Research Article

Psychological Science
XX(X) 1–8
© The Author(s) 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797613486485
pss.sagepub.com

A *bafri*, un *pafri*: Bilinguals' Pseudoword Identifications Support Language-Specific Phonetic Systems

Kalim Gonzales[1] **and Andrew J. Lotto**[2]
[1]Department of Psychology and [2]Department of Speech, Language and Hearing
Sciences, University of Arizona

Abstract

Bilinguals perceptually accommodate speech variation across languages, but to what extent this flexibility depends on bilingual experience is uncertain. One account suggests that bilingual experience promotes language-specific processing modes, implying that bilinguals can switch as appropriate between the different phonetic systems of the languages they speak. Another account suggests that bilinguals rapidly recalibrate to the unique acoustic properties of each language following language-general processes common to monolinguals. Challenging this latter account, the present results show that Spanish-English bilinguals with exposure to both languages from early childhood, but not English monolinguals, shift perception as appropriate across acoustically controlled English and Spanish contexts. Early bilingual experience appears to promote language-specific phonetic systems.

Keywords

bilingualism, speech perception, cognitive processes, language

Received 3/30/12; Revision accepted 3/25/13

The extent of the mind's capacity to accommodate contextual influences on a complex physical signal is a central question in the psychological sciences. In speech perception, such accommodation is seen as critical for negotiating the lack of invariance in how speech categories are realized acoustically. For example, a variety of contextual factors can result in different phonemes having the exact same acoustic pattern (see Magnuson & Nusbaum, 2007). Most research has focused on within-language factors such as speaker characteristics (e.g., gender) or coarticulation with surrounding speech. But for the majority of listeners worldwide who communicate in more than one language, another factor is the language itself. This is because different languages commonly distinguish phonemes along the same acoustic dimensions but differ in the placement of the phoneme boundary.

An example is the distinction between syllable-initial *b* and *p* in English words such as *bowl* and *pole* versus Spanish words such as *boca* and *poca* (meaning "mouth" and "little," respectively). In both languages, the voicing distinction between these phonemes is signaled by a change in the duration of voice-onset time (VOT), or the time between release of the occlusion (at the lips) and the onset of vocal-fold vibration. However, the duration of VOT at which a sound is produced as *p* is shorter in Spanish than in English, such that the duration of VOT for the Spanish *p* actually overlaps with that for the English *b*. Accordingly, English and Spanish monolinguals differentially identify stops with short-lag voicing as *b* and *p*, respectively (see Fig. 1). This difference is evident in second-language learners, who often retain their native-language boundary when trying to perceive their newly obtained language (Williams, 1979), which sometimes leads to confusions between words differing only in the voicing of the initial consonant (Ortega-LLebaria, Faulkner, & Hazan, 2001).

Presumably, more competent bilinguals overcome this clash of phonetic systems. The general intuition is that

Corresponding Author:
Kalim Gonzales, Department of Psychology, University of Arizona,
Tucson, AZ 85721
E-mail: kalimg@email.arizona.edu

