Journal of Phonetics 65 (2017) 45–59

Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model

Stephen J. Tobin [a,b,c,*], Hosung Nam [b,d,*], Carol A. Fowler [a,b]

[a] Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020, USA
[b] Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA
[c] Universität Potsdam, Department Linguistik, Haus 14, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany
[d] Department of English Language and Literature, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 136-701, South Korea

ARTICLE INFO

Article history:
Received 6 February 2015
Received in revised form 9 May 2017
Accepted 26 May 2017

Keywords:
Speech accommodation
Gestural drift
Phonetic drift
Bilingualism
Self-organization
Dynamical systems
Computational model

ABSTRACT

Studies of speech accommodation provide evidence for change in use of language structures beyond the critical/sensitive period. For example, Sancier and Fowler (1997) found changes in the voice-onset-times (VOTs) of both languages of a Portuguese-English bilingual as a function of her language context. Though accommodation has been studied widely within a monolingual context, it has received less attention in and between the languages of bilinguals. We tested whether these findings of phonetic accommodation, speech accommodation at the phonetic level, would generalize to a sample of Spanish-English bilinguals. We recorded participants reading Spanish and English sentences after 3–4 months in the US and after 2–4 weeks in a Spanish speaking country and measured the VOTs of their voiceless plosives. Our statistical analyses show that participants' English VOTs drifted towards those of the ambient language, but their Spanish VOTs did not. We found considerable variation in the extent of individual participants' drift in English. Further analysis of our results suggested that native-likeness of L2 VOTs and extent of active language use predict the extent of drift. We provide a model based on principles of self-organizing dynamical systems to account for our Spanish-English phonetic drift findings and the Portuguese-English findings.

1. Introduction

When language users speak with one another, whether they are native speakers of the same language (Pardo, 2006), of different varieties of a language (Babel, 2010; Kim, Horton, & Bradlow, 2011), different idiolects (Pardo, 2013) or of different languages altogether (Sancier & Fowler, 1997), adaptation takes place at every level of language (Bock, 1986; Dunstan, 2010; Hohenstein, Eisenberg, & Naigles, 2006; Liberman, 2012). For example, Pardo (2006), found that the speech of native speakers of American English became more similar during conversational interaction, as measured by listeners' judgments. Babel (2010) showed convergence of vowel quality of New Zealand English speakers towards those of Australian English speakers in a word repetition task. Sancier and Fowler (1997) found that the voiceless-stop voice onset times (VOTs) of a Portuguese-English bilingual converged to those of the ambient language as the speaker traveled between Brazil and the U.S. Likewise, it has been shown that syntactic (Bock, 1986), morphological (Dunstan, 2010), and semantic (Liberman, 2012) structures present in a speaker-hearer's ambient language influence subsequent language use.

These findings are interesting because the adaptations are often subcategorical or involve a change in the frequency of usage of a particular construction. Adaptations of this kind cannot be fully accounted for by a system based purely on discrete categories (Chomsky & Halle, 1968; McCarthy, 2003). They are better explained on the basis of *continuous* adaptation (Roon & Gafos, 2015; Spivey, Grosjean, & Knoblich, 2005). Additionally, it has been claimed that speaker-hearers are not able to acquire a second language (L2) authentically beyond the end of a critical or sensitive period (Lenneberg, 1967). Findings of ongoing modification of linguistic units over the course of spoken interactions suggest that these units and

* Corresponding authors at: Universität Potsdam, Department Linguistik, Haus 14, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany (S.J. Tobin). Department of English Language and Literature, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul 136-701, South Korea (H. Nam).

E-mail addresses: tobin@uni-potsdam.de, stephen.tobin@yale.edu (S.J. Tobin), hnam@korea.ac.kr (H. Nam), carol.fowler@yale.edu (C.A. Fowler).

http://dx.doi.org/10.1016/j.wocn.2017.05.006
0095-4470/

the speaker-hearers who exchange them have a degree of flexibility that the critical period hypothesis does not account for, which raises questions about how well the sensitive period claim generalizes across different language learners, as well as the factors underlying observed age-related declines in language-learning success (Flege, 1987b; Flege, Yeni-Komshian, & Liu, 1999; Krashen, 1975; Muñoz & Singleton, 2011). These observations motivate us to consider additional factors to explain between-speaker differences and to propose a dynamic approach to account for patterns of phonetic accommodation.

We briefly summarize a selection of studies focused on the phonetic level of this broad area of research, beginning with the more short-term and passive manipulations and leading to the more long-term and interactive ones, with the aim of identifying manipulations most likely to yield measurable accommodation. We then describe an investigation of learning at the phonetic level in bilingual speakers who spend weeks or months alternately in a predominantly L1 or L2 language community.

There is considerable variation in terminology applied to the kind of phonetic change we review and report here. Choice of terminology depends on whether the primary focus of the investigators is within or between languages, a first language (L1) or a second language (L2), and spontaneous or intentional imitation, among other factors (Tobin, 2015). For clarity, we define our terms here. We use the term *speech accommodation* (cf. Bourhis & Giles, 1977) to refer very broadly to these perceptually guided changes in speech production at any level of spoken language, whether within or between a speaker's languages. We use the term *phonetic accommodation* (or simply *accommodation*, because the phonetic level is our primary focus) to refer generally to speech accommodation at the phonetic level, whether within or between languages. In the case of people who travel back and forth between language environments, effects are more appropriately attributed to changes in ambient language rather than to the unique impact of one ambient language or another, rendering moot the question of within- vs. between-language effects. When relevant, we use the term *phonetic drift* to refer specifically to the subset of these findings that demonstrate cross-language effects of phonetic accommodation, that is, effects of exposure to an ambient language on another language that a speaker-hearer is familiar with but is not using at the time of exposure. However, we still use the terms *convergence* and *divergence* to denote the direction of an effect, irrespective of the language(s) involved in the effect. Both in monolingual, within-language contexts, and in between-language contexts, these terms refer to adaptation of a speaker's phones *towards* or *away from* those of interlocutors, or model speakers. The bilingual, between-language context, however, has the additional complexity that both L1 and L2 phones may converge towards or diverge from those of interlocutors or model speakers.

We begin with a study by Nielsen (2011), which illustrates the basic finding of within-language phonetic accommodation. She investigated whether speakers' VOTs in [p]-initial words would accommodate after hearing artificially shortened or lengthened VOTs in words of their native language, American English. She found that they did, that this accommodation generalized to new [p]-initial words that speakers had not heard, and to a novel place of articulation, [k]. However, her participants only converged to longer VOTs, not to shorter ones, perhaps because the latter accommodation pattern could affect the voiced-voiceless category boundary.

The authors of two different investigations of code switching, in which speakers change rapidly from one language to another within a single utterance, report reliable drift of Spanish-English and Greek-English bilingual participants' English VOTs towards Spanish (Bullock & Toribio, 2009) and Greek (Antoniou, Best, Tyler, & Kroos, 2011) VOT, respectively.[1] Evidence for convergence of the L1 VOTs towards English VOT, on the other hand, was much less reliable. Among the Spanish speakers, convergence was only observed in the early Spanish-English bilingual group ($n = 8$). Among the Greek speakers ($n = 16$) no convergence of Greek VOT towards English VOT was observed at all.

Likewise, investigations of word shadowing have produced a variety of findings of accommodation. For example, Babel (2012) reported that participants who rated images of a speaker as attractive converged towards the speaker in vowel quality when they shadowed words ostensibly produced by the speaker more than participants who rated the speaker as less attractive. Goldinger (1998) found that greater immediacy of shadowing, higher numbers of instances of shadowing the same word token, and lower frequency of shadowed words yielded greater accommodation as judged by listeners in a holistic word-similarity judgment task.

Using a more naturalistic task than shadowing, in which one member in each pair of interlocutors described a route across a map while the other tried to draw the route, Pardo (2006) observed the effects of a variety of social and conversational factors on phonetic accommodation in American English over the time scale of a conversational interaction. She found that interlocutors' vowel quality converged, as quantified in vowel $F1 \times F2$ space. In a perceptual and acoustic follow-up study using the same map task (Pardo, 2010), she found that a significant portion of the variance in similarity judgments of listeners who judged the words the speakers had produced, could be attributed to word duration and to the female participants. That is, words spoken by participants were more likely to be judged similar to the model's speech if words were longer and if participants were female. The investigation also showed that not all phonetic segments or other aspects of speech accommodate to the same extent; for example, high vowels accommodated more than low vowels. It also showed that accommodation in these different segments is also not necessarily equally perceptible, and that it is possible to observe both convergence and divergence within the same conversation (high vowels converged, while low vowels diverged).

Kim et al. (2011) extended Pardo's study by manipulating the language "distance" between members of participant pairs. Pair members shared or differed in dialect or in native language. Using a task in which interlocutors had to identify differences in two versions of a depicted scene through conversation, they found bidirectional convergence between same-dialect interlocutors of a common L1, but no clear evidence for accommodation between different-dialect interlocutors or different-L1 interlocutors. Although it is possible that

[1] Both Spanish and Greek, compared to English, have short-lag voiceless VOTs.

acoustic measurements would uncover convergence too small to be perceptible, it is interesting that evidence for accommodation was greater for interlocutors who were more similar to begin with, since these interlocutors have less room to converge than those who were less similar to begin with. From a rudimentary dynamical approach, a greater initial distance from the target would predict greater accommodation, assuming the interlocutors' productions are sufficiently close to one another for any influence to be exerted (see, for example Kopecz & Schöner, 1995).

In one investigation of phonetic learning during linguistic immersion, Chang (2012) recorded native English learners of Korean enrolled in an intensive language course. He found that they drifted towards Korean in their English voiceless-stop VOTs and their voiced- and voiceless-stop F0 onset frequencies in a word reading task. All of these changes constitute phonetic drift of the L1 towards the L2. This investigation confirms and extends Sancier and Fowler's (1997) and Major's (1992) observations that a speaker-hearer's L1 is not immune to the influence of later acquired languages.

