this instruction was similar between groups (M_{LU} = 45 h, M_{HU} = 47 h), and participants reported spending little time on self-regulated Korean study (M_{LU} = 22 min/wk, M_{HU} = 32 min/wk). Crucially, additional Korean instruction tended to occur early in the year, such that, on average, more than three months had elapsed between participants' most recent Korean class and the final study session (M_{LU} = 3.7 mo, M_{HU} = 5.1 mo).

4.1.3. Procedure

In the initial five-week study, participants completed two production experiments (one in English, one in Korean) at the end of each of the first five weeks of their language program, generally in a quiet room in their dormitory. Instructions were provided in English, and the experiments were usually completed in one session (in the order of Korean followed by English, with an intervening break). The task was isolated word reading: participants were shown a target item, spelled in the target language orthography, on screen and asked to say the item out loud upon seeing a subsequent visual cue. This task was meant to elicit a relatively formal register providing a strong test of L2 influence on L1 speech, as formal registers have been shown to resist L2 influence in comparison to more casual registers (Major, 1992). The experiments were administered on a Sony Vaio PCG-TR5L laptop computer running DMDX (Forster, 2014). In both experiments, items were randomized and presented once in each of four blocks, such that four tokens were collected of each item. Recordings were made at 44.1 kHz and 16 bps using an AKG C420 or C520 head-mounted condenser microphone, connected either to the computer via an M-AUDIO USB preamp or to a Marantz PMD660 recorder.

In the additional study session that took place a year later, participants completed the two production experiments one more time in a quiet office in Seoul. All other aspects of the procedure, materials, equipment, and recording specifications were the same as in the initial five-week study.

4.1.4. Materials

The speech materials for the English production experiment consisted of 24 monosyllabic English words: 16 critical and 8 filler items. Six critical items were used to measure VOT of stops and onset f_0; these items contained the same vowel /ɑ/ to control for the effect of vowel environment on VOT and facilitate comparison with VOT norms based on similar contexts (Morris, McCrea, & Herring, 2008). Eleven critical items (including one item used to measure VOT/f_0) were used to measure F_1 and F_2 of vowels; these items began with /h/ or an otherwise aspirated onset to control for coarticulatory perturbations from an initial consonant and facilitate comparison with formant norms based on similar contexts (Hagiwara, 1997; Hillenbrand, Getty, Clark, & Wheeler, 1995; Peterson & Barney, 1952; Yang, 1996). The full set of critical items, the same in every iteration of the experiment, is shown in Table 1.

The speech materials for the Korean production experiment consisted of 22 monosyllabic Korean items: 15 critical items and 7 fillers. All critical items consisted of an open syllable comprising one consonant and one vowel. The items used to measure VOT of stops and onset f_0 comprised a stop (one of /p p* pʰ t t* tʰ k k* kʰ/[4]) followed by /a/, while those used to measure vowel formants comprised a vowel (one of /i ɨ u ɛ o ʌ a/) preceded by /h/ (or an otherwise aspirated onset).

4.1.5. Acoustic analysis

The four acoustic measures were VOT in word-initial stops, onset f_0 in the following vowel, and F_1 and F_2 at vowel midpoint. All measurements were taken in Praat (Boersma & Weenink, 2016) on the waveform or a wide-band Fourier spectrogram with a Gaussian window shape (window length: 5 ms, dynamic range: 50 dB, pre-emphasis: 6.0 dB/oct).

The measures related to stops were VOT and onset f_0. VOT was calculated by subtracting the time at the beginning of the release burst interval from the time at voicing onset (the first point at which a voicing bar with clear glottal striations appeared in the spectrogram). Onset f_0 was calculated by taking the combined wavelength of the first three regular glottal periods in the vowel and converting to Hertz (Hz). The interval of three periods was demarcated on the waveform, with an initial period being skipped if it was more than 33% longer or shorter than the following period. Tokens in which the earliest interval of three regular periods occurred more than five periods into the vowel were considered to have an irregularly phonated vowel onset and were thus discarded.[5]

Table 1
Critical items used in the L1 (English) production experiment, by dependent measure.

Measures	Items
VOT, f_0	*bot*, *pot*, *dot*, *tot*, *got*, *cot*
F_1, F_2	*heed*, *hid*, *hate*, *head*, *had*, *who'd*, *hood*, *hoed*, *hut*, *hawk*, *pot*

The measures related to vowel quality were F_1 and F_2. Both formants were measured automatically over the middle 50 ms of a vowel interval, which was demarcated manually at the first and last glottal striations showing formant structure in the spectrogram. The analysis method was linear predictive coding, using the Burg algorithm (Childers, 1978) in Praat. Parameters for the formant analysis (frequency range, number of formants) were determined by visually inspecting a few spectrograms from the given participant and adjusting the defaults until formant tracking was smooth and closely followed the formants visible in the spectrogram. To further check the accuracy of the formant measurements, they were inspected for outliers by vowel, potential errors were flagged, and spectrograms of all tokens were inspected individually. When formant tracking was inaccurate, the analysis parameters were adjusted; if this did not fix the tracking, then measurements were taken manually on an average spectrum of the middle 50 ms of the vowel.[6]

Intra-rater reliability was examined via Pearson's correlations, which 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. This second round of measurements was closely correlated with the first round for all measures [$r = .92$ to $r = .98, ps < .001$]. The average difference between paired VOT measurements was 3 ms; between paired f_0 measurements, 4 Hz; between paired F_1 measurements, 7 Hz; and between paired F_2 measurements, 15 Hz.

4.1.6. Statistical analysis

Prior to statistical analysis, the acoustic data were reorganized in two ways to achieve a valid comparison of values across the LU and HU groups. First, stop tokens were binned into three phonetic categories of stop voicing ('prevoiced', 'short-lag', 'long-lag') according to VOT boundaries estimated from the literature (Keating, 1984; Lisker & Abramson, 1964; Lisker, Liberman, Erickson, Dechovitz, & Mandler, 1977), with the most common phonetic voicing category for each stop type submitted to statistical analysis: short-lag (VOT of 0–30 ms) for English voiced (and Korean fortis) stops, and long-lag (VOT > 30 ms) for English voiceless (and Korean aspirated) stops.[7] Second, frequency values (f_0, F_1, F_2) were standardized by participant, by calculating the participant's mean for the given frequency component during the initial five-week study and then expressing each of the participant's raw values for that

[4] The Korean stops are indicated here using conventional transcriptions for Korean laryngeal categories. Note, however, that in Chang (2012, 2013), these stops are transcribed with the extended IPA diacritics for weak and strong articulations as, respectively, /p͉ p͈ pʰ t͉ t͈ tʰ k͉ k͈ kʰ/.

[5] A total of 1.2% of English tokens and 2.1% of Korean tokens were discarded for this reason or because of other pronunciation anomalies such as coughing.

[6] A total of 1.3% of English tokens and 0.7% of Korean tokens were discarded because of pronunciation anomalies or speech errors.

[7] Stop tokens were divided in this way primarily to exclude prevoicing (VOT < 0 ms) from the analysis of voiced stops. Although relatively infrequent, prevoiced tokens represented a different phonetic voicing category than short-lag tokens, so to obtain a clear picture of within-group change and between-group differences in voiced stop production—one that did not simply reflect change in the frequency or robustness of closure voicing—the analysis of voiced stops was limited to tokens representing their typical short-lag realization. In the interest of consistency, the analysis of voiceless stops was also limited to tokens representing their typical long-lag realization (resulting in hardly any exclusions because nearly all tokens of voiceless stops had VOT longer than 30 ms).

frequency component as *z*-scores about the mean. This standardization allowed for longitudinal analyses within individuals that could be compared across individuals (of both genders) on the same scale.

