



## Research Article

# Perceptual categorization and bilingual language modes: Assessing the *double phonemic boundary* in early and late bilinguals <sup>☆</sup>

Joseph V. Casillas <sup>a</sup>, Miquel Simonet <sup>b,\*</sup><sup>a</sup> Rutgers University, United States<sup>b</sup> University of Arizona, United States

## ARTICLE INFO

*Article history:*

Received 16 January 2018

Received in revised form 26 June 2018

Accepted 20 July 2018

*Keywords:*

Speech perception

Bilingualism

Double phonemic boundary

Bilingual language modes

Voice onset time

## ABSTRACT

In the present study, Spanish-English bilinguals' perceptual boundaries between voiced and voiceless stops (*a* / *b* / *-p* / continuum including pre-voiced, voiceless unaspirated, and voiceless aspirated tokens) are shown to be modulated by whether participants are "led to believe" they are classifying Spanish or English sounds. In Experiment 1, simultaneous Spanish-English bilinguals and beginner second-language learners of Spanish labeled the same acoustic continuum in two experimental sessions (Spanish mode, English mode), and both groups were found to display language-specific perceptual boundaries (or session effects). In Experiment 2, early bilinguals and late second-language learners of various levels of proficiency participated in a single session in which, in random order, they labeled nonwords that were designed to prime either Spanish or English language modes. Early bilinguals and relatively proficient second-language learners, but not less proficient learners, displayed mode-specific perceptual normalization criteria even in conditions of rapid, random mode switching. Along with similar ones, the experiments reported here demonstrate that bilinguals are able to exploit language-specific perceptual processes (or norms) when processing speech sounds, which entails some degree of separation between their sound systems.

© 2018 Elsevier Ltd. All rights reserved.

## 1. Introduction

People who speak two languages often pronounce and perceive speech sounds differently from those who speak one language. For instance, a bilingual speaker of English and French is likely to produce and perceive the sounds of both English and French differently from how monolingual speakers of either English or French produce and perceive such sounds (Flege, 1987b; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008, among many others). This seems true of simultaneous,

early, and late bilinguals, but such differences are particularly noticeable in bilinguals who learned their second language as adults rather than as children (Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; Flege, 1987a; Flege & Hillenbrand, 1984; Fowler et al., 2008; Oyama, 1976; Pallier, Bosch, & Sebastián-Gallés, 1997; Piske, MacKay, & Flege, 2001; Sundara & Polka, 2008). The extant evidence suggests that the differences in phonetic behavior between bilingual and monolingual speakers are most importantly due to the fact that, sharing a common representational network, bilinguals' first- and second-language sounds interact (Best & Tyler, 2007; Flege, 1995, 2007; Simonet, 2016; Van Leussen & Escudero, 2015).

While the evidence in favor of cross-linguistic interaction in bilinguals is robust, it is still the case that bilinguals are not prevented from using language-specific sounds in a language-specific manner. For instance, /*p*/ is pronounced differently in French than it is in English—among other things, English /*p*/ is aspirated and French /*p*/ is not—and this difference is typically observed in bilingual, as well as in monolingual, speech. Caramazza et al. (1973), to cite one example, showed that early French-English bilinguals produced aspirated /*p*/ in their

<sup>☆</sup>Note: The experimental procedures utilized in the present study have been approved by a Human Subjects Protection Program supervised by an Institutional Review Board. All participants signed an informed consent form prior to their participation in the tasks reported in the present study. The authors declare that they do not have any conflicts of commitment. The first experiment reported in this study was part of the doctoral dissertation of the first author, which was supervised by the second author. The second experiment is the result of the collaborative effort of both authors. The present study benefited from a Comanche Nation Higher Education Grant and a University of Arizona Graduate and Professional Student Council (GPSC) Research Grant, both provided to the first author.

\* Corresponding author at: Modern Languages 545, 1423 E. University Blvd., Tucson, AZ 85721, United States.

E-mail addresses: [joseph.casillas@rutgers.edu](mailto:joseph.casillas@rutgers.edu) (J.V. Casillas), [simonet@email.arizona.edu](mailto:simonet@email.arizona.edu) (M. Simonet).

English productions, but not in their French productions (see also Antoniou, Best, Tyler, & Kroos, 2010; Magloire & Green, 1999; Olson, 2013). Moreover, adult second-language learners can form new phonetic categories specific to their second language. This does not mean that, in the bilingual mind, language-specific phonetic categories are truly independent from each other. Fowler et al. (2008) reported that, in the productions of French-English bilinguals, the length of the aspiration period of English /p/ was shorter than it was in the productions of English monolinguals (see also Flege, 1987b; Flege & Hillenbrand, 1984). In other words, bilinguals seem to be able to limit the production of language-specific sounds to the appropriate language, but this does not impede cross-linguistic interactions, which often take the form of low-level phonetic convergence across languages.

The broad research questions that motivate the present study are as follows: Bilinguals have been found to manifest language-specific production patterns (at least to some extent), but do they also demonstrate language-specific perceptual routines? In other words, **do bilinguals have separate perceptual categorization habits for the sounds of their two languages?** In this study, we report on two experiments in which early and late Spanish-English bilinguals were asked to categorize a /b-/p/ acoustic continuum under **two experimental conditions**, one in which they were led to believe they were listening to Spanish sounds and one in which they were led to believe they were listening to English sounds. We operationalize the questions above with the following, narrower questions: Do our experimental settings, which are designed to prime **language-specific processing modes in bilinguals**, modulate the way bilingual listeners classify an acoustic /b-/p/ continuum? And does bilingual language experience (early vs. late bilingualism, and linguistic proficiency) interact with experimental settings effects (if any)?

### 1.1. Review of the literature

The present study is particularly concerned with how Spanish-English bilinguals categorize English and Spanish voiced and voiceless bilabial stops, /b/ and /p/ (García-Sierra, Diehl, & Champlin, 2009; Gonzales & Lotto, 2013; Williams, 1977). In English (in utterance-initial position), both /b/ and /p/ are voiceless—typically, there is no voicing during closure—but, while /p/ is aspirated, /b/ is not. Thus, in English, the difference between /b/ and /p/ depends on aspiration, [p]-[p<sup>h</sup>] (Beckman, Jessen, & Ringen, 2013; Iverson & Salmons, 1995; Lisker & Abramson, 1964). In Spanish (in utterance-initial position), both /b/ and /p/ are unaspirated but, while /p/ is voiceless, /b/ is voiced—voicing begins during the closure in /b/ but not /p/. Thus, in Spanish, the difference between /b/ and /p/ is one of “true” voicing, [b]-[p] (Abramson & Lisker, 1972; Kirby & Ladd, 2016; Lisker & Abramson, 1964; Rosner, López-Bascuas, García-Albea, & Fahey, 2000). This asymmetry leads to the following scenario: A Spanish speaker is likely to perceptually categorize an unaspirated, voiceless bilabial stop, [p], as /p/ while an English speaker is likely to categorize it as /b/ (Elman, Diehl, & Buchwald, 1977). In other words, the perceptual (phonemic) boundary between /b/ and /p/ varies as a function of the language spoken by the listener because the phonetic implementation of these phonemes—but

perhaps also their phonological composition (Beckman et al., 2013; Iverson & Salmons, 1995)—is language specific.

What about Spanish-English bilinguals? The hypothesis we explore in the present study is that, in bilinguals, **perceptual categorization depends on language modes**. The term *language modes* refers to the state of activation of the bilingual’s languages (and their processing mechanisms) at a given point in time (Grosjean, 1985, 1989, 1998a, 1998b). According to Grosjean (1985), mode activation is determined by psychosocial and linguistic factors modulated by the communicative context of an interaction. For instance, a Spanish-English bilingual may be in Spanish unilingual mode when the interlocutor or the situation requires that only Spanish be used, whereas she may be in English unilingual mode when the situation requires that only English be used. In particular, the hypothesis is that, when in Spanish mode, Spanish-English bilinguals are more likely to categorize [p] as phonemically voiceless while, when in English mode, they are more likely to do so as phonemically voiced.

Exploring the perceptual routines of bilinguals and their potential interaction with language modes is not a recent trend. The first study about this investigated a sample of early, proficient French-English bilinguals from Québec, Canada (Caramazza et al., 1973; Caramazza, Yeni-Komshian, & Zurif, 1974). In one session, participants were placed in French mode by talking to them in French before the experiment, giving them experimental instructions in French, and asking them to pronounce a set of French materials. In another session, they were placed in English mode by talking to them in English, providing them instructions in English, and asking them to produce a list of English materials. Importantly, participants were asked to categorize the same acoustic /b-/p/ continuum in the two experimental sessions—the stimuli varied as a function of the shape and length of their voice onset times (VOT) and included pre-voiced, voiceless unaspirated, and voiceless aspirated tokens (Abramson & Lisker, 1972; Lisker & Abramson, 1964). The hypothesis was that, if language modes facilitate the activation of language-specific perceptual strategies, the same acoustic continuum would be categorized differently as a function of the experimental session. Interestingly, bilinguals’ classification patterns were found to be unaffected by language modes. Williams (1977), who replicated the studies of Caramazza and colleagues with a sample of Spanish-English bilinguals recruited in the United States, also produced null results. Williams, however, acknowledged that this “finding does not constitute proof that a language-specific set does not influence the perception of speech. It only indicates that the conditions of this experiment do not elicit such an effect” (Williams, 1977, p. 295). Since then, researchers have focused their effort on creating (and understanding) the conditions that (might) elicit this effect. Three subsequent studies were successful in doing so (Elman et al., 1977; Flege & Eefting, 1987; Hazan & Boulakia, 1993).

