Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode?

Mark Antoniou\textsuperscript{a,}\textsuperscript{,}, Michael D. Tyler\textsuperscript{a,}\textsuperscript{b}, Catherine T. Best\textsuperscript{a,}\textsuperscript{c}

\textsuperscript{a} Marcs Institute, Building 5, Bankstown Campus, University of Western Sydney, Locked Bag 1797, Penrith, NSW 2751, Australia
\textsuperscript{b} School of Social Sciences and Psychology, University of Western Sydney, Australia
\textsuperscript{c} Haskins Laboratories, New Haven, CT, USA

\section*{A R T I C L E   I N F O}

Article history:
Received 9 June 2011
Received in revised form 12 April 2012
Accepted 11 May 2012

\section*{A B S T R A C T}

How listeners categorize two phones predicts the success with which they will discriminate the given phonetic distinction. In the case of bilinguals, such perceptual patterns could reveal whether the listener’s two phonological systems are integrated or separate. This is of particular interest when a given contrast is realized differently in each language, as is the case with Greek and English stop-voicing distinctions. We had Greek–English early sequential bilinguals and Greek and English monolinguals (baselines) categorize, rate, and discriminate stop-voicing contrasts in each language. All communication with each group of bilinguals occurred solely in one language mode, Greek or English. The monolingual groups showed the expected native-language constraints, each perceiving their native contrast more accurately than the opposing nonnative contrast. Bilinguals’ category-goodness ratings for the same physical stimuli differed, consistent with their language mode, yet their discrimination performance was unaffected by language mode and biased toward their dominant language (English). We conclude that bilinguals integrate both languages in a common phonetic space that is swayed by their long-term dominant language environment for discrimination, but that they selectively attend to language-specific phonetic information for phonologically motivated judgments (category-goodness ratings).

\section*{1. Introduction}

When a perceiver is bilingual, performance on cross-language speech perception tasks may indicate whether the first language (L1) and second language (L2) are integrated or kept separate. Performance equal to that of native, monolingual listeners in each language has been interpreted as evidence that bilinguals have developed separate L1 and L2 phonetic categories, whereas intermediate discrimination ability has been taken to indicate that the bilinguals rely on a set of merged categories for both languages (Flege, 2002).

Bilinguals may differ in their perception from monolinguals because of their L1–L2 sensitivity and the interaction that occurs between their languages. Even fluent, early bilinguals who use their L2 daily may show some L1 influence on their perception of the L2. For example, Spanish–Catalan\textsuperscript{1} bilinguals retain perceptual dominance from their L1 (Spanish) when categorizing and discriminating Catalan contrasts that Spanish lacks, despite years of daily experience with both languages in a bilingual environment (Pallier, Colomène, & Sebastián-Gallés, 2001; Sebastián-Gallés & Soto-Faraco, 1999). Conversely, however, there is also much evidence that learning an L2 may change the way the listener perceives the L1 (Flege, 1993; Flege, Schirru, & MacKay, 2003; Flege, Yeni-Komshian, & Liu, 1999).

Several theoretical models have been proposed to account for L2 users’ perception of speech segments. The Speech Learning Model (Flege, 1995, 2002) was designed to account for L2 speech production, but assumes that it is strongly influenced by perceptual biases and therefore seeks to explain the perceptual reorganization that occurs due to L2 acquisition. The Native Language Magnet model (Kuhl, 1992, 1993) seeks to explain how early attunement to the native language (L1) warps the perceptual space, such that learning after early childhood is influenced by that L1. The Perceptual Assimilation Model (PAM; Best, 1994, 1995) accounts for naive listeners’ discrimination of unfamiliar nonnative contrasts, and its extension to L2 perceptual learning, PAM-L2 (Best & Tyler, 2007) predicts the success with which learners should discriminate L2 contrasts. According to PAM-L2, L2 phones are first assimilated into already existing L1 categories...

\textsuperscript{1} By convention, naming order specifies which language is the L1 versus L2. For example, Spanish–Catalan bilinguals learned Spanish first and Catalan second, and the Greek–English bilinguals in our studies learned L1 Greek from birth and acquired L2 English second when they attended school.

© 2012 Elsevier Ltd. All rights reserved.
or dissimilated from existing categories and established as new categories firstly on a phonetic level. As the learner’s L2 vocabulary expands they attune to the higher-order organization of the L2 phonology, and phones come to be discriminated on the basis of meaningful categorical differences that are lexically relevant in the L2.

However, none of the aforementioned models were designed to specifically address the performance of early fluent bilinguals. Here, our purpose is not to contrast the various models of nonnative speech perception, but rather to focus on testing PAM-L2 predictions, which appear to be best suited to the bilingual case as only that model addresses and distinguishes between two levels in cross-language perception: the phonetic level of sensitivity to gradient, physical details and the abstract phonological level of perceiving meaningful distinctions.

The original PAM framework addresses nonnative speech perception by naive adults. But whereas monolingual listeners assimilate unfamiliar nonnative phones to L1 categories, bilingual listeners face a different challenge, as their perceptual system must accommodate to two known languages (Grosjean, 2008; Romaine, 1989). When perceiving speech in a foreign language, the naive monolingual listener often (but not always) perceptually assimilates nonnative phones to existing native language categories. Based on the possible assimilation patterns to the L1, PAM predicts the success with which naive listeners will discriminate nonnative contrasts. According to PAM predictions, two nonnative phones perceived as exemplars of two different native categories will result in two-category (TC) assimilation, and discrimination will be excellent. When one nonnative phone is perceived as speech, but not as reliably belonging to any single native category, and another is assimilated to a native category, that nonnative contrast will form an uncategorized–categorized pair (UC). Two nonnative phones perceived as equally good or poor exemplars of a native category will result in single category (SC) assimilation, and discrimination will be poor. Category goodness (CG) assimilation will result when two nonnative phones are assimilated to the same native category but one is perceived as a better exemplar of that category than the other. The size of this category-goodness difference will result in discrimination ranging from good to very good (for descriptions of all possible assimilation types see Best, 1995).

PAM-L2 extends the original PAM to explain effects of L2 perceptual learning and predict the success with which learners will discern L2 contrasts depending on how they are assimilated to the L1 on the basis of both phonetic and phonological similarities, and predicts on this basis which new L2 phonological categories will be established. Phonetic categories describe clusters of individual phones, which differ from each other in a gradient fashion. Phonological categories are comprised of one or more phonetic category and specify the linguistic function of phonetic segments within the lexicon of a language, that is, which segments are distinctive and which are not. Bilinguals may thus perceive phonetic differences between L1 and L2 segments (e.g., Greek versus English realizations of /p/) while simultaneously recognizing properties that indicate a shared (or ‘common’) phonological contrast (e.g., /b/ vs. /p/). That is, a common L1/L2 phonological category can include two distinct regions of phonetic space for each language, and those language-specific phonetic categories are free to evolve without necessarily influencing one another.

From a PAM-L2 perspective, language users never stop learning, and yet, there is a qualitative difference between individuals who are actively acquiring an L2 versus stable bilinguals (early state vs. end state). PAM-L2 focuses on learners who are actively acquiring their L2, and identifies factors that are important in the formation of L2 speech categories (Best & Tyler, 2007). For instance, the contribution of age of L2 learning is modulated by the relative quantity and quality of input from native L2 speakers (Flege & Liu, 2001; Jia & Aaronson, 2003; Jia, Strange, Wu, Collado, & Guan, 2006), length of residence in an L2-speaking environment, and relative L1:L2 usage (Flege, 1999, 2002; Flege & MacKay, 2004). Based on these assumptions, it follows that early bilinguals who have used their L2 so extensively that they have become L2-dominant should be the most likely to have developed L2 categories because they have strong L2 biases on all of these factors. Specifically, the type of L2-dominant bilingual we focus on here acquired both their L1 and L2 from native speakers, they were born and have lived the entirety of their lives in the country in which their L2 is spoken, and their L1:L2 usage ratio is heavily skewed toward the L2 as it is used in a much wider variety of social situations. Thus, this type of bilingualism could be highly informative to our understanding of L2 speech category formation. However, little to no research has focused on this type of bilingual. Most bilingual and L2-learner speech perception research has instead focused on other questions suited to the populations that have much more often been studied: L1-dominant bilinguals, and bilinguals living in more balanced bilingual environments where both languages are spoken widely and on a daily basis.