The acoustic data were then modeled with mixed-effects linear regression using the `lme()` function in R (R Development Core Team, 2018). All of the final models on the full dataset contained a random intercept for Participant (no random slopes because these usually caused a model to fail to converge) and two treatment-coded fixed effects: Time (i.e., weeks after the start of the initial Korean language program: 1–5, 52; reference level = 1) and Group (LU, HU; reference level = LU). In addition, models of stop-related measures (VOT, f_0) included a deviation-coded (meaning the contrast estimate is against the grand mean, rather than the reference level) fixed effect for Place (of articulation: bilabial, velar, alveolar), while models of vowel-related measures (F_1, F_2) included a deviation-coded fixed effect for Vowel (/iɪeɛæuʊoʌɑɔ/), since these factors have been shown to influence the given dependent variables (Ladefoged, 2005; Nearey & Rochet, 1994). All models further included a deviation-coded fixed effect for Block (in the experiment: 1–4), representing the potential influence of fatigue or practice, and all possible interactions among predictors. Block and interactions with Block did not have a significant effect in any model and are thus not discussed further below.

Due to the unbalanced distribution of talker gender and the uncertain status of the /ɑ/-/ɔ/ merger in participants' vowel systems, models of the frequency measures underwent additional scrutiny to check that the results were robust. In regard to gender, this factor was observed to have a significant effect on change in frequency measures in Chang (2012) and Chang (2013); for example, females showed a significant change in f_0, but males did not. However, gender could not be entered into the models in the current study because there was only one male participant in each group. Thus, two models of each of the frequency measures were compared: one built on the full dataset, and one built on female-only data. In regard to the vowel /ɔ/, most participants (even those whose native dialect purportedly shows an /ɑ/-/ɔ/ merger) did not show a clear merger of /ɔ/ with /ɑ/. Thus, two models of each formant measure were compared: one including both /ɑ/ and /ɔ/ (i.e., assuming that participants maintained a contrast between these vowels), and one including only /ɑ/. Both cases of model comparison showed no significant effect of the relevant data exclusion on the results, so the models reported below are those built on the full dataset without exclusions. All data (including both the acoustic data collected in the production experiments and the background data coded from questionnaires) are publicly accessible at https://osf.io/u7864/.

4.2. Results

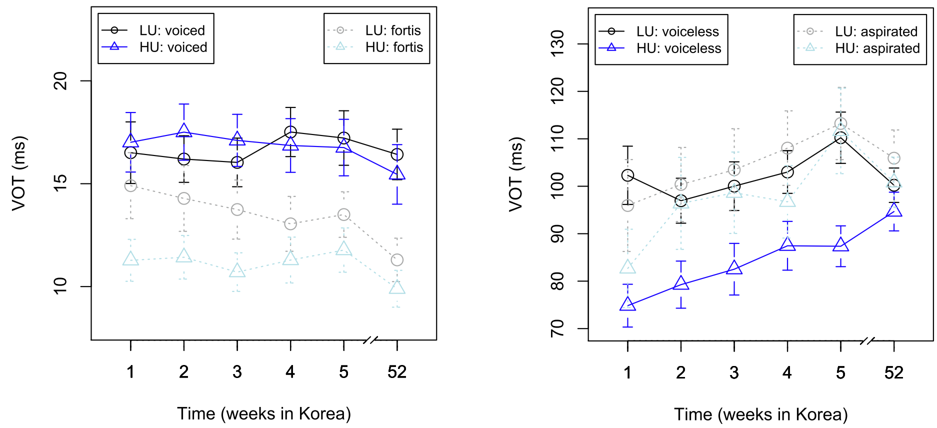
4.2.1. Phonetic drift in VOT

Consistent with the findings of Chang (2012, 2013), the VOT of English voiced stops did not drift significantly in the LU or HU group (see Fig. 1a). The model of VOT in voiced stops is shown in the Supplementary material (Table 2), which also provides model summaries for all of the main models discussed below. An analysis of variance (ANOVA) on the model of VOT in voiced stops (using `Anova()` in the `car` package in R; Fox et al., 2018) revealed a significant effect of Place [$\chi^2(2) = 538.789, p < .0001$], which was due to bilabials having shorter-than-average VOT [$\beta = -3.944, t = -5.575, p < .0001$] and velars having longer-than-average VOT [$\beta = 4.346, t = 5.901, p < .0001$]. However, there was no effect of Time [$\chi^2(5) = 7.987, p = .157$] or Group [$\chi^2(1) = 0.009, p = .923$]. Furthermore, no interactions were significant, including the Time × Group interaction [$\chi^2(5) = 7.498, p = .186$]. As a point of comparison, Fig. 1a plots the VOT of the perceptually similar Korean fortis stops as well, showing that, at nearly every time point, both groups produce a distinction between the English and Korean stop series.

In contrast to the VOT of voiced stops, the VOT of English voiceless stops lengthened over time. This drift in VOT was found in both groups, but persisted through week 52 only in the HU group (Fig. 1b). An ANOVA on the model of VOT in voiceless stops revealed a significant effect of Place [$\chi^2(2) = 23.919, p < .0001$], Time [$\chi^2(5) = 43.794, p < .0001$], and Group [$\chi^2(1) = 7.570, p = .006$]. The effect of Place here was similar to the Place effect for voiced stops, while the effect of Group was due to the HU group showing significantly shorter VOTs overall than the LU group [$\beta = -26.733, t = -3.815, p = .002$]. The effect of Time reflected a longitudinal lengthening of VOT, which for the LU group was significant in week 5 [$\beta = 8.265, t = 2.790, p = .005$] but not week 52 [$\beta = -1.637, t = -0.553, p = .581$]. The only significant interaction was the Time × Group interaction [$\chi^2(5) = 28.314, p < .0001$]; this interaction arose because VOT drifted more in the HU group than in the LU group in weeks 2, 3, 4, and 52 [$\beta s > 8.999, ts > 2.087, ps < .05$], although this may be due to the fact that the HU group started with shorter VOT than the LU group in week 1.[8] The end result was thus that only the HU group's mean VOT remained longer in week 52 than in week 1. This group disparity did not appear to be due to differences in L2 development: as shown in Fig. 1b, the LU group showed a trajectory for the similar Korean aspirated stops that resembled the HU group's, yet only the HU group showed a VOT increase from week 5 to 52 in the English voiceless stops.

In short, LU and HU participants were consistent in showing no drift in VOT of voiced stops, but significant drift in VOT of voiceless stops during L2 instruction (i.e., weeks 1–5). Only HU participants, however, continued to produce voiceless stop

[8] Because the LU and HU groups were based on self-reported L2 use after week 5, there was no particular expectation regarding how the two groups would compare before week 5. However, the early divergence between groups here, particularly in week 1, merits an explanation. This may have to do with a global pattern evident across all of the dependent variables: in general, the HU group seems to establish more of a distance between the L1 and L2 than the LU group. Consequently, it is possible that the between-group difference in English voiceless stop VOT evident in week 1 is due at least in part to the fact that, whereas the HU group produces the English voiceless stops with shorter VOT than the Korean aspirated stops (in line with norms, given that the English stops are supposed to have shorter VOT than the Korean stops), the LU group produces them with exceedingly long VOT that appears to be "piggybacking" on the long VOTs produced for the Korean stops. On the other hand, the convergence of the groups at week 52 appears to be due to the confluence of two developments from week 5 to week 52: (1) the LU group's decrease in VOT (for both the Korean and the English stops), and (2) the HU group's continued increase in VOT for the English stops (despite a similar decrease in VOT for the Korean stops as seen in the LU group). Thus, it may be the case that, between week 5 and week 52, the HU group became more like the LU group in terms of conflating the English voiceless and Korean aspirated stops.



(a) English voiced & Korean fortis (b) English voiceless & Korean aspirated

Fig. 1. Change in VOT of (a) English voiced and Korean fortis, and (b) English voiceless and Korean aspirated stops. The low (LU) and high active L2 use (HU) groups are shown in circles and triangles, respectively. Error bars indicate 95% confidence intervals of the mean over participants. For reference, mean VOT norms for the Korean fortis and aspirated stops (averaging over all places of articulation) are, respectively, 11–17 ms and 90–97 ms (Chang, 2012, p. 253).

VOTs in week 52 that were significantly longer than in week 1. These results thus suggest that prolongation of drift in VOT is driven not simply by continued L2 exposure, but the combination of L2 exposure and active use.