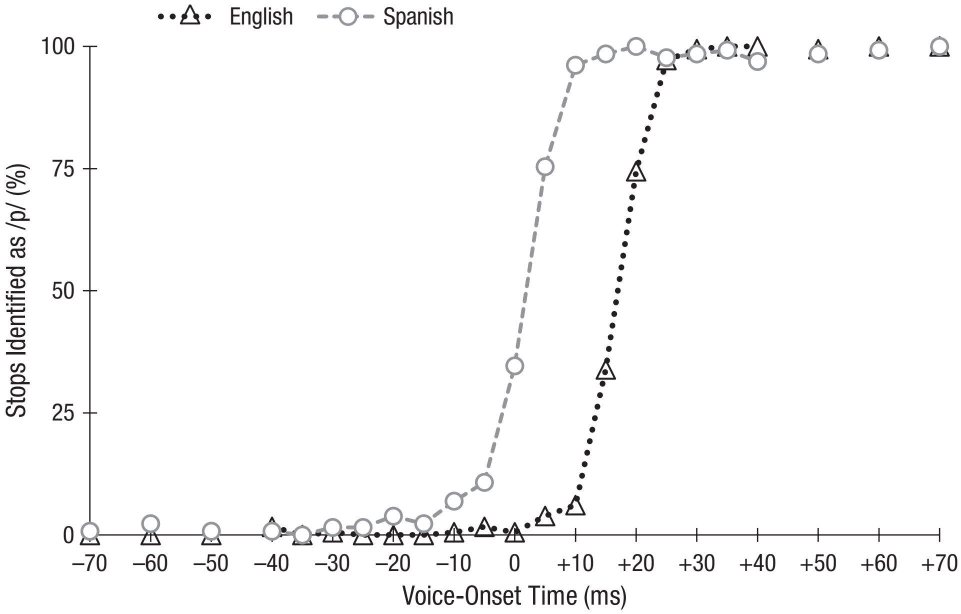

Fig. 1. Percentage of stops with short-lag voicing correctly identified as /p/ in a continuum from /ba/ to /pa/ as a function of voice-onset time and the language background of participants. Adapted from Hay (2005).

they accomplish this by learning two separate phonetic systems and facilely switching between these systems to utilize the appropriate phonetics depending on the language spoken by their interlocutor. This resonates with Grosjean's (2001) more general view that bilinguals operate in different language modes, meaning that the relative activation levels of bilinguals' two languages change dynamically as a function of the language context. These changes permit the selection of language-relevant representations over otherwise equally probable candidates in the other language.

The view that bilinguals have language-specific speech perception is also broadly compatible with the findings of studies outside the bilingual literature, in which expectations about the nature of an auditory signal influence its processing. For example, when highly unnatural synthetic speech was presented to listeners naive to the sounds being produced, they interpreted the speech as computer bleeps or other nonspeech sounds, whereas the same sounds were partially or fully intelligible to listeners expecting to hear a speech replica (Remez, Rubin, Pisoni, & Carrell, 1981). More closely related to the present study, other findings have shown that listeners' expectations about the age, gender, and dialect of a speaker produce shifts in vowel identifications consistent with accommodating the within-language variation associated with these sociolinguistic factors (see Drager, 2010).

However, none of these perceptual effects require representing the phonetic systems of different languages. Accommodation to gender and dialect could reflect representation of within-language variation. Yet the relatively large overlap across varieties of the same language—in terms of speech sounds, vocabulary, and higher linguistic processes—is thought to facilitate perceptual learning in speech (Tuinman, Mitterer, & Cutler, 2011). Thus, representation of within-language variation is viewed as less challenging than representation of between-language variation. Indeed, several investigators have recently posited strong constraints on representing speech in more than one language (e.g., Navarra, Sebastián-Gallés, & Soto-Faraco, 2005). For example, Dupoux, Peperkamp, and Sebastián-Gallés (2010) proposed that even when two languages are learned simultaneously from birth, one language will inevitably be processed like a late-acquired second language. Still, other investigators contend that, at least with early bilingual exposure, robust nativelike representation of both languages is possible (Curtin, Byers-Heinlein, & Werker, 2011; Flege, 1995; Kuhl, 2008; Sundara & Polka, 2008).

Beginning with Caramazza and his colleagues (Caramazza, Yeni-Komshian, & Zurif, 1974; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973), a number of researchers have directly probed for language-specific phonetic systems by asking whether bilinguals shift their phonetic perception across experimentally constructed language contexts, distinguished minimally by the language spoken with participants. In particular, these studies have concentrated on the language-specific differences relating to voicing perception as a function of VOT. The question, then, is whether bilinguals represent the voicing boundary of each language (English and either Dutch, French, or Spanish) and switch modes according to the communication context.

Although three initial studies (Caramazza et al., 1974; Caramazza et al., 1973; Williams, 1977) found no evidence that bilinguals shift, evidence of a shift has been obtained in all subsequent studies of which we are aware (e.g., Bohn & Flege, 1993; Elman, Diehl, & Buchwald, 1977; Flege & Eefting, 1987; Garcia-Sierra, Diehl, & Champlin, 2009; García-Sierra, Ramírez-Esparza, Silva-Pereyra, Siard, & Champlin, 2012; Hazan & Boulakia, 1993). One difference between initial and more recent studies is that the latter not only provided pretask instructions in the relevant language, but also provided salient language cues throughout the identification task itself. For example, Elman et al. (1977) appended target syllables to context-relevant English or Spanish phrases (e.g., "*Escriba la palabra* . . . "). Several researchers (Elman et al., 1977; Flege & Eefting, 1987; Garcia-Sierra et al., 2009; Grosjean, 2001) have speculated that these more elaborate cues are necessary for language-specific perception.