Elman et al. (1977) observed two limitations with the methodology of past studies. First, auditory stimuli were synthetic, which may have encouraged listeners to fall back on a language-general, rather than language-specific, perceptual routine. Second, while listeners may have been placed on the “correct” language mode towards the beginning of the labeling task (because they had just heard the instructions in

the appropriate language, among other things), they may have gradually slid back to a language-general routine as the experiment progressed. To overcome the two limitations, Elman and colleagues used naturally-produced stimuli, and they also reinforced mode activation by playing mode-appropriate filler words interspersed with the nonsense syllables (/ba/, /pa/) listeners were to label. Language modes were induced similarly to how they had been induced in past studies: in two sessions (Caramazza et al., 1973, 1974; Williams, 1977). Indeed, using this methodology, Elman and colleagues showed that their sample of Spanish-English bilinguals, recruited in central Texas, were more likely to classify voiceless unaspirated stops as phonemically voiced while in English mode but as phonemically voiceless while in Spanish mode.

Subsequently, Hazan and Boulakia (1993) and Flege and Eefting (1987) replicated these positive results with two different bilingual populations, French-English (Hazan & Boulakia, 1993) and Dutch-English (Flege & Eefting, 1987) bilinguals. In one study, /b/-/p/ auditory stimuli were resynthesized—VOT was manipulated artificially, but base materials had been selected from naturally-produced French or English words, thus creating language-specific “modes” in themselves (Hazan & Boulakia, 1993). The other study returned to using synthetic stimuli, but mode activation was reinforced throughout the task by asking questions in the mode-appropriate language throughout the length of the task. Both experiments found that mode manipulations led to differences in the way listeners labeled acoustic continua. Finally, a study on the perceptual behavior of Greek-English bilinguals found that category-goodness ratings, but not discrimination patterns, were affected by language modes (Antonioniou, Tyler, & Best, 2012).

At least one recent study has returned to this issue by focusing on pre-attentive listening rather than labeling, discrimination, or category-goodness ratings (García-Sierra, Ramírez-Esparza, Silva-Pereyra, Siard, & Champlin, 2012). García-Sierra and colleagues examined electrophysiological data gathered from a sample of Spanish-English bilinguals, recruited in central Texas. In this study, bilinguals’ discrimination of stop consonants varying in VOT was explored by analyzing their mismatch negativity response patterns in two conditions, one in which participants were silently reading a Spanish-language magazine and one in which they were reading an English-language magazine. It was demonstrated that the difference between [p] and [p<sup>h</sup>], contrastive in English but not Spanish, triggered a response when the listeners were silently reading an English-language magazine but not in the other condition. Relatedly, the difference between [b] and [p], contrastive in Spanish but not English, provoked a response in the Spanish-language condition but not the other one.

An important limitation of most of these studies lied on how language modes were induced. In all cases, participants were tested in two sessions. Listeners were exposed to an abundance of uncontrolled language-specific auditory input during and before each session (e.g., chatting before the experiment, listening to instructions, etc.), and this additional auditory input could have affected their subsequent auditory classifications in a manner unanticipated by the researchers. Studies have demonstrated that the results of perceptual categorization tasks are influenced by factors such as acoustic range effects

(Brady & Darwin, 1978; Keating, Mikoś, & Ganong, 1981) and the presence of adaptors or contrastive elements (Diehl, Elman, & McCusker, 1978; Samuel, 1982); that is, phonetic context can alter the way sounds are processed (Holt & Lotto, 2002, among many others). Bohn and Flege (1993) noted that the mode effects captured in the bilingual studies cited above could have been due in actuality to acoustic range, anchor, or contrast effects (Brady & Darwin, 1978; Diehl et al., 1978; Samuel, 1982), not to the activation of language modes and their dependent perceptual routines. In order to further investigate their observation, Bohn and Flege (1993) recruited two groups of participants, a bilingual and a monolingual group. Indeed, this study demonstrated that *both* groups of listeners, the monolinguals included, seemed to produce the “mode” effects observed before. This suggested that the results of prior studies, such as Elman et al. (1977), Flege and Eefting (1987), and Hazan and Boulakia (1993), were not incontestably due to the activation of language-specific perceptual routines, but could be attributed to general auditory factors induced by the presence of uncontrolled, environmental auditory input: the conversation preceding the task.

An important lesson of Bohn and Flege’s (1993) study is that, in order to demonstrate that bilinguals possess the ability to utilize different perceptual routines to parse speech in their two languages, one of at least two things need to be true. First, one must demonstrate that “mode” effects are larger in bilinguals than in monolinguals, or that at least bilingual linguistic proficiencies interact with the size of such effects (Bohn & Flege, 1993; e.g., Elman et al., 1977; García-Sierra et al., 2009). Second, one must induce appropriate language modes in a manner that does not depend on providing participants more (environmental) acoustic input in one language than the other and is thus not suspicious of phonetic context effects (Gonzales & Lotto, 2013). Both strategies have been pursued, and we do so here as well.

García-Sierra et al. (2009) recruited, in central Texas, a group of Spanish-English bilinguals and a group of English-speaking monolinguals. This study fundamentally replicated past studies (Bohn & Flege, 1993; Caramazza et al., 1973, 1974; Elman et al., 1977; Flege & Eefting, 1987; Hazan & Boulakia, 1993; Williams, 1977). Along with Bohn and Flege (1993), García-Sierra and colleagues found that “mode” effects were present in both bilinguals and monolinguals. However, unlike in Bohn and Flege (1993), it was shown that the size of such effects was modulated by self-assessed linguistic proficiency—that is, more proficient bilinguals were shown to display a larger effect size than less proficient bilinguals. Thus, above and beyond any effects of phonetic context, found in both bilinguals and monolinguals, proficient bilinguals seemed to display language-specific perceptual norms.

A less ambiguous demonstration of this was reported in Gonzales and Lotto (2013). These researchers managed to create the conditions in which bilinguals, but not monolinguals, would be induced to display the effects of language modes. A group of Spanish-English bilinguals and English-speaking monolinguals from southern Arizona were recruited for a single session in which they were asked to label as phonemically voiced or voiceless a /b/-/p/ continuum varying in VOT. The stimuli were formed by a nonce syllable (/ba/-/pa/), which was to be labeled, and a sequence of sounds that was

appended to each token syllable. Such nonce words took an “English” form by using an English-specific appendix and a “Spanish” form by using a Spanish-specific appendix. The syllable *-fri*, which was appended to all target tokens, could take the form [fri], a combination that exists in Spanish but not in English, or [fɹi], a combination that exists in English but not in Spanish. The *bafri-pafri* tokens were played in two conditions. Some participants, both bilinguals and monolinguals, received the “Spanish” version (*ba*[fri]-*pa*[fri]) and others, both bilinguals and monolinguals, received the “English” version (*ba*[fɹi]-*pa*[fɹi]). Importantly, all interactions with all of the participants, bilinguals and monolinguals, took place in English. This minimized any potential effects of phonetic context, such as range and anchor effects, that could be due to pre-task conversations. The mode manipulations, therefore, were based exclusively on the shape of the appendix syllable, *-fri*. The results demonstrated that the perceptual categorization patterns of the bilinguals, but not those of the monolinguals, differed as a function of whether they had been exposed to the “Spanish” or the “English” version of the *bafri-pafri* continuum. This finding, in our view, strongly suggests that bilinguals—and, obviously, not monolinguals—utilize language-specific perceptual categorization routines to process incoming speech. This finding might constitute the least ambiguous demonstration that, in bilingualism, language modes affect perceptual categorization.

The present study takes off from that of [Gonzales and Lotto \(2013\)](#) and extends it in three important ways. First, noting that their bilinguals had all acquired both English and Spanish as children, the authors mention that “an important question for future research is whether these participant’s results generalize to bilinguals who acquired their second language later in life. One possibility is that such ‘late bilinguals’ do not develop language specific phonetic systems, because representing a second language is more difficult once native-language categories are firmly established” ([Gonzales & Lotto, 2013, p. 2140](#)). Some studies suggest that linguistic experience (or proficiency) could interact with mode effects (e.g., [Elman et al., 1977](#); [García-Sierra et al., 2009](#)), but others do not ([Flege & Eefting, 1987](#)). At any rate, such demonstrations might have been driven by phonetic context effects, as discussed above. It is, therefore, important to test participants across a spectrum of linguistic experiences using [Gonzales and Lotto’s \(2013\)](#) methodology. Second, [Gonzales and Lotto \(2013\)](#) used a between-listener design—that is, they showed that bilinguals induced in Spanish mode differed, in their labeling patterns, from bilinguals induced in English mode. To clearly demonstrate that bilinguals switch back and forth between two sets of perceptual norms, one must show mode effects in a within-listener design. Finally, one should wonder about the speed at which bilinguals may switch between one set of perceptual criteria and another. All prior demonstrations of this effect have compared data collected in different experimental sessions. What about switching modes within the same experimental session?

### 1.2. The present study

The present study investigates the *double phonemic boundary effect* in bilinguals, a term we borrow from [García-Sierra](#)

[et al. \(2009\)](#). We are concerned with how Spanish-English bilinguals categorize a /b/-/p/ acoustic continuum varying in VOT, and we ask whether bilinguals utilize different categorization criteria to label the continuum, a set of criteria when in Spanish mode and another when in English mode. Even though mode-specific perceptual categorization has been explored using various methodologies (e.g., [Elman et al., 1977](#); [García-Sierra et al., 2012](#)), we follow the majority and compare the labeling (fitted sigmoid) functions obtained from the responses to a continuum. Therefore, in our study, demonstrating the presence of mode-specific perceptual norms is akin to demonstrating the presence of a *double phonemic boundary*, a significant difference between the labeling functions obtained when listeners find themselves in Spanish mode and those obtained when they are in English mode ([Bohn & Flege, 1993](#); [Caramazza et al., 1973, 1974](#); [Flege & Eefting, 1987](#); [García-Sierra et al., 2009](#); [Gonzales & Lotto, 2013](#); [Hazan & Boulakia, 1993](#); [Williams, 1977](#)). To avoid the potential experimental artifacts (e.g., phonetic context effects) of most comparable studies, we settled on the method utilized by [Gonzales and Lotto \(2013\)](#). Unlike in [Gonzales and Lotto \(2013\)](#), however, mode effects are demonstrated in within-listener comparisons.