In the present investigation we will focus on phonetic accommodation that is naturalistic in the sense that it reflects effects of ordinary linguistic interactions that took place outside the laboratory in different ambient language settings. We obtained spoken utterances from our speakers on two occasions after they had spent weeks or months in two different ambient language environments. In this way, we hoped to see the summed effect of their verbal interactions in each language environment. While lacking the immediacy of impact of accommodation induced in a laboratory task, this paradigm is likely more reflective of systematic changes to established phonetic categories, given that it shows the results of a longer period of interaction.

We aimed to replicate and extend a study by Sancier and Fowler (1997). They studied the speech of a bilingual speaker of Brazilian Portuguese and American English, measuring her voiceless-stop VOTs after periods of 2–4 months in the US and Brazil. They found accommodation in both languages toward the ambient language, whether English or Portuguese.

Flege (1987a) and Major (1992) also investigated patterns of phonetic accommodation in speakers who had lived in an L2-environment. However, their participants had been immersed in a language environment for longer periods than the participant in Sancier and Fowler's study (on the order of 12 years in the case of Major [1992]). Moreover, their measures were between-group comparisons of monolinguals and bilinguals. Following Sancier and Fowler (1997), we are interested in change in language forms within individual speaker-hearers.

In his Speech Learning Model (henceforth SLM), Flege (1995, 2007) proposes that internal representations of those phones of bilingual speaker-hearers' L1 and L2 that are sufficiently similar to one another (for example, the voiceless stops of their two languages), influence each other because they share a common acoustic–phonetic or phonological space (Flege, 1987a, 1995, 2007). Major (1992), Sancier and Fowler (1997), and broader inquiries into language by Cook (1991, 2003) demonstrate that these effects are bidirectional, as confirmed by Chang (2012).

The array of findings summarized above merit further exploration because they show that modification of phonetic categories occurs after the end of a critical or sensitive period for language acquisition, whether through improvement in the authenticity of L2 categories or through an effect of L2 accent on the L1 categories (Chang, 2012; Major, 1992; Sancier & Fowler, 1997), and because the modifications are of a continuous rather than a categorical nature. Thus, we are motivated to confirm and further specify the nature of phonetic category modification beyond the critical period and the conditions under which it occurs within a dynamical systems framework.

1.1. Present investigation

The purpose of the present investigation was threefold. First, we planned to replicate Sancier and Fowler's findings of phonetic accommodation caused by temporary migration. Though it is tempting to consider all of the different conditions under which phonetic accommodation may take place (listening, map task-based, shadowing, code-switching) to be equivalent in their effects on accommodation, investigators have only begun to scratch the surface of this field, and there may be systematic differences between them. Temporary migration fell at the longer-term and more intensive end of the range of conditions of exposure we reviewed that was also practical to implement within the timeframe of this investigation. Thus, short of long-term or permanent migration, we expected that these conditions would be as conducive as possible to phonetic accommodation.

Second, we intended to extend the coverage of studies of migration-induced phonetic accommodation. Sancier and Fowler's (1997) study was based on just one speaker of Brazilian Portuguese and English. We intended to test their finding of migration-induced accommodation, this time considering multiple participants with a different L1, Spanish, which has similar VOTs to Portuguese. Both Spanish and Portuguese have voiceless-stop VOTs on the order of 5–15 ms (cf. Rosner, López-Bascuas, García-Albea, & Fahey, 2000; Sancier & Fowler, 1997). We chose to investigate Spanish to test the generalizability of the findings. We aimed to look for phonetic accommodation on a time scale of weeks to months, to approximate the conditions of exposure of Sancier and Fowler's (1997) speaker. Sancier and Fowler (1997) found evidence for phonetic accommodation on a time scale of 2–4 months of exposure. Due to the constraints of our participants' travel itineraries we were only able to obtain data from participants after the shorter period of 2–4 weeks of continuous exposure to their L1. However, we still expected that continuous exposure to the L1 would yield greater accommodation than we could induce in a short experimental session.

Third, we proposed to apply principles of dynamical systems theory to account for patterns of cross-linguistic phonetic drift observed in this investigation as well as those of Sancier and Fowler (1997). We propose that participants' voiceless-stop VOTs in both of their languages will accommodate to those of their ambient language. Grosjean (1998) introduced the concept of language mode, asserting that bilinguals' activation of their L1 and L2 varies along a continuum, and that at any time, one of their languages may be predominantly active (monolingual mode) or both of them may be active (bilingual mode). We propose that patterns of articulatory coordination corresponding to participants' L1 and L2 phones are interdependent, and that, particularly in an experimental task in which they use both languages, participants are likely to be in bilingual mode, thus increasing the likelihood of between-language influence (cf. Grosjean, 1998). We propose that a speaker-hearer's production of vocal tract gestures with particular coordination patterns (for example, those characteristic of different VOTs) influences their production of similarly timed coordination patterns. Within a dynamical framework, a speaker-hearer's ongoing exposure to a particular distribution of VOTs (such as those of Spanish) through production and perception establishes this coordination pattern as a stable and attractive one, and exerts a degree of attractive force on similarly timed articulatory gestures. Thus the laryngeal-oral coordination patterns associated with a speaker-hearer's Spanish voiceless-stop VOTs would be predicted to influence those of the same speaker-hearer's English voiceless-stop VOTs as well as those of other languages (s)he may speak.

Although our focus in this investigation is exclusively the *voiceless* stops of Spanish and English, we acknowledge that Spanish voiceless-stop VOT values are closer to English *voiced*-stop VOT values than they are to English *voiceless*-stop VOT values, which might suggest a cross-language influence of Spanish voiceless stops on English voiced stops. The question of how these cross-language correspondences are established is not settled, although considerable attention has been given to it (e.g., Best, 1994; Best & Tyler, 2007). However, following Flege (1987a) and Chang (2012), we assume a correspondence or linkage between similar phonetic categories of bilinguals' L1 and L2, such that Spanish-English bilinguals associate their Spanish voiceless stops with their English voiceless stops. Such correspondences cannot be attributed to acoustic similarity alone, given findings that cross-linguistic segments are not always assimilated into a listener's closest L1 category (Strange, Levy, & Lehnholf, 2004) and are likely formed on the basis of perceptual similarity.

Our predictions and theoretical explanations are predicated on dynamical principles, between-language influence and ongoing exposure to the language and phones in question. The predictions and theoretical explanations of other investigators of phonetic accommodation and drift overlap with ours to some extent, but also differ in a number of ways. Although the communicative and experimental settings among these investigations vary widely, we consider the predictions of the four alternative approaches that most directly relate to this investigation in terms of cross-language effects on VOT. In his investigation of L1-English learners of Korean, Chang (2012) made parallel predictions to ours, namely, of bidirectional influence of L1 on L2 and of L2 on L1 in participants' VOTs. The results supported this prediction and he explained that a ubiquitous effect of an L2 on an L1 can be attributed to the establishment of a variety of cross-language linkages between the L1 and the L2 early in L2 learning, but without drawing on dynamical concepts. Antoniou et al. (2011), in distinction to Chang (2012) and the authors of this paper, predicted and found asymmetric cross-language effects in Greek-English bilinguals performing a code-switching task. They found no reliable effect of the L2 (English) on the L1 (Greek), but they did find that, even among bilinguals who were dominant in their L2 (English), the L1 (Greek) still had a pervasive effect on the L2. They explain that this is because the phonetic categories of the L1 are established earlier than those of the L2. Along the same lines as Antoniou et al. (2011) and Bullock and Toribio (2009) hypothesized that the L2 (English for some participants, Spanish for others) would be more vulnerable to influence from the L1 (conversely, Spanish or English), given the earlier establishment of L1 phonetic categories. However, they found that English VOTs exhibited drift, whether it was an L1 or an L2. They suggest that the greater variation among English voiceless VOTs (hence, larger phonetic target) may allow more drift because there is a larger range within which a speaker's category has the possibility to drift, while the lesser variation among Spanish VOTs (hence, smaller phonetic target) may prevent it because of this smaller range. They also suggested that variation among the observed patterns of cross-language influence may be due to variation in participants' language proficiency. Nielsen (2011) predicted and found VOT lengthening but not VOT shortening in her L1-English listening/reading task, explaining that the absence of shortening was due to phonological constraints. In particular, the short-positive VOT of the abutting /b/ category prevents accommodation in this direction.

Ignoring the finer distinguishing features of these investigations, we can extend their predictions to the present experiment as follows. Chang (2012) predicts a bidirectional effect of VOT between L1 (Spanish) and L2 (English), that is, shortening of participants' English VOTs and lengthening of their Spanish VOTs. Antoniou et al. (2011) and Bullock and Toribio (2009) both predict a primarily unidirectional influence of the L1 (Spanish) on the L2 (English; i.e., shortening of participant's English VOTs), the former because of the earlier established and pervasive effect of L1 phonetic categories, and the latter because of language-specific factors, namely, the restrictive VOT ranges of Spanish compared with the tolerant VOT ranges of English. Finally, Nielsen predicts a unidirectional influence of the L2 (English) on the L1 (Spanish; i.e., lengthening of the Spanish VOTs), but not of the L1 on the L2 (i.e., no shortening of the English VOTs) because lengthening of Spanish VOTs has no phonological consequences given the absence of a longer VOT phonetic category in Spanish, while shortening of English VOTs may lead to ambiguity between the /b/ and /p/ categories of English.[2]

In addition to investigating phonetic accommodation in voiceless-stop VOTs experimentally, we have created a model of the observed data to show that they can be accounted for within a dynamical, self-organizing framework. In this model, language mode is implemented by means of a variable *r*, which controls the extent to which the units of speech that a bilingual speaker-hearer hears influence similar speech units that (s)he produces. Thus, considering the within-language case, this variable constitutes the proclivity of bilingual speaker-hearers to accommodate in one of their languages, while in the between-language case of phonetic drift, it constitutes the

[2] We acknowledge that the prediction we make based on Nielsen (2011) is somewhat extrapolative, given that her investigation was focused on phonetic imitation in a monolingual context, while the other investigations we have considered take cross-language effects as their focus.

extent to which their productions of L1 and L2 phonetic segments influence each other.