Though these shifts in labeling behavior based on language context may seem like prima facie evidence for language-specific phonetic systems, it is also possible that bilinguals have only one phonetic system, which rapidly recalibrates to the unique acoustic properties of each language based on language-general processes common to monolinguals. For example, Bohn and Flege (1993) suggested that language contexts could produce different patterns of range effects, which refer to the tendency to identify a stimulus with reference to its position within a range of preceding stimuli. Thus, range effects reflect sensitivity to changing acoustic distributions not dependent on dual language representation or higher-order linguistic information about which language is being spoken. To illustrate, Brady and Darwin (1978) found that English monolinguals' voicing boundaries shifted to maintain a more central position on each of several range-varying voicing continua, all presented in the same English context. Bohn and Flege suggested that different language contexts could produce range effects depending on the stop consonants within the experimental instructions and carrier sentences. Further, because speech-production data (Caramazza et al., 1973; Flege & Eefting, 1987; Williams, 1977) indicate that the range of stops in English speech cues would likely center at a longer VOT than the range in the other languages tested, range effects naturally explain a more Englishlike voicing boundary (at a longer VOT) in the English context. However, range effects would not explain truly native voicing perception in both languages because these shifts are generally much smaller than the phonetic distance between the languages' different voicing boundaries (Bohn & Flege, 1993).

The evidence to date cannot distinguish these accounts. Because the language-general account attributes cross-language shifts to recalibration effects produced by acoustically distinct language contexts, this account predicts shifts in monolinguals—and of a magnitude equal to that in bilinguals (Bohn & Flege, 1993). Indeed, comparisons between Spanish-English bilinguals and English monolinguals have twice confirmed these predictions (Bohn & Flege, 1993; Garcia-Sierra et al., 2009). However, Garcia-Sierra et al. (2009) argued that a language-general account is inconsistent with positive correlations that they and Elman et al. (1977) found between shift magnitude and bilingual proficiency. Unfortunately, contexts in these studies were partly cued by the language spoken with participants, which may have differed as a function of participant proficiency. Finally, García-Sierra et al. (2012) recorded event-related potentials from Spanish-English bilinguals who silently read a Spanish or an English magazine. Results showed preattentive sensitivity to a voicing contrast in one or the other language, according to the context. Though impressive, these results again could reflect recalibration effects created by speaking with bilinguals in the context-relevant language (before recording).

The Present Experiment

To examine whether separate phonetic systems contribute uniquely to cross-language shifts in perception, we asked in the present experiment whether Spanish-English bilinguals shift voicing perception across acoustically controlled language contexts. As indicated previously, a challenge to probing for language-specific phonetic systems is to cue language contexts without introducing acoustic differences that might by themselves modulate shifting patterns. To address this challenge, we provided recorded English instructions in both language contexts and presented target stimuli in isolation rather than in language-relevant phrases. Contexts were instead cued by the ostensible language membership of the target stimuli themselves—a continuum from *bafri* to *pafri*, pseudowords in both languages. Following evidence that language-specific phonetic information constrains bilingual lexical activation (Ju & Luce, 2004; but see Lagrou, Hartsuiker, & Duyck, 2011), we conveyed the language membership of these pseudowords not only with explicit instructions, but also by manipulating the phonetic makeup of pseudoword endings. This resulted in perceptually salient, language-specific realizations of the *r* segment. These endings were closely matched on acoustic properties that might modulate a voicing boundary and did not differentiate the VOT ranges of the contexts.

If bilinguals have language-specific phonetic systems, then shifts in phonemic labeling should occur as a function of language across these acoustically controlled contexts; however, monolinguals should show a much smaller or nonexistent shift (because they presumably have only one language-general phonetic system). To test this hypothesis, we included English monolinguals who labeled target words in the exact same contexts. To increase power for detecting differences between language groups, we compared highly proficient bilinguals who had extensive exposure to both languages from early childhood with monolinguals who had minimal Spanish exposure and proficiency.

Method

Participants

Participants were 32 Spanish-English bilinguals and 32 English monolinguals, all of whom were undergraduates at the University of Arizona. We randomly assigned half the participants in each language group to the English condition and the other half to the Spanish condition.