4.2.2. Phonetic drift in f_0

In accordance with the prediction of upward drift in f_0, onset f_0 following English voiced stops increased from week 1 to 5 in both the LU and HU groups; from week 5 to 52, however, f_0 decreased in both groups, although it tended to decrease less in the HU than the LU group (Fig. 2a). An ANOVA on the model of f_0 following voiced stops showed a significant effect of Place [$\chi^2(2) = 7.028, p = .030$] and Time [$\chi^2(5) = 67.640, p < .0001$], but no effect of Group [$\chi^2(1) = 1.648, p = .199$]. The effect of Place was unexpected and not apparent in the coefficients of the main model; however, a single-predictor model treating Place as a treatment-coded factor (reference level 'alveolar') revealed that the source of the effect was bilabials showing a lower f_0 than alveolars [$\beta = -0.130, t = -2.251, p = .025$].[9] The effect of Time reflected an inverse U-shaped pattern of f_0 drift, which for the LU group resulted in higher onset f_0 (relative to week 1) in weeks 2–5 [βs > 0.285, ts > 2.462, ps $< .05$] as well as week 52 [$\beta = 0.369, t = 3.224, p = .001$]. The only significant interaction was the Time × Group interaction [$\chi^2(5) = 18.871, p = .002$], which arose primarily due to a tendency for the HU participants' f_0 in weeks 3–4 to evince less drift relative to week 1 than LU participants' [βs < -0.356, ts < -2.108, ps $< .05$]. However, a follow-up model built just on the HU group's data confirmed that, as in the LU group, the HU group's f_0 remained higher in week 52 than in week 1 [$\beta = 0.275, t = 2.282, p = .023$], by a margin of about 7 Hz. At every time point and in both groups, the similar Korean fortis stops were produced with much higher f_0 than the English voiced stops, thus providing the impetus for the voiced stops to drift upwards in f_0 (Fig. 2a).

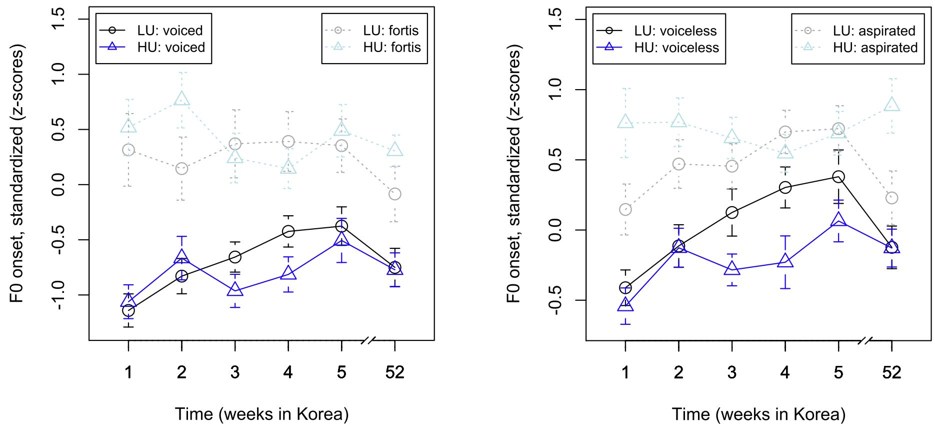
As with voiced stops, onset f_0 following English voiceless stops increased from week 1 to 5, and then decreased from week 5 to 52, in both the LU and HU groups; however, the week 5-to-52 decline was smaller in the HU than the LU group (Fig. 2b). An ANOVA on the model of f_0 following voiceless stops showed a significant effect of Time [$\chi^2(5) = 120.115, p < .0001$], which reflected a general pattern of upward drift in weeks 2–5; for the LU group, this drift resulted in f_0 being significantly higher (compared to week 1) in all following weeks including week 52 [βs > 0.289, ts > 3.120, ps $< .01$]. There was again an effect of Place [$\chi^2(2) = 6.051, p = .049$], similar to that found for voiced stops, but no effect of Group [$\chi^2(1) = 1.328, p = .249$]. The only significant interaction was the Time × Group interaction [$\chi^2(5) = 27.211, p < .0001$], which, as for voiced stops, was due to HU participants' f_0 in weeks 3–4 drifting less relative to week 1 than LU participants' [βs < -0.292, ts < -2.168, ps $< .05$]. Crucially, however, a follow-up model built just on the HU group's data confirmed that the HU group's f_0 remained significantly higher in week 52 than in week 1 [$\beta = 0.406, t = 4.395, p < .0001$], by a margin of about 9 Hz. At every time point and in both groups, especially the HU group, the similar Korean aspirated stops were produced with higher f_0 than the English voiceless stops (Fig. 2b).

Thus, the LU and HU groups both showed drift in onset f_0 during L2 instruction, as well as a decline in this drift after L2 instruction. In both groups, however, the decline was incomplete, with f_0 remaining elevated above week 1 levels in week 52. Given that sustained drift in VOT was found only in the HU group, these results therefore suggest that drift in features related to f_0 level may be more persistent than drift in VOT. In particular, prolongation of drift in onset f_0 does not appear to require extensive active L2 use as in the HU group.

4.2.3. Phonetic drift in F_1 and F_2

The evolution of the English vowel space over time is shown in Figs. 3 and 4 (omitting weeks 2–4 for clarity) for the LU and HU groups, respectively. These figures show that

[9] A similar effect of place of articulation is observed in some, but not all, tones in Taiwanese, where the clearest effect of place is velars showing the highest f_0 of all (Lai, Huff, Sereno, & Jongman, 2009). Lai et al. (2009) hypothesize that the higher f_0 after velars may be due to larynx raising associated with tongue back raising; however, since alveolars do not involve tongue back raising, this is unlikely to cause f_0 to be higher after alveolars. Thus, the cause of the Place effect observed here remains unclear.



(a) English voiced & Korean fortis (b) English voiceless & Korean aspirated

Fig. 2. Change in f_0 following (a) English voiced and Korean fortis, and (b) English voiceless and Korean aspirated stops. The low (LU) and high active L2 use (HU) groups are shown in circles and triangles, respectively. Error bars indicate 95% confidence intervals of the mean over participants. For reference, onset f_0 norms for the Korean fortis and aspirated stops, due to the f_0 elevation associated with laryngeally marked stop types in Korean, are expected to be higher than for the English stops.

although longitudinal shifts in individual vowels were generally subtle, there was systematic change at the level of the system with respect to both F_1 and F_2; however, this change was more evident in the LU than the HU group. An ANOVA on the model of F_1 revealed the expected effect of Vowel [$\chi^2(10) = 69954.072, p < .0001$], as well as an effect of Time [$\chi^2(5) = 18.154, p = .003$]; however, there was no effect of Group [$\chi^2(1) = 1.945, p = .163$]. The effect of Time reflected an overall pattern of F_1 decrease from week 1 to week 5, followed by F_1 increase from week 5 to 52, which tracked quite closely the pattern in Korean (Fig. 5a). For the LU group, this drift pattern ended in mean F_1 being significantly higher in week 52 than in week 1 [$\beta = 0.072, t = 3.980, p < .001$]. Note, however, that the higher F_1 in week 52 does not necessarily represent an overcompensation for the downward drift in F_1 since the true baseline corresponds to week 0, which was not observed (i.e., it is possible that F_1 in week 52, although higher than F_1 in week 1, does not differ from the baseline F_1 in week 0).

In addition to the main effects of Vowel and Time, there were three significant interactions: Vowel × Group [$\chi^2(10) = 26.091, p = .004$], Time × Group [$\chi^2(5) = 21.469, p < .001$], and Vowel × Time × Group [$\chi^2(50) = 89.987, p < .001$]. The Vowel × Group interaction was primarily due to the vowel /ɑ/, produced with higher F_1 (relative to the center of the vowel space) in the HU than the LU group [$\beta = 0.118, t = 2.074, p = .038$]. The Time × Group interaction reflected the relatively flat pattern of F_1 drift in the HU group—in particular, the smaller F_1 increase between weeks 5 and 52 compared to the LU group [$\beta = -0.062, t = -2.430, p = .015$]. To explore this interaction further, an additional model (with the same structure as the main model but no Group factor) was built on just the HU group's data. An ANOVA on this model indicated that the Time × Group interaction arose because, unlike the LU group, the HU group did not show a significant effect of Time on F_1 [$\chi^2(5) = 4.684, p = .456$]. Finally, the Vowel × Time × Group interaction was due to several vowels (/i æ o ɑ/) patterning differently in week 2 in the HU group compared to the LU group.