This is the first application of the dynamical self-organizing framework to phonetic drift. There are a number of advantages of using this framework to model phonetic drift. First, a dynamical approach allows for accurate modeling of continuously varying articulatory variables and constraints, which is not possible with a purely categorical approach to linguistic phenomena. Second, and relatedly, a self-organizing approach reduces the need to stipulate units of linguistic structure, because in this approach, higher-level structural units form spontaneously as a result of interactions at lower levels of structure (Haken, 1999). Third, this approach has previously been applied in language processing modeling at the syntactic (Tabor & Hutchins, 2004), lexical-semantic (Li, 2009), and segmental phonological (Gafos & Benus, 2006; Spivey et al., 2005) levels, as well as to speech production (Saltzman, 1986) and speech perception (Tuller, 2004). Accordingly, the framework presents the possibility of an integrated approach to processing across levels of language and varieties of language use.

1.2. Predictions

We predict that Spanish-English bilinguals' voiceless-stop VOTs will accommodate bidirectionally with exposure to different ambient languages, in the same way that Sancier and Fowler's (1997) speaker accommodated. Specifically, their voiceless-stop VOTs in both of their languages will increase with exposure to English and will decrease with exposure to Spanish. While exposure to a language in code-switching or within a single experimental session appears to induce unidirectional effects due to the sequence in which languages were acquired and the size of phonetic targets, longer periods of exposure (i.e., weeks to months, cf. Chang [2012], Sancier & Fowler [1997]) appear to induce bidirectional effects. We will develop a model with an architecture consistent with established findings on speech learning, and we will simulate the effects obtained in the experiment, in order to show how a self-organizing dynamical approach can account for patterns of speech learning.

2. Method

2.1. Participants

Eleven native speakers of Spanish participated in the experiment. All of them were graduate students at the University of Connecticut or Yale University. They received $10 for each of the two 45-min sessions. Data from one participant were excluded because she primarily used English during her time in a Spanish-speaking country and used Spanish to a large extent while in the U.S. because of academic and language teaching commitments. This yielded a useable sample of $N = 10$. Table 1 provides a summary of the participants' linguistic backgrounds, some elements of which were collected with the help of the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007). Participants were encouraged to consider their past few months in ambient English when providing their current average exposure to each language. Although participants are native speakers of different varieties of Spanish, which have different mean voiceless-stop VOTs, the variation that has been reported among Spanish VOTs is relatively small in comparison with the difference between Spanish and English voiceless-stop VOTs. Rosner et al. (2000), for example, report ranges on the order of 5 ms for [p], and on the order of 10 ms for [t] and [k] between mean VOTs in varieties of Spanish. This is in contrast to between language differences on the order of 50 ms at each place of articulation (see, for example, Lisker & Abramson, 1964).

2.2. Stimuli

Twelve English sentences and 12 Spanish sentences were constructed. Each one contained at least one instance of the three voiceless stops [p], [t], [k] to be measured. Stimulus words were selected such that high-front, high-back, central and low vowels were represented among the post-consonantal vowels. The complete set of sentences is listed in Appendix A with target consonants in bold.[3]

2.3. Procedure

Participants were first recorded reading after they had spent three months or more[4] in Connecticut (i.e., outside of ambient Spanish), though most of them had already spent a year or more in Connecticut (see Table 1 for more details of participants' language experience). They were recorded for a second time after a period of 2–4 weeks in a Spanish-speaking country. The second recording was made within 24 h after returning to Connecticut. On both occasions, participants read five repetitions of the English sentences followed by five repetitions of the Spanish sentences. We chose this order of languages to maximize the possibility of recording participants speaking their 'best' most native-like English on the first occasion, and to avoid the possibility of confounding effects of the temporary migration with a possible effect of the Spanish reading task on their English speech. Participants were instructed in English to read at a comfortable, conversational pace.

2.4. Acoustic analysis

The first author measured the VOTs by hand using Wavesurfer (Sjölander & Beskow, 2000). Measurements were taken from the beginning of aperiodic energy in the waveform, corresponding to the end of the stop closure silence. The periodic energy attributed to vibration of the vocal folds for the vowel was selected as the end of the VOT. In cases of multiple stop bursts, measurements were taken from the beginning of the earliest release burst that was followed by continuous frication up to the beginning of periodic energy. In order to assess the reliability of the measures, the first author also remeasured 10% of the VOTs ($N = 720$). The reliability analysis yielded an intraclass correlation coefficient of 0.962 with a 95% CI of (.955, 0.967).

[3] A few instances of duplicate words were found after data collection. We measured the VOTs of non-duplicate voiceless stop-initial words which had already been recorded among the sentences of the stimulus set to substitute for these duplicates.

[4] This period of time was not selected as an inclusion criterion. It simply characterizes the time scale of participants' exposure to ambient English prior to their first recording.

Table 1
Summary of participants' linguistic background.

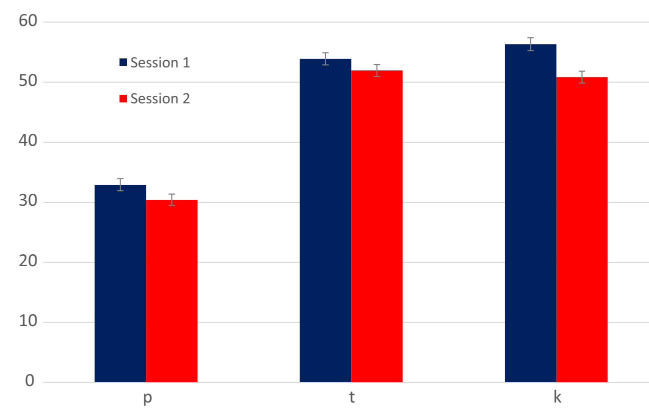
Participant	Country of origin	Total time in ambient English (years; months)	Age of arrival in the U.S. (years)	Age of English acquisition (years)	Current average exposure to English	Current average exposure to Spanish
1	Puerto Rico	9;0	21	5	85%	15%
2	Mexico	4;5	28	8	30%	70%
3	Ecuador	3;5	29	16	50%	50%
4	Spain	1;5	22	5	65%	30%
5	Argentina	1;3	25	8	55%	35%
6	Spain	2;1	25	7	50%	30%
7	Peru	5;6	28	11	48%	50%
8	Nicaragua	2;0	1	5	69%	30%
9	Puerto Rico	2;2	23	3	45%	45%
10	Mexico	1;6	27	10	39%	60%

Summary of some relevant aspects of participants linguistic background, collected at the time of the first recording (i.e., Session 1 = Ambient English). Participants were encouraged to consider their past few months in ambient English when providing their current average exposure to each language.

In order to assess the possibility that speaking rate varied systematically between conditions, the first author also measured the durations of the vowels following each of the VOTs under investigation.

3. Results

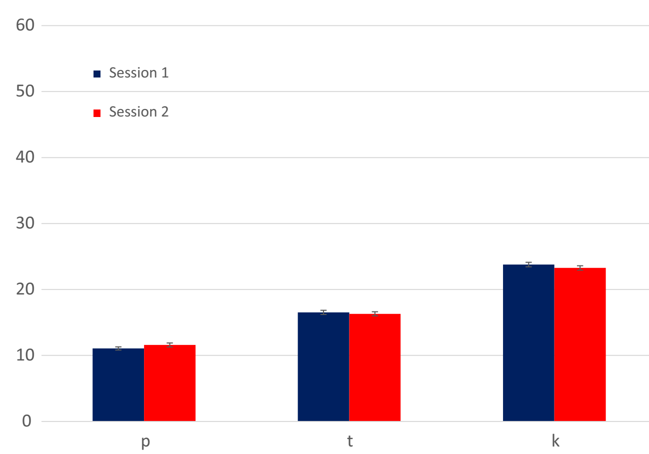
Findings of the experiment are shown in Fig. 1 for English and Fig. 2 for Spanish. We used linear mixed modeling in the R statistical package for the statistical analysis of the data (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2013). Linear mixed modeling is a type of regression analysis that includes both fixed and random effects (see, e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008). The fixed effects are the independent variables that are built into the design of the experiment (e.g., Recording Session [henceforth referred to as Session]: Ambient English, Ambient Spanish). The random effects are sources of variability that are not manipulated or experimentally controlled (e.g., Participant). We compared a series of models, adding independent variables to determine whether they improved the fit of the model to the data (Bates et al., 2015). We began with a model including Language and Consonant as fixed effects, a random intercept for Subject and by-subject random slopes for Language, Consonant and Session. Language and Session were both dummy coded with English and Session 1 as the default levels of these variables, respectively, and Consonant was coded as an ordinal variable ([k] = 0, [t] = 1, [p] = 2). The dependent variable was VOT. In our second model, the addition of Session as a fixed effect (thus, Language, Consonant and Session as fixed effects) significantly improved the fit of the model, as measured by the difference in deviance[5], or ΔD ($\Delta D = 33$, $p < 0.05$). We also tried adding the factor Age of Acquisition (widely reported as a predictor of language skill and accentedness [cf. Flege et al., 1999; Patkowski, 1980], and which had been collected during the experimental session), and then we tried replacing Session with this factor. However, neither adding the factor ($\Delta D = 14$, $p > 0.1$) nor replacing Session with it ($\Delta D = 9$, $p > 0.1$) improved the model fit. Thus, we report the results of the model including Language, Consonant and Session as fixed effects, a random intercept for Subject and by-subject random slopes for Language, Consonant and Session. In Table 2, we report the results of the model run with default factor levels as described above. However, in the text we also report the results of additional models in which the default factor levels were rotated in order to observe all simple effects and interactions.