Bilinguals and monolinguals were distinguished on the basis of their responses to a detailed questionnaire on language background and Spanish and English self-evaluation scores. The self-evaluation score for each language was created by averaging self-ratings on separate 5-point scales (1 = *very poor*, 5 = *almost perfect*) of speaking and comprehension skills. English monolinguals were native English speakers whose Spanish scores did not exceed 1.5 (M = 1.22) and whose experience with a second language, Spanish or otherwise, was limited to 1 year or less of formal classroom instruction. Spanish-English bilinguals received at least 10% of their input from one or more native speakers of each language before the age of 8 years (mean Spanish age = 0.27 years; mean English age = 4.4 years) and had a self-evaluation score of at least 3.5 in each language (mean Spanish score = 4.4; mean English score = 4.83). Further, Spanish-English bilinguals' experience with a third language was limited to 1 year or less of formal classroom instruction. The criterion for bilinguals' age of second-language acquisition was based on a variety of second-language neural and behavioral differences between bilinguals divided at around this age of acquisition (e.g., Johnson & Newport, 1989; see Silverberg & Samuel, 2004).

Materials

An instruction module controlled by the computer program DMDX (Forster & Forster, 2003) presented instructions simultaneously in spoken and written form. All instructions in both contexts were in English. Spoken instructions (duration = 67 s) were recorded by a male native speaker of English. Depending on the context (i.e., Spanish vs. English), instructions stressed that *bafri* and *pafri* were real Spanish or English words spoken by a native speaker of that language. Pseudowords were never spoken in the recording; they appeared only in the written text. Instructions for the Spanish context were identical to those for the English context, with the exception that all occurrences of the word *English*, both in the recording and text, were replaced with *Spanish*.

The Spanish and English continua from *bafri* to *pafri* were created from natural speech recorded from a female Spanish-English bilingual. Each continuum was created by appending to every token on a base continuum from *baf_* to *paf_* the ending *ri* portion from either a Spanish or English pronunciation of *pafri* (Spanish /ɾi/ or English /ɹi/), as appropriate. These endings differed primarily because of the *r* segment and its coarticulation effects in the *i* segment. Spanish /ɾ/ (as in *frío*, meaning "cold") and English /ɹ/ (as in *freedom*) both have an alveolar place of articulation, but the former is produced as a tap, whereas the latter is an approximant. Spanish /ɾ/ is perceptually more similar to flapped /d/ and /t/ in American English (as in *latter*), but English speakers do not strongly identify this segment with any in English (Rose, 2010). Similarly, English /ɹ/ is described as foreign sounding to Spanish speakers (Dalbor, 1980). These endings were resynthesized in Praat software (Boersma & Weenink, 2010) to control properties that could potentially influence voicing perception in the pseudoword-initial stops—namely, duration (both = 141 ms; Summerfield, 1981), F0 onset (English ɹi = 179 Hz, Spanish ɾi = 181 Hz), and contour (Abramson & Lisker, 1985; Haggard, Ambler, & Callow, 1970).

To create the base continuum, we digitally stripped a Spanish pronunciation of *pafri* (/pafɾi/) of its final two segments, leaving *paf_*. This truncated token was then used to create 14 variants whose VOT values increased in equal steps of 5 ms from −35 to +35, skipping 0 VOT. The +5 VOT *paf_* variant was created first by digitally excising the entire voiceless portion of the initial stop consonant except the initial 5 ms consisting of the release burst. Then, between the release burst and the onset of voicing, voiceless portions of the *p* in other Spanish *pafri* tokens were successively inserted to create the remaining six voiceless *paf_* tokens. To create the seven voiced tokens, we successively added prevoicing intervals from the *b* in Spanish pronunciations of *bafri* (/bafɾi/) to the +5 VOT *paf_* token prior to the release burst.

Procedure

At the beginning of each session, participants in both language contexts were given an overview of the experiment by an English-speaking experimenter. They were then individually seated in front of a computer and fitted