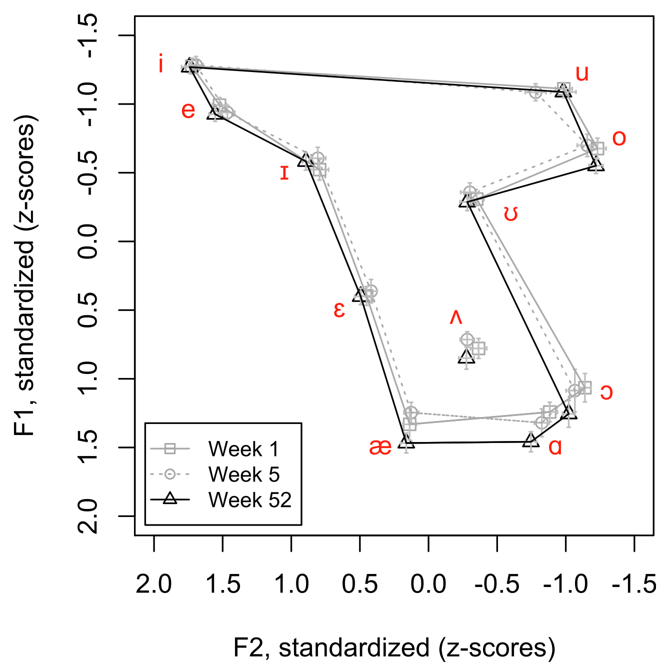


Fig. 3. F_1 by F_2 of English vowels over time for the low active L2 use (LU) group. Week 1 means are shown with squares and solid gray lines; Week 5 means, with circles and dotted gray lines; and Week 52 means, with triangles and solid black lines. Error bars indicate ±1 mean standard error.

As for F_2, here, too, there was significant drift, but again only in the LU group; furthermore, the pattern of drift was increasing (as opposed to mostly decreasing for F_1). An ANOVA on the model of F_2 again revealed a significant effect of Vowel [$\chi^2(10) = 98327.083, p < .0001$] and Time [$\chi^2(5) = 20.437$, $p = .001$], but no effect of Group [$\chi^2(1) = 1.667, p = .197$]. The effect of Time was due to a longitudinal increase in F_2 for LU learners, which resembled their trajectory in Korean (Fig. 5b) and resulted in significantly higher F_2 (relative to week 1) in all following weeks [βs > 0.032, ts > 2.099, ps $< .05$]. By contrast, an additional model built on just the HU group's data (with the same structure as the main model but no Group factor) showed no significant effect of Time [$\chi^2(5) = 8.300$,

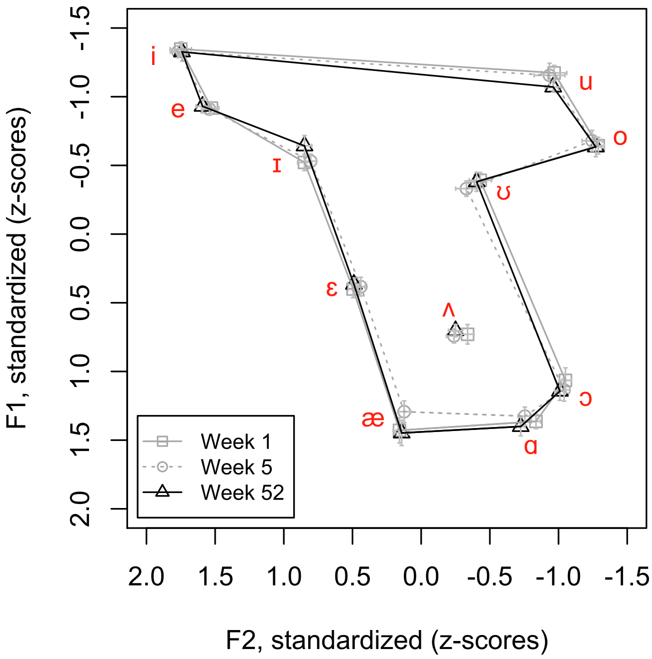


Fig. 4. F_1 by F_2 of English vowels over time for the high active L2 use (HU) group. Week 1 means are shown with squares and solid gray lines; Week 5 means, with circles and dotted gray lines; and Week 52 means, with triangles and solid black lines. Error bars indicate ±1 mean standard error.

$p = .141$], reflecting the fact that F_2 in the HU group did not significantly differ in weeks 2–52 from week 1 levels [βs < 0.029, ts < 1.893, ps $> .05$]. The only significant interaction in the main model was the Vowel × Group interaction [$\chi^2(10) = 36.204, p < .0001$], due to /ʊ/ being produced with lower F_2 (relative to the center of the vowel space) in the HU than the LU group [$\beta = -0.101, t = -2.082, p = .037$].

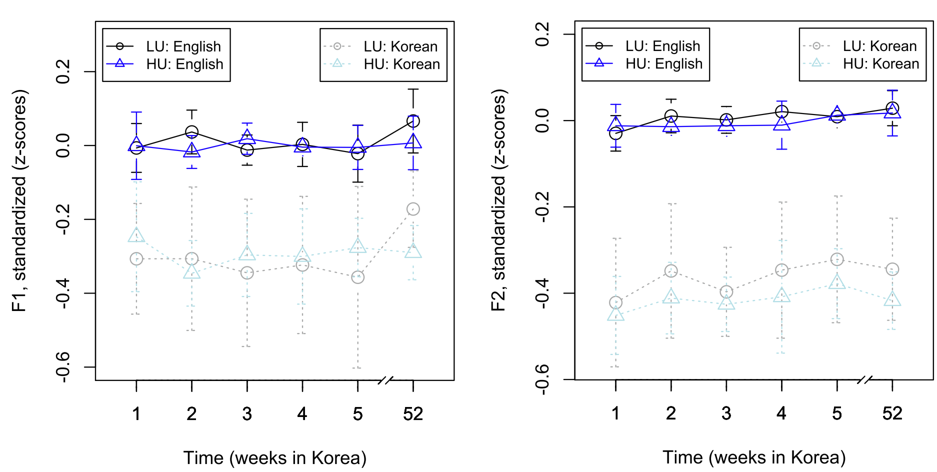
In sum, the LU group, but not the HU group, showed drift in F_1 and F_2. Drift in F_1 occurred via a decrease between weeks 1 and 5, followed by an increase between weeks 5 and 52, whereas drift in F_2 occurred overall via an increase from week 2 onwards. Although not all vowels moved in a manner consistent with the overall patterns, the observed effects were not isolated to just a few vowels, as reflected in the non-significance of the Vowel × Time interaction in all models. Notably, these results, vis-a-vis the VOT results, show the reverse group disparity, suggesting that prolongation of drift in vowel formants is not dependent on frequent active L2 use and, moreover, that frequent active L2 use might actually play a role in increasing the stability of the L1 vowel space in the face of ambient L2 exposure.

5. General discussion

5.1. Synthesis of the findings

Results of the longitudinal study were consistent with the Incidental Processing Hypothesis (IPH) that ambient input in a familiar L2 would tend to be processed, thereby promoting high L2 activation: phonetic drift of the L1 during L2 instruction persisted post-instruction within the L2 environment. However, the specific predictions based on findings in Chang (2012) and Chang (2013) were only partially supported. As expected, voiceless stops drifted in VOT and both voiced and voiceless stops drifted in onset f_0 during initial L2 instruction; while the VOT drift persisted only among frequent L2 speakers (cf. predictions (1) and (4)), the f_0 drift persisted among both frequent and less frequent L2 speakers (cf. prediction (2)). Unexpectedly, vowels underwent drift in both F_1 and F_2 among less frequent L2 speakers, and this drift persisted after L2 instruction (cf. prediction (3)). Also unexpected was the fact that extensive active L2 use was associated with sustained drift in VOT of consonants, but with resistance to drift in F_1 and F_2 of vowels (cf. prediction (4)). Together, these results provide evidence that, in one or more ways, the L1 production of L2 learners tends to diverge from monolingual L1 norms during L2 instruction, and then tends to stay that way in an L2 environment, even when learners do not continue to speak the L2 very frequently.