Experimental results exhibiting a significant effect of Ambient Language on participants English VOTs.

Session 1 = Ambient English, Session 2 = Ambient Spanish.

Fig. 1. Effect of Ambient Language on English voiceless-stop VOT production.



Experimental results exhibiting no effect of Ambient Language on participants Spanish VOTs.

Session 1 = Ambient English, Session 2 = Ambient Spanish.

Fig. 2. Effect of Ambient Language on Spanish voiceless-stop VOT production.

[5] Deviance is −2LL or negative two times the log likelihood. See Bates et al. (2015) for further details.

Table 2
Summary of optimal mixed-effects model demonstrating phonetic drift.

Fixed Factor	Estimate	Std. Error	*t* value
(Intercept)	56.35	4.78	11.79
Lang	−32.58	4.60	−7.08*
Cons(k/t)	−2.45	1.19	−2.07*
Cons(k/p)	−23.42	1.55	−15.10*
Session	−5.50	1.24	−4.44*
Lang:Cons(k/t)	−4.80	1.19	−4.03*
Lang:Cons(k/p)	10.70	1.19	8.99*
Lang:Session	5.01	1.19	4.20*
Cons(k/t):Session	3.57	1.29	2.77*
Cons(k/p):Session	3.00	1.37	2.18*
Lang:Cons(k/t):Session	−3.24	1.69	−1.92
Lang:Cons(k/p):Session	−1.87	1.69	−1.11

Results of a mixed linear effects analysis of the VOT data obtained in the experiment. The default levels of the variables are as follows: Language = English (vs. Spanish), Consonant = [k] (vs. [t], vs. [p]), Session = Session 1 (vs. Session 2).
* Denotes $p < 0.05$.

As is to be expected, we observe a significant effect of Language such that Spanish has a shorter mean VOT than English for all three consonants across both Sessions, and a significant effect of Consonant, such that [t] and [p] have shorter mean VOTs than [k] in both English and Spanish and across both Sessions, with the exception that English [t] and [k] do not differ significantly in mean VOT in Session 2. Additionally, we observe a significant effect of Session on VOT, such that Session 2 yields a shorter mean VOT than Session 1 for the English consonants [k] and [p], but not for English [t] or for the Spanish consonants. The significant interaction of Language × Consonant indicates a greater mean VOT difference between [k] and [t] in Spanish than in English and a greater mean VOT difference between [k] and [p] in English than in Spanish. The significant interaction of Language × Session indicates a difference in English VOT means between Sessions 1 and 2, but not in Spanish VOT means. Also significant is the interaction of Consonant × Session at the default level of Language (English), indicating that VOT shortening we observe for [k] is greater than that observed for [t] or [p] and that the difference between the VOT means for [k] and [t] are significant in Session 1 but not in Session 2. This interaction is not significant for Spanish. Neither of the three-way interactions achieved significance, no matter how the levels of each factor were sequenced.

In order to test whether the individual variation in the drift in English induced by ambient Spanish might be due to varying levels of proficiency in producing native-like VOTs in L2 English, we conducted a further linear mixed model, regressing extent of drift in English on baseline VOT discrepancy from L1 English, using the same random effects structure as in our VOT analysis above. To quantify participants' VOT discrepancy (henceforth VOTD), we subtracted the mean VOTs of [p], [t] and [k] of Fowler, Sramko, Ostry, Rowland, and Hallé's (2008)[6] sample ($n = 32$) of native English speakers from our participants' VOTs in their first, ambient English, recording. On this scale, values closer to zero signify English VOTs typical of native English speakers, while values in the negative range are typical of Spanish-influenced English VOTs and values in the positive

[6] Fowler et al. (2008) stimuli were comparable to those used in the present investigation. Both sets of stimuli consisted of full sentences with target words containing voiceless stops with a range of following vocalic contexts.

Table 3
Summary of optimal mixed-effects model demonstrating effect of VOTD.

Fixed Factor	Estimate	Std. Error	*t* value
(Intercept)	14.16	2.513	−6.66*
VOTD	−0.44	0.03	−16.87*
Cons(k/t)	4.05	1.15	3.54*
Cons(k/p)	2.01	1.21	1.66
VOTD:Cons(k/t)	0.03	0.04	0.95
VOTD:Cons(k/p)	0.05	0.04	1.34

Results of a mixed linear effects regression of participants' drift in English L2 VOT on their discrepancy from typical American English L1 VOT (VOTD).
* Denotes $p < 0.05$.

range are typical of English clear speech VOTs. To quantify magnitude of drift, we subtracted participants' English VOTs in the Ambient Spanish condition from their English VOTs in the Ambient English condition, such that negative values indicate a reduction in VOT between recording sessions, and positive values indicate an increase in VOT. We added the factor Consonant to the model and found that it significantly improved the model fit ($\Delta D = 17$, $p < 0.05$). This second model is the one we report in Table 3. The intercept of the regression line is −14.16 (baseline consonant = [k]), indicating that the majority of cases show a reduction in VOT during participants' time in ambient Spanish. The negative value of the slope for VOTD shows that the extent of this reduction was greatest for large positive values of VOTD (VOTs typical of English clear speech) than for small positive or small negative values of VOTD (typical English VOTs). The regression line crosses the x-axis at approximately 30 ms of VOTD, such that at the largest negative values of VOTD there is a slight positive drift. The effect of consonant shows that the slope for [t] is significantly shallower than for [k] and that the slope for [p] is between those for [t] and [k] but does not differ significantly from either.

In order to test for possible effects of speaking rate on drift, we also initially ran additional parallel analyses of the durations of the vowels following each VOT, comparing the same series of model structures as for the VOT analysis, but with vowel duration as the dependent variable. We considered running an additional analysis using the VOT/vowel duration quotient as the dependent measure, but opted to use vowel duration as a measure of speaking rate for a number of reasons. First, we expected we would observe an effect of speaking rate more readily in vowel duration than in VOT/vowel duration quotient, given that English vowel durations are typically greater than English VOTs (see, e.g., Smiljanic & Bradlow, 2008). Second, given that some degree of error is an inevitable element of measurement, dividing these values would compound this error. This compounding of error coupled with the reduction in scale of magnitude in moving from VOTs to VOT/vowel duration quotients could dramatically reduce the systematic variability in the dependent measure. Given that Sancier and Fowler (1997) found phonetic drift on the order of 5 ms, we decided that it would be unwise to conduct such an analysis. Finally, although the VOT/vowel duration quotient has been used as a means to normalize VOT (e.g., Kang, 2014), VOT varies proportionally with vowel duration (Boucher, 2002; Kessinger & Blumstein, 1998), and both of these measures vary inversely with speaking rate. Thus, if effects of speaking rate were present in our data, we would not expect to be able to detect them with the VOT/vowel duration quotient.

We began our analysis of vowel duration with a model including Language and Consonant as fixed effects, a random intercept for Subject and by-subject random slopes for Language, Consonant and Session. Adding Session as a fixed effect did not significantly improve the fit of the model, ($\Delta D = 4$, $p > 0.1$), thus we fail to reject the null hypothesis that no effect of speaking rate was present in the data. Given that the factor under consideration did not improve the fit of the model we omit the results of the optimal model, which does not contain this factor. However, we provide the mean vowel durations by Language and Session. For English, the mean vowel duration in the first session was 122 ms and in the second session was 123 ms. For Spanish the respective means were 73 ms and 75 ms.

4. Discussion

We hypothesized that our results would parallel those of Sancier and Fowler (1997), that is, we expected that VOT in both of our bilingual participants' languages would accommodate towards the values of the ambient language. We corroborated this result for English voiceless stops but not for Spanish.[7] Our focus in this investigation is on change in voiceless-stop VOT as a function of ambient language as opposed to the specific VOT values themselves. In this regard, it may be of interest to return to the findings of Rosner et al. (2000). These authors reported variation on the order of 5–10 ms in mean voiceless-stop VOT among monolingual native speakers of different varieties of Spanish. Given that our sample is diverse with respect to their native variety of Spanish, and given the potential for speaker-specific patterns of variation in VOT (cf. Theodore, Miller, & DeSteno, 2009; Theodore, Myers, & Lomibao, 2015), focusing on change in VOT is the best approach to the present data.

The results reported here are not consistent with the predictions based on Chang (2012), who found bidirectional transfer of VOT between English and Korean, or with the predictions based on Nielsen (2011), who found accommodative lengthening of VOT within English, but not shortening, attributing this difference to the phonological constraint of the VOT boundary for /b/. However, the results are consistent with predictions based on Antoniou et al. (2011) and Bullock and Toribio (2009), the former attributing an observed absence of drift in the L1 to the earlier establishment of these categories, and the latter attributing it to the smaller VOT range among the shorter-lag Spanish voiceless stops.

The fact that VOT for English [k] exhibited the greatest reduction on exposure to Ambient Spanish could also be attributed to the fact that [k] has longer VOTs and, thus, has more room within which to accommodate (cf. Babel, 2012; Bullock & Toribio, 2009; Kim et al., 2011). However, the pattern of results observed for VOTD casts some doubt on this explanation (see p. 31). In contrast with Sancier and Fowler's (1997) findings in a Portuguese speaker, we found no effect of ambient language on participants' L1 Spanish VOTs. That is, the participant's articulation of their L2 voiceless stops changed but there was no significant change in their L1 voiceless stops. At one level this appears to contradict the findings of Flege (1987a), Sancier and Fowler (1997) and Chang (2012), all of whom report significant changes in participants' L1 and L2. Although the changes observed in Flege's (1987a) participants could be attributed to factors other than within-participant phonetic accommodation or drift because of its cross-sectional design, the latter two investigations were both longitudinal in design, and the consistency of effects in both languages among these three investigations is notable. The absence of an effect of ambient language on Spanish VOTs in this investigation suggests that the shorter VOTs may somehow be more stable and resistant to accommodation than longer VOTs (c.f., Tobin, 2015). It is conceivable that the absence of this effect may be in part due to the shorter period of exposure to the L1 (weeks) in the present investigation than in Sancier and Fowler (1997) (months). However, this pattern of resistance to accommodation in voiceless stops with shorter VOTs appears among the results of a variety of other investigations, whether explicitly commented on by the authors or not (e.g., Antoniou et al., 2011; Bullock & Toribio, 2009; Chang, 2012).