The objectives of the present study are twofold. Firstly, we want to assess whether both early and late bilinguals display the *double phonemic boundary* effect in their categorization functions and whether there are differences in the size of such an effect as a function of linguistic experience or proficiency. Secondly, we want to investigate how fast can bilingual listeners switch from a set of mode-specific perceptual routines to another. In this study, we explore the behavior of simultaneous and early Spanish-English bilinguals as well as that of late second-language learners of Spanish whose first language is English. The latter group includes learners at the very early stages of their acquisition as well as intermediate learners. We report on two experiments. In the first experiment, like in all comparable studies, the two modes are induced in separate experimental sessions—one session for Spanish mode and one for English mode. In the second, the two modes are induced within the same session. Thus, in the second experiment, we assess whether bilinguals can switch back and forth, in real time, between mode-specific perceptual categorization habits.

## 2. Experiment 1

The purpose of our first experiment was twofold. Our first goal was to assess whether simultaneous Spanish-English bilinguals demonstrate the *double phonemic boundary* effect with a methodology that fundamentally replicates that of [Gonzales and Lotto \(2013\)](#)—i.e., the cue for language mode is provided exclusively in the auditory stimuli and not in the environment. We move beyond an exact replication in that we utilize a within-subjects design. Our second goal was to investigate whether late second-language learners of Spanish whose first language is English display the *double phonemic boundary* effect even at the very initial stages of their learning. This extends prior findings to a new population. The studies that have recruited late second-language learners seem to have been focused on relatively advanced learners.



## 2.1. Method

### 2.1.1. Participants

A total of 20 people participated in a perceptual identification experiment involving two sessions. The participants were classified in one of two groups. The first group was comprised of ten simultaneous Spanish-English bilinguals born and raised in southern Arizona. The second group was comprised of ten adult second-language learners of Spanish whose native language is English. The learners were enrolled in a seven-week introductory Spanish language course that took place in Middlebury, Vermont. All 20 listeners participated in the same perceptual categorization experiment and, thus, labeled the same auditory stimuli.

The ten simultaneous bilinguals recruited for the study reported speaking both English and Spanish for as long as they could remember, and that their parents were also bilingual. All participants also stated that they used both Spanish and English on a daily basis with friends and family. They were asked to complete two of the four modules of the Bilingual Language Profile (BLP) questionnaire (Birdsong, Gertken, & Amengual, 2012). The BLP is comprised of four modules related to language history, use, competency, and attitudes. Each module is scored for the two languages in question. The totals are then summed, and a composite language dominance score is derived by subtracting one language's score from that of the other. The two modules we focused on were the language history (range [−120, 120]) and the proficiency (range [−24, 24]) portions. For the present investigation, a negative score indicates dominance in Spanish and a positive score indicates dominance in English. A score near zero is taken as an indication of balanced bilingualism or balanced bilingual proficiency. The group mean for language history was 4.63 (SD = 7.65, 95% c.i. = −0.11, 9.37), and for proficiency it was 0.45 (SD = 6.22, 95% c.i. = −3.40, 4.31). The BLP score for this group was 5.08 (SD = 13.25, 95% c.i. = −3.13, 13.30), suggesting “balanced” bilingualism or comparable Spanish-English competencies.

Ten beginning learners of Spanish with English as their native language were recruited. The learners were students in the Middlebury College Language School, which is a domestic immersion program in which all participants sign a formal pledge promising to use only the target language (Spanish, in this case) throughout a seven-week course. Students live in the residence halls on the campus of Middlebury College with other students and professors, and they attend class for four hours in the morning and participate in co-curricular activities in the afternoon. Typically, students and faculty take the language pledge seriously, as a failure to speak in any language other than the target language can result in expulsion from the program. The participants in our study completed a background questionnaire (not the BLP) in order to ensure that their native language was indeed English, that they had not previously matriculated in more than one semester of Spanish (or any other second language), and that they had not spent a significant amount of time (more than one month) in a Spanish-speaking country. Even though the program offers pronunciation clinics, these are not mandatory, and none of the participants in this experiment attended any of the clinic sessions. The ten participants in this group were, therefore, late

second-language learners of Spanish at the initial stages of their linguistic development — they had very little experience with Spanish. The learners' data were collected during the last week of their stay in the language program.

### 2.1.2. Procedure

Two resynthesized acoustic continua were used in the two experiments reported in the present study. The acoustic continua manipulated the voice onset times (VOTs) of a set of bilabial stop consonants in the context of vowel [a], /ba-/pa/. The continua were created using Praat (Boersma, 2001) in, essentially, a two-step process. First, we created a base VOT continuum (detailed below) and, second, we made two copies of the base continuum and appended Spanish-like or English-like segments to them to create “Spanish” and “English” acoustic continua, respectively. The “Spanish” continuum was created by appending the sequence [fri], a sequence that is phonotactically legal in Spanish but not in (American) English. The “English” continuum was created by appending the sequence [fɹi], a combination of sounds that occurs in English but not in Spanish. Appending the aforementioned segment sequences to the bases resulted, in both languages, in phonotactically-legal nonwords, *bafri* and *pafri*. We refer to the /bafri-/pafri/ continuum as the “Spanish” continuum, and to the /bafɹi-/pafɹi/ one as the “English” continuum. Note that the /ba-/pa/ portion of the two continua was identical.

A two-alternative forced-choice labeling task was administered in two sessions. The “English” and the “Spanish” continua were presented in separate sessions a minimum of 24 h apart. The order was not counterbalanced—i.e., the “Spanish” version was presented first to all listeners. The interactions between all of the participants and the experimenter took place in Spanish. The experimental instructions were also given in Spanish. This was true of both experimental sessions. This was important to avoid the possibility of acoustic range effects, explained above. Thus, we made an effort to place all of our participants in unilingual Spanish mode in both sessions, and we relied exclusively on the auditory stimuli to activate the Spanish and English modes. This design is essentially a replication of Gonzales and Lotto (2013), except that our participants were tested twice and our base language was Spanish rather than English.

Participants completed the task in a quiet room, individually. The experiment was run in PsychoPy (Peirce, 2007), based in Python. The program presented one stimulus per trial in a single randomized block, drawing from only one continuum per session, either /bafri-/pafri/ or /bafɹi-/pafɹi/. Simultaneous with the auditory stimuli, visual renderings of the target nonwords (*bafri*, *pafri*) appeared on the screen. The participants were instructed to choose, using a DirectIn Rotary controller (Empirisoft Corporation, New York, NY), which nonword they thought they had heard. A red cross appeared in the middle of the screen between trials. There was no set time limit for each trial, but participants were instructed to respond as quickly as possible. The inter-trial interval was set at 500 ms. The participants finished the task in approximately 16 min, eight minutes per session.

### 2.1.3. Stimuli

A 23-year old female Spanish-English simultaneous bilingual born and raised in Arizona provided natural productions

of disyllabic words containing bilabial stops in utterance initial position to be used as stimuli. The talker grew up in southern Arizona, in a border town, in a Spanish-English bilingual household. Her parents and siblings are bilingual. The talker is a phonetician, and she is familiar with the phonetics of both Spanish and English. We used an AKG C520 (Vienna, Austria) condenser microphone to record the talker and a Sound Devices USBPre 2 (Reedsburg, Wisconsin) audio interface to digitize the signal at 44.1 kHz and 16-bit quantization. The signal was recorded onto a laptop computer running Praat (Boersma, 2001). For resynthesis, we then selected the best token of [p<sup>h</sup>] (aspirated, long-lag /p/), which was determined by the absence of clicks or extraneous noise.

The VOT of the stimuli ranged from -60 to 60 ms in 10 ms increments; the range thus includes tokens with pre-voiced, short-lag, and long-lag VOTs. For the stimuli with positive VOT, the aspirated portion of the stop was manipulated via the Time-Domain Pitch-Synchronous-Overlap and Add algorithm (TD-PSOLA) available in Praat. Only the duration of the aspiration period was modified: it was shortened. For the stimuli with negative VOT, periods of prevoicing were pasted into the signal at zero-crossings before the release of the stop. The pre-voiced portions were taken from phonetically voiced stop productions of the aforementioned talker. The stimuli were then normalized for peak intensity.

The talker also provided natural productions of the sequences [fri] and [fi]. These recordings were used to create the appendices to be pasted to the two continua used in the experimental task—i.e., one considered to be Spanish-like ([fri]) and one English-like ([fi]). The sequences were recorded with the aforementioned audio setup.

Using the same criteria detailed above, the best tokens of each *fri* sequence ([fri], [fi]) were appended to each stimulus from the original /ba/-/pa/ base continuum. Thus, two separate continua were created, a “Spanish” continuum ranging from [bafri] to [p<sup>h</sup>afri] and an “English” one ranging from [bafi] to [p<sup>h</sup>afi]. Crucially, the nonword continua ranged from *bafri* to *pafri* based solely on VOT, while *mode* (Spanish, English) was induced solely by the penultimate segment, *r*. To be clear, there was only one acoustic VOT continuum, and this continuum was used across both *mode* conditions: the appendices had been copy-pasted “to the right” of the target continuum.

#### 2.1.4. Analyses

The data, consisting of the participant’s responses to the auditory stimuli, were analyzed using a series of generalized linear mixed effects models. Given the categorical nature of the participants’ responses, the models utilized a binomial distribution with a logit link function. The criterion, ‘voiceless’ responses, was modeled as a function of the hypothesized *mode* of the continuum (“Spanish,” “English”), and VOT ([-60, 60] ms). Causal priority was given to *mode*. Our main research question is whether bilinguals and late learners demonstrate the double phonemic boundary effect (García-Sierra et al., 2009), which, in our design, results from a significant difference in the participants’ responses as a function of the factor *mode*.