While researchers have interpreted bilinguals’ performance in speech perception tasks as evidence for their stable phonological organization for the two languages, bilingual performance on such tasks may instead depend on what language-specific information a bilingual attends to. This may be especially important to consider in light of possible differences between L2-dominant bilinguals and L1-dominant or bilingual–environment bilinguals. The Language Mode framework (Grosjean, 1998, 2001, 2008) posits that, depending on the communicative situation, just one of a bilingual’s languages may be activated (called monolingual mode, although the other language is never completely deactivated), or both languages may be activated (bilingual mode), thereby directing attention to language-appropriate information. As we will be comparing (L2-dominant) bilinguals with monolinguals of each language, we instead refer to monolingual mode as unilingual mode to avoid confusion. These language modes are thought to influence all aspects of language processing (Grosjean, 1998); however, research has barely touched on the perceptual effects of language mode in bilinguals. Although clear language mode effects have been observed for speech production (Antoniou, Best, Tyler, & Kroos, 2010, 2011; Hazan & Boulakia, 1993; Magloire & Green, 1999), language mode effects on speech perception are less clear.

Despite the lack of a theory of segmental speech perception that explicitly considers the contribution of the situational language context in bilinguals, the effect of language mode on perception has been investigated in a number of studies. Language mode has not always resulted in shifts in perception. In one study, Canadian-French–English bilinguals identified synthetic /p, t, k/ in separate French and English unilingual language modes (Caramazza, Yeni-Komshian, Zurif, & Carmone 1973). The bilinguals were unaffected by language mode and showed intermediate identification patterns relative to the monolingual French and English listeners, in both English and French language modes. Similarly, language mode failed to influence the identification of Spanish–English bilinguals’ discrimination of /b, p/ (Williams, 1977). Their category boundaries were at an intermediate point relative to those of monolingual listeners of Spanish and English. These combined findings suggested that bilinguals differ from monolinguals in their perceptual sensitivity.

Yet, some other studies have demonstrated language mode effects on bilinguals’ perceptual sensitivity, as evident in phoneme category shifts for both language modes. For example, Dutch–English bilinguals shifted their phonemic boundaries (by
the bilinguals were L2 dominant, and this was reflected in their 
language use. They used English (60%) more than Spanish (40%) on a daily basis, 
(5.6). All were students at the University of Texas and reported that 

a dr e p o r t e dt h a tt h e yw e r em o r eproficient at speaking, reading, 

though 12 of the 15 bilinguals had acquired Spanish as their L1. All 

English in English mode. Despite the small shift in boundary with 

b o u n d a r i e ss h i f t e dt o w a r dt h ev o i c i n gl e a d( n e g a t i v eV O T) e n do f 

(Garcia-Sierra, Diehl, & Champlin, 2009). The bilinguals' category 

stimuli from a / 

started learning their L2 between 3 and 12 years of age) categorized 

gualism', and was determined from each bilingual's language of 

identification patterns that reflected their dominant language, as 

defined by those researchers. However, we must note that in that 

study 'language dominance' was equated with 'strength of bilin-

gualism', and was determined from each bilingual's language of 

exposure during their early years. Thus, it did not necessarily 

represent their dominant language at the time of testing.

In a more recent study, Spanish–English early bilinguals (who 

started learning their L2 between 3 and 12 years of age) categorized 

stimuli from a /g̊a/–/ka/ continuum in separate language modes 

(Garcia-Sierra, Diehl, & Champlin, 2009). The bilinguals' category 

boundaries shifted toward the voicing lead (negative VOT) end of 

the continuum in Spanish mode and toward lag VOT values of 

English in English mode. Despite the small shift in boundary with 

the two language modes, the bilinguals nonetheless showed an 

English-like phonemic boundary overall, and this was observed even 

though 12 of the 15 bilinguals had acquired Spanish as their L1. All 

had reported that they were more proficient at speaking, reading, 

writing and understanding in English (6.5 out of 7) than in Spanish 

(5.6). All were students at the University of Texas and reported that 

they used English (60%) more than Spanish (40%) on a daily basis, 

with their teachers, classmates and friends. Therefore, it is likely that 

the bilinguals were L2 dominant, and this was reflected in their 

perceptual performance.

This review of the bilingual perception literature demonstrates that 

language mode influences bilinguals' perception, and language 

dominance appears to be an important factor that mediates whether 

language mode effects will emerge. However, none of these perceptu-

als studies explicitly controlled for the environment in which the 

bilinguals acquired their two languages, or their language domi-

nance at the time of testing. Therefore, in order to clearly under-

stand the influence of language dominance and language mode on 

bilingual perception, it is necessary to use strict selection criteria 

test the perception abilities of bilinguals who are clearly L2 

dominant, and to do so in L2 versus L1 language modes. If both 

language mode and language dominance affect bilinguals' speech 

perception, theories of L2 segmental perception (e.g., PAM-L2) will 

need to be extended to account for these factors.

2. The present study

To test the effects of language mode and language dominance on 
native speech perception, we recruited L2-dominant, early bilinguals 

who were born in Australia, acquired their L1 (Greek) from their 
native-speaking parents and family members in Australia, and first 

began speaking their L2 (English) when they attended school. These 

participants became L2-dominant and now have larger L2 vocabu-

laries and use their L2 in a wider variety of social contexts, but 
maintain their L1 in certain contexts on a regular basis. We 

investigate whether this L2-dominance is reflected in their percep-
tion of L1 and L2 stop voicing contrasts by comparing them to Greek 

and English monolinguals, whose assimilation patterns of the 

opposing, nonnative language are predicted by PAM.

In previous work, we investigated the production of stop voicing 

by L2-dominant Greek–English bilinguals relative to monolinguals 

(Antoniou et al., 2010). Greek and English are known to differ in 

their specifications for VOT, which describes the timing relationship 

between the moment of release of the stop constriction and the 

onset of vocal fold vibration (Lisker & Abramson, 1964). We found 

that, consistent with prior reports, English monolinguals produced 

voiced stops with very short VOTs (e.g., /b/, /d/, range = −10 to 

20 ms), and voiceless stops with long-lag VOT (e.g., /p/, /t/, 

range = 65–100 ms). Also consistent with the existing literature, 

the Greek monolinguals produced voiced stops with long voicing 

lead (e.g., /b/, /d/, range = −100 to −160 ms), and the voiceless stops 

with short-lag VOT (e.g., /p/, /t/, range = 8–25 ms). Therefore, English 

/b/ and /d/ occupy a similar VOT region to Greek /p/ and /t/. 

Importantly, bilinguals' stop productions in Greek and English were 

consistent with those of their monolingual counterparts in spite of 

the phonetic–phonological mismatch. This raises the question of 

whether bilinguals are sensitive to the overlapping VOTs of the 

Greek and English stop consonants in perception as well.

The language pairing of English and Greek allows for comparison 

between monolinguals' and bilinguals' perception of stop-

voicing distinctions that differ in their phonetic realization, 

phonological status, and orthographic representation. In addition 

to the VOT overlap outlined earlier, English and Greek coronal 

stops differ in place of articulation (respectively, alveolar vs. 

dental; Cox & Palethorpe, 2007; Newton, 1972), which is known to 

influence VOT as place moves from anterior to posterior in the 

vocal tract (Cho & Ladefoged, 1999). We investigated coronal and 

bilabial stops so as to observe any differences in perception when 

place of articulation differs between L1 and L2 (coronals) versus 

when it is the same (bilabials), as well as how bilinguals deal with 

these language-specific double (coronals: VOT and place) versus 
singular (bilabials: VOT only) articulatory differences.

Secondly, English and Greek differ phonologically. It is well 

accepted that the English voiced–voiceless stop distinction is 

phonologically contrastive. On the other hand, the phonological 

status of Greek voiced stops is a topic of longstanding debate 

(Arvaniti, 1999; Arvaniti & Joseph, 2000, 2004; Malikouti-

Drachman, 1993; Newton, 1961; Viechnicki, 1996). The debate 

concerns whether Greek voiced stops are phonological units or 

are underlingly sequences of nasal+voiceless stop (as they are 

still represented in Greek orthography). There is considerable 

variability across speakers and regional accents in the realization 

of Greek voiced stops in medial position, which may be produced 

with voicing lead and prenasalization, or without nasalization. 

Due to these differences in English and Greek phonology, it was 

necessary to test perception of stops in both initial and medial 

positions, the results of which we have presented separately in 

Experiments 1 and 2, respectively. We predict that, regardless of 

the phonological status of voiced stops for monolingual Greek 

listeners, the bilinguals will have developed L1/L2 phonological 

categories for both voiced and voiceless stops, due to their 

experience with those phonological contrasts in English.