However, this does not account for the fact that Sancier and Fowler (1997) *did* observe accommodation towards the typical values of the ambient language in Portuguese voiceless-stop VOTs. Notably, in their code switching studies, Bullock and Toribio (2009) and Antoniou et al. (2011) both report more systematic change in the longer VOTs of English than in the shorter VOTs of Spanish or Greek, respectively. Thus, the fact that we only observed significant changes in the longer VOTs of English is supported by some patterns of phonetic accommodation and drift reported in the literature, though the authors do not explicitly make this observation. One explanation that could account for the difference in the pattern of results in the present investigation and those of Sancier and Fowler (1997) is our Spanish-English participants' reports on their language use in each of the two ambient language environments. They reported that they were engaged in the Spanish-speaking community while at the University of Connecticut (see Table 1 for details of self-reported exposure to English and Spanish over the few months prior to their first recording, spent in ambient English), and thus spoke both English and Spanish in the U.S., but that they spoke Spanish almost exclusively while in Spanish-speaking countries. That is to say, our participants alternated between bilingual Spanish-English mode, in which they were ready to use or switch into either language, while in Connecticut, and monolingual Spanish mode in the Spanish-speaking countries they visited. Sancier and Fowler's (1997) participant, on the other hand, spoke primarily English in the U.S. and primarily Portuguese in Brazil, that is, she alternated between languages but whether she was in the US or Brazil, she was in monolingual mode (Grosjean, 1998). We suggest that active verbal interaction in a language maintains the stability of its phonetic categories and minimizes potential drift in parallel with similar phonetic categories of other languages.

Another globally consistent explanation is that participants vary in the magnitude of their accommodation. One of our participants exhibited accommodation across all of the voiceless

[7] We note briefly, that participant eight arrived in the U.S. at a much earlier age than any of the other participants (see Table 1). This participant's mean VOTs, however, did not stand out from those of other participants. Although his mean English VOT actually increased slightly between recording sessions, he was not unique in doing so (participant seven's mean English VOT increased to a greater extent, and he acquired English at an age comparable to that of the remaining participants).

stops in both languages, with ambient language yielding mean differences of up to 15 ms, while two other participants exhibited no accommodation or drift at all. One source of this between-participant variation may be discrepancy from native-like English VOT in L2 English voiceless stops (VOTD). We found a significant effect of VOTD on drift magnitude. Larger, more positive values for VOTD yielded greater negative convergent drift towards Spanish VOT among the English voiceless stops, while large negative values of VOTD yielded a degree of positive divergent drift away from Spanish VOT. That is to say, voiceless stops with VOTs more typical of English, and especially those with longer VOTs typical of English clear speech, were those that drifted most markedly towards Spanish VOT, and that the most Spanish-accented English voiceless stops diverged somewhat away from Spanish VOT. In an investigation of accommodation strategies among English-Arabic bilingual children in semi-structured interactions with their mothers, Khattab (2013) reports findings that some bilinguals show sensitivity to fine phonetic detail, using Arabic-accented vs. English-accented variants of English phones and words as a means to express degrees of conformity/nonconformity with an Arabic-language task or interaction. In this way, Khattab identifies this capacity to accommodate or drift between accents or varieties of a language as a marker of linguistic skill and refers to it as *phonetic-switching* (cf. *code-switching*). It is a systematic source of language-specific variability that may be explicitly or implicitly recognized by an interlocutor. Our participants' task did not involve social interaction, as Khattab's (2013) participants' task did. However, we take the capacity of some of our participants to vary their productions of English voiceless stops between VOTs in the range of those of English clear speech and shorter VOTs closer to those of Spanish, indicating an influence of Spanish, as a prerequisite for this phonetic skill. The subsequent dramatic reduction of these participants' English VOTs in the ambient Spanish condition shows a capacity to modify the implementation of these stops, without which this phonetic skill could not be acquired. While there are likely to be many other sources of between-participant variation, variation in VOTD, as a precursor of the skill of phonetic-switching, is one highly likely candidate. Notably, this contradicts the prediction based on Bullock and Toribio (2009). According to that prediction, shorter VOTs have smaller targets and therefore afford too narrow a range to allow accommodation, but we found divergent accommodation among stops with VOTs that were considerably shorter than the English VOT norm.

Sancier and Fowler (1997) measured just one talker, and so it is not possible to know whether their findings of accommodation in both languages were due to her being consistently in monolingual mode (speaking either exclusively Portuguese or exclusively English). However, the influence of bilingual mode is plausible in spite of the single-speaker design of Sancier and Fowler (1997), given the self-reported differences in language use of the participants. Of the groups that Flege (1987a) investigated, those who were immersed in a foreign language showed the greatest divergent drift in their L1 from the respective L1 monolingual VOT norms (Americans in Paris, French speakers in Chicago). Likewise, Chang's (2012) L1-English participants were immersed in a Korean language environment and showed considerable drift in their L1. Importantly, Chang's participants had traveled to their destination with the specific intention of learning Korean from scratch, while this was not the intention of participants in the present investigation, who had already learned English to some degree of proficiency before traveling to the US. Thus differences in educational goals and motivations of these groups of participants, and concomitant differences in openness to linguistic influence may be a source of this variation.

Additional data are needed to fully answer the question of whether the absence of a significant effect of Ambient Language on Spanish VOTs in the present investigation, as compared to the Portuguese VOTs of Sancier and Fowler's (1997) experiment, is primarily attributable to bilingual language mode, the native language of participants' interlocutors, or to some other factor.

5. Self-organizing model

5.1. Dynamical systems

Nonlinear dynamic studies of human movement (e.g., Turvey, 1990) have provided a basis for understanding and modeling the temporal coordination patterns exhibited by action systems. The approach has been applied to patterns in speech production. For example, articulatory actions (or gestures) in discrete, nonrepetitive utterances can be viewed as motions of oscillating systems with critical damping (Browman & Goldstein, 1988; Saltzman & Munhall, 1989). Further, temporal coordination among such articulatory actions can be interpreted in terms of the interaction of multiple oscillatory systems (Nam & Saltzman, 2003; Saltzman & Byrd, 2000). Dynamical systems have also been used to model patterns in speech perception. Tuller, Case, Ding, and Kelso (1994) and Case, Tuller, Ding, and Kelso (1995) have provided a model that accounts for within-category stability and a variety of between-category transitions in the categorical perception of speech continua.

Self-organizing models are, by definition, dynamical models because they account for the evolution of phenomena over time. Self-organization is the emergence of higher-order structure that cannot be attributed to the internal structure of the elements of which a system is made (Browman & Goldstein, 2000; Oudeyer, 2002). The structure emerges from the interactions among the elements. The principles of self-organization and dynamical systems have previously been applied in studies of the emergence of inventories of speech sounds in communities of language users. For example, de Boer (2000) provides a self-organizing model that accounts for some of the observed cross-linguistic vowel patterns, such as the preference for symmetry in vowel systems and the preference for dispersion of vowels within a system.

We propose to develop a self-organizing model using Hebbian learning that captures the phonetic learning effects obtained in the experiments summarized above, among others. To provide a model of phonetic drift alone without the possibility of any wider applicability may not be especially revealing. As such, we propose an architecture that can, in principle, account for a range of findings on L1 and L2. The model acquires the speech categories of a first language followed by those of a second language from input from an ambient language, and then accommodates these categories to subsequent variation in input from the ambient languages.

5.2. Motivation

We developed a model to show that principles of self-organization can account for our results. The self-organizing model that we use for our simulations not only produces the effects of phonetic accommodation and drift observed in this investigation and in Sancier and Fowler (1997), but also provides a comprehensive tool for understanding speech learning in general. It can account for a variety of findings from the field. The capacity of the model to account for findings relating to phonetic accommodation follows naturally from the self-organizing principles of the model.

5.3. Learning algorithm

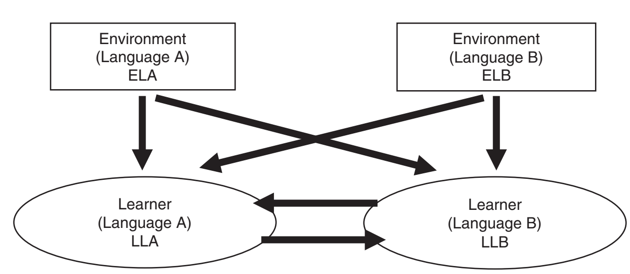
The model acquires L1 and L2 speech categories and accommodates them to new input by means of a single mechanism: Hebbian learning. Hebbian learning is an unsupervised learning algorithm, based on the dynamics of biological systems (Hebb, 1949). In this algorithm, (synaptic) connection strengths between nodes (neural units) are adjusted such that simultaneous activation of nodes leads to increases in the strength of the connections between those nodes. The algorithm only affects local nodes, unlike the algorithms of other neural network models in which the learning algorithm is applied over all nodes. This allows for continuous variation in learning rather than categorical application of the algorithm across all nodes, irrespective of distance.