In evaluating the current findings, it is worth noting that these conclusions are on the conservative side, since weeks



(a) F_1 (standardized) (b) F_2 (standardized)

Fig. 5. Mean (a) F_1 and (b) F_2 of the English and Korean vowel systems over time, by group. The low (LU) and high active L2 use (HU) groups are shown in circles and triangles, respectively. Error bars indicate 95% confidence intervals of the mean over participants. For reference, mean F_1 and F_2 norms for Korean vowels (averaging over the vowel inventory) are consistently lower than for English vowels due to the smaller number of open and front vowels in Korean (Chang, 2012, p. 254).

2–52 were compared to week 1 (i.e., L1 production after one week of L2 learning), not to week 0 (i.e., true baseline L1 production). For the one case of apparent reversal of drift by week 52 (i.e., VOT in the LU group), it is therefore possible that the data from week 52 represent only partial reversal (returning to a week 1 level already significantly different from baseline) rather than full reversal (returning all the way to baseline). The fact that true baseline L1 production was not able to be observed is a limitation of this study that does not allow for the conclusion of full reversal of phonetic drift. Crucially, however, the data observed from week 1 onwards support the conclusion that even L2 learners who report limited active use of the L2 tend to differ from L1 monolinguals while in an L2 environment, which argues against the methodological conflation of monolingual and multicompetent 'native speakers'.

An additional limitation of this study is the fact that the LU and HU groups differed in at least two ways besides active L2 use after week 5. Recall from §4.1.1 that, although the groups were matched along a variety of demographic and experiential dimensions, they differed in terms of prior exposure to Korean and bilingualism. The fact that the HU group contained more individuals with prior exposure to Korean, including Korean Americans who were adopted from Korea at an early age, as well as some bilinguals (in contrast to the LU group) is relevant given the evidence that international adoptees retain knowledge of their birth language, which may confer an advantage in (re)learning (Bowers, Mattys, & Gage, 2009; Oh, Au, & Jun, 2010; Choi, Broersma, & Cutler, 2017), and that bilinguals also may have an advantage in phonetic learning (Antoniou, Liang, Ettlinger, & Wong, 2015). Post-hoc analyses did not actually reveal systematic differences between adoptee or bilingual members of the HU group and the other members, suggesting that they were not solely responsible for the patterning of the HU group. Nevertheless, the group differences observed in §4—in particular, in week 52—should be taken with the proverbial grain of salt, as it cannot be guaranteed that they were solely due to the basis of the group division (i.e., self-reported active L2 use after week 5).

Although the current findings are consistent with those reported in Chang (2012, 2013) for drift in VOT and f_0, they differ with respect to drift in vowel formants. In particular, the (mostly female) LU group in this study showed drift in both F_1 and F_2, whereas the female learners in Chang (2012, 2013) showed drift in F_1, but not in F_2. In fact, the upward drift in F_2 exhibited by the LU group in this study resembles the upward drift in F_2 exhibited by the male learners in Chang (2012, 2013); however, whereas the drift exhibited by those male learners can be interpreted as convergence toward the L2 at an acoustic level (in approximation to an L2 model based on female instructors), the drift exhibited by the LU group in the current study cannot be interpreted as convergence, since the increase in F_2 had the effect of taking the L1 vowel system further away from the lower F_2 level of the target L2 vowel system (Fig. 5b; see also Chang, 2012, p. 254).

Why, then, did the learners in this study show a pattern of drift in F_2 that was effectively dissimilatory vis-a-vis the L2? Further research is needed to answer this question, but one factor that may be playing a role is the crosslinguistic convergence resulting from the drift in F_1. In other words, perhaps F_2 drifted in the observed manner in response to the drift that occurred in F_1, diverging from the L2 in order to keep a certain amount of distance between the L1 and L2 vowel systems; such a concern for maintaining crosslinguistic contrast would be consistent with the SLM principle of a shared phonetic space for L1 and L2 sounds (Flege, 1995; see also Lang & Davidson, in press). Regardless of how this drift pattern is interpreted, however, the basic finding of F_2 drift among less frequent L2 speakers—sustained well after L2 instruction—lends further support to the claim that L1 users with L2 experience are different users of the L1 than monolinguals.

To my knowledge, this is the first study to track L1 phonetic developments in L2 learners in relation to temporal separation from L2 instruction and frequency of continued active L2 use. As such, apart from the IPH, there is no established theory that applies directly to all of the different acoustic properties examined here. The main contribution of the current set of results, therefore, is not in testing the predictions of an existing theory, but in paving the way for further work in this area, which may lead ultimately to an elaborated theory of phonetic drift as a function of variables such as acoustic dimension, amount/type/timing of L2 experience, and cognitive and affective factors. At this point in time, one can appeal to certain explanations for some of the variation in drift seen in this study; for example, a control mechanism for f_0 that is distinct from segment-level control mechanisms and shared across languages might be behind the VOT-f_0 disparity (cf. Chang, 2010). However, given the modest sample size as well as participants' unique backgrounds as language teachers, it should be borne in mind that more research is needed to understand the extent to which the current results will generalize to other L2 users.[10]

5.2. Language change over the lifespan

In the context of a growing body of research in language variation and change showing L1 developments occurring well after childhood (Harrington, Palethorpe, & Watson, 2000; Sankoff & Blondeau, 2007; Wagner & Sankoff, 2011; Wagner, 2012; Rickford & Price, 2013), the contribution of the present study is in highlighting the role of L2 experience in lifespan change. The data in §4.2 suggest that L2 contact does not need to build to a high level of proficiency or involve extensive active use in order to have a detectable effect on the L1. On the contrary, once L2 learners have acquired a modicum of L2 knowledge, continued ambient exposure to the L2

[10] An anonymous reviewer wondered, for instance, about participants' exposure to Korean-accented English as language teachers in L2 English classrooms; English teachers in Korea may interact with L2 English interlocutors outside of the classroom as well. Could contact with Korean-accented interlocutors have caused the observed drift of participants' English, which usually converged toward acoustic properties of Korean? While the potential role of such contact cannot be completely excluded, research suggests, on the contrary, that native interlocutors tend to diverge from nonnative interlocutors, at least those who are strongly accented and who are not particularly close to them socially (Kim, 2009; Kim, Horton, & Bradlow, 2011). More generally, there are also characteristics of language teachers that disfavor their assimilating an L2 accent resulting in deviation from L1 norms. L1 users who gravitate toward language teaching may do so because they have an "instructional orientation" toward the L1, which may be related to, and/or enhanced by, high metalinguistic awareness and explicit knowledge of rules, norms, and standards. Additionally, teaching one's L1 involves the unique production experience of repeatedly articulating the L1 in a clear, careful, standard manner. Consequently, although the need to examine the generalizability of these results remains, I regard it as unlikely for the current participants to have assimilated the observed drift directly from Korean-accented English speakers.

exerts a significant influence on some (but not all) aspects of L1 production, and continued active L2 use appears to enhance this effect in certain cases. These findings support the dynamic view of L1 knowledge that follows from the multi-competence framework (Cook, 1992; Cook, 2003) and the dynamic systems approach to lifespan development (de Bot, 2007), arguing in favor of giving thoughtful consideration to language background, including recent L2 experience, as a component of behavioral linguistic research. To provide a more comprehensive view of the persistence of L1 drift, future studies could track developments in L1 production after L2 learners have returned to an L1 environment (cf. Sancier & Fowler, 1997).

Of course, lifespan change is not necessarily limited to language knowledge, but may extend to socio-affective dimensions such as identity and group affiliation. Thus, it is worth noting that, apart from the cognitive pressure favoring L2-influenced change in the L1, such change may also be influenced by the social signaling potential of manifesting L2-like features (see, e.g., Sharma & Sankaran, 2011; Alam & Stuart-Smith, 2011; Alam & Stuart-Smith, 2014). Given that 'Korean English' did not have the status of a regional ethnic English within Korea at the time of the current study and, moreover, participants were often the only L1 English speaker in their locality, it is not clear that participants' manifesting Korean-like features in their English would have served a coherent socio-indexical function in this case. Nevertheless, one can imagine how the right conditions could arise such that L1 English teachers in Korea begin to identify as part of a unique, 'L1 English expat in Korea' community, resulting in Korean-colored English acquiring social meaning (e.g., 'localness') that encourages the increasing use of Korean-like features as this community sets itself apart from more short-term English-speaking visitors.