All statistical models included the maximal error structure justified by our design (Barr, Levy, Scheepers, & Tily, 2013)—i.e., participants were given random intercepts with

random slopes for *mode*. Main effects and higher order interactions were assessed by hierarchically partitioning the variance using nested model comparisons. Unless otherwise noted, the fixed factor *mode* was treatment-coded with “Spanish” set as the baseline. Marginal  $R^2$  ( $R^2_m$ ) and conditional  $R^2$  ( $R^2_c$ ) provided an indication of goodness-of-fit for each model (Nakagawa & Schielzeth, 2013).  $R^2_m$  specified a measure of variance explained without mixed effects, and  $R^2_c$  included them. The perceptual responses of the simultaneous bilinguals and the late second-language learners were analyzed separately, since it is our understanding that the two groups of participants have qualitatively different linguistic experiences. We are not interested here in whether there is an interaction between *mode* and experience (early bilinguals vs. late learners), but on whether both groups of participants display the effects of *mode* or not.

## 2.2. Results

### 2.2.1. Simultaneous bilinguals

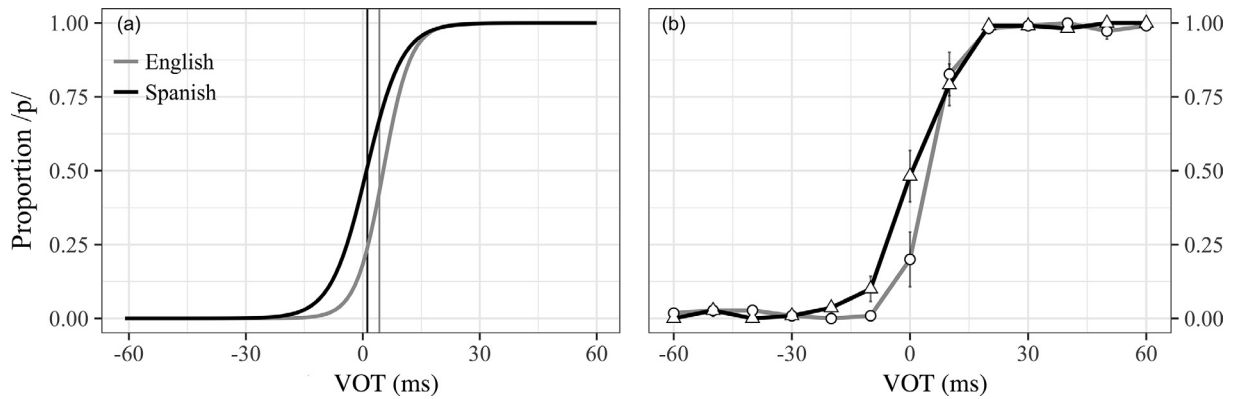
Fig. 1, left panel, plots the best-fit logit curve of responses to the *bafri-pafri* continua as a function of VOT and *mode* for the simultaneous bilinguals. The sigmoid functions represent the rate of change from *bafri* to *pafri* responses in the probability space. The vertical bars indicate the boundary for each *mode*, the 50% cross-over point. The difference between the vertical bars indicates how much (in ms) the boundary shifted between modes. One can observe that, relative to the black line, the gray line appears to be phase-shifted to the right, starting around the -20 ms point. The right panel plots the raw data to which the sigmoid function was fitted.

The analyses revealed a main effect of *mode* ( $\chi^2(1) = 7.15$ ;  $p < 0.008$ ) and, unsurprisingly, also of VOT ( $\chi^2(1) = 16.85$ ;  $p < 0.001$ ), but there was no interaction between the two predictors ( $\chi^2(1) = 0.16$ ;  $p > 0.05$ ). Subsequently, parameter estimates were obtained from a model that did not include the higher order variable, the interaction. This model aptly fit the data ( $R^2_m = 0.78$ ;  $R^2_c = 0.95$ ). Specifically, a shift from “Spanish” to “English” was associated with a decrease in the log odds of ‘voiceless’ responses of 0.83 ( $\pm 0.29$  SE) (95% c.i. [-1.38, -0.27];  $z = -2.89$ ;  $p < 0.004$ ). In other words, the simultaneous bilinguals were less likely to identify the “English” stimuli as voiceless overall.

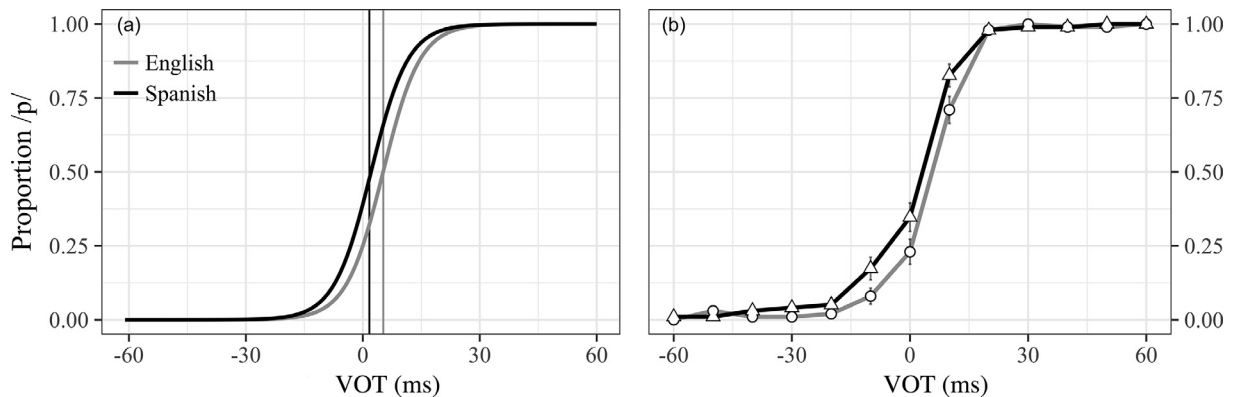
### 2.2.2. Late learners

Fig. 2 plots the predicted model fits (left panel) and the raw data (right panel) corresponding to the late learners’ responses. Relative to the sigmoid function associated with responses to the “Spanish” continuum (black line), one can observe that the sigmoid associated with responses to the “English” continuum is phase-shifted to the right. The raw data suggests that, towards the middle of the continuum, the learners were slightly more likely to respond ‘voiceless’ in the “Spanish” than in the “English” condition.

The corresponding model yielded a main effect of *mode* ( $\chi^2(1) = 12.55$ ;  $p < 0.01$ ) and of VOT ( $\chi^2(1) = 12.66$ ;  $p < 0.001$ ), and the two factors did not interact ( $\chi^2(1) = 1.79$ ;  $p > 0.05$ ). Parameter estimates were subsequently acquired from a model that excluded the interaction term. The multilevel model provided a good fit to the data ( $R^2_m = 0.74$ ;  $R^2_c = 0.96$ ),



**Fig. 1.** Perceptual data gathered from ten simultaneous Spanish-English bilinguals. The data are plotted as sigmoidal curves in the probability space (a) and the corresponding raw data (b) for Experiment 1. In panel (a), the vertical bars indicate the boundary crossover point for each continuum, “Spanish” and “English” modes. In panel (b), the shapes represent the mean proportion of ‘voiceless’ responses and the error bars stand for the standard error of the mean. In both panels, the black line indicates responses to the “Spanish” stimuli, and the grey line indicates responses to the “English” stimuli.



**Fig. 2.** Perceptual data gathered from ten late, beginner second-language learners of Spanish whose native language is English. The data are displayed as sigmoidal curves in the probability space (a) and the corresponding raw data (b) for Experiment 1. In panel (a), the vertical bars indicate the boundary crossover point for each continuum, “Spanish” and “English” modes. In panel (b), the shapes represent the mean proportion of ‘voiceless’ responses and the error bars stand for the standard error of the mean. In both panels, the black line indicates responses to the “Spanish” stimuli, and the grey line indicates responses to the “English” stimuli.

and it indicated that the participants’ responses varied as a function of which version of the continuum they heard. The participants were more likely to identify the stimuli as voiceless as VOT increased, and, crucially, they were less likely to respond ‘voiceless’ when listening to the “Spanish” version than the “English” version. The change from “Spanish” to “English” modes was associated with a decrease in the log odds of responding ‘voiceless’ of  $0.65 (\pm 0.18 \text{ SE})$  (95% c.i.  $[-1.01, -0.29]$ ,  $z = 3.55$ ,  $p < 0.001$ ). In sum, the late learners were less likely to label the “English” stimuli as voiceless overall.

### 2.3. Interim discussion

The findings of the present experiment replicate and extend those of [Gonzales and Lotto \(2013\)](#) and others. First, alongside those of [Gonzales and Lotto \(2013\)](#), our findings suggest that simultaneous bilinguals can shift between mode-specific perceptual routines as a function of environmental cues to language identity. Our experiment extended that of [Gonzales and Lotto \(2013\)](#) in that we tested our participants in a within-subjects design. Together with the authors of comparable studies ([Elman et al., 1977](#); [García-Sierra et al., 2009, 2012](#); [Gonzales & Lotto, 2013](#)), therefore, we conclude that early Spanish-English bilinguals demonstrate the double phonemic

boundary effect in their perceptual categorization routines. Perceptual routines are modulated by environmental cues to language identity in that, in our experiment, language-specific “modes” relied maximally on the phonetics of the rhotic following, in the same nonce word, the target stimulus.

Our experiment extended the findings of [Gonzales and Lotto \(2013\)](#), and others, in a second, more meaningful way. Along with simultaneous bilinguals, our study found that late second-language learners—including those at the initial stages of their development—can demonstrate the double phonemic boundary effect. Specifically, a group of language learners who had participated in a domestic immersion program displayed the effect after only seven weeks of intensive exposure to their second language. This fact suggests that these learners had developed a set of perceptual routines specific to their second language, which, in turn, suggests that they had formed phonetic categories specific to their second language.