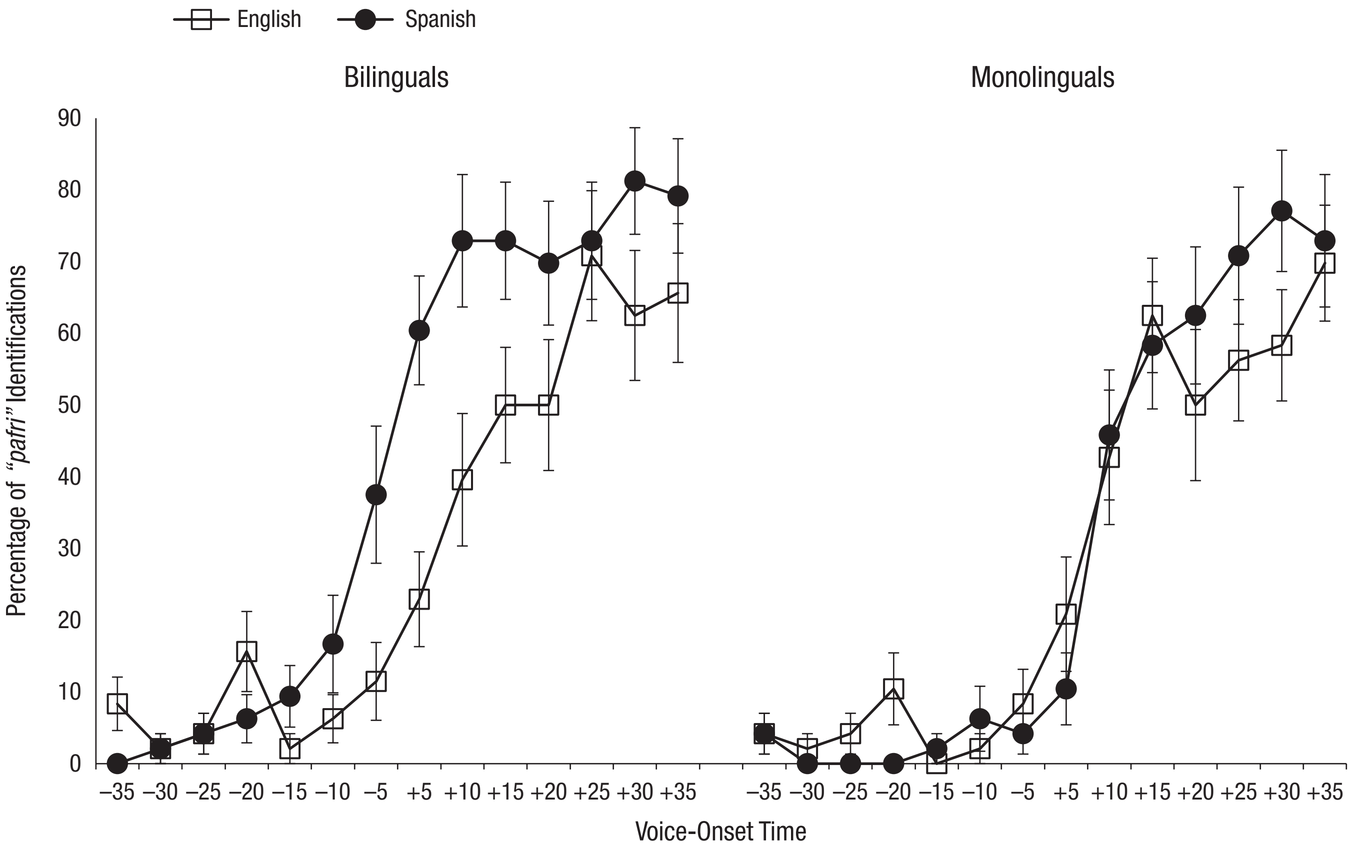

Fig. 2. Percentage of correct "pafri" identifications as a function of voice-onset time and the language background of participants. Results are shown separately for bilinguals and monolinguals. Error bars show standard errors of the mean.

with headphones to begin the English instruction module.

The experimental task immediately followed the module. Each trial began with a fixation cross for 1 s, followed by the appearance of the stimuli *pafri* and *bafri* on either side of the screen (side order was counterbalanced across participants) for the remainder of the trial. There was 1 s of silence between the appearance of the pseudowords and the target stimulus, which was delivered binaurally through headphones at a comfortable intensity. Participants were instructed to indicate which of the two "words" was spoken by pressing the appropriate button on a response box. They had 4.1 s to provide a response before the program automatically initiated the next trial. There were three initial practice trials and three blocks of 14 randomly ordered experimental trials corresponding to the 14 VOT variants.