Hebbian models of communicating agents with a set of neurons have been used to investigate various aspects of phonological and phonetic structure, including the partitioning of physical continua into discrete phonetic categories (Goldstein, 2003; Oudeyer, 2002, 2005, 2006), the structure of vowel systems (de Boer, 2001; Oudeyer, 2006; Wedel, 2003), consonant–vowel differentiation (Oudeyer, 2005), sequentiality in consonant sequences (Browman & Goldstein, 2000), and the emergence of syllable structure (Nam, Goldstein, & Saltzman, 2009). In this study, we employ an algorithm similar to that of Oudeyer (2002) as part of a generalized language acquisition and accommodation model to account for the acquisition of categories in two languages and the interaction of these categories across languages in a unified framework. Learning in this model is accomplished by self-organization under conditions imposed by attunement to the patterns implicit in, and presumably recoverable from, the phonetic environment structured by the ambient language. Such models have one or more agents, each of which has a set of nodes associated with numerical values on a physical continuum for a linguistic variable, here, VOT. Proximity between nodes along the continuum can be understood as connection strengths between them. The density distribution of the nodes along the continuum is equivalent to the probability of producing particular values along the range of the phonetic variable. As agents interact with one another based on the Hebbian algorithm, the connection strengths or proximity between nodes will be adjusted resulting in a change in the distribution; consequently, more global properties of the system emerge. For example, a mode may grow in the distribution where nodes are crowded, shift to a different location, or disappear as they get dispersed. For clarity, we state that here we are referring to statistical modes, not to language mode.

We propose a model of the phonetic skill of a bilingual who receives input from two language environments. We model the two knowledge systems of the learner and the two language environments with a network of four virtual agents, as shown in Fig. 3. The agents corresponding to the distribution of phonetic variables in a bilingual's knowledge of L1 and L2 are denoted by ellipses. These agents are connected to each other bidirectionally, which indicates that, irrespective of the ambient language(s) a bilingual interacts with, their L1 and L2 speech categories may affect each other, a finding supported by cross-language investigations of Flege (1995, 2007) and code-switching experiments of Antoniou et al. (2011). However, they are modeled as separate agents to allow for greater flexibility in the influence of each language. This approach is consistent with the assertions of Grosjean (1998) that bilinguals vary along a continuum from bilingual to monolingual mode in their daily lives.

The agents corresponding to the ambient languages with which the bilingual interacts are denoted with rectangles. The agents for the ambient languages represent the collective/average probability that language speakers will produce a phonetic variable of a particular value. We model these language agents without direct connections between them, because in general the community of speakers of one language (e.g., Spanish) that is geographically separated from the community of speakers of another language (e.g., English), do not affect the community of speakers of that other language. Likewise, arrows are unidirectional between ambient language agents and learners, because, whereas the language community is assumed to affect learning by individuals, individuals are assumed to have a negligible effect on the larger community, relatively speaking.

In a simulation of the learning of a speech category distributed along a VOT continuum, for example, the individual learner and the ambient language environment are modeled as agents. Each agent contains a set of nodes, V_i, each of which represents some value of VOT. Before learning has commenced, the beginning learner will have a uniform distribution of VOTs (all VOTs have equal probability of occurring) whereas the ambient language will show a probability distribution with a fully developed mode. One simulation iteration corresponds to the learner's learning process (or cycle) of a VOT. First, a VOT value is randomly sampled from the probability



Architecture of a model of a language learner exposed to two language environments.

Fig. 3. Language learning model of two languages for a single learner.

distribution of an ambient language. The learner hears this VOT veridically. At this stage, the learner's own VOT probability distribution is not updated. Next, the learner randomly selects a VOT value (V^{SEL}) from his/her probability distribution of VOTs. The attunement of the learner to the language environment is modeled by comparing the VOT heard in the environmental input to that selected from his/her own probability of VOT. This process would not necessarily involve actual production, which is reflective, in particular, of the early stage of phonological development. The learner's probability distribution is not intended to exclusively correspond to production or to perception, but to both, for simplicity and in recognition of the existence of a link between production and perception (Fowler, 1986; Liberman & Mattingly, 1985; Liberman & Whalen, 2000). If the difference between these two VOTs falls within a criterion tolerance, the selected VOT used by the learner on that trial is gated into the tuning (or learning) process, such that the learner tunes its VOT density distribution to increase the likelihood of producing that VOT. Otherwise, nothing happens. Those of the learner's VOT nodes that have values V_i similar to the values of V^{SEL} respond by increasing their level of activation as a function of the distance between V_i and V^{SEL} in VOT space. The "receptive field" of each unit-i is a Gaussian function with the position of the center of the peak, V^{SEL}, and the width of the function, σ. The Gaussian function defines the unit's activation level, G_i, as a function of this distance, according to (1).

$$G_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{V_i - V^{SEL}}{\sigma}\right)^2} \tag{1}$$

The values (i.e., VOTs) of all the units, V_i, are then attracted to V^{SEL} in proportion to their activation levels, according to the parameter-dynamics Eq. in (2), where V_i' is the new unit value, and r is a learning rate parameter.

$$V_i' = V_i - rG_i(V_i - V^{SEL}) \tag{2}$$

The result of this parameter-dynamic tuning process is an evolution of the density distribution of units along the VOT continuum in the direction of the distribution of VOTs in the input.

5.4. Architecture

The presence, absence, and strength of connections between agents, and the initial conditions of agents' internal states can be used as constraints in simulations. Depending on the choice of constraints, the system may evolve to have quite different properties. The importance of a constraint in the evolution of some property of interest can be evaluated by contrasting the results of simulations as the constraints vary. Note that the models are not meant to be faithful simulations of the detailed process of the ontogenesis of some property, but rather a way of testing the natural attractors of a simple system.

5.5. Simulations 1 and 2

We first describe how the acquisition of L1 and L2 categories can be modeled in our proposed model to show its generality. For the acquisition of L1 categories, we have no need to stipulate pre-existing categories within the model learner because they are formed on the basis of the L1 environment (ELA in Fig. 3), nor do we need to provide separate explanations for category formation on the one hand and category modification on the other. We initiated two agents: one for the learner and one for the ambient language environment. Each agent has a probabilistic distribution of VOTs. Initially, the language learner (LLA) has no speech categories, so the units representing VOT values are uniformly distributed across the VOT range, and the value of V^{SEL} is completely random. That is, at the outset of the simulation, values of VOT were assigned to these neural units in one-millisecond increments from −199 ms to 200 ms. Thus, $V_1 = -199$ ms, $V_2 = -198$ ms, ... $V_{400} = 200$ ms, for a total of 400 units. 200 ms distance in both directions, center at 0 ms VOT provides a large enough range to prevent simulation artifacts that are likely to occur at boundaries. On the other hand, the ambient language (ELA) had a mode at 0 ms, an approximation of the target VOT of Spanish and Portuguese voiceless stops. A connection between agents corresponds to a learning (or tuning) process described earlier, which is independently defined by the tuning parameters, σ and r. The parameter labeled σ determines the extent to which a perceived VOT value influences the surrounding VOT values, while r determines the strength of a connection. We set σ and r to 0.05 and 0.005, respectively, for simulations 1 and 2.[8] As learning progresses through attunement, the values of the neural units become clustered, with most units having values near the VOT mode, or most typical VOT, of the ambient language (L1).

Having introduced the parameters of these simulations, we return to some of the findings we will simulate. According to Flege (1987a), when producing an L2 sound for which there exists a "similar," corresponding sound in their L1 (e.g., /u/ and /t/ of French and English), L2 learners produce the L2 sound similarly to the L1 counterpart or intermediately between the corresponding sounds of L1 and L2. On the other hand, they are able to learn to produce a "new" L2 phone accurately, when there is no counterpart in their L1 (e.g., French /y/ for native English speakers).

Our model was extended to simulate the two cases ("similar" vs. "new" L2 phone): the formation of new L2 categories on the basis of the acquired L1 categories. The simulation includes four agents as in Fig. 3: two of these denote the language user's L1 and L2 phonetic categories (LLA, LLB), and the other two denote the L1 and L2 language environments (ELA, ELB). We used, as initial conditions, ELA and LLA distributions obtained from the first simulation. In ELB, we approximate English voiceless-stop VOTs as a distribution with a mode at 100 ms. The amount of exposure to a given language was modeled by the strengths of the connections between (i) the language environments of the first (ELA) and second (ELB) language and (ii) the language learner (LLA, LLB). These connections are displayed as vertical and diagonal arrows in Fig. 3. We note that, in the case of a monolingual language learner without exposure to any other language, the strength of the connection to the second language environment (ELB) would be zero. We set the simulation iteration to 10,000.

[8] The parameter values set for these simulations are arbitrary.

For the former case ("similar" L2), ELA and ELB have different modes at 0 ms and 100 ms respectively, established by an L1 acquisition simulation. These values approximate the short VOTs of Spanish voiceless stops and the longer VOTs of English voiceless stops and are sufficiently close to each other to yield mutual influence.[9] In addition, it is more reasonable to assume that LLB starts out with the final distribution of LLA with a mode at 0 ms rather than a flat distribution such that L2 speakers can utilize their existing L1 modes. This is similar to the Full Transfer in the Full Transfer Full Access hypothesis (FTFA) (e.g. Schwartz & Sprouse, 1994, 1996; White, 1989, 2003), in which L2 learners are proposed to initially transfer all of their L1 skill to their L2 (full transfer) under the constraints of Universal Grammar (full access to UG). Simulation results show that a final mode in LLB ended up at 50 (approx.) ms as a result of competition between ELA and ELB.

For the latter case ("new" L2 phone), ELB has a mode at 100 ms but ELA does not have a mode at 0 ms because in this case there is no similar phone in the L1 (ELA). In fact, since the new L2 phone does not exist in the L1, LLB is assumed to start with a flat distribution (i.e., no VOT mode at all). Results show that a new mode emerges at approximately 100 ms in LLB because LLB is affected by ELB alone. Neither case is entirely equivalent to Full Transfer Full Access (FTFA) in the sense that the self-organizing model does not assume any innate abilities.