Taken together, the present experiment suggests that simultaneous bilinguals and late second-language learners exploit mode-specific perceptual routines when encoding incoming speech. Such mode-specific routines seem to be primed by environmental phonetic cues—the pronunciation of a nearby rhotic, in this case. Since all interactions preceding the exper-

iment were in Spanish, in both sessions, any potential effects of immediate acoustic recalibration or phonetic context must be understood to be minimal (Bohn & Flege, 1993; Gonzales & Lotto, 2013). It is still possible, however, that the nature of the experiment lends itself to a type of acoustic recalibration not fully based on bilinguals' language modes, and begs the question as to whether the double perceptual boundary effect can also occur in real time. Note that, in this experiment, listeners participated in two sessions. It cannot be fully ruled out that they recalibrated their perceptual routines as a response to stimuli external to our task—i.e. stimuli they had heard between one session and the next. Since the order in which the two “modes” was presented was not counterbalanced in our study, this is particularly troubling. In a second experiment, we test early and late bilinguals in a single experimental session, playing all stimuli in a single block, in random order. To increase the generalizability of our results, a larger group of participants is tested ( $N = 64$ ), 21 early Spanish-English bilinguals and 43 late second-language learners of Spanish.

### 3. Experiment 2

The current experiment addresses the question of whether (early and late) bilinguals demonstrate the *double phonemic boundary* effect also when placed in “bilingual” mode—that is, when both of their languages are simultaneously activated. The preceding experiment compared listeners' behaviors across two unilingual modes, while the current one focuses on bilingual mode (Grosjean, 1989, 1998a, 1998b). It has been found that settings that require bilinguals to use both of their languages simultaneously are more likely to lead to enhanced cross-linguistic phonetic interactions, at least in production (Goldrick, Runnqvist, & Costa, 2014; Olson, 2013; Simonet, 2014). If bilinguals switch between mode-specific perceptual routines as a response to environmental cues, how fast can they do so? How fast can bilinguals go “in and out of languages” (Grosjean & Miller, 1994, p. 201)? Is this switching between routines neutralized when the environment presents cues consistent with both of their languages' phonologies? We address such questions here. An important difference between our second and first experiments is that, here, both acoustic continua (“Spanish,” “English”) are tested simultaneously, within the same block, in random order. This neutralizes the ordering confound present in our first experiment.

#### 3.1. Method

##### 3.1.1. Participants

Data from a total of 64 participants were included in our analyses. Participants were divided into two groups. The first group comprised early, highly proficient Spanish-English bilinguals born and raised in southern Arizona or southern California. Note that, for the preceding experiment, we took great care to recruit *simultaneous* bilinguals. For the current experiment, on the other hand, our early bilingual group included both simultaneous and sequential bilinguals. At any rate, participants in this group were exposed to both Spanish and English in early childhood.

The second group comprised late second-language learners of Spanish with English as their native language. The par-

ticipants in this group were taking introductory- or intermediate-level Spanish language courses at a large land-grant university in southern Arizona. The length of their exposure to Spanish was presumably much more significant in this group of language learners than in that recruited for the preceding experiment, since they were tested towards the end of a semester in which they had been taking Spanish classes and since they lived in a community with a significant presence of Spanish. To the extent that it may serve as an indication of the presence of Spanish in the community, we may add that, according to data compiled in 2010 by the *Bureau of the Census* (census.gov) of the United States, 41.6% of Tucson's population self-identify as Hispanic or Latino, with 36.1% self-identifying as Mexican American. On the other hand, whereas the learners in the preceding experiment were enrolled in an immersion program that “banned” the use of their native language, the ones in the current study were not, which means that the latter had been using their native language very often prior to participating in the task. Due to the nature of the programs, there could be very significant differences in terms of their motivation between the Middlebury students and the late learners explored in the second experiment. And both programs claim to employ the methods of the communicative language teaching paradigm. The following two paragraphs provide some details about the two groups of participants, the early and the late bilinguals.

A total of 33 adult Spanish-English early bilinguals were initially recruited; however, data from 12 participants had to be excluded because initial recruitment had not been very prudent. Data exclusion was determined by one of three factors: (i) researcher error during the administration of the experiment or the archiving of the data, resulting in lost or incomplete data ( $N = 4$ ); (ii) participant misunderstanding of task instructions or inadequate engagement with the task—e.g., selecting ‘left’ throughout the experiment ( $N = 4$ ); and (iii) extreme dominance in English or Spanish ( $N = 4$ ) as determined by the BLP (Birdsong et al., 2012). The “extremely dominant” participants we excluded had BLP scores in the hundreds, further than 2.5 standard deviations from the group mean. Even though we did not chose a specific cut off point, the excluded people were obvious outliers. As a consequence, the present analysis utilizes data from a total of 21 early-bilingual participants. The group of participants whose data were retained for analysis scored an average of  $-11.30$  ( $SD = 18.70$ , 95% c.i. =  $-17.30$ ,  $-5.35$ ) for language history,  $-2.57$  ( $SD = 12.60$ , 95% c.i. =  $-6.57$ ,  $1.43$ ) for proficiency, and  $-13.88$  ( $SD = 28.58$ , 95% c.i. =  $-23.00$ ,  $-4.79$ ) on the composite score derived from two components of the Bilingual Language Profile questionnaire (Birdsong et al., 2012). These scores indicate moderately balanced bilingualism. Note, once again, that we make no claim about simultaneity of exposure for this group, just early exposure.

A total of 50 adult second-language learners of Spanish were initially recruited for their participation in this experiment. Data from seven participants had to be excluded due to experimenter error during data collection ( $N = 2$ ), or to lost or incomplete data ( $N = 5$ ). As a consequence, data from 43 learners were retained in the final analysis. The learners had been born and raised in either Arizona or southern California, they were all attending the University of Arizona, and they were enrolled



in one of three college Spanish courses. These courses corresponded with 2nd-, 4th-, and 5th-semester Spanish, which comprise a range of proficiency levels, from introductory to intermediate levels. The more advanced speakers were more likely to have taken some pronunciation classes, but we did not ask them about this. All participants completed the BLP (Birdsong et al., 2012). A composite Spanish score was calculated for each participant using their responses to the language history and language competence modules of the questionnaire. For this group, the composite score served as a continuous predictor in a linear model (see below). The composite score was normed—i.e., instead of using the scores provided by the BLP, we scaled the values to z-scores, which allowed us to calibrate the scores around a mean of zero. In sum, in this experiment we explored the possibility that a composite score based on language experience and self-assessed competence modulated the probability of capturing the *double phonemic boundary* effect in a population of adult second-language learners of Spanish. This is a sensible procedure in this case because the language learner group includes participants across a spectrum of competency levels.

### 3.1.2. Procedure and analysis

We utilized the same resynthesized acoustic continua that we used in the preceding experiment. The experimental procedure was identical to that of the previous experiment, a two-alternative forced-choice perceptual categorization task, with two noteworthy differences.

Firstly, in order to reduce the overall duration of the experiment (note that the BLP survey was administered in the same session), four stimuli were removed, two from each end of the *bafri-pafri* continua. Therefore, each acoustic continuum (“English,” *ba[ɸi]-pa[ɸi]*; “Spanish,” *ba[frɪ]-pa[frɪ]*) consisted of nine steps ranging from  $-40$  ms to  $40$  ms in  $10$  ms increments. This allowed us to retain the crucial data, which are located in the ambiguous region around  $VOT = 0$ . Secondly, the number of repetitions of each continuum step was reduced from ten to five. In total, the experiment drew from 18 stimuli, each of which was presented randomly in a single block five times, for a total of 90 trials. Instructions, as well as all interactions between the participants and the experimenter, were in Span-

ish. Note, once again, that there was a single experimental session in which tokens from both continua (“English,” “Spanish”) were presented in random order.

The early bilingual data were collected individually, participants were tested one by one in a quiet room. The learner data, on the other hand, were collected in a group session—i.e., participants were tested simultaneously in a computer laboratory. The experiment was run in PsychoPy (Peirce, 2007) with either a laptop (early bilinguals) or a desktop (second-language learners) computer and a headphone set.

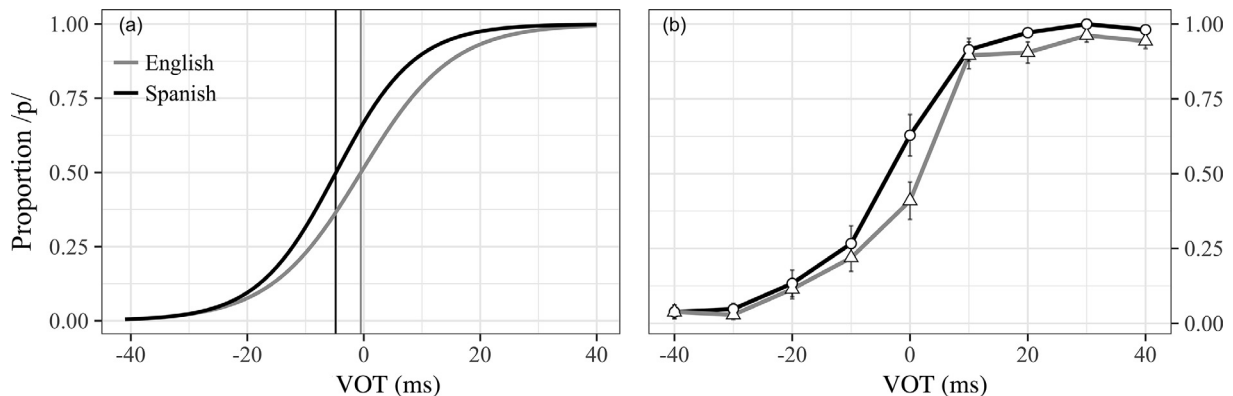
We submitted the binomial response data to a series of generalized linear mixed effects models. The criterion (‘voiceless’ responses) was modeled as a function of the hypothesized *mode* of the stimuli (“Spanish,” “English”), and *VOT* ( $-40$  to  $40$  ms). Causal priority was given to *mode*.