Results

If language-specific phonetic systems contribute to cross-language shifting, bilinguals should show a larger effect of context than monolinguals only for stops that fall between the voicing identification boundaries of English and Spanish monolinguals—namely, short-lag stops that these groups differentially identify as voiced (*b*) and voiceless (*p*), respectively (Hay, 2005; Williams, 1977); bilinguals should not show a larger effect of context than monolinguals for stops outside these boundaries, including long-lead stops that both monolingual groups identify as voiced and long-lag stops that both identify as voiceless. We tested this prediction by constructing a token-type factor with three levels: long-lead tokens (−30 and −35 VOT), short-lag tokens (+5 and +10 VOT), and long-lag tokens (+30 and +35 VOT). The VOT values selected were conservative estimates, consistent with previous studies (Hay, 2005; Williams, 1977), of which tokens on our continuum would fall within versus outside English and Spanish monolingual-voicing boundaries.

Figure 2 plots each group's percentage of voiceless (*pafri*) identifications for each of the 14 VOT tokens in the Spanish and English contexts. Submitting these percentages to a three-way analysis of variance (ANOVA) revealed a significant interaction between language group, language context, and token type, $F(2, 120) = 4.28$, $p < .05$. Planned ANOVAs revealed a significant two-way interaction between language group and context within short-lag tokens, $F(1, 60) = 7.65$, $p < .01$, $r = .34$, but not within long-lead tokens, $F(1, 60) = 0.80$, $p > .05$, $r = .11$, or within long-lag tokens, $F(1, 60) = 0.11$, $p > .05$, $r = .04$. The main effect of context was not significant within long-lead tokens, $F(1, 60) = 2.22$, $p > .05$, $r = .19$, or within long-lag tokens, $F(1, 60) = 3.09$, $p > .05$, $r = .22$. Following up the two-way group-by-context interaction within short-lag tokens, *t* tests revealed that bilinguals' percentage of voiceless identifications to these particular tokens was significantly higher in the Spanish context ($M = 66.67$, $SE = 7.76$) than in the English context ($M = 31.25$, $SE = 6.78$), $t(30) = -3.44$, $p = .001$, $r = .53$, whereas that of monolinguals was not significantly different across these contexts (Spanish: $M = 28.13$, $SE = 6.03$; English: $M = 31.77$, $SE = 7.54$), $t(30) = 0.38$, $p > .05$, $r = .06$. In sum, only bilinguals shifted voicing perception across language contexts, and they did so only for short-lag stops that Spanish and English monolinguals identify differently.

Discussion

The present results support the hypothesis that bilinguals shift perception across language contexts by switching between language-specific phonetic systems rather than solely by recalibrating, like monolinguals do and as the language-general account assumes, to the unique acoustic properties of each context. Spanish-English bilinguals differentially identified the exact same pseudoword-initial stop consonants as *b* and *p* across acoustically controlled English and Spanish contexts, respectively. This shift, which reflects these languages' different *b* and *p* boundaries, cannot be explained by the small acoustic difference between contexts created by the language-specific pseudoword endings: Recall that these endings were resynthesized to control for acoustic properties that might influence voicing perception. The interpretation that bilinguals' shift instead reflects language-specific phonetic systems is buttressed by our additional finding that English monolinguals failed to shift across the exact same contexts.

The present results conflict with those of two previous studies showing shifts in both bilinguals and monolinguals. This suggests that our acoustic control over language contexts successfully prevented some recalibration effect that previously elicited shifts in both language groups and masked bilinguals' unique capacity to shift based on language-specific phonetic systems.[1] Unlike in the present study, for example, language contexts in these previous studies were cued by speech materials with different VOT ranges, which possibly elicited shifts due to range effects (Bohn & Flege, 1993).

Because the bilinguals tested here all learned both languages before the age of 8 years, an important question for future research is whether these participants' 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 (Flege, 1995; Kuhl, 2008). Instead, late bilinguals may develop language-general representations or cue-weighting strategies that constitute a compromise between similar cross-language categories (Flege, 1995; Lotto, Sato, & Diehl, 2004).