In addition to the role of sociolinguistic factors, another direction for future research on L2-influenced L1 change is its time course. Although some results in §4.2 indicate sustained drift, a portion of the results also indicate that L2-influenced changes in the L1 can be short-lived, dissipating in the absence of frequent active L2 use. This type of finding is consistent with some of the results reviewed in §2.1 as well as other results, such as the lack of effect of variation in L2 Spanish use on L1 Quichua accent (Guion, Flege, & Loftin, 2000; cf. Yeni-Komshian, Flege, & Liu, 2000, de Leeuw et al., 2010). The existence of weak or null effects of L2 experience thus brings us back to the question of constraints on L2 influence. Given that previously proposed constraints do not seem to hold consistently (see §2.1), much more research will be required to understand when L2 effects persist, where they come from (e.g., learning, use, and/or exposure), and how they differ according to the type of L1 structure or property at issue.

In connection with the latter question, a central concern for future studies of L2-influenced L1 change will be accounting for when L2 influence manifests as an ostensibly 'negative' effect (i.e., resulting in divergence from monolingual norms) and when it does not. Complementing the abundance of 'negative' L2 effects summarized in §2.1, 'positive' effects resulting in some kind of advantage over monolinguals are also reported in the literature. In addition to the domain-general and metalinguistic benefits of bilingualism, L2 learning has been linked to production of greater L1 complexity (Kecskes, 1998) as well as *less* apparent attrition of the L1. For example, L1 Russian immigrants in Israel proficient in L2 Hebrew perform more similarly to Russian monolinguals in judging the correctness of complex grammatical constructions in Russian than do Russian immigrants who do not know Hebrew (Laufer & Baladzhaeva, 2015). Similarly, in the present study, frequent L2 speakers had an advantage over less frequent L2 speakers in L1 vowel stability. Therefore, it would be a gross oversimplification to say that L2 learning necessarily 'interferes' with the L1, because it is clear that the consequences of multicompetence show a variability that is not yet fully understood.

5.3. Best practice in treatment of language background

The methodological review in §3 revealed a tendency for behavioral studies in linguistics to contain vague definitions of target populations and/or mismatches between target populations and participant samples, reflecting an overreliance on nativeness to define language background despite the fact that the term 'native' is not a precise descriptor. Because language histories can be complex (in particular, multilingual), omission of this information from a study report implies that it is irrelevant to the aims or results of the study. In light of the findings in §2, however, it is not clear that any type of linguistic behavior can be safely assumed to remain unaffected by multilingualism. On the contrary, an abundance of evidence—not only the phonetic data in the present study but also the extensive findings reviewed in §2.1—suggests that L1 knowledge remains, to some degree, plastic across the lifespan and, in particular, responsive to changes in the L1 user's circumstances.

This view of the adult L1 system as dynamic, as opposed to immutable, highlights both the empirical inadequacy of the category 'native' as well as the need to move toward more informative descriptions of language users. Because results found with one kind of native speaker may not generalize to a different kind (e.g., one with a different language background or belonging to a different speech community), specifying a target population only in terms of the broad category of 'native' is likely to complicate attempts at replication as long as the L1 is assumed to be unchanging. For example, if a result reportedly obtained with 'native' speakers of Canadian English fails to be replicated with Canadian English monolinguals, it is difficult to know how to interpret this: is it actually evidence against the original result or just the product of a sampling difference with respect to the original study (which, given the trends discussed in §3.2, may have been based on English-French bilinguals or speakers residing outside Canada)?

Importantly, it should be noted that the kind of considered treatment of language background that will aid future research, including replication, does not have to be complicated and is already supported by a number of published resources. There are, for example, several well-described instruments for collecting data about language history and background, such as the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), the Bilingual Language Profile (Birdsong, Gertken, & Amengual, 2012), the Bilingual Language Experience Calculator (Unsworth, 2013), the Language History Questionnaire (Li, Zhang, Tsai, & Puls, 2014), and the custom-designed questionnaire for heritage Korean speakers used by Ahn, Chang, DeKeyser, and Lee-Ellis (2017). Such questionnaires do not take very long to administer and provide a wealth of data about participants' language backgrounds, which are useful for determining whether study participants are representative of the target population. However, only a fraction of this information (in particular, language proficiencies and acquisition profile) typically needs to be presented for the reader to understand the nature of the participant sample. That is, showing that language background has been controlled for, especially when the target population is monolinguals, does not require an undue amount of time or journal space.

The imperative to appropriately control and describe language background invites the question of what specific variables related to language background are crucial to consider in linguistic research. In short, the answer to this question will depend on the nature of the research question(s), but it is worth bearing in mind that there are two reasons why a researcher might consider collecting and then reporting certain data about language background. The first reason is to ensure that the results obtained in the study can in fact address the research question posed. The second is to aid replication: for other researchers to run the same study (targeting the same population), they need to know enough about the participant sample from the original study to be able to put together a similar sample.

Thus, apart from general information about participants' multicompetence (i.e., knowledge and acquisition of languages beyond the target language), the aspects of language background which are relevant to focus on will differ across studies and must be identified by thinking about whether/how the L1 phenomenon under investigation might be affected by L2 experience. In the easy (and probably uncommon) case, there is already strong evidence that the given phenomenon is not affected by L2 experience. However, in the absence of such evidence, there is a need to understand both the nature of participants' L2 experience, as well as the manner in which the target phenomenon might be affected by this L2 experience (e.g., whether the potential effect goes in the direction of or against the hypothesis), which underscores the importance of research in L2 acquisition and L1 attrition in furthering the theory of crosslinguistic congruence (i.e., the points of overlap or similarity that may exist between different languages), convergence, and divergence, across multiple levels of language.

6. Conclusion

Although the monolingual model of the language user is prevalent in linguistic research, research on L2 learning and L1 attrition, including the study in §4, suggests that the monolingual model, in conjunction with assuming an unchanging L1, may lead to an inaccurate picture of the target speech community. Actual language users are often not monolingual, and when they are competent in additional languages, their multicompetence cannot be ignored because multicompetence is transformative, not merely additive; that is, previously acquired linguistic systems are, to some degree, plastic, rather than fixed. This view is at odds with current methodological practices in linguistic research related to participant sampling and description of monolinguals, which tend to be underinformed regarding the language background of so-called 'native' speakers.

Thus, the present study serves as a call to the field to address the variable of language background in sufficient detail to allow behavioral research findings focused on monolinguals to be interpreted and replicated transparently. As shown in §4, L2 learning exerts a rapid and, in some cases, persistent effect on the L1 even when the L2 is not spoken very frequently; consequently, this is an issue about L2 contact in general, not about high levels of L2 proficiency or frequent active L2 use in particular. For research on monolingual speakers, the way forward is to give language background the same kind of considered treatment that one sees only occasionally in research on monolinguals (e.g., studies cited in §3.3) but regularly in the research on bi-/multilinguals.

In closing, it cannot be overemphasized that the recommendation for updating methodological practices in research on monolinguals should not be construed as a recommendation for conducting research according to the monolingual model. For a language typically spoken by multilinguals, there may be fine research questions that directly engage only one of these speakers' languages, and controlling for language background appropriately provides a reasonable way of addressing such questions; however, there are also many interesting questions to be asked about these speakers' multingualism. Thus, whereas one approach to examining a language in this type of multilingual context is to control for language background, another approach, poised to provide broader insight into such multilingual speakers, is to examine these multilinguals as multilinguals. In fact, the type of diasporic or expatriate community examined in this study is not only worthy of holistic investigation, but uniquely positioned to improve our understanding of crosslinguistic interaction, language change, and language stability in a mobile, multilingual world.