A third difference between English and Greek is the orthogra-

phy. English stops are represented by a unique orthographic 

character (p, t, b, d). Greek voiceless stops are represented by a 

unique orthographic character (/p/, /t/), but voiced stops are represented 

by digraphs (e.g., /b/ = μτ, /d/ = ντ), reflecting their Classical Greek 

origin as sequences of nasal+voiceless stop.

These phonetic, phonological and orthographic differences pose 

challenges for bilinguals. Consistent with PAM-L2, we 

posited that L2-dominant, early bilinguals have established L2 

Note that three bilinguals had acquired English as their L1, meaning that 

these three bilinguals were L1-dominant.
phonetic categories for English realizations of /b, p, d, t/ and that these are perceptually linked to their pre-existing L1 (Greek) phonological categories /b, p, d, t/ during language acquisition. Thus, bilinguals possess common phonological categories for the two languages, which are composed of the phonetic categories of the L1 and L2 (e.g., both Greek and English realizations correspond to the interlanguage phoneme /p/, both Greek and English realizations correspond to /b/, and so on).

We extend PAM-L2 predictions to the situation of early fluent bilinguals by incorporating the factor of selective attention posited by the Language Mode framework. If the bilinguals have developed language-specific phonetic categories for stop voicing distinctions, we should observe language mode effects for tasks that require language-specific phonological judgments about stimuli. That is, their judgments should be relevant to the phonetic properties of the language that is currently mostly activated, the contextually-established language mode. For instance, categorization and goodness rating require the listener to make a judgment concerning the goodness-of-fit mapping between the phonetic properties of the stimulus and the phonological category of their native language(s). We had bilinguals in Greek mode and Greek monolinguals select from category labels presented in Greek orthography whereas bilinguals in English mode and English monolinguals selected from English labels, when making perceptual judgments about nonce stimuli containing voiced and voiceless stops recorded in English and in Greek. If bilinguals' category choices and/or goodness ratings for Greek versus English stops vary according to language mode, this would be consistent with the PAM-L2 notion that they have developed common L1/L2 phonological categories with separate L1 and L2 phonetic categories. Specifically, it is expected that higher ratings will be assigned to mode-congruent stimuli.

Participants also performed a discrimination task. According to PAM, discrimination should be most accurate when two phones are separated by a native phonological category boundary. As such, Two Category (TC) assimilation should result in the most accurate discrimination, followed by UC, then CG and finally least accurate discrimination for SC assimilations. Specifically, based on the known VOT differences between English and Greek, English monolinguals should perceive the English stops as excellent exemplars of their native stop-voicing categories, whereas they should assimilate the Greek voiced (long lead) and voiceless (short-lag) stops as either differing in goodness-of-fit as exemplars of English (short-lag) voiced stops, or as not reliably categorizable as voiced or voiceless, resulting in CG or uncategorized-voicing assimilations of the Greek bilabial and coronal contrasts, respectively. Greek monolinguals should perceive the Greek stops as excellent exemplars of their native categories, and should assimilate English voiced (short-lag) and voiceless (long-lag) stops, respectively, as good and poor exemplars of Greek voiceless (short-lag) stops, resulting in CG assimilations of the two voicing contrasts. Given the place difference between Greek and English coronals, we anticipated that monolingual listeners would be sensitive to this and show less consistency in categorization and assign lower ratings to the nonnative coronals than the nonnative bilabials. We compared bilinguals’ categorization responses and discrimination performance across language mode and against monolinguals of each language to evaluate PAM-L2 predictions.

3. Experiment 1: Perception of word-initial stop-voicing contrasts

3.1. Method

3.1.1. Participants

Twenty Australian-English monolinguals ($M_{\text{age}}$ = 23.4 years), 20 Greek monolinguals ($M_{\text{age}}$ = 26.1 years), and 40 Greek–English bilinguals participated. The bilinguals were randomly assigned to two groups that were placed in different unilingual language modes (English mode $M_{\text{age}}$ = 25.9 years; Greek mode $M_{\text{age}}$ = 25.0 years; 10 males, 10 females in each group). The English monolinguals were undergraduate students at the University of Western Sydney and participated in exchange for course credit. The Greek monolinguals were residents of Athens, Greece, and recruited from the general population by word-of-mouth. Some had knowledge of English from school, but none had extended immersion in an English-speaking country. The bilinguals were recruited from the Greek-Australian community in Sydney. All were born in Sydney, Australia, had been exposed to Greek since birth, began learning English by age six, and were dominant in English, their L2 as evidenced by their self-ratings of fluency ($M_{\text{Greek}}$: 3.8 out of 5; $M_{\text{English}}$: 5 out of 5) and daily language use ($M_{\text{Greek}}$: 32%; $M_{\text{English}}$: 88%). All continued to use both Greek and English in their everyday lives, and were literate in Greek, although English was used in a wider variety of social situations the majority of the time. They were paid $20 for their participation.

3.1.2. Stimuli

The stimuli consisted of English and Greek bilabial and coronal initial stops in consonant–vowel (CV) syllables /pa, ta, ba, da/. The stimuli were recorded from two Greek (Athens) and two English (Sydney) male native speakers. Two tokens of each target were selected from each speaker, matched as closely as possible across tokens, speakers and stop targets, for overall duration, fundamental frequency and intensity. Thus, there was a total of four tokens of each stop per language. The mean VOTs for /ba/, /da/, /pa/, and /ta/ were, respectively, 13.3 ms, 16.8 ms, 77.3 ms, and 87.2 ms in English and −118.3 ms, −103.2 ms, 13.5 ms, and 13.5 ms in Greek.

3.1.3. Procedure

All participants, both monolingual and bilingual, were tested by the same Greek–English simultaneous bilingual experimenter (MA). Care was taken to ensure that interactions with each participant occurred in only one language: English monolinguals and the bilinguals in English mode were only spoken to in English, whereas Greek monolinguals and the bilinguals in Greek mode were only spoken to in Greek. Any written and onscreen instructions were presented solely in that same language. All participants were presented with the same acoustic stimuli, regardless of language mode. Once the experiment had ended, participants were debriefed and completed self-ratings and language use questionnaires for their other language.

Participants completed a categorization task with goodness ratings using a laptop computer, Roland UA 25-EX audio interface, and Sennheiser HD 650 headphones. A closed-response set was presented onscreen for each language condition (English: {b, d, p, t, m, n, m+p, n+t}; Greek [corresponding order]: {μ, ντ, π, τ, μ+π, ντ+π, μ+π+ντ}). Following the presentation of a token, the participant selected a label by clicking on one of the options (presented in either English or Greek orthography, consistent with their language mode) using a mouse. The same token was then presented again and participants assigned a category-goodness rating ranging from 1 very strange to 7 perfect with respect to the goodness of fit to their selected category label. Participants were required to respond within a 5 s time limit. The pairs of trials (categorization followed by goodness-rating) were presented in

---

3 We asked participants to estimate their daily L1 and L2 use at separate times in the test session so as not to interfere with our language mode procedure. Thus, the percentages do not sum to 100%. Nevertheless, the figures clearly illustrate the bilinguals’ L2 dominance.
phonological match across speakers rather than a simple acoustic
speaker 2 served as the X tokens, requiring participants to make a
speaker 1 served as the A and B tokens and those recorded from
testing environments. 

\[ \text{(as to control the signal-to-noise ratio across participants} \]
\[ \text{sound-level meter was used to adjust the headphone volume so} \]
\[ \text{monolinguals were tested in a quiet room in their homes. A} \]
\[ \text{testing booth at Marcs, or in a quiet room in their homes. Greek} \]
\[ \text{AXB trial was presented again at the end of the contrast block.} \]
\[ \text{respond within 3.5 s. If the participant exceeded that limit, the} \]
\[ \text{blocks (e.g., Greek /ba/–/pa/ could not be preceded or followed by} \]
\[ \text{the same contrast were not permitted to occur in consecutive} \]
\[ \text{blocks} \]
\[ \text{were missed.} \]
\[ \text{random order. Missed trials were not presented again (2% of all trials} \]
\[ \text{were missed).} \]

Participants also completed a categorical AXB discrimination


4. Results and discussion

4.1. Categorization of word-initial stops (CV)

Mean percentage categorization of initial-position bilabial and coronal stops are presented in Tables 1 and 2, respectively. Italics indicate the speakers’ intended category. In past work, assimilation patterns have been computed for individual listeners, but this was not feasible here due to the modest number of categorization trials. Due to the relatively small number of response options, we adopted a categorization criterion of 70%, rather than 50% or 90%, both of which have been applied in other studies in the literature (for a discussion, see Bundgaard-Nielsen, Best, & Tyler, 2011). That is, if no single label was selected for (at least) 70% of responses, that phone was considered uncategorized. Pairs of stimulus categories were then compared to determine assimilation patterns. The categorization criterion and comparison of the category-goodness ratings (where relevant) were sufficient for determining the assimilation patterns for a given contrast by a given group of listeners. Analyses of variance (ANOVAs) could not be conducted because, due to the highly consistent categorization behavior of our participants, some cells had no variance (i.e., unanimous agreement in labeling).