## 3.2. Results

### 3.2.1. Early bilinguals

Fig. 3, which plots the predicted model fits (left panel) and the raw data (right panel) corresponding to the early bilinguals, shows that, relative to the sigmoid function associated with responses to the “Spanish” continuum (black line), the one associated with the “English” one (gray line) is phase-shifted to the right.

The analyses of the early bilingual response data revealed a main effect of *mode* ( $\chi^2(1) = 12.40$ ;  $p < 0.001$ ) and, unsurprisingly, also of *VOT* ( $\chi^2(1) = 46.63$ ;  $p < 0.001$ ). The two predictors did not interact ( $\chi^2(1) = 2.98$ ;  $p > 0.05$ ). We then refitted the data with a model that dropped the higher order interaction. This model explained 81% ( $R^2c$ ) of the variance with the random effects structure and 72% ( $R^2m$ ) without it. There was a positive correlation between *VOT* and the proportion of ‘voiceless’ responses: the longer the *VOT*, the higher the probability of responding ‘voiceless.’ More importantly, responses were also modulated by *mode*; that is, relative to the responses to the “Spanish” continuum, there was a decrease of  $0.60$  ( $\pm 0.16$  SE) (95% c.i. [ $-0.91$ ,  $-0.28$ ],  $z = -3.66$ ,  $p < 0.001$ ) in the log odds of responding ‘voiceless’ in the “English” one. In sum, the listeners were less likely to label the “English” stimuli as voiceless overall.



**Fig. 3.** Perceptual data gathered from 21 early Spanish-English bilinguals for Experiment 2. Data are shown as sigmoidal curves in the probability space (a) and the corresponding raw data (b). In panel (a), the vertical bars indicate the boundary crossover point for each continuum, “Spanish” and “English” modes. In panel (b), the shapes represent the mean proportion of ‘voiceless’ responses and the error bars stand for the standard error of the mean. In both panels, the black line indicates responses to the “Spanish” stimuli, and the grey line indicates responses to the “English” stimuli.

### 3.2.2. Late learners

The learner data were analyzed using a generalized linear mixed effects model with listener's label ('voiceless') as the criterion and *mode* ("Spanish," "English"), scaled *proficiency score*, and *VOT* (−40 ms to 40 ms) as fixed effects. For this analysis, we utilized sum coding for the fixed effect *mode*. As a result, the model output provided an assessment of the main effect, as well as the simple effects of two three-way interactions: (i) "English" continuum  $\times$  *proficiency score*  $\times$  *VOT*, and (ii) "Spanish" continuum  $\times$  *proficiency score*  $\times$  *VOT*. (This was used in conjunction with the nested model comparisons, and the more conservative of the two tests is reported.)

The model comparisons revealed a main effect of *VOT* ( $\chi^2(1) = 79.33$ ;  $p < 0.001$ ), such that 'voiceless' responses increased as *VOT* increased. There were no other significant main effects (*mode*:  $\chi^2(1) = 0.45$ ,  $p > 0.05$ ; *proficiency*:  $\chi^2(1) = 0.19$ ,  $p > 0.05$ ), nor was there a *mode* by *proficiency* interaction ( $\chi^2(1) = 0.72$ ;  $p > 0.05$ ). There was, however, a three-way (*mode*  $\times$  *proficiency*  $\times$  *VOT*) interaction ( $\chi^2(2) = 7.92$ ;  $p < 0.02$ ). The model output revealed that, for the "Spanish" continuum, the effect of *VOT* on the rate of 'voiceless' responses was modulated by *proficiency*. Specifically, in this mode, an increase of one standardized *proficiency* unit resulted in a decrease in the log odds of responding 'voiceless' of 0.02 ( $\pm 0.01$  SE) (95% c.i. [−0.03, −0.01];  $z = -2.84$ ;  $p < 0.005$ ).

To further examine the three-way interaction, we used the random effects output from the omnibus model in order to calculate the boundary crossover point for each participant, in each mode, grouped by proficiency. We cut the normalized proficiency scores into three quantile groups: low, low-intermediate, and intermediate proficiency levels. The resulting grouping variable coincided with the language class in which the participants were enrolled: by coincidence, all the 2nd-semester learners were grouped as low proficiency, the 4th-semester learners were grouped as mid-low proficiency, and the 5th-semester learners were grouped as intermediate proficiency. Then, for each participant and mode we calculated the perceptual crossover boundaries, two 50% crossover *VOT* boundaries per listener. The crossover boundary represents the exact value in ms (*VOT*) for which the probability of responding 'voiced' or 'voiceless' is exactly 50%. The crossover function utilizes the formula below, where the intercept

for each individual ( $\beta 0_i$ ) is divided by their respective slope ( $\beta vot_i$ ) and multiplied by  $-1$ .

$$CO_i = \frac{\beta 0_i}{\beta vot_i} \times -1$$

This allows for a more fine-grained look at the learner's behavior in the perceptual labeling task. The final step was to standardize the resulting crossover values and submit them to a linear mixed effects model with *mode* as a fixed predictor and a *proficiency group* (low, low-intermediate, intermediate) by *mode* interaction term. (Note that this model does not include a fixed effect for *proficiency group*, as this effect was partialled out via z-scoring. Nonetheless, the *proficiency group* by *mode* interaction term captures group deviations from 0 for each mode, essentially supplying the desired information in an easily interpretable way.) Each participant was given a random intercept. The main effect of *mode* and the *proficiency group* by *language mode* interaction were tested using nested model comparisons.

Fig. 4, left panel, plots the standardized scores derived from two modules of the BLP (Birdsong et al., 2012) as a function of participant *proficiency group* (low, low-intermediate, intermediate). The middle panel in Fig. 4 plots the listeners' 50% crossover points (ms) as a function of *proficiency group* and *mode*. Finally, the right panel in Fig. 4 plots the effect of *mode* as a function of participant group. One can observe the likelihood of an interaction between *mode* and *group*. The plots suggest that the most advanced learners among the participants, but not the other learners, display separate /b/-/p/ perceptual boundaries when labeling the "English" and the "Spanish" continua.

Indeed, the generalized linear model revealed a main effect of *mode* ( $\chi^2(1) = 5.26$ ;  $p < 0.03$ ) and a *group* by *mode* interaction ( $\chi^2(4) = 11.99$ ;  $p < 0.003$ ). Individual model outputs showed that the crossover boundaries for the low and low-intermediate groups did not vary as a function of *mode* (low proficiency group:  $\beta = 0.02$ , SE = 0.16, 95% c.i. [−0.29, 0.33],  $t(11) = 0.13$ ,  $p > 0.05$ ; low-intermediate proficiency group:  $\beta = -0.01$ , SE = 0.08, 95% c.i. [−0.17, 0.14],  $t(17) = -0.18$ ,  $p > 0.05$ ); but rather the *mode* effect was found solely in the intermediate proficiency group ( $\beta = -0.43$ , SE = 0.06, 95% c.i. [−0.55, −0.31],  $t(15) = -6.99$ ,  $p < 0.001$ ), the group with higher proficiency in this experiment.

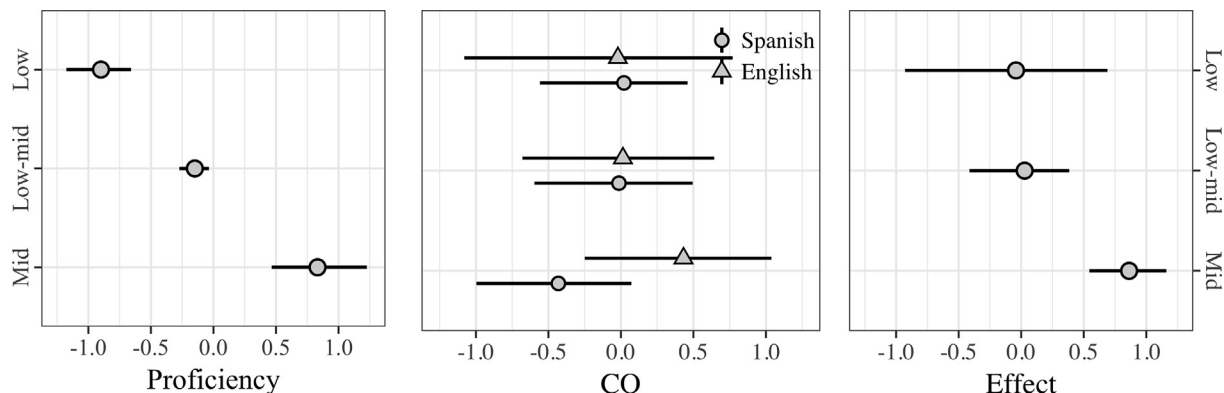
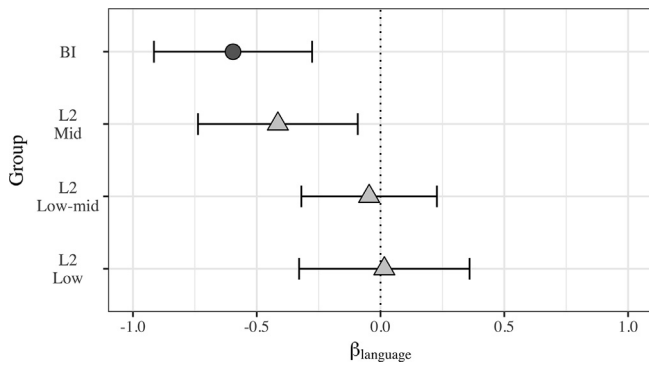


Fig. 4. Standardized proficiency scores (left panel), crossover boundaries (CO) (middle panel), and effect sizes of factor *language mode* ("Spanish," "English") as a function of *proficiency group* (Low, Low-mid [low-intermediate], Mid [intermediate]) in Experiment 2. Data were gathered from 43 second-language learners of Spanish whose native language is English.