Acknowledgments

The author gratefully acknowledges financial support from the National Science Foundation (BCS-0922652) and the Center for Korean Studies and Department of Linguistics at UC Berkeley; research assistance from Daiana Chang and Kevin Sitek; and logistical assistance from the Fulbright Korean-American Educational Commission. The research reported here benefited from the feedback of many individuals, including Ocke-Schwen Bohn, Ann Bradlow, Chiara Celata, Taehong Cho, Lisa Davidson, Esther de Leeuw, Susanne Gahl, Carla Hudson Kam, Sharon Inkelas, Keith Johnson, John Ohala, Kathryn Pruitt, three anonymous reviewers, and audiences at UC Berkeley, Yonsei University, NYU, the CUNY Graduate Center, Rice University, the 11th International Symposium on Bilingualism (ISB 11), and meetings of the Linguistic Society of America and the Acoustical Society of America. Any errors, however, are those of the author.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.wocn.2019.03.001.

References

Abdelli-Beruh, N. B. (2012). Voicing assimilation of French /t/. *Journal of Phonetics, 40*(3), 521–534.

Ahn, S., Chang, C. B., DeKeyser, R., & Lee-Ellis, S. (2017). Age effects in first language attrition: Speech perception by Korean-English bilinguals. *Language Learning, 67*(3), 694–733.
Alam, F., & Stuart-Smith, J. (2011). Identity and ethnicity in /t/ in Glasgow-Pakistani high-school girls. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 216–219). Hong Kong: City University of Hong Kong.
Alam, F., & Stuart-Smith, J. (2014). Identity, ethnicity and fine phonetic detail: An acoustic phonetic analysis of syllable-initial /t/ in Glaswegian girls of Pakistani heritage. In M. Hundt & D. Sharma (Eds.), *English in the Indian Diaspora* (pp. 29–53). Amsterdam, The Netherlands: John Benjamins Publishing.
Antoniou, M., Liang, E., Ettlinger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition, 18*(4), 683–695.
Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics, 40*(4), 582–594.
Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M., & Yoon, T.-J. (2018). The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis. *Journal of Phonetics, 66*, 120–144.
Bassetti, B., & Cook, V. J. (2011). Relating language and cognition: The second language user. In V. J. Cook & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 143–190). Hove, UK: Psychology Press.
Beinhoff, B. (2008). Looking for the "real" native speaker: The perception of native and non-native English accents by non-native speakers of English. In E. Waniek-Klimczak (Ed.), *Issues of Accents in English* (pp. 120–139). Cambridge, UK: Cambridge Scholars Publishing.
Bergmann, C., Nota, A., Sprenger, S. A., & Schmid, M. S. (2016). L2 immersion causes non-native-like L1 pronunciation in German attriters. *Journal of Phonetics, 58*, 71–86.
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.
Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science, 19*(1), 19–23.
Birdsong D., Gertken L.M., Amengual M. (2012). Bilingual language profile: An easy-to-use instrument to assess bilingualism. *Center for Open Educational Resources and Language Learning*, University of Texas at Austin, https://sites.la.utexas.edu/bilingual/
Boersma, P. & Weenink, D. (2016). Praat: Doing phonetics by computer. Version 6.0.19. http://www.praat.org.
Bournot-Trites, M., & Tellowitz, U. (2002). *Report of current research on the effects of second language learning on first language literacy skills*. Halifax, Canada: The Printing House.
Bowers, J. S., Mattys, S. L., & Gage, S. H. (2009). Preserved implicit knowledge of a forgotten childhood language. *Psychological Science, 20*(9), 1064–1069.
Brown, A. (2008). Gesture viewpoint in Japanese and English: Crosslinguistic interactions between two languages in one speaker. *Gesture, 8*(2), 256–276.
Brown, A., & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of Path among Japanese learners of English. *Bilingualism: Language and Cognition, 14*(1), 79–94.
Brown, A., & Gullberg, M. (2012). Multicompetence and native speaker variation in clausal packaging in Japanese. *Second Language Research, 28*(4), 415–442.
Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2*(3), 239–245.
Celata, C. (in press). Phonological L1 attrition. In M.S. Schmid & B. Köpke (Eds.), *The Oxford Handbook of Language Attrition*. Oxford, UK: Oxford University Press.
Chang, C. B. (2010). *First language phonetic drift during second language acquisition* PhD thesis. Berkeley, Berkeley, CA: University of California.
Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics, 40*(2), 249–268.
Chang, C. B. (2013). A novelty effect in phonetic drift of the native language. *Journal of Phonetics, 41*(6), 491–504.
Chang, C.B. (in press). Phonetic drift. In M.S. Schmid & B. Köpke (Eds.), *The Oxford Handbook of Language Attrition*. Oxford, UK: Oxford University Press.
Childers, D. G. (1978). *Modern Spectrum Analysis*. New York, NY: John Wiley & Sons.
Cho, M.-H., & Lee, S. (2016). The impact of different L1 and L2 learning experience in the acquisition of L1 phonological processes. *Language Sciences, 56*, 30–44.
Choi, J., Broersma, M., & Cutler, A. (2017). Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences of the United States of America, 114*(28), 7307–7312.
Chomsky, N. (1957). *Syntactic Structures*. The Hague, The Netherlands: Mouton de Gruyter.
Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, NY: Praeger Publishers.
Cook, V. J. (1991). The poverty-of-the-stimulus argument and multicompetence. *Second Language Research, 7*(2), 103–117.
Cook, V. J. (1992). Evidence for multicompetence. *Language Learning, 42*(4), 557–591.
Cook, V. [J.] (1997). The consequences of bilingualism for cognitive processing. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 279–300). Mahwah, NJ: Lawrence Erlbaum Associates.
Cook, V. [J.] (2003). The changing L1 in the L2 user's mind. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 1–18). Clevedon, UK: Multilingual Matters.
Davies, A. (2003). *The native speaker: Myth and reality*. Clevedon, UK: Multilingual Matters.
de Bot, K. (2007). Dynamic systems theory, lifespan development and language attrition. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition: Theoretical Perspectives* (pp. 53–68). Amsterdam, The Netherlands: John Benjamins Publishing.
de Bot, K., Gommans, P., & Rossing, C. (1991). L1 loss in an L2 environment: Dutch immigrants in France. In H. W. Seliger & R. M. Vago (Eds.), *First Language Attrition* (pp. 87–98). Cambridge, UK: Cambridge University Press.
de Leeuw, E. (in press). Phonetic L1 attrition. In M.S. Schmid & B. Köpke (Eds.), *The Oxford Handbook of Language Attrition*. Oxford, UK: Oxford University Press.
de Leeuw, E., Mennen, I., & Scobbie, J. M. (2013). Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism, 17*(6), 683–700.
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.
de Leeuw, E., Tusha, A., & Schmid, M. S. (2018). Individual phonological attrition in Albanian-English late bilinguals. *Bilingualism: Language and Cognition, 21*(2), 278–295.
Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics, 38*(3), 483–492.
Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition, 10*(1), 101–116.
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 9* (pp. 353–382). Berlin, Germany: Walter de Gruyter.
Forster, J.C. (2014). DMDX. Version 5.1.0.0. http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm.
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.
Fox, E. (1996). Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory and Language, 35*(3), 353–370.
Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., Ogle, D., Ripley, B., Venables, W., Walker, S., Winsemius, D., Zeileis, A., & R Core Team. (2018). car: Companion to applied regression. R package version 3.0-2. https://cran.r-project.org/web/packages/car/. Last retrieved February 20, 2019.
Freedman, S. E., & Barlow, J. A. (2012). Using whole-word production measures to determine the influence of phonotactic probability and neighborhood density on bilingual speech production. *International Journal of Bilingualism, 16*(4), 369–387.
Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language, 16*(1), 161–179.
Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development, 6*(6), 467–477.
Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3–15.
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.
Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28*(1), 27–42.
Hagiwara, R. (1997). Dialect variation and formant frequency: The American English vowels revisited. *Journal of the Acoustical Society of America, 102*(1), 655–658.
Harrington, J., Palethorpe, S., & Watson, C. I. (2000). Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association, 30*(1–2), 63–78.
Helgason, P., & Ringen, C. (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics, 36*(4), 607–628.
Herd, W. J., Walden, R. L., Knight, W. L., & Alexander, S. N. (2015). Phonetic drift in a first language dominant environment. *Proceedings of Meetings on Acoustics, 23* 060005.
Higby, E., & Obler, L. (2014). The adaptation of native language construal patterns in second language acquisition. *AFinLA-e: Soveltavan kielitieteen tutkimuksia, 6*, 32–44.
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.
Holliday, J. J. (2015). A longitudinal study of the second language acquisition of a three-way stop contrast. *Journal of Phonetics, 50*, 1–14.