First, we consider categorization by the monolinguals to


### Table 1

Mean percent categorization (and goodness rating 1–7) of English and Greek word-initial bilabial stops by monolinguals and bilinguals in English (En) and Greek (Gr) Modes (Experiment 1). Values in italics indicate speakers’ intended category.

<table>
<thead>
<tr>
<th>Group</th>
<th>Consonant stimuli</th>
<th>Category label</th>
<th>Criterion met&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Categorized as</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>70% criterion is uncategorized (U), whereas &lt;70% criterion was categorized (C).</td>
<td></td>
</tr>
<tr>
<td>Monolingual English</td>
<td>Gr /ba/</td>
<td>100.0 (4.5)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>Gr /pa/</td>
<td>53.8 (4.3)</td>
<td>U</td>
<td>/p/</td>
</tr>
<tr>
<td></td>
<td>En /ba/</td>
<td>87.5 (5.0)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>En /pa/</td>
<td>100.0 (5.4)</td>
<td>C</td>
<td>/p/</td>
</tr>
<tr>
<td>Monolingual Greek</td>
<td>Gr /ba/</td>
<td>98.3 (6.1)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>Gr /pa/</td>
<td>100.0 (6.3)</td>
<td>C</td>
<td>/p/</td>
</tr>
<tr>
<td></td>
<td>En /ba/</td>
<td>5.0 (2.0)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>En /pa/</td>
<td>100.0 (3.7)</td>
<td>C</td>
<td>/p/</td>
</tr>
<tr>
<td>Bilingual EnMode</td>
<td>Gr /ba/</td>
<td>90.0 (5.3)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>Gr /pa/</td>
<td>43.8 (4.7)</td>
<td>C</td>
<td>/p/</td>
</tr>
<tr>
<td></td>
<td>En /ba/</td>
<td>81.0 (4.8)</td>
<td>U</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>En /pa/</td>
<td>100.0 (5.6)</td>
<td>U</td>
<td>/p/</td>
</tr>
<tr>
<td>Bilingual GrMode</td>
<td>Gr /ba/</td>
<td>93.2 (6.0)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>Gr /pa/</td>
<td>17.5 (6.0)</td>
<td>C</td>
<td>/p/</td>
</tr>
<tr>
<td></td>
<td>En /ba/</td>
<td>53.8 (3.1)</td>
<td>C</td>
<td>/b/</td>
</tr>
<tr>
<td></td>
<td>En /pa/</td>
<td>1.3 (1.0)</td>
<td>U</td>
<td>/p/</td>
</tr>
</tbody>
</table>

<sup>a</sup> > 70% was considered categorized (C), whereas < 70% criterion was uncategorized (U).

random order. Missed trials were not presented again (2% of all trials were missed).

when they were shared between the voiced and voiceless stop categories,
/pa/, /ta/ to either native stop voicing category. Instead, responses
were assimilated voiced Greek /ba/, /da/ to their native English /ba/, /da/ categories. Consistent with one of our possible predictions,
the English monolinguals did not consistently categorize Greek
/ba/, /da/ categories. Consistent with one of our possible predictions, when presented with Greek stimuli, the English monolinguals assimilated voiced Greek /ba/, /da/ to their native English /ba/, /da/ categories. Consistent with one of our possible predictions, the English monolinguals did not consistently categorize Greek /pa/, /ta/ to either native stop voicing category. Instead, responses were shared between the voiced and voiceless stop categories, which partially overlap in VOT with English /b/ and /d/.

Whereas English monolinguals consistently categorized the Greek voiced stops to their native categories, but not the voiceless stops, the Greeks consistently assimilated English voiceless stops.
coronal, monolinguals. More sensitive to coronal constriction location differences than (Greek: dental; English: alveolar), leading bilinguals’ ratings to be places of articulation for the coronal stops in their two languages monolinguals. A possible explanation might be due to different bilinguals in Greek mode rated Greek coronal stops as slightly p action of Stimulus Language vs. coronal). The analysis revealed a significant three-way inter-

greek (1) = 3.1, p = .006, presumably because English /b/ and Greek /p/ have short-lag VOT, whereas English /p/ has long-lag. This highlights the problem that bilinguals face when accommodating Greek and English in their perceptual system: English /b/ is perceived as a very good Greek /p/ to Greek monolinguals, but Greek /p/ lies between English /b/ and /p/ for English monolinguals. Furthermore, Greek monolinguals did not consistently categorize English /da/ tokens, perhaps because the place difference between Greek (dental) and English (alveolar) coronals made them less acceptable exemplars of Greek /da/. The two monolingual groups were thus constrained differently in their perception of stop voicing from each other’s languages.

Having established baseline response rates for monolingual listeners, we turn to the bilinguals’ responses to the same stimuli. The bilinguals categorized all mode-congruent stimuli. As shown in Tables 1 and 2, the bilinguals in Greek versus English modes show clear differences in categorization and/or goodness ratings for the same target stimuli. Therefore, due to our language context manipulation, the bilinguals attended to language-specific phonetic information when making their judgments about the voicing category (phonological identity) of the stimulus tokens.

We analyzed the rating scores using a 2 × 2 × (2 × 2) ANOVA with between-subjects factors of stimulus language (English vs. Greek) and lingualism (monolinguals vs. bilinguals), and within-subjects factors of voicing (voiced vs. voiceless) and place (bilabial vs. coronal). The analysis revealed a significant three-way interaction of Stimulus Language × Lingualism × Place, F(1, 75) = 30.5, p < .001, η² = .289. The simple two-way Stimulus Language × Lingualism interaction (Winer, 1971) was significant for coronal, F(1, 75) = 5.4, p = .023, but not bilabial stops. That is, bilinguals in Greek mode rated Greek coronal stops as slightly poorer exemplars of their respective categories than Greek monolinguals. A possible explanation might be due to different places of articulation for the coronal stops in their two languages (Greek: dental; English: alveolar), leading bilinguals’ ratings to be more sensitive to coronal constriction location differences than monolinguals. Importantly, bilinguals made language-appropriate phonological judgments in categorization in spite of overlapping VOTs for Greek voiceless and English voiced stops. We argue that in order to achieve this, bilinguals must have developed common L1–L2 phonological categories as well as finely tuned language-specific phonetic specifications for those categories.

Comparison of bilinguals and monolinguals on mode-incongruent/nonnative stops reflects their perceptual learning histories. Tables 1 and 2 show that bilinguals’ categorization and rating of mode-incongruent stimuli did not always match those of monolinguals to nonnative stimuli. Importantly, performance was similar for stimuli at the extremes of the VOT continuum, Greek /b/, /d/, and English /p/, /t/, but there were a few differences for stops with overlapping VOTs for the two languages. The most noticeable difference is on the categorization of English /b/, which Greek monolinguals identified clearly as Greek /p/, but for Greek mode bilinguals it was uncategorized.

From a PAM perspective, categorization of nonnative consonants determines the assimilation types that predict discrimination success, which reveals systematic functional relationships in the native phonology (Best, 1994, p. 261). For monolinguals, discrimination will be best when nonnative phones fall on opposite sides of a native phonological category boundary (i.e., TC and UC assimilations). Monolinguals should show excellent discrimination of their native language contrasts, and less good discrimination of the nonnative contrasts. It remains unclear exactly how PAM discrimination predictions would apply to bilingual listeners, as they were listening to their two native languages. Bilinguals may show cumulative L1–L2 sensitivity and an advantage in discrimination relative to monolinguals, or competition between the L1 and L2 may hinder their discrimination, or they may perform differently in the two language modes, equivalent to each monolingual group. These possibilities for discrimination are especially interesting given the language mode effects observed for categorization. If language mode has an effect on discrimination, then bilinguals’ assimilation patterns should predict discrimination success, consistent with PAM-L2. The assimilation types are presented in the top row for each group in Table 3, which should predict relative discrimination success according to PAM (i.e., TC > UC > CG > SC). It is clear that the bilinguals did not perform in a perfectly monolingual-like way in their assimilation patterns. For example, they did not show any