**Fig. 5.**  $\beta$  weights for the effect of *language mode* ( $\pm$ SE) as a function of *proficiency group* (early bilinguals [BI], intermediate-proficiency second-language learners [L2 Mid], low-intermediate-proficiency second-language learners [L2 Low-mid], low-proficiency second-language learners [L2 Low]). Values represent the change in the log odds for 'voiceless' responses between "Spanish" and "English" mode acoustic (VOT) continua.

### 3.3. Interim discussion

The current experiment investigated the perceptual categorization routines (of a /b-/p/ acoustic continuum) of Spanish-English bilinguals, which were classified as a function of their linguistic experience and self-rated linguistic proficiencies. In two different analyses, we explored the behavior of a group of early Spanish-English bilinguals and that of a group of second-language learners of Spanish. The learners were late bilinguals at various stages of Spanish-language proficiency.

The listeners labeled a /b-/p/ continuum in two experimental conditions, one that was hypothesized to prime their "Spanish" processing mode and one that was thought to prime their "English" mode. Unlike in the preceding experiment, the participants in the current one labeled the stimuli belonging to both conditions (or language modes) in the same session.

The findings were as follows. First, for the early bilinguals, the perceptual boundary between /b/ and /p/ differed significantly as a function of mode. The facts regarding the second-language learners were more nuanced. We found an interaction between mode and proficiency in Spanish such that more proficient learners were more likely to display the double phonemic boundary effect than less proficient ones. The listeners in the two lower proficiency levels displayed a single /b-/p/ perceptual boundary. In the case of the listeners in the higher level proficiency group, on the other hand, the /b-/p/ perceptual boundary pertaining to the English-mode continuum was shifted to the right relative to that pertaining to the Spanish-mode continuum. In Fig. 5, one can observe a progression of mode effects from the two lower proficiency levels (no mode effect) to the early bilinguals (strong mode effect) through the intermediate proficiency group (moderate, but significant, effect).

## 4. Discussion

### 4.1. Summary of findings

In the present study, our goal was to assess whether bilinguals utilize perceptual routines specific to each of their two languages to map acoustic events into phonemic categories and to investigate how rapidly they can switch between such two processing modes (or sets of perceptual routines). Our

method was based on that of [Gonzales and Lotto's \(2013\)](#) study—i.e., bilinguals labeled a VOT continuum in two experimental conditions, which were hypothesized to activate different language modes ([Grosjean, 1998b](#)), a "Spanish" and an "English" mode. The activation of mode-specific perceptual routines was based on the presence of language-specific sound sequences alongside the target stimuli, not in the language used in the context of the task.

For our first experiment, we recruited a group of simultaneous bilinguals and a group of late second-language learners at the very initial stages of their development (beginners). The stimuli were presented in two sessions, and the materials in the sessions varied solely in the phonetics of a sequence that had been appended to the target tokens to form bisyllabic nonce words, *bafri-pafri* ([Gonzales & Lotto, 2013](#)). To the extent that the two conditions primed separate processing modes in these bilinguals, our results suggest that both early and late bilinguals are able to employ mode-specific perceptual routines to map incoming speech into phonemic categories. The experiment, however, had a serious limitation. Listeners were tested in two sessions, always in the same order.

For our second experiment, we recruited a group of early bilinguals and one of late second-language learners; the latter included learners at various stages of their development, from beginning to intermediate learners. In this experiment, the stimuli corresponding to the two language modes were presented within the same experimental block, in random order, rather than in two separate sessions. This ensured that, if a mode effect were found, it could not easily be attributed to factors extraneous to the experimental materials but to the nature of the stimuli. It was found that the early bilinguals and the more advanced learners, but not the lower proficiency learners, displayed the effects of a *double phonemic boundary* ([García-Sierra et al., 2009](#)). This suggests that bilinguals can rapidly switch between language-specific perceptual routines and that their cues to switch from one to the other are evaluated in real time.

### 4.2. Interpretation and implications

This section of the Discussion centers around three points we believe the results of our study allow us to make. First, as a group, bilinguals can switch between perceptual categorization modes specific to each of their two languages when some cue present in the environment suggests that doing so might lead to more efficient categorization. Second, bilinguals' switching between perceptual categorization modes is rapid and may be based on shallow, prelexical environmental evidence. Third, bilinguals who can switch between mode-specific routines include early as well as late bilinguals—i.e., adult second-language learners can learn to take advantage of this perceptual strategy. After elaborating on these three points, we conclude by positing a connection between phonetic-category formation and the development of mode-specific perceptual routines in adult speech learning.

In the present study, we replicated [Gonzales and Lotto's \(2013\)](#) experimental method but did so with a within-subjects design. First, simultaneous bilinguals displayed a double phonemic boundary effect when labeling functions correspond-

ing to data gathered in two sessions, a “Spanish” session and an “English” session, were compared. This finding is in line with that of [Gonzales and Lotto \(2013\)](#), further validating their results in a within-subjects design. Second, early sequential bilinguals displayed a double phonemic boundary effect when we compared labeling functions corresponding to data which, gathered during a single session, belonged to two experimental conditions hypothesized to place bilinguals into different language modes. The results of neither of our two experiments, but especially not those of the second one, are likely to have been driven by extraneous acoustic input, such as acoustic context and range effects derived from pre-task conversations between the experimenter and the participant ([Bohn & Flege, 1993](#); [Gonzales & Lotto, 2013](#)). Note also that, even though we did not test a monolingual group, a group of beginning second-language learners who participated in our second experiment yielded null results—a point that further confirms our conclusion that our positive findings were due to bilingual language switching rather than task conditions. We wish to conclude, therefore, that our study, in line with comparable ones ([Elman et al., 1977](#); [García-Sierra et al., 2009, 2012](#); [Gonzales & Lotto, 2013](#)), has found strong support for the hypothesis that bilinguals can switch between mode-specific perceptual categorization routines.

An aspect of our second experiment suggests that, as it relates to language-specific categorization routines, bilingual language switching is fast, and it can be driven by shallow environmental cues. In our experiments, the effects of bilingual language modes were brought about by appending language-specific sound sequences to the target stimuli. By hypothesis, listeners were placed in Spanish mode when the sequence [fri] was appended to the target tokens, and they were placed in English mode when the appended sequence was [fi] ([Gonzales & Lotto, 2013](#)). There were no other cues to language mode. Consider two points that follow from this fact. On the one hand, it is known that lexical information can modulate perceptual categorization ([Ganong, 1980](#); [Samuel, 1997, 2001](#)). Therefore, it would not be surprising if a study found that bilinguals exploit lexical information to activate language-specific perceptual routines. The results of the present study, however, demonstrate that bilinguals can switch between language-specific speech perception modes even in the absence of lexical evidence, as we used nonwords. This suggests that the switching between mode-appropriate routines is (at least in part) a prelexical process or that it can be cued by information available prelexically. On the other hand, research shows that multiple sources of environmental information, such as sociolinguistic expectations and physical location ([Hay & Drager, 2010](#); [Hay, Podlubny, Drager, & McAuliffe, 2017](#)), may interact with speech processing routines. In the context of such research, the present study suggests that environmental evidence of various sources can lead bilinguals to switch between mode-appropriate perceptual behavior. Investigating the nature (and mapping the limits) of such environmental evidence must be left for future research. Indeed, we still need an explicit explanation of the mechanism by which [fri] places bilinguals in “Spanish” mode and [fi] places them in “English” mode. We tentatively hypothesize that, in bilinguals, connections between sounds (and phonotactic regularities) belonging to the same language system have some sort of primacy over

those that connect sounds across systems. While the evidence clearly suggests that bilinguals’ phonological representations share a single representational network, it is obvious that representational nodes comprise some sort of tagging system that includes language identity.

Bilingual language switching is fast and it occurs in real time. Our second experiment suggests that bilinguals are able to switch between mode-appropriate categorization routines even between trials in a single perceptual task when each trial contains bisyllabic nonce words. It is not the case, therefore, that bilinguals, once in a given language mode, will remain in it by inertia for a considerable amount of time. Thus, even though language switching is likely to have carry-over effects (leading to cross-linguistic convergence) across trials, as it has been found for speech production ([Goldrick et al., 2014](#); [Olson, 2013](#); [Simonet, 2014](#)), such effects do not seem to neutralize language specificity entirely. We hypothesize that bilinguals can assess environmental information of multiple types in a fast and efficient manner and calibrate their perceptual expectations accordingly. Such rapid calibration may make use of the same cognitive resources and architecture that have been captured in other research paradigms (e.g., [Clarke & Garrett, 2004](#)), such as the family of findings known as *perceptual learning for speech* ([Eisner & McQueen, 2005, 2006](#); [Kraljic & Samuel, 2005, 2006](#); [Norris, McQueen, & Cutler, 2003](#); [Samuel & Kraljic, 2009](#)).

The third point on which we would like to focus here concerns the fact that not only early bilinguals, but also adult second-language learners, seem to be able to manifest the double phonemic boundary effect in their perceptual labeling patterns. Prior demonstrations of the effect concerned early or simultaneous bilinguals ([Elman et al., 1977](#); [García-Sierra et al., 2009](#); [Gonzales & Lotto, 2013](#); [Hazan & Boulakia, 1993](#)). At least one study had shown the effect with highly proficient second-language learners ([Flege & Eefting, 1987](#)), but one must take into consideration that some of these findings had been tainted by the suspicion that extraneous factors had driven the results ([Bohn & Flege, 1993](#)). Our first experiment suggested that beginner language learners at the early stages of their development can manifest the effect. It could be that the positive finding in our first experiment were due to the fact that the participants comprised people who had just been immersed in an intensive second-language environment for seven weeks. Perhaps the fact that their exposure to their first language had been severely limited modulated the effect, but it could also be that some unknown extraneous factor caused the results. Note that they were tested in “Spanish” mode first and, the following day, they were tested in “English” mode. The anticipation of using English for the first time after seven weeks could have partially driven the findings. Our second experiment, less likely to be affected by such extraneous factors, demonstrates that late learners possess mode-specific categorization routines and that the likelihood of their displaying the effects of such routines is a function of their linguistic development or exposure. We conclude that bilingual language switching, the ability of switching between language-specific perceptual modes, can be learned in adulthood.