5.6. Simulations 3 and 4

Next, we simulated phonetic accommodation in bilinguals due to changes in ambient language by modulating the connection weights from ELA and ELB to LLA and LLB (see Fig. 3). In our experiment, we found only English drift among a group of Spanish-English speakers who switched from bilingual language mode in the U.S. to monolingual L1 mode in Spanish-speaking countries. On the other hand, Sancier and Fowler (1997) found accommodation in both languages of a Portuguese-English speaker who switched from speaking only her L2 in the US (monolingual English mode) to speaking only her L1 in Brazil (monolingual Portuguese mode). To simulate the different patterns in phonetic accommodation, we used different initial settings of the connection strengths between ELA/ELB and LLA/LLB in order to reflect the different patterns of language use between the ambient language environments. In the Spanish-English case, only the L1 ambient language (ELA) affects speakers' L1 knowledge (LLA) in monolingual L1 mode (i.e., when the participants were speaking almost exclusively Spanish in Ambient Spanish) whereas both the L1 ambient language (ELA) and the L2 ambient language (ELB) affect speakers' L2 knowledge (LLB) in bilingual mode (i.e., when the participants were speaking both English and Spanish in Ambient English), as reported by the participants in this investigation (see also Table 1). On the other hand, in the Portuguese-English case, the L1 ambient language (ELA) affects speakers' L1 knowledge (LLA) only in monolingual L1 mode (i.e., when the participant spoke Portuguese in Brazil) while the L2 ambient language (ELB) affects speakers' L2 knowledge (LLB) only in monolingual L2 mode (i.e., when the participant spoke English in the U.S.). Initially, ELA and ELB have a mode at 0 and 100 ms, respectively. LLA and LLB have a mode near those of the corresponding ambient languages, which is obtained through an L1 acquisition simulation. For all the connections from language environments to language learners, σ is set to 200, while r is set differently depending on the connection type: within-language connections are stronger than between-language connections. Thus, r for ELA to LLA and ELB to LLB connections is set to 0.0125 whereas that for ELA to LLB and ELB to LLA connections is 0.00025. For the connections between LLA and LLB, r remained at 0.005. The simulation iteration time is set to 10,000. For the Portuguese-English symmetrical accommodation, we delinked the outgoing connections from the non-ambient language to simulate moving into a monolingual environment (delinking the connections out of ELB [English in this case] and delinking the connections out of ELA [Portuguese in this case]). For the Spanish-English asymmetrical drift, we delinked the connections out of ELB (English in this case) but did not delink ELA (Spanish in this case). By altering the connections in this way, we are able to simulate variation in language mode to some degree of approximation. After 10,000 simulation iterations, we measured the difference of the mode location between the different ambient language environments given a category. The Portuguese-English monolingual mode simulation yielded a 4.2 ms accommodation in Portuguese VOTs and a 3.9 ms drift in English VOTs. The Spanish-English monolingual to bilingual mode simulation yielded 0.6 ms of accommodation in Spanish VOTs and a 4.1 ms drift in English VOTs. See Fig. 4 for a comparison of the experimental and simulated data for the Spanish-English case.

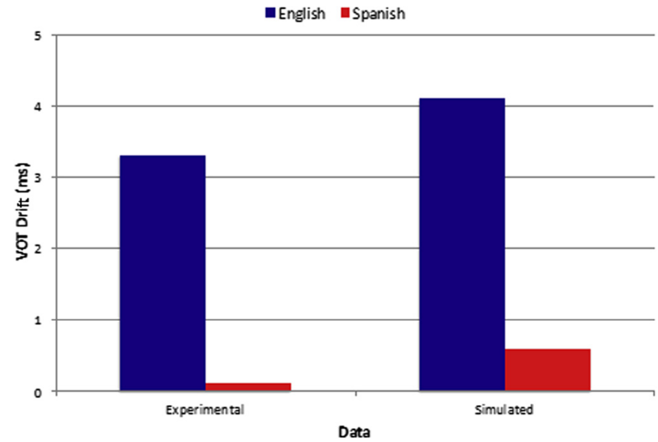


Fig. 4. Comparison of Drift in Experimental and Simulated Data Comparison of experimental and simulated effects of Ambient English vs. Spanish on speaker-hearers' English and Spanish VOTs.

6. General discussion

In this investigation we intended to replicate and extend findings of phonetic accommodation, and to show how these patterns of accommodation can be accounted for within a dynamical, self-organizing model. We replicated one of Sancier and Fowler's (1997) findings, namely, drift on the order

[9] Although these values differ somewhat from those obtained in the data set, they are used for simplicity of exposition in exemplifying the model. The difference from the observed means in the production data has no impact on the simulation results.

of 5 ms in participants' L2 English when speakers moved from the L2 language environment to the L1 environment. However, there was no corresponding accommodation in participants' L1 Spanish VOTs, possibly because of differences between these investigations in terms of participants' use of their L1.

The observed pattern of results is consistent with the predictions of Antoniou et al. (2011) and Bullock and Toribio (2009) since drift was observed among the L2 English VOTs but not among the L1 Spanish VOTs. The results are not consistent with the proposed predictions of Chang (2012) or of Nielsen (2011). Chang (2012) proposed and found a pervasive effect of an L2 (Korean) on the L1 (English) of language learners, while our results showed no impact of the bilingual participants' L2 on their L1. We might consider attributing this inconsistency to the greater experience of participants in the present experiment than those of Chang (2012). Chang (2013), for example, reports less drift among learners who are more familiar with an L2 than among those who are less familiar with it. However, adding age of acquisition as a factor in our mixed linear model did not improve the fit. Additionally, we found greatest drift among participants who produced typical English voiceless-stop VOTs and clear speech English VOTs when speaking English, that is, among those participants whose English VOTs suggest a considerably high degree of familiarity with English. Nielsen (2011) found increases in VOT but no parallel reductions in VOT among her American English participants, attributing these differences to an effect of the phonological category boundary between voiced and voiceless stops, while we found reductions in VOT but no parallel increases. These differences may, in part, be due to differences in the location of the voiced-voiceless category boundary between the two populations (native speakers of English vs. Spanish-English bilinguals). Garcia-Sierra, Diehl, and Champlin (2009), for example, report language-specific differences in category boundaries among Spanish-English bilinguals that also vary as a function of participants' confidence in each of their languages.

Further investigation of individual participants' VOTs showed that there was considerable between-participant variation in drift. One participant accommodated in both languages, while two others showed no accommodation at all. This between-participant variation has many possible sources. Our results suggest that discrepancy of L2 VOT from native-like English VOT accounts for some of this variation. English voiceless-stop VOTs that were within the range typical of native English speakers drifted towards VOT values typical of Spanish voiceless stops. Those of participants who produced VOTs typical of English clear-speech voiceless stops drifted most, while participants whose English voiceless-stop VOTs were considerably shorter than those typical of English voiceless stops drifted to a lesser extent and diverged from VOTs typical of Spanish voiceless stops. These differences likely reflect variation in phonetic skill. The divergent drift, away from the typical short Spanish voiceless-stop VOTs, exhibited by those speakers with strongly Spanish-influenced English voiceless-stop VOTs may reflect an effort to achieve a basic distinction between the phonetic categories of these two languages. Conversely, the convergent drift, towards typical short Spanish voiceless-stop VOTs, exhibited by those speakers with voiceless-stop VOTs in the range typical for native English may reflect greater skill and the capacity to adjust their VOTs for the purposes of adaptively distinguishing phonetic categories from one another when potential conflicts arise or to produce more subtle sociolinguistically meaningful phonetic variation. Khattab (2013), for example, has proposed that bilinguals can harness sensitivity to fine phonetic detail as a means to express their attitude towards a task or interlocutor. In a study of code-switching among English-Arabic bilingual children, she reports cases of switches to English during an Arabic story-telling task as a means to express disagreement or nonconformity with the task. Additionally, she describes cases of English words produced with Arabic phonetics as a means to conform to the expectation that the response to a question should be produced in Arabic. In some cases these are reported as instances of efforts to avoid producing an Arabic word, whereas in other cases, some of the children appear to use these Arabic-accented English words as Arabic word forms. Other possible sources of variation that should be considered in future investigations include the amount of variability in the category distribution. Spanish VOTs showed less within-category variability than English VOTs, and this tighter clustering, along with other characteristics of the distribution, may affect the likelihood that they will accommodate. The novelty vs. familiarity of the L2 to participants (Chang, 2013) may also affect the likelihood that participants will show accommodation, though in the present investigation all of the participants would be considered to be highly familiar with L2 English (see Age of Acquisition in Table 1). Finally, the active use of a language may also affect the likelihood of drift, more active languages undergoing less drift than less active ones. We propose to conduct further experiments to determine whether language activity/use and the amount of variability exhibited by a measure are factors that influence extent of accommodation.

Whatever motivating factors drive these observations of phonetic category change, this investigation adds to a growing body of literature documenting such changes. Taken together, they present a challenge to the strong form of the critical period hypothesis (Lenneberg, 1967) and suggest that more constrained claims of a sensitive period for language may need further elaboration to acknowledge the extent of phonetic category change reported in these investigations and the factors that modulate such change (Hakuta, Bialystok, & Wiley, 2003).

With our simulations, we have shown that a model based on self-organizing principles can account for the symmetrical and asymmetrical patterns of phonetic accommodation observed in Sancier and Fowler (1997) and in the present data set. By adapting the connections in the model to match participants' experience of ambient language, we observed symmetrical phonetic accommodation in the case of the Portuguese-English bilingual's sequence of monolingual modes (English, Portuguese, English) and asymmetrical drift in the case of the Spanish-English bilinguals' sequence of bilingual and monolingual modes (English + Spanish, Spanish). In running these simulations, we did not set the model to have normative English, Spanish and Portuguese VOTs. Rather, we allowed the model to acquire L1 and L2 categories in a plausible way, forming L1 speech categories by starting without any phonetic categories, then tuning to the categories of an ambient language, and learning a set of L2 categories on the basis of the L1 categories, all through Hebbian learning.

In summary, we have partially replicated findings of Sancier and Fowler (1997) that bilingual speakers VOTs accommodate to those of their ambient language. We have suggested that participant-specific patterns of language use may account for variation in accommodation, and have presented a model that can account for findings of phonetic accommodation and drift in one or both of a bilingual's languages.