<table>
<thead>
<tr>
<th>Group</th>
<th>Consonant stimuli</th>
<th>Category label</th>
<th>Criterion met</th>
<th>Categorized as</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual English</td>
<td>Gr /da/</td>
<td>98.8 (4.7)</td>
<td>1.3 (5.0)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Gr /ta/</td>
<td>58.2 (3.9)</td>
<td>41.8 (3.5)</td>
<td>U</td>
</tr>
<tr>
<td></td>
<td>En /da/</td>
<td>82.2 (4.5)</td>
<td>17.5 (3.9)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /ta/</td>
<td>100.0 (5.3)</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Monolingual Greek</td>
<td>Gr /da/</td>
<td>89.9 (6.1)</td>
<td>2.5 (5.5)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Gr /ta/</td>
<td>3.8 (4.3)</td>
<td>95.0 (6.3)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /da/</td>
<td>37.5 (2.6)</td>
<td>53.8 (4.0)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /ta/</td>
<td>100.0 (3.5)</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Bilingual EnMode</td>
<td>Gr /da/</td>
<td>93.8 (5.0)</td>
<td>3.8 (5.0)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Gr /ta/</td>
<td>38.8 (4.6)</td>
<td>58.8 (4.2)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /da/</td>
<td>81.3 (4.9)</td>
<td>15.0 (3.7)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /ta/</td>
<td>100.0 (5.5)</td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Bilingual GrMode</td>
<td>Gr /da/</td>
<td>92.4 (5.3)</td>
<td>1.3 (5.0)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>Gr /ta/</td>
<td>13.9 (4.1)</td>
<td>82.3 (5.7)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /da/</td>
<td>62.5 (2.5)</td>
<td>31.3 (2.6)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>En /ta/</td>
<td>98.8 (2.5)</td>
<td>1.3 (1.0)</td>
<td>C</td>
</tr>
</tbody>
</table>

* > 70% was considered categorized (C), whereas < 70% criterion was uncategorized (U).
and Greek contrasts, the monolingual groups differed in their discrimination of English language, coronal). Most importantly, a significant main effect of stimulus of stimulus language (English or Greek) and place (bilabial or coronal). The Greek monolinguals' TC assimilations of the Greek contrasts resulted in excellent discrimination. The Greek monolinguals' TC assimilations of Greek /ba/–/pa/, /da/–/ta/ resulted in excellent discrimination on their two languages, there are three possible patterns that they might exhibit: (a) discrimination akin to Greek monolinguals, consistent with the reported persistence of L1 effects on bilingual speech perception (e.g., Pallier et al., 2001; Sebastián-Gallés & Soto-Faraco, 1999), (b) discrimination akin to English monolinguals, that is, a bias from L2-dominance, or (c) they may differ from both monolingual groups and show intermediate discrimination performance on the stops of both languages due to their unique language-learning histories.

4.2. Discrimination of word-initial voiced versus voiceless stop contrasts (CV–CV)

As predicted by PAM, the English monolinguals' TC assimilations of English /ba/–/pa/, /da/–/ta/ resulted in excellent discrimination. However, their UC assimilations of the Greek contrasts resulted in fairly poor discrimination. The Greek monolinguals' TC assimilations of Greek /ba/–/pa/, /da/–/ta/ resulted in excellent discrimination, whereas their CG and UC assimilations of the English contrasts resulted in only moderate discrimination. Mean percent correct discrimination is shown in Table 3 on the row below the assimilation pattern for each group.

We first analyzed the monolinguals’ discrimination performance via a $2 \times (2 \times 2)$ ANOVA with the between-subjects factor of native language (English or Greek) and within-subjects factors of stimulus language (English or Greek) and place (bilial or coronal). Most importantly, a significant main effect of stimulus language, $F(1, 38)=45.9, p<.001$, $n_p^2=.547$, and a significant Native Language $\times$ Stimulus Language interaction indicated that the monolingual groups differed in their discrimination of English and Greek contrasts, $F(1, 38)=91.2, p<.001$, $n_p^2=.706$. Simple effects tests (Winer, 1971), confirmed that English listeners discriminated the English contrasts more accurately than the Greek, $F(1, 38)=43.6, p<.001$, and Greek listeners discriminated Greek more accurately than English, $F(1, 38)=51.5, p<.001$, consistent with their TC assimilations of native contrasts.

The bilinguals’ discrimination performance did not support PAM’s predictions based on assimilation patterns. Both groups showed excellent discrimination of the English contrasts and good discrimination of the Greek, regardless of language mode. The Greek mode bilinguals did not show the excellent discrimination for Greek contrasts predicted from their TC assimilations. Rather, their performance was consistent with PAM’s predictions for bilinguals in English mode. From Table 3, it seems that the language mode did not affect the bilinguals’ discrimination. To analyze the bilinguals’ discrimination data we conducted another $2 \times (2 \times 2)$ ANOVA using language mode in place of the native language factor. A significant main effect of stimulus language indicated that they more accurately discriminated the English contrasts (94%) than the Greek (80%), $F(1, 38)=129.6, p<.001$, $n_p^2=.773$. Remarkably, we found no significant main effect or any interactions involving language mode. Placing the bilinguals in different unilingual language modes did not result in any subgroup differences in discrimination for any of the word-initial contrasts. This was very surprising given the strong and clear language mode effect observed for categorization. The lack of a language mode effect on discrimination suggests that the L2-dominant early bilinguals have undergone a perceptual reat-tunement, and have come to resemble monolinguals of the L2 in their discrimination performance, at least more so than monolinguals of the L1, despite many years of continued L1 use.

From the discrimination results for bilinguals, it appears that they are performing equivalently to monolingual English listeners. To test this we conducted an additional $2 \times 2 \times (2 \times 2)$ ANOVA to compare the monolinguals' and bilinguals' overall discrimination, with between-subjects factors of language context (English vs. Greek) and lingualism (monolingual vs. bilingual), and a within-subjects factor of stimulus language (English vs. Greek). As this analysis was post-hoc, a Bonferroni-adjusted $\alpha$ of .025 was employed. Importantly, there was a significant three-way Language Context $\times$ Lingualism $\times$ Stimulus Language interaction, $F(1, 76)=43.9, p<.001$, $n_p^2=.366$. Simple two-way Lingualism $\times$ Stimulus Language analyses of that interaction revealed that the two-way interaction was significant both for participants tested in Greek context (i.e., Greek monolinguals and Greek mode bilinguals), $F(1, 76)=41.3, p<.001$, and English context, $F(1, 76)=10.5, p=.002$. Thus, although the bilinguals’ discrimination was unaffected by language mode in discrimination, their performance was not identical to either group of monolinguals. It is clear from Fig. 1 that their performance was much closer to English than to Greek monolinguals. The bilingual lines in Fig. 1 are almost parallel, illustrating that language mode had no effect on their discrimination.

The bilinguals’ discrimination performance thus reflects their unique language-learning histories. Although their discrimination patterns were very similar to the English monolinguals, the bilinguals possessed an advantage for the Greek contrasts relative to the English monolinguals because of their many years of experience with Greek.

Having observed that the bilinguals approximate English (L2) but not Greek (L1) monolinguals in discrimination, we next examined whether this L2 bias was more pronounced in a context

Table 3: Monolinguals' and bilinguals' assimilation types and mean percent correct discrimination of English and Greek word-initial stop-voicing contrasts (Experiment 1). Standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Group</th>
<th>English Assimilation type</th>
<th>Greek Assimilation type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/ba/–/pa/</td>
<td>/da/–/ta/</td>
</tr>
<tr>
<td>English M discrimination</td>
<td>TC</td>
<td>94.9</td>
</tr>
<tr>
<td>SEM</td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Greek M discrimination</td>
<td>CG</td>
<td>74.1</td>
</tr>
<tr>
<td>SEM</td>
<td>(0.6)</td>
<td>(0.7)</td>
</tr>
<tr>
<td>EmMode M discrimination</td>
<td>TC</td>
<td>95.9</td>
</tr>
<tr>
<td>SEM</td>
<td>(0.3)</td>
<td>(0.4)</td>
</tr>
<tr>
<td>GrMode M discrimination</td>
<td>UC</td>
<td>93.1</td>
</tr>
<tr>
<td>SEM</td>
<td>(0.4)</td>
<td>(0.3)</td>
</tr>
</tbody>
</table>

5 Although not directly relevant here, we also found a significant main effect of place, $F(1, 38)=32.5, p<.001$, $n_p^2=.461$. Native Language $\times$ Place interaction,

7 There was also a significant main effect of stimulus language, $F(1, 76)=158.2, p<.001$, $n_p^2=.675$, and significant two-way interactions of Language Context $\times$ Stimulus Language, $F(1, 76)=61.0, p<.001$, $n_p^2=.445$ and Lingualism $\times$ Stimulus Language, $F(1, 76)=5.4, p=.022$, $n_p^2=.067$. 