To conclude this section, we would like to make a point regarding the relevance of our findings in the greater context



of second-language speech research. A very important question in the context of this research concerns the establishment of *new* phonetic categories, categories specific to the second language. This issue is at the core of current second-language speech research (Best & Tyler, 2007; Escudero, 2005; Flege, 1995, 2007; Van Leussen & Escudero, 2015). Production and perception of second-language sounds is, to a great extent, determined by the characteristics of the phonology of the first language. In order to investigate whether second-language learners can establish categories specific to their second language, learners are often compared with monolingual speakers of the target language. As long as learners are found to differ from monolinguals of the target language in the direction one might expect from the characteristics of the phonology of their first language, one usually claims that learners have *not* successfully established second-language-specific phonetic categories. A more illuminating perspective, however, is obtained when comparing the behavior of second-language learners, not with that of monolinguals, but across their two languages—a between-language, within-speaker comparison rather than a within-language, between-speaker one. We would like to surmise that a very important proof that bilinguals and second-language learners can indeed develop language-specific categories concerns the findings of our study. In our view, a within-subject demonstration that bilingual listeners label the same acoustic continua differently as a function of the language mode in which they are placed offers important proof that separate phonetic categories have been formed for the second language. Here is how.

Switching between perceptual routines is likely driven by perceptual expectations, and such expectations, it would seem, are likely driven by previous phonetic knowledge (Apfelbaum, Bullock-Rest, Rhone, Jongman, & McMurray, 2014; McMurray & Jongman, 2016). Evidently, in bilinguals, this includes knowledge about the characteristics of the sounds of their two languages. It would follow that such knowledge must be in place before mode-specific categorization can be manifested. In sum, we postulate that, in order for bilinguals and language learners to demonstrate the effects of mode-specific perceptual categorization, they must have developed language-specific phonetic representations (of some sort)—in other words, phonetic differences across languages must be part of a bilingual's grammatical knowledge. One way in which this scenario could be modeled is offered by Escudero's *Second Language Linguistic Perception* model (L2LP) (Escudero, 2005; Van Leussen & Escudero, 2015), which postulates that second-language learners, upon being exposed to the sounds of their second language for the first time, create a copy of the phonology of their first language. This copy, which is assigned to the second language, is then allowed to change gradually (and independently from the first language) as phonetic evidence specific to the second language becomes available. We would further say that, once a second-language-specific "copy" has become sufficiently different from the most similar first-language category, mode-specific perceptual categorization leads to more efficient speech processing and is, therefore, exploited by the learner. It is not important whether the modified copy matches that of monolingual speakers of the target language or not; what matters is that the copy is no longer identical to the closest first-language sound, but is a sound

specific to the second language. In such a case, one may speak of new-category formation.

## 5. Conclusion

Bilinguals are typically able to utilize language-specific sounds in a language-specific manner—e.g., they produce different sets of sounds in their two languages (when there is a reason to do so, of course). Do bilinguals have separate perceptual categorization habits and expectations for the phonetic systems of their two languages? The first goal of the present study was to address these questions directly. By doing so, our study contributes to a body of literature that has pursued these ideas from the nineteen seventies to this day (see Bohn & Flege, 1993; Caramazza et al., 1973, 1974; Elman et al., 1977; Flege & Eefting, 1987; García-Sierra et al., 2009, 2012; Gonzales & Lotto, 2013; Hazan & Boulakia, 1993; Williams, 1977). Following such studies, we analyzed the perceptual categorization routines of various groups of Spanish-English bilinguals as a function of two experimental conditions designed to activate (or prime) the processing routines of the bilinguals' two languages (separately).

The findings of the present study allow us to tentatively reach the following conclusions: First, when bilinguals have some evidence that incoming speech takes the form of one of their languages (and not the other) they automatically utilize perceptual parsing routines that are appropriate for that language (and not the other). By "appropriate" we mean that such parsing routines are specific to one language (i.e., that bilinguals use them to process sounds of only one language), not that they are identical to those of monolinguals or that they are optimally efficient. Such cues can be rather shallow, such as the presence of language-specific sounds (or sound sequences) in nonce words present in the environment (Gonzales & Lotto, 2013). It appears that prelexical processing is aided by such a strategy. Second, bilinguals can switch between mode-specific categorization routines rather rapidly. In conditions of rapid (and random) switching, bilinguals adjust their expectations in real time and respond in a mode-appropriate manner. And, third, bilinguals who are able to display the effects of mode-specific categorization routines include, not only early (highly proficient) bilinguals, but also late second-language learners of intermediate proficiency.

To conclude, we postulate that the development of mode-specific categorization routines is closely tied with that of language-specific phonetic categories, and we surmise that, in order to demonstrate that bilinguals possess separate phonetic categories (and systems) for their two languages, it should be very beneficial to *first* investigate whether they display (or not) mode-specific processing routines.

## 6. Declaration of interest

None.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.wocn.2018.07.002>.

## References

- Abramson, A. S., & Lisker, L. (1972). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1, 1–8.
- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, 38, 640–653.
- Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, 40, 582–594.
- Apfelbaum, K. S., Bullock-Rest, N., Rhone, A. E., Jongman, A., & McMurray, B. (2014). Contingent categorization in speech perception. *Language, Cognition and Neuroscience*, 29, 1070–1082.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Beckman, J., Jessen, M., & Ringen, C. (2013). Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics*, 49, 259–284.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam, The Netherlands: John Benjamins.
- Birdsong, D., Gerken, L. M., & Amengual, M. (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL: University of Texas at Austin.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341–345.
- Bohn, O.-S., & Flege, J. E. (1993). Perceptual switching in Spanish-English bilinguals. *Journal of Phonetics*, 21, 267–290.
- Brady, S. A., & Darwin, C. J. (1978). Range effect in the perception of voicing. *Journal of the Acoustical Society of America*, 63, 1556–1558.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, 54, 421–428.
- Caramazza, A., Yeni-Komshian, G., & Zurif, E. B. (1974). Bilingual switching: The phonological level. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 28, 310.
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, 116, 3647–3658.
- Diehl, R. L., Elman, J. L., & McCusker, S. B. (1978). Contrast effects on stop consonant identification. *Journal of Experimental Psychology. Human Perception and Performance*, 4, 599–609.
- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, 67, 224–238.
- Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America*, 119, 1950–1953.
- Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, 62, 971–974.
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization* (Ph.D.). Utrecht, The Netherlands: Utrecht University. LOT Dissertation Series 113.
- Flege, J. E. (1987a). A critical period for learning to pronounce foreign languages? *Applied Linguistics*, 8, 162–177.
- Flege, J. E. (1987b). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, Maryland: York Press.
- Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. I. Hualde (Eds.), *Laboratory phonology 9* (pp. 353–380). Berlin, Germany: Mouton de Gruyter.
- Flege, J. E., & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185–202.
- Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, 76, 708–721.
- 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, 649–663.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology. Human Perception and Performance*, 6, 110–125.
- García-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, 51, 369–378.
- García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language*, 121, 194–205.
- Goldrick, M., Runqvist, E., & Costa, A. (2014). Language switching makes pronunciation less native-like. *Psychological Science*, 25, 1031–1036.
- Gonzales, K., & Lotto, A. J. (2013). A Bafri, un Pafri: Bilinguals' pseudoword identifications support language-specific phonetic systems. *Psychological Science*, 24, 2135–2142.
- Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual & Multicultural Development*, 6, 467–477.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3–15.
- Grosjean, F. (1998a). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131–148.
- Grosjean, F. (1998b). Transfer and language mode. *Bilingualism: Language and Cognition*, 1, 175–176.
- Grosjean, F., & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological Science*, 5, 201–206.
- Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, 48, 865–892.
- Hay, J., Podlubny, R., Drager, K., & McAuliffe, M. (2017). Car-talk: Location-specific speech production and perception. *Journal of Phonetics*, 65, 94–109.
- Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, 36, 17–38.
- Holt, L. L., & Lotto, A. J. (2002). Behavioral examinations of the level of auditory processing of speech context effects. *Hearing Research*, 167, 156–169.
- Iverson, G. K., & Salmans, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, 12, 369–396.
- Keating, P. A., Mikoš, M. J., & Ganong, W. F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America*, 70, 1261–1271.
- Kirby, J. P., & Ladd, D. R. (2016). Effects of obstruent voicing on vowel F0: Evidence from “true voicing” languages. *Journal of the Acoustical Society of America*, 140, 2400–2411.
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51, 141–178.
- Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, 13, 262–268.
- Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384–422.
- Magloire, J., & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, 56, 158–185.
- McMurray, B., & Jongman, A. (2016). What comes after /f/? Prediction in speech derives from data-explanatory processes. *Psychological Science*, 27, 43–52.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods in Ecology and Evolution / British Ecological Society*, 4, 133–142.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47, 204–238.
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41, 407–420.
- Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5, 261–283.
- Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64, B9–B17.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191–215.
- 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.
- Samuel, A. G. (1982). Phonetic prototypes. *Perception & Psychophysics*, 31, 307–314.
- Samuel, A. G. (1997). Lexical activation produces potent phonemic percepts. *Cognitive Psychology*, 32, 97–127.
- Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, 12, 348–351.
- Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception & Psychophysics*, 71, 1207–1218.
- Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, 43, 26–37.
- Simonet, M. (2016). The phonetics and phonology of bilingualism. In *Oxford handbooks online: Scholarly research reviews* (pp. 1–25). Oxford, UK: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199935345.013.72>.
- Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition*, 106, 234–258.
- Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and acquire a second language: The L2LP model revised. *Frontiers in Psychology*, 6, 1–12.
- Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Attention, Perception & Psychophysics*, 21, 289–297.