The present results may seem somewhat at odds with recent claims that even "early bilinguals" (i.e., those who acquired both of their languages at an early age) are strongly constrained in representing speech in both languages (Dupoux et al., 2010; Navarra et al., 2005). However, these claims derive from evidence that early bilinguals do not always show nativelike perception in their nondominant (weaker) language. Note that bilinguals may deviate behaviorally from native norms in one language yet still represent this language separately from the other language to some extent, which permits cross-language shifting. Alternatively, this deviation may reflect a failure to sufficiently deactivate the other language in the language context provided, especially because these studies did not report any deliberate attempt to cue the context throughout testing. As noted previously, there is speculation that bilinguals rely on elaborate contextual cues for language-specific processing. Thus, although our results may reflect some fundamental difference in perceptual flexibility, perhaps because of some difference in the bilingual population or the speech materials tested,[2] we suspect that our results simply offer a different window on similar underlying flexibility.

Our results likewise complement theories that assume that humans possess a robust capacity to represent two languages, because these theories have not addressed whether the languages are represented in language-specific systems (Curtin et al., 2011; Flege, 1995; Kuhl, 2008; Sundara & Polka, 2008). Thus, these theories are compatible with a language-general system in which the properties intrinsic to each speech sound (e.g., the long VOT in the English *p*, contrasting with the short VOT in the Spanish *p*) by themselves determine the representation onto which that sound is mapped, independent of extrinsic language cues. Our results contradict this strictly intrinsically cued mapping inasmuch as bilinguals identified the exact same pseudoword-initial stops very differently depending on the language context. The implication is that bilinguals can use extrinsic language cues to map onto distinct representations sounds between languages that might overlap too much in phonetic space for their own properties to mediate this mapping (e.g., Spanish *p* and English *b*, both with short VOTs).

Note, however, that it remains uncertain which extrinsic language cues bilinguals used. Reliance on the pronunciation of the pseudoword endings would accord with evidence that phonetic cues can constrain bilingual lexical activation (Ju & Luce, 2004). Reliance on the instructions alone, the only other cue, would indicate a shift driven solely by expectations about the language being spoken. Supporting this latter possibility are the sociolinguistic studies showing that speech perception can be influenced by expectations about the speaker's age, gender, and dialect (see Drager, 2010). For example, when instructed to match word-medial vowels to the perceptually closest token on a synthesized vowel continuum, monolinguals produce "matches" in the direction of the visually cued dialect (Niedzielski, 1999).

However, the resemblance of these within-language shifts to those in the present study should be interpreted cautiously, because they were all obtained with words (not pseudowords) and thus may reflect postlexical judgments rather than a remapping at the phonetic level (McQueen, Cutler, & Norris, 2006). Thus, it could be that separate phonetic systems are established only for phonetic variation across languages, not within them (in a natural learning environment). This would contradict the view that learning speech is more difficult in another language than in another dialect, because of the greater overall differences between languages than between dialects (Tuinman et al., 2011). However, emerging evidence indicates that distinct linguistic environments, whether in terms of lexical context (Feldman, Myers, White, Griffiths, & Morgan, 2013) or rhythmic distance (Sundara & Scutellaro, 2011), can actually facilitate the acquisition of phonetic variants by cuing their differentiation.

In conclusion, the present results provide evidence within the domain of speech perception for the general view that bilinguals operate in different language modes (Grosjean, 2001), which allows them to avoid potential confusion of incompatible surface structure across languages. As our results here show, the mind is capable of accommodating the phonetic systems of different languages.

Author Contributions

K. Gonzales developed the study concept and design in consultation with A. J. Lotto. Data collection and analysis were performed by K. Gonzales. Both authors interpreted the results. K. Gonzales drafted the manuscript, and A. J. Lotto provided critical revisions. Both authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Rebecca Gómez, LouAnn Gerken, and three anonymous reviewers for helpful comments, and Jessica Maye and Olimpia Rosenthal for technical assistance.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. Garcia-Sierra, Diehl, and Champlin (2009) suggested that monolingual shifts reflect some minimal Spanish proficiency. Following this account, however, one would expect the monolinguals in previous studies to be higher in Spanish proficiency than those tested here. In fact, monolinguals in Garcia-Sierra et al. reported comparably low self-ratings on Spanish listening and speaking skills, and monolinguals in Bohn and Flege (1993) reported no knowledge of Spanish whatsoever. Thus, Spanish proficiency cannot easily account for the monolingual shifts observed in these previous studies.
2. Preliminary results indicate that Chinese-English bilinguals' sensitivity to lexical tone is language-specific (Quam & Creel, 2013).