Footnote continued
that is not only phonetically more complex, but also in which the phonological status of L1 (Greek) segments is unclear. The ambiguous status of Greek voiced stops is evident in the phonetically variable medial positions, where the nasal that underlyingly precedes the voiced stop may be overtly realized by some native Greek speakers. The unique language pairing of Greek and English provides us with an opportunity for a clear test of whether L2 phonetic categories have been established and perceptual reattunement has occurred. Therefore, in order to understand phonological organization in Greek monolinguals, it is essential to examine the perception of stops in the phonetically variable medial contexts. Once the monolinguals’ performance is known, we will have a clearer understanding of the challenge facing bilinguals and how L2 learning has shaped their perception.

5. Experiment 2: Word-medial stop-voicing contrasts in stressed syllables

We presented English and Greek voiced and voiceless stops in intervocalic (VCV) context to test whether Greek listeners are sensitive to voicing differences between stops in medial position. We presented word-medial stops as the onsets of stressed syllables in iambic disyllables (second syllable stress) because they have similar VOTs to initial-position stops, allowing straightforward comparison with the results of Experiment 1. ⁶ In English, voiceless stops in medial stressed syllable-initial position are long-lag aspirated, as in word-initial position (Cox & Palethorpe, 2007), but medial voiced stops may be produced with voicing lead VOT (Klatt, 1976). In contrast, Greek medial voiceless stops are produced with short-lag and voiced stops with lead VOT, as in word-initial position (Fourakis, 1986).

According to PAM, monolinguals should perceive native stops as excellent exemplars of the native categories. Based on differences in VOT, it was predicted that English monolinguals would assimilate medial, stress-initial Greek voiced and voiceless stops to English voiced categories as good and moderate (or unclassified) exemplars, respectively. Greek monolinguals were expected to categorize the English medial, stress-initial voiced and voiceless stops as good and poor exemplars of the Greek voiceless and voiced stops, respectively, yielding CG assimilations. However, from the results of Experiment 1 it was unclear whether they would reliably categorize English /a’d/ as a /d/.

Consistent with PAM-L2, it was hypothesized that the bilinguals would have developed common phonological categories and separate phonetic categories for their L1 and L2 stops, that is, bilinguals were expected to show different categorization and ratings in each language mode. Specifically, English mode bilinguals should meet our categorization criterion for the English stops, and assimilate Greek stops to English categories. Conversely, Greek mode bilinguals, because of their experience with English, were expected to have established phonological categories for English voiced stops, and therefore should meet our categorization criterion for both Greek voiced and voiceless stops even if Greek monolinguals fail to consistently do so and assimilate English stops to Greek categories.

6. Method

The participants were the same groups of monolingual and bilingual listeners tested in Experiment 1. The experimental testing procedure was identical to that employed in Experiment 1. The medial stops were presented randomly among the word-initial stops reported in Experiment 1.

We controlled the variability regarding the prenasalization of Greek medial stops by testing the listeners’ perception of stops in medial intervocalic context in iambic (second syllable stress) disyllables (VCV). The stimuli were recorded from the same native speakers of Greek and English as in Experiment 1, and matched as closely as possible for duration, fundamental frequency, and intensity. For categorization and goodness ratings, the stimuli were English and Greek biliteral /b, p/ and coronal stops /d, t/ in intervocalic position (VCV) /a’pa, a’ta, a’ba, a’d/.

For discrimination, we presented AXB trials (as described in Experiment 1) of within-language word-medial contrasts of voiced versus voiceless bilabial and coronal stops in stressed-syllable medial contexts (VCV–VCV, e.g., /a’ba/–/a’pa/, /a’d/–/a’ta/).

7. Results and discussion

7.1. Categorization and rating of intervocalic stops in stressed position (VCV)

As in Experiment 1, monolinguals consistently categorized stop consonants of their native language. Categorization and goodness rating data for intervocalic biliteral and coronal stops are presented in Tables 4 and 5, respectively.

When presented with nonnative stops, unlike in Experiment 1, and consistent with our predictions, English monolinguals met the categorization criterion for the intervocalic Greek stops. In contrast, Greek monolinguals failed to reliably categorize English /a’d/ (consistent with our observations for English /d/) and categorized both English /a’ba/ and /a’pa/ as Greek /p/, but with a significant difference in category-goodness ratings, t(19)=2.1, p=.05.

We now turn to the bilinguals. As in Experiment 1, the bilinguals reliably categorized all mode-congruent stops. Bilinguals assigned higher goodness-of-fit ratings when presented with mode-congruent stimuli. In English mode, they assigned higher ratings than those in Greek mode for English /a’ba/, t(35)=4.8, p < .001, /a’d/, t(38)=6.5, p < .001, /a’pa/, t(38)=7.8, p < .001, and /a’ta/, t(38)=6.4, p < .001. Similarly, the bilinguals in Greek mode assigned higher ratings for Greek than did bilinguals in English mode, but only for the Greek voiceless stops /a’pa/, t(38)=4.3, p < .001, and /a’ta/, t(37)=2.0, p=.05.

To assess the effect of language mode on categorization, we compared the bilinguals’ mode-congruent goodness ratings with the monolinguals’ ratings of native stops. A 2 x 2 x (2 x 2) ANOVA with between-subjects factors of stimulus language and bilingualism, and within-subjects factors of voicing and place revealed
The bilinguals’ assimilation types suggest that the two bilingual groups had not affected the bilinguals’ discrimination in Experiment 1. Whether the monolinguals’ discrimination performance was as contrasts (see top row for each group in Table 6), to again examine the assimilation patterns from the categorization data for these 1) that Greek mono- and bilingual listeners are sensitive to VOT and a Stimulus Language of their respective native stops. Gr /a/ and Greek (Gr) modes (Experiment 2). Values in italics indicate speakers’ intended category.

<table>
<thead>
<tr>
<th>Group</th>
<th>Consonant stimuli</th>
<th>Category label</th>
<th>Criterion met</th>
<th>Categorized as</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual English</td>
<td>Gr /a/</td>
<td>91.0 (3.5)</td>
<td>7.7 (3.2)</td>
<td>1.3 (2.0)</td>
</tr>
<tr>
<td>Gr /a/</td>
<td>10.0 (3.0)</td>
<td>73.8 (3.4)</td>
<td>1.3 (2.0)</td>
<td>15.0 (2.3)</td>
</tr>
<tr>
<td>En /a/</td>
<td>85.0 (2.7)</td>
<td>1.3 (3.0)</td>
<td>11.3 (2.6)</td>
<td>2.5 (3.0)</td>
</tr>
<tr>
<td>En /a/</td>
<td>91.1 (4.5)</td>
<td>8.9 (3.3)</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Monolingual Greek</td>
<td>Gr /a/</td>
<td>90.0 (5.6)</td>
<td>5.0 (4.3)</td>
<td>5.0 (5.0)</td>
</tr>
<tr>
<td>Gr /a/</td>
<td>100.0 (5.6)</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>En /a/</td>
<td>88.8 (2.7)</td>
<td>71.3 (2.8)</td>
<td>1.3 (4.0)</td>
<td>8.8 (2.5)</td>
</tr>
<tr>
<td>En /a/</td>
<td>96.3 (3.3)</td>
<td>8.9 (4.4)</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Bilingual EnMode</td>
<td>Gr /a/</td>
<td>85.0 (4.5)</td>
<td>13.8 (3.5)</td>
<td>1.3 (4.0)</td>
</tr>
<tr>
<td>Gr /a/</td>
<td>6.3 (4.8)</td>
<td>82.5 (4.3)</td>
<td>1.3 (6.0)</td>
<td>8.8 (3.9)</td>
</tr>
<tr>
<td>En /a/</td>
<td>77.2 (4.2)</td>
<td>10.1 (3.1)</td>
<td>8.9 (3.9)</td>
<td>3.8 (4.5)</td>
</tr>
<tr>
<td>En /a/</td>
<td>1.3 (4.0)</td>
<td>89.9 (5.2)</td>
<td>8.9 (4.4)</td>
<td>C</td>
</tr>
<tr>
<td>Bilingual GrMode</td>
<td>Gr /a/</td>
<td>88.6 (5.1)</td>
<td>7.6 (3.8)</td>
<td>3.8 (4.3)</td>
</tr>
<tr>
<td>Gr /a/</td>
<td>1.3 (3.0)</td>
<td>93.7 (5.8)</td>
<td>3.8 (4.5)</td>
<td>C</td>
</tr>
<tr>
<td>En /a/</td>
<td>68.8 (2.4)</td>
<td>17.5 (2.8)</td>
<td>8.8 (3.3)</td>
<td>5.0 (1.8)</td>
</tr>
<tr>
<td>En /a/</td>
<td>3.8 (3.0)</td>
<td>93.8 (2.4)</td>
<td>1.3 (4.0)</td>
<td>C</td>
</tr>
</tbody>
</table>

**Table 4**
Mean percent categorization (and goodness ratings 1 – 7) of English and Greek biliteral stops in intervocalic stress-initial position by monolinguals and bilinguals in English (En) and Greek (Gr) modes (Experiment 2). Values in italics indicate speakers’ intended category.