Huffman, M. K., & Schuhmann, K. (2015). Individual differences in phonetic drift by English-speaking learners of Spanish. *Journal of the Acoustical Society of America, 137*(4), 2268.
Jarvis, S. (2003). Probing the effects of the L2 on the L1: A case study. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 81–102). Clevedon, UK: Multilingual Matters.
Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York, NY: Routledge.
Jessner, U. (1999). Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language Awareness, 8*(3–4), 201–209.
Joh, J., Ko, H., Lim, I., & Lee, S. (2010). The impact of English language learning on the acquisition of native language among Korean children. *Eungyong Eoneohag [Korean Journal of Applied Linguistics], 26*(3), 101–141.
Kartushina, N. (2015). *Second language phonological acquisition in adults*. PhD thesis. Geneva, Switzerland: University of Geneva.
Kartushina, N., Frauenfelder, U. H., & Golestani, N. (2016). How and when does the second language influence the production of native speech sounds: A literature review. *Language Learning, 66*(S2), 155–186.
Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2016). Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training. *Journal of Phonetics, 57*, 21–39.
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.
Kecskes, I. (1998). The state of L1 knowledge in foreign language learners. *Word, 49*(3), 321–340.
Kecskes, I., & Papp, T. (2000). *Foreign language and mother tongue*. Mahwah, NJ: Lawrence Erlbaum Associates.
Kim, J.-H., Montrul, S., & Yoon, J. (2010). Dominant language influence in acquisition and attrition of binding: Interpretation of the Korean reflexive *caki*. *Bilingualism: Language and Cognition, 13*(1), 73–84.
Kim, K. H. S., Relkin, N. R., Lee, K.-M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*(6638), 171–174.
Kim, M. (2009). Phonetic accommodation in conversations between native and nonnative speakers. *Journal of the Acoustical Society of America, 125*(4), 2764.
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.
Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science, 23*(3), 159–163.
Ladefoged, P. (2005). *Vowels and consonants* (2nd ed.). Malden, MA: Blackwell Publishing.
Lai, Y., Huff, C., Sereno, J., & Jongman, A. (2009). The raising effect of aspirated prevocalic consonants on f_0 in Taiwanese. In J. Brooke, G. Coppola, E. Görgülü, M. Mameni, E. Mileva, S. Morton, & A. Rimrott (Eds.), *Proceedings of the 2nd International Conference on East Asian Linguistics*. Vancouver, Canada: Simon Fraser University.
Lang, B. & Davidson, L. (in press). Effects of exposure and vowel space distribution on phonetic drift: Evidence from American English learners of French. *Language and Speech*.
Laufer, B., & Baladzhaeva, L. (2015). First language attrition without second language acquisition: An exploratory study. *International Journal of Applied Linguistics, 166*(2), 229–253.
Lewis, P.M., Simons, G.F., & Fennig, C.D. (2015). Ethnologue: Languages of the world. 18th edition.http://www.ethnologue.com. Last retrieved March 8, 2015.
Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition, 17*(3), 673–680.
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.
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.
Lüpke, F., & Storch, A. (2013). *Repertoires and choices in African languages*. Boston, MA: Walter de Gruyter.
Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal, 76*(2), 190–208.
Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967.
Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language, 86*(1), 70–82.
Mayr, R., Price, S., & Mennen, I. (2012). First language attrition in the speech of Dutch-English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition, 15*(4), 687–700.
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.
Namjoshi, J., Tremblay, A., Spinelli, E., Broersma, M., Martínez-García, M. T., Connell, K., ... Kim, S. (2015). Speech segmentation is adaptive even in adulthood: Role of the linguistic environment. In The Scottish Consortium for ICPhS 2015. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: University of Glasgow (paper number 0676).
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.
Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics, 39*(2), 156–167.
Oh, J. S., Au, T. K., & Jun, S.-A. (2010). Early childhood language memory in the speech perception of international adoptees. *Journal of Child Language, 37*(5), 1123–1132.
Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism, 5*(1), 19–38.
Parmentier, F. B. R., Elsley, J. V., & Ljungberg, J. K. (2010). Behavioral distraction by auditory novelty is not only about novelty: The role of the distracter's informational value. *Cognition, 115*(3), 504–511.
Pavlenko, A. (2000). L2 influence on L1 in late bilingualism. *Issues in Applied Linguistics, 11*(2), 175–205.
Pavlenko, A. (2003). 'I feel clumsy speaking Russian': L2 influence on L1 in narratives of Russian L2 users of English. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 32–61). Clevedon, UK: Multilingual Matters.
Pavlenko, A., & Jarvis, S. (2002). Bidirectional transfer. *Applied Linguistics, 23*(2), 190–214.
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.
Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics, 6*(2), 179–206.
R Development Core Team (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.r-project.org. Last retrieved February 20, 2019.
Rickford, J., & Price, M. (2013). Girlz II women: Age-grading, language change and stylistic variation. *Journal of Sociolinguistics, 17*(2), 143–179.
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., & Blondeau, H. (2007). Language change across the lifespan: /r/ in Montreal French. *Language, 83*(3), 560–588.
Schmid, M. S. (2013). First language attrition. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(2), 117–123.
Schmid, M. S., & Köpke, B. (2009). L1 attrition and the mental lexicon. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon: Interdisciplinary Approaches* (pp. 209–238). Bristol, UK: Multilingual Matters.
Schuhmann, K. S., & Huffman, M. K. (2015). L1 drift and L2 category formation in second language learning. In The Scottish Consortium for ICPhS 2015. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: University of Glasgow (paper number 0850).
Schwanenflugel, P. J., & Rey, M. (1986). Interlingual semantic facilitation: Evidence for a common representational system in the bilingual lexicon. *Journal of Memory and Language, 25*(5), 605–618.
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.
Sharma, D., & Sankaran, L. (2011). Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change, 23*(3), 399–428.
Stolberg, D., & Münch, A. (2010). "Die Muttersprache vergisst man nicht" – or do you? A case study in L1 attrition and its (partial) reversal. *Bilingualism: Language and Cognition, 13*(1), 19–31.
Tice, M., & Woodley, M. (2012). Paguettes and bastries: Novice French learners show shifts in native phoneme boundaries. *UC Berkeley Phonology Lab Annual Report, 2012*, 72–75.
Tobin, S. J., Nam, H., & Fowler, C. A. (2017). Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics, 65*, 45–59.
Tsimpli, I., Sorace, A., Heycock, C., & Filiaci, F. (2004). First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism, 8*(3), 257–277.
Tucker G.R. (1999). *A global perspective on bilingualism and bilingual education*. ERIC Digest EDO-FL-99-04, Center for Applied Linguistics; Washington, DC.
Ulbrich, C., & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic/r/. *Journal of Phonetics, 45*, 26–42.
Unsworth, S. (2013). Assessing the role of current and CUMULATIVE exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition, 16*(1), 86–110.
van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review, 9*(4), 780–789.
Wagner, S. E. (2012). Real-time evidence for age grad(ing) in late adolescence. *Language Variation and Change, 24*(2), 179–202.
Wagner, S. E., & Sankoff, G. (2011). Age grading in the Montréal French inflected future. *Language Variation and Change, 23*(3), 275–313.

Ward, N., Sundara, M., Conboy, B., & Kuhl, P. (2009). Consequences of short-term language exposure in infancy on babbling. *Journal of the Acoustical Society of America, 126*(4), 2311.
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.
Yelland, G. W., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied Psycholinguistics, 14*(4), 423–444.
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.
Yoon, S.-Y. (2015). Acoustic properties of Korean stops as L1 produced by L2 learners of the English language. *Communication Sciences and Disorders, 20*(2), 178–188.