References

Abramson, A. S., & Lisker, L. (1985). Relative power of cues: F0 shift versus voice timing. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 25–33). New York, NY: Academic Press.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer (Version 5.1.44) [Computer program]. Retrieved from www.fon.hum.uva.nl/praat/

Bohn, O. S., & Flege, J. E. (1993). Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments. *Journal of Phonetics*, *21*, 267–290.

Brady, S. A., & Darwin, C. J. (1978). A range effect in the perception of voicing. *Journal of the Acoustical Society of America*, *63*, 1556–1558.

Caramazza, A., Yeni-Komshian, G., & Zurif, E. B. (1974). Bilingual switching: The phonological level. *Canadian Journal of Psychology*, *28*, 310–318.

Caramazza, A., Yeni-Komshian, G., 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.

Curtin, S. A., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*, 492–504.

Dalbor, J. (1980). *Spanish pronunciation; Theory and practice: An introductory manual of Spanish phonology and remedial drill*. New York, NY: Holt, Rinehart, and Winston.

Drager, K. (2010). Sociophonetic variation in speech perception. *Language & Linguistics Compass*, *4*, 473–480.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: Stress "deafness" in simultaneous French–Spanish bilinguals. *Cognition*, *114*, 266–275.

Elman, J. L., Diehl, R. L., & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, *62*, 971–974.

Feldman, N. H., Myers, E. B., White, K. S., Griffiths, T. L., & Morgan, J. L. (2013). Word-level information influences phonetic learning in adults and infants. *Cognition*, *127*, 427–438.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

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

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.

Garcia-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. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain & Language*, *121*, 194–205.

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 117–133). Malden, MA: Blackwell Publishing.

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

Hay, J. F. (2005). *How auditory discontinuities and linguistic experience affect the perception of speech and non-speech in English- and Spanish-speaking listeners* (Doctoral dissertation). Available from Proquest Dissertations and Theses database. (UMI No. 3203519)

Hazan, V. L., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, *36*, 17–38.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Ju, M., & Luce, P. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, *15*, 314–318.

Kuhl, P. K. (2008). Linking infant speech perception to language acquisition: Phonetic learning predicts language growth. In P. McCardle, J. Colombo, & L. Freund (Eds.), *Infant pathways to language: Methods, models, and research directions* (pp. 213–243). New York, NY: Erlbaum.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 952–965.

Lotto, A. J., Sato, M., & Diehl, R. L. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. In J. Slifka, S. Manuel, & M. Matthies (Eds.), *From sound to sense: 50+ years of discoveries in speech communication* (pp. C181–C186). Cambridge, MA: MIT.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 391–409.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*, 1113–1126.

Navarra, J., Sebastián-Gallés, N., & Soto-Faraco, S. (2005). The perception of second language sounds in early bilinguals: New evidence from an implicit measure. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 912–918.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*, 62–85.

Ortega-LLebaria, M., Faulkner, A., & Hazan, V. (2001). Auditory-visual L2 speech perception: Effects of visual cues and acoustic–phonetic context for Spanish learners of English. In D. W. Massaro, J. Light, & K. Geraci (Eds.), *Proceedings of AVSP 2001: International Conference on Auditory-Visual Speech Processing* (pp. 149–154). Aalborg, Denmark: ISCA Archive.

Quam, C., & Creel, S. C. (2013). *Both lifelong language experience and immediate language context affect lexical-tone processing*. Manuscript submitted for publication.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, *212*, 947–950.

Rose, M. (2010). Differences in discriminating L2 consonants: A comparison of Spanish taps and trills. In Y. Watanabe, M. Prior, & S.-K. Lee (Eds.), *Selected proceedings of the 2008 Second Language Research Forum* (pp. 181–196). Somerville, MA: Cascadilla Proceedings Project.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, *51*, 381–398.

Summerfield, A. Q. (1981). On articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1074–1095.

Sundara, M., & Polka, L. (2008). Discrimination of coronal stops by bilingual adults: The timing and nature of language interaction. *Cognition*, *106*, 234–258.

Sundara, M., & Scutellaro, A. (2011). Rhythmic distance between languages affects the development of speech perception in bilingual infants. *Journal of Phonetics*, *39*, 505–513.

Tuinman, A., Mitterer, H., & Cutler, A. (2011). Perception of intrusive /r/ in English by native, cross-language and cross-dialect listeners. *Journal of the Acoustical Society of America*, *130*, 1643–1652.

Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, *21*, 289–297.

Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception & Psychophysics*, *26*, 95–104.