**Table 5**
Mean percent categorization (and goodness ratings 1 – 7) of English and Greek coronal stops in intervocalic stress-initial position by monolinguals and bilinguals in English (En) and Greek (Gr) modes (Experiment 2). Values in italics indicate speakers’ intended category.

Due to the phonetic, phonological, and orthographic differences between Greek and English, we were interested in testing listeners’ discrimination of medial contrasts in Experiment 2 (e.g., /a/β/–/a/β/), to confirm our finding with word-initial stops (Experiment 1) that Greek mono- and bilingual listeners are sensitive to VOT versus voiceless stops (VCV–VCV). The analysis revealed a significant main effect of stimulus language, $F(1, 76)=15.5, p<.001, \eta_p^2=.170$, and voicing, $F(1, 76)=14.1, p<.001, \eta_p^2=.157$, and a Stimulus Language × Voicing interaction, $F(1, 76)=7.1, p=.010, \eta_p^2=.085$. There were no significant effects involving linguistic, suggesting that the bilinguals labeled the mode-congruent stops in a manner comparable to the monolinguals’ labeling of their respective native stops.

Mean percent correct discrimination for VCV–VCV contrasts is shown in Table 6. As predicted by TC assimilations, English monolinguals showed excellent discrimination of all intervocalic contrasts. Greek monolinguals showed excellent discrimination of the intervocalic Greek contrasts, but only good discrimination of English contrasts, consistent with their CG and UC assimilations.

Due to the phonetic, phonological, and orthographic differences between Greek and English, we were interested in testing listeners’ discrimination of medial contrasts in Experiment 2 (e.g., /a/β/–/a/β/), to confirm our finding with word-initial stops (Experiment 1) that Greek mono- and bilingual listeners are sensitive to VOT versus voiceless stops (VCV–VCV). The analysis revealed a significant main effect of stimulus language, suggesting that overall, the Greek contrasts
Importantly, we found a significant main effect of native language, $F(1, 38) = 18.5, p < .001, \eta^2_p = .327$, and a Native Language $\times$ Stimulus Language interaction, which indicates that the English monolinguals more accurately discriminated the English contrasts than the Greek, and the Greek monolinguals discriminated the Greek contrasts better than the English, $F(1, 38) = 55.5, p < .001, \eta^2_p = .393$. Simple effects tests confirmed that stimulus language affected both the English, $F(1, 38) = 6.4, p = .016$, and the Greek listeners’ discrimination, $F(1, 38) = 64.2, p < .001$. Given that the difference between native versus nonnative discrimination was larger for the Greek listeners, this sets an interesting scenario for our bilingual participants. If the bilinguals have developed L2 phonetic categories, we would expect them to outperform the Greek monolinguals on the English contrasts.

We compared the bilinguals’ discrimination in a separate $2 \times (2 \times 2)$ ANOVA and the analysis yielded a significant Stimulus Language $\times$ Place interaction, $F(1, 38) = 7.7, p = .008, \eta^2_p = .169$. Simple effects tests revealed that the bilinguals more accurately discriminated the English (94%) than the Greek coronal contrasts (91%), $F(1, 38) = 4.3, p = .045$, whereas there was no difference for the biliteral contrasts. Importantly, there were no significant effects or interactions involving language mode. The bilinguals once again discriminated all intervocalic contrasts very similarly to the English monolinguals, reflecting their L2 dominance.

To compare the monolinguals’ and bilinguals’ overall discrimination, we conducted an additional $2 \times 2 \times (2)$ ANOVA with between-subjects factors of language context and lingualism, and a within-subjects factor of stimulus language, with a Bonferroni adjusted $\alpha$ level of .025. The analysis yielded a significant three-way Language Context $\times$ Lingualism $\times$ Stimulus Language interaction, $F(1, 76) = 16.9, p < .001, \eta^2_p = .182$. Simple two-way Lingualism $\times$ Stimulus Language interactions confirmed that the two-way interaction was significant for participants tested in a Greek context, $F(1, 76) = 28.0, p < .001$, but not for those tested in English context. As shown in Fig. 2, the English mode bilinguals performed nearly identically to the English monolinguals.

The results of Experiment 2 confirm that language mode affects bilinguals’ categorization and goodness ratings, but not their discrimination performance for stress-initial medial stops just as for word-initial stops. In categorization, the bilinguals showed language-specific sensitivity by assigning ratings relative to the categories of the language activated by the language mode. In discrimination, however, the bilinguals in both language modes were indistinguishable from English monolinguals. This L2 bias, originally observed in Experiment 1, was more pronounced in the medial stops for which the phonological status of the Greek voiced stops remains unclear. These findings provide strong evidence that the bilinguals’ language dominance predicts their discrimination success for contrasts not just in the dominant language (L2), but in the nondominant language as well, even when the nondominant language is the L1.

8. General discussion

These experiments demonstrate that Greek–English L2-dominant early bilinguals are very accurate, relative to monolinguals in either language, in identifying L1 and L2 phones in two phonetic environments (Experiment 1: word-initial, Experiment 2: word-medial stress-initial). We found no evidence that the bilinguals confuse the phonological status in each language of the phonetically similar short-lag stops of Greek (which they identified as phonologically similar, English (identifying appropriately as phonologically voiceless, appropriate to Greek) and English (identified appropriately as phonologically voiced in English). The bilinguals instead were obviously capable of attending to higher level, abstract, phonological information. This is evident as well in the complementary finding that they identified phonetically dissimilar (in terms of VOT) English and Greek phones as phonological cognates (e.g., English short-lag /b/ and Greek prevoiced /b/ were both identified as /b/). English long-lag /t/ and Greek short-lag /t/ were both identified as /t/). These observations are compatible with the PAM-L2 hypothesis that they developed
separate L2 phonetic categories for English stops and that these are linked to the existing L1 phonetic categories via the more abstract level of the phonological units [b, d, p, t].

The bilinguals nonetheless remained sensitive to the different L1 and L2 phonetic realizations of the same phoneme (e.g., English vs. Greek realizations of /p/). The data also show that by placing bilinguals in different unilingual language modes, they judge the category goodness of phones relative to the phonetic categories of that language. For example, the voiceless unaspirated stops /p/ and /t/ were categorized as excellent exemplars of /p/ and /t/ by the bilinguals in Greek mode, but only as a moderately good /p/ and /t/ by the bilinguals in English mode, reflecting the language-specific phonetic VOT settings of Greek and English.

In spite of these clear language mode effects in categorization, the bilinguals were unaffected by language mode in discrimination. This effect was robust across word-initial and medial stress-initial positional contexts. For the simple word-initial contrasts, the bilinguals showed very good discrimination of the Greek contrasts and near-native discrimination of English contrasts. In the phonetically more complex medial contexts, this L2 bias was even more pronounced: the bilinguals mirrored the performance of the monolingual listeners of English (their L2). This occurred even though in a previous production study (Antoniou et al., 2010) we demonstrated that bilinguals from the same population produced initial stops with VOTs that did not statistically differ from those of monolinguals, in both languages. In that study, only very modest cross-language interference was observed. That modest interference was restricted to only some of the English word-medial voiced stops, which bilinguals produced with longer (more Greek-like) voicing lead than monolinguals. We clearly have not found any such L1 interference on the L2 in bilinguals' perception of stop voicing in their two languages.