Acknowledgement

This work was supported by NIH-NIDCD grant DC-002717 to Haskins Laboratories, and the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A2A03926788). We are also grateful to Alexander Demos for guidance on conducting statistical analyses using mixed effects models.

Appendix A

Stimulus list

A – English

1. Those **p**eople were **t**al**k**ing about my baseball **c**ap and shoes.
2. We **p**aid two thousand dollars for two cool widescreen **T**Vs last August.
3. John had soup and **t**ea at Bob & **P**ete's **c**afé.
4. Joyce **c**alled Anthony **t**oday but Luke picked up the phone and **p**assed it to Marie.
5. Kelsey **t**aught the **p**iano and flute to Adam, Pete and **K**aren.
6. Our vacations in Cape **C**od were happy **t**imes, **p**ossibly because we had good food.
7. Southern **C**alifornia's national **p**arks got **t**oo hot last week.
8. She **t**old a neat story about a **c**at who always **p**urred.
9. The weather **t**omorrow will be **c**ool so **p**ack some thermal underwear.
10. The **P**etersons lived in a small but beautiful house with a **q**uaint **t**errace on the back.
11. The basketball **t**eam **c**ame home having lost two games in **P**ennsylvania.
12. The **p**ost office has told its employees **t**o **k**eep their uniforms clean.

B – Spanish

1. Javier me ha **p**edido **q**ue no tarde mucho en **t**erminar el proyecto.
2. El señor del **p**elo rubio está sentado al lado del **t**ío **C**arlos.
3. La noticia de **q**ue se había rechazado su solici**t**ud le tomó **p**or sorpresa.
4. Es más sano comer fruta y vege**t**ales que azú**c**ar y grasa **p**ero a mí me gustan los caramelos.
5. En el supermercado se venden tomates, lechuga, **p**asteles y varios **t**ipos de **c**arne.
6. Julio a**p**agó la lucecí**t**a y se dio un golpe **c**ontra la **p**ared.
7. El sonido de la or**q**uesta rusa resonó por **t**oda la sala.
8. El **p**adre de **T**eresa me ha dicho que le bus**q**ue en la biblioteca.
9. El **p**erro ladra porque no le han dado de **c**omer desde la **t**arde del primer día de octubre.
10. **C**omo no sabía **c**ocinar bien, **P**ablo solía comer **t**ostadas con alubias.
11. Ana nos mandó una **t**arjeta **p**ostal con un dibujo precioso desde **C**osta Rica.
12. Al darse **c**uenta de que le habían visto el ladrón **p**uso **t**odo lo que podía en el bolso y se fue corriendo.

Bolded letters correspond to the voiceless-stop consonant VOTs that were measured as part of the investigation.

References

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39*(4), 558–570.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society, 43*(4), 437–456.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*(1), 177–189.

Bates, B., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). Retrieved from http://arxiv.org/pdf/1406.5823.pdf on October 6, 2015.

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

Best, C., & Tyler, M. (2007). Non-native and second language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 15–34). Amsterdam: John Benjamins.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Boucher, V. J. (2002). Timing relations in speech and the identification of voice-onset times: A stable perceptual boundary for voicing categories across speaking rates. *Perception & Psychophysics, 64*(1), 121–130.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (1st ed.. London: Academic Press. 119-15.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica, 45*, 140–155.

Browman, C. P., & Goldstein, L. G. (2000). Competing constraints on intergestural coordination and the self-organization of phonological structures. *Bulletin de la Communication Parlée, 5*, 25–34.

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

Case, P., Tuller, B., Ding, M., & Kelso, J. A. S. (1995). Evaluation of a dynamical model of speech perception. *Perception & Psychophysics, 57*, 977–988.

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*, 520–533.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Cook, V. (1991). The poverty-of-the-stimulus argument and multi-competence. *Second Language Research, 7*(2), 103–117.

Cook, V. (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.

de Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics, 28*(4), 441–465.

de Boer, B. (2001). *The origins of vowel systems*. Oxford: Oxford University Press.

Dunstan, S. B. (2010). Identities in transition: The use of AAVE grammatical features by Hispanic adolescents in two North Carolina communities. *American Speech, 85*(2), 185–204.

Flege, J. E. (1987a). 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. (1987b). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8*, 162–177.

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

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: York Press.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78–104.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics, 14*, 3–28.

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.

Gafos, A. I., & Benus, S. (2006). Dynamics of phonological cognition. *Cognitive Science, 30*, 905–943.

Garcia-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication, 51*(4), 369–378.

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

Goldstein, L. (2003). Emergence of discrete gestures. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th international congress of phonetic sciences*. Barcelona: Universitat Autònoma de Barcelona.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*, 131–149.

Haken, H. (1999). *Information and Self-organization: A macroscopic approach to complex systems*. New York: Springer.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science, 14*(1), 31–38.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Oxford England: Wiley.

Hohenstein, J., Eisenberg, A., & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish-English bilingual adults. *Bilingualism: Language and Cognition, 9*(3), 249–261.

Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics, 45*, 76–90.

Kessinger, R. H., & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *Journal of Phonetics, 26*, 117–128.

Khattab, G. (2013). Phonetic convergence and divergence strategies in English-Arabic bilingual children. *Linguistics, 51*(2), 439–472.

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.

Kopecz, K., & Schöner, G. (1995). Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biological Cybernetics, 73*, 49–60.

Krashen, S. D. (1975). The critical period for language acquisition and its possible bases. In D. Aaronson & R. W. Rieber (Eds.), *Developmental psycholinguistics and communication disorders*. New York: New York Academy of Sciences.

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

Li, P. (2009). Lexical organization and competition in first and second languages: Computational and neural mechanisms. *Cognitive Science, 33*(4), 629–664.

Liberman, K. (2012). Semantic drift in conversations. *Human Studies, 35*, 263–277.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1–36.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Science, 4*(5), 187–196.

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

Major, R. C. (1992). Losing English as a first language. *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.

McCarthy, J. J. (2003). OT constraints are categorical. *Phonology, 20*, 75–138.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching, 44*(1), 1–35.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. In F. Pellegrino, E. Marisco, & I. Chitoran (Eds.), *Approaches to phonological complexity* (pp. 299–328). Berlin, New York: Mouton de Gruyter.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. In P. Romeas (Ed.), *Proceedings of the XIIth international congress of phonetic sciences* (pp. 2253–2256). Aix-en-Provence: Université de Provence.

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

Oudeyer, P.-Y. (2005). The self-organization of combinatoriality and phonotactics in vocalization systems. *Connection Science, 17*(3–4), 325–341.

Oudeyer, P.-Y. (2006). *Self-Organization in the evolution of speech: Studies in the evolution of language*. Oxford University Press.

Oudeyer, P.-Y. (2002). Phonemic coding might be a result of sensory-motor coupling dynamics. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, & J.-A. Meyer (Eds.), *Proceedings of the 7th international conference on the simulation of adaptive behavior* (pp. 406–416). Cambridge: MIT Press.

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

Pardo, J. S. (2013). Phonetic convergence in college roommates. *Journal of Phonetics, 40*, 190–197.

Pardo, J. S. (2010). Expressing oneself in conversational interaction. In E. Morsella (Ed.), *Expressing oneself/expressing one's self: communication, cognition, and identity* (pp. 183–196). Hove, England: Psychology Press/Taylor & Francis.

Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30*(2), 449–468.

R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org/.

Roon, K. D., & Gafos, A. I. (2015). Perceptuo-motor effects of response-distractor compatibility in speech: beyond phonemic identity. *Psychonomic Bulletin and Review, 22*, 242–250.

Rosner, B. S., López-Bascuas, L. E., García-Albea, J. E., & Fahey, R. P. (2000). Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics, 28*, 217–224.

Saltzman, E. L. (1986). Task dynamic coordination of the speech articulators: A preliminary model. *Experimental Brain Research Series, 15*, 129–144.

Saltzman, E. L., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science, 19*. 499-27.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology, 1*(4), 332–382.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*(4), 421–436.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research, 12*(1), 40–72.

Schwartz, B. D., & Sprouse, R. A. (1994). Word order and nominative case in nonnative language acquisition: a longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 317–368). Amsterdam/Philadelphia: John Benjamins.

Sjölander, K., & Beskow, J. (2000). Wavesurfer—an open source tool. In X. Tang, B. Yuan, & T. Huang (Eds.), *Proceedings of the sixth international conference on spoken language processing, IV* (pp. 464–467). Beijing: China Military Friendship Publishing.

Smiljanic, R., & Bradlow, A. R. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics, 36*, 91–113.

Spivey, M., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences United States of America, 102*(29), 10393–10398.

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

Tabor, W., & Hutchins, S. (2004). Evidence for self-organized sentence processing: digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 431–450.

Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *Journal of the Acoustical Society of America, 125* (6), 3974–3982.

Theodore, R. M., Myers, E. B., & Lomibao, J. A. (2015). Talker-specific influences on phonetic category structure. *The Journal of the Acoustical Society of America, 138* (2), 1068–1078.

Tobin, S. J. (2015). *Phonetic accommodation in Spanish-English and Korean-English bilinguals: A dynamical account* (Unpublished doctoral dissertation). Storrs, CT: University of Connecticut.

Tuller, B. (2004). Categorization and learning in speech perception as dynamical processes. In M.A. Riley & G.C. Van Orden (Eds.), *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*. Retrieved from http://www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.

Tuller, B., Case, P., Ding, M., & Kelso, J. A. S. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance, 20*(1), 3–16.

Turvey, M. T. (1990). Coordination. *American Psychologist, 45*(8), 938–953.

Wedel, A. (2003). Self-organization and categorical behavior in phonology. In P.M. Nowak, C. Yoquelet, D. Mortensen (Eds.), *Proceedings of the 29th annual meeting of the berkeley linguistics society*, 611–622.

White, L. (1989). *Universal grammar and second language acquisition*. Philadelphia/Amsterdam: John Benjamins.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.