Although previous research suggests a link between perception and production in monolinguals and L2 learners (e.g., Fleger, 1999; Strange, 1995), we have observed somewhat different patterns of L1–L2 interaction in bilinguals in these two domains that depends on the task involved: in production, a modest L1-influence on L2 (Antoniou et al., 2010); in discrimination, the L2 influenced the L1; in categorization/rating, a bidirectional language mode bias was observed for perception of both L1 and L2 stimuli. We speculate that the reasons for this difference in speech perception and production lie in the bilinguals' language learning histories. Our bilingual participants were exposed to Greek from birth, and then became immersed in English when they attended school. They quickly acquired English as their L2 and became dominant in it, as English remains the primary language immersion environment for people living throughout most of Australia. We reason that when faced with the challenge of keeping up with their native-speaking peers, and accommodating a rapidly expanding L2 vocabulary, early bilinguals in such contexts attuned to the dominant L2 (English) so much so that they came to resemble native listeners, with the important exception that they could still also speak and comprehend their (now nondominant) L1. Our participants continued to maintain their L1 (Greek) in some contexts with family members, other speakers of their L1, and other bilinguals. In production, conversely, the bilinguals faced no communicative pressure to reduce voicing lead in their English voiced stops, especially in intervocalic positions in which lead VOT is permissible in English, which accounts for the modest L1-interference on L2 productions of medial stops as observed by Antoniou et al. (2010). We hypothesize that this difference in communicative pressure contributed to the different asymmetrical patterns of L1–L2 interaction for speech perception and production. Necessarily, further research will be needed to determine whether our intuitions are correct regarding the source of the intriguing L1 versus L2 performance-bias differences among the perception and production tasks.

Bilinguals' performance on speech perception tests is thought to indicate whether they operate like two independent monolingual listeners or as uniquely configured language users whose L1 and L2 are accommodated in a common phonetic space. Our findings offer partial support for both views, depending on whether the focus is on categorization or discrimination. The results suggest that a bilingual is a single (dominant-language) listener with respect to discrimination, but behaves more like a monolingual of the activated language with respect to categorization judgments.

When the language mode requires that bilinguals categorize and rate stimuli relative to their Greek categories, they do so. In fact, the findings suggest that perceptual flexibility, which may be seen as responsiveness to one's linguistic context, is a crucial element in bilingual language use. Findings of (remarkable) perceptual flexibility, in terms of ability to shift phonetic/phonological boundaries to accommodate to a given speaker after even short periods of exposure, have also been reported for monolinguals (Norris, McQueen, & Cutler, 2003) and may indeed be indispensable for second language learning beyond early childhood. This flexibility may also be central to language change over the lifespan, and also to the patterns of speech perception in bilinguals. It is important to note that neither PAM nor the Speech Learning Model (nor other models of cross-language speech perception or production, for that matter) have addressed how speech perception (and production) may be affected by such developmental adjustments, nor the factors that facilitate/drive those flexible adjustments.

In contrast, when asked to discriminate contrasts, the bilinguals relied on their cumulative L1–L2 sensitivity, which reflected their L2 dominance in English. Therefore, a bilingual's detection of phonetic differences is heavily influenced by the dominant language at the time of testing, even if that language is the L2, and even if the listener is currently operating in L1 language mode. We suggest that when faced with a taxing task, such as discrimination, the dominant language is the default setting. Indeed, the bilinguals who had been maintained in Greek mode reported they felt very tired at the conclusion of the perceptual experiment, whereas those in English mode did not make any such comments. The observation that the bilinguals reverted to their dominant language for discrimination of not only L2, but also L1 contrasts, is a testament to how well they have acquired and mastered their L2. This L2 bias in discrimination (perceptual sensitivity to phonetic distinctions) must be a direct result of the bilinguals' unique language learning and usage histories, that is, bilinguals face a pervasive pressure to perceive the speech of a wide variety of speakers in an immersive L2 context.9

Our findings complement those of Pallier, Bosch, and Sebastián-Gallés (1997) who reported that early bilinguals poorly discriminated L2-only contrasts that fit an SC pattern with respect to their L1, which suggest a pervasive L1 influence on discrimination. The results of the present series of experiments are consistent instead with the possibility that it is language dominance, rather than the order in which their languages were acquired, that influences bilinguals' discrimination performance. The Spanish–Catalan bilinguals tested by Pallier et al. (1997) were not L2-dominant, unlike the bilinguals tested here. Unfortunately, we

9 We recognize that this point could be made even more strongly if it were possible to compare the bilinguals of the present experiment against complementary groups of L1-English/L2-dominant-Greek bilinguals. However, the realities of bilingual contexts globally make this logical comparison virtually impossible to carry out. We know of no sizeable L1-English communities embedded within Greek primary-language environments.
cannot make a direct interpretive comparison to that study because we did not find any SC assimilations and they did not examine any UC, TC or CG assimilations. Therefore, future studies should examine whether Pallier et al.'s (1997) findings for SC type L2-to-L1 contrasts extend to L2-dominant bilinguals, and/or whether our findings for the other assimilation types extend to L1-dominant bilinguals, and/or to bilinguals living in more balanced bilingual environments.

The discrimination predictions based on PAM were largely supported by the monolinguals’ discrimination performance in that the participants’ overall discrimination performance was predicted by their assimilation patterns in their categorization responses. PAM may need to be extended to take account for phones that are un categorized but overlap in phonetic space with native language phonological categories (e.g., monolingual English listeners identified Greek /pa/ as English /ba/ 53.8% and /pa/ 46.3% of the time). A more detailed discussion of this may be found in Tyler, Best, Faber, and Levitt (submitted for publication). With regard to the bilinguals, however, discrimination of the word-initial contrasts appears compatible with PAM-L2. Bilinguals in Greek mode performed better on the English contrasts than the Greek monolinguals, but did not reach the ceiling levels of the Greek listeners on the Greek contrasts. Those in English mode performed similarly to the English monolinguals, although they discriminated the Greek word-initial contrasts more accurately. Despite differing from both monolingual groups, the bilinguals discriminated in a manner more similar to the English than Greek monolinguals. Overall, the PAM/PAM-L2 predictions were mostly accurate for the bilinguals in English mode, whereas the Greek mode bilinguals’ discrimination was not clearly linked to their categorization, but was better predicted by English mode assimilation patterns. Neither bilingual group exhibited an L1 influence on the L2, as would be predicted by the Speech Learning Model (Flege, 1995; Flege et al., 2003) or Native Language Magnet Model (Kuhl, 1992, 1993).

Existing models on L1–L2 speech perception, including PAM-L2, cannot explain why language mode effects emerged for categorization but not discrimination, nor can they account for why the bilinguals’ dominance in their L2 predicted their discrimination success. It appears that the experimental task determines what kind of judgment is required, and thus whether language mode has an effect. Language mode appears to influence performance on tasks that specify which language is to be used as the basis for judgments, as was the case for categorization where distinct English and Greek orthographies were used to present response choices. On the other hand, if response choices are not overtly language specific, as was the case for discrimination, bilinguals may have reverted to their dominant language (here, their L2). The observed language mode effects thus pose a challenge for PAM-L2 and the Speech Learning Model (as well as for the Native Language Magnet model). In order to explain our results, these theories would require extension. Thus, we recommend that some notion of language dominance needs to be incorporated into models of bilingual speech production (Antoniou et al., 2010) and perception.

Although we interpret the present findings as evidence that bilinguals have developed separate phonetic categories but maintain common L1/L2 phonological categories, we cannot rule out the alternative explanation that separate L1 and L2 phonological categories have also been established. It would be possible to examine this in future by testing bilinguals on a task to match L1-accented speech to L2 categories or vice versa. If bilinguals possess L1 and L2 phonetic categories that are linked at the phonological level, we would expect that they would possess an advantage over native monolingual listeners of the L2 when discriminating L1-accented L2 speech.

Future research should also examine bilinguals’ perception of nonnative contrasts from unfamiliar languages to determine whether the influence of the L2 on the L1 reported here occurs even when bilinguals are discriminating unfamiliar contrasts from a foreign language. For example, do bilinguals exhibit language-mode-specific categorization and rating performance, but L2-dominant discrimination performance, when presented with nonnative stop contrasts that involve nonnative laryngeal (glottic) distinctions, such as the plosive–implosive stop distinction /b/-/ɓ/ as found in certain African languages?

In conclusion, bilingual listeners were able to attend to language-specific VOT differences in their categorization and goodness ratings, yet they discriminated in a manner comparable to the English monolinguals. These findings provide evidence that bilinguals integrate both language systems in a common phonetic space that is influenced by their dominant language environment. Yet the bilinguals are sensitive to the situational language context and selectively attend to language-specific phonological information when phonological judgments are called for, that is, when making categorization and goodness-rating judgments.

Acknowledgments

This research was supported by a UWS Postgraduate Research Award, Marc Institute Fieldwork Travel grant, NIH grant DC00403 (PI: C. T. Best), and Australian Research Council grant DP0880913 (CI: M. Tyler).

References


