Focusing the lens of language experience: Perception of Ma'di stops by Greek and English bilinguals and monolinguals

Mark Antoniou[a]

The Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, Illinois 60208

Catherine T. Best

MARCS Institute, University of Western Sydney, Locked Bag 1797, Penrith, New South Wales 2751, Australia

Michael D. Tyler

School of Social Sciences and Psychology, University of Western Sydney, Locked Bag 1797, Penrith, New South Wales 2751, Australia

Monolingual listeners are constrained by native language experience when categorizing and discriminating unfamiliar non-native contrasts. Are early bilinguals constrained in the same way by their two languages, or do they possess an advantage? Greek–English bilinguals in either Greek or English language mode were compared to monolinguals on categorization and discrimination of Ma'di stop-voicing distinctions that are non-native to both languages. As predicted, English monolinguals categorized Ma'di prevoiced plosive and implosive stops and the coronal voiceless stop as English voiced stops. The Greek monolinguals categorized the Ma'di short-lag voiceless stops as Greek voiceless stops, and the prevoiced implosive stops and the coronal prevoiced stop as Greek voiced stops. Ma'di prenasalized stops were uncategorized. Greek monolinguals discriminated the non-native voiced-voiceless contrasts very well, whereas the English monolinguals did poorly. Bilinguals were given all oral and written instructions either in English or in Greek (language mode manipulation). Each language mode subgroup categorized Ma'di stop-voicing comparably to the corresponding monolingual group. However, the bilinguals' discrimination was unaffected by language mode: both subgroups performed intermediate to the monolinguals for the prevoiced-voiceless contrast. Thus, bilinguals do not possess an advantage for unfamiliar non-native contrasts, but are nonetheless uniquely configured language users, differing from either monolingual group.
© 2013 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4792358]

I. INTRODUCTION

Monolingual listeners have difficulty discriminating many non-native distinctions that are not contrastive in their native language (L1). For this reason, the L1 has historically been likened to a "filter" or "sieve" because it appears to interfere with perception of segmental phonetic differences in a non-native language that are not meaningful in the L1 (Polivanov, 1931; Trubetzkoy, 1939). More recent accounts of non-native perception (e.g., Best, 1995; Flege, 1995) characterize the process of perceptual assimilation as an ongoing and active process in which the difficulty encountered is determined by perceived phonetic similarity to L1 categories and may change over time as a result of the individual's immediate and/or longer-term language use. It is this more active view that drives the present investigation. The metaphor of a "lens" is more suitable for such active accounts as it indicates an act of directing or focusing attention to particular pieces of information and letting other irrelevant speech information fall outside of the focus. Monolingual listeners who are perceptually biased by their L1 show "accented"

non-native perception (Jenkins *et al.*, 1995). A well-known example of this interference from L1 perceptual assimilation is the difficulty that Japanese listeners experience when categorizing and discriminating English /r/ and /l/ (Miyawaki *et al.*, 1975). But whereas monolingual listeners assimilate phones to the categories of their native language, fluent early bilingual listeners have developed categories relevant to not just one, but two languages. What effect might this have on bilinguals' perceptual assimilation of unfamiliar non-native speech contrasts? As we will argue, the answer is likely to provide new insights into both bilingual speech perception and into the effects of language experience on speech perception more generally. Yet the issue of non-native speech perception in bilinguals has thus far received little to no direct attention in speech perception research.

A number of prior findings do, however, address how bilinguals' language experience affects perception of their two languages. We have recently shown that early sequential bilinguals' dominance in their second language (L2) enhances their ability to perceive L2 minimal pair contrasts that differ from the corresponding contrasts in their L1 (Antoniou *et al.*, 2012). However, it is unclear whether early bilingualism affects perception of *unfamiliar* speech contrasts that are non-native to both of a bilingual's languages. If the native

[a]Author to whom correspondence should be addressed. Electronic mail: mark.antoniou@northwestern.edu

language serves as a lens that focuses monolinguals' perception, does this mean that bilinguals have two lenses, one for each language? Or do they instead use a single lens that is brought into focus by their first language, or by their dominant language, or instead by some combined influence from both their languages? Or, if we take seriously an active and dynamic language-specific process, might the lens that is used depend on the concurrent activation of one or the other language of the bilingual? And how would each of these possibilities impact on their perception of entirely non-native contrasts that do not occur in either of their languages?

For monolingual listeners, the L1 "filter" refers to the difficulties they experience when discriminating phonetic distinctions that do not occur as a native phonological contrast. These difficulties may be predicted from the range of ways in which they perceptually assimilate the contrasting non-native phones to native segmental categories. The perceptual assimilation model (PAM) (Best, 1995) makes explicit predictions about assimilation and discrimination differences for non-native contrasts by taking into account both contrastive phonological and non-contrastive phonetic properties of native speech segments. We take phonetic categories to refer to functionally equivalent sets of individual phones, e.g., tokens of a position-dependent allophone of /p/, which differ from each other in a gradient fashion within the category. By comparison, phonological categories specify which segments are used contrastively to support lexical distinctions within the language. A phonological category may be comprised of one or more phonetic categories (for example, the phonological category of voiceless stop /p/ in English is comprised of the position-dependent allophonic categories of aspirated [pʰ], unaspirated voiceless [p], and unreleased [pˀ]).

According to PAM, a non-native phone will be perceptually assimilated to a monolingual's phonological system in one of three ways: (1) categorized as belonging to a native phoneme category, (2) falling in between multiple native phonemes, and therefore, as an uncategorized speech segment, or (3) perceived as a nonspeech sound if it deviates substantially from all native phonemes and is therefore non-assimilable to the L1 phonological system. Discrimination performance will depend on the assimilation pattern of the two phones in the non-native contrast in question. For example, the contrasting non-native phones may be perceived as similar to two separate native phonemes, termed two category assimilation (TC). Alternatively, each may assimilate equally well or poorly to a single native phoneme, called single category assimilation (SC). If one is perceived as a better exemplar than the other for that single native category, a category goodness difference (CG) in assimilation results instead. If one of the non-native phones is uncategorized, as defined above, the contrast will form an uncategorized-categorized pair (UC). If both phonemes are uncategorized, they will form an uncategorized-uncategorized pair (UU). PAM predicts that the native language phonology enhances discrimination when the two phones comprising the non-native contrast are separated by a native phonological boundary, meaning that TC contrasts should be discriminated best, as may many though not all UC assimilations

(see discussion below). Conversely, discrimination will be hindered when both phones assimilate to the same native phoneme, particularly if they are perceived as equivalent in goodness-of-fit (SC assimilation). As a result, PAM predicts the following gradient for discrimination performance: TC > CG > SC, where TC assimilation will result in the highest levels of discrimination, and SC assimilation the poorest. Importantly for the present research, recent developments of PAM (see Bohn et al., 2011) provide preliminary evidence that uncategorized assimilations lie along a continuum of partial assimilations. A partial assimilation refers to when categorization responses for an uncategorized phone are shared across a number of native categories. The continuum described above ranges from partial assimilation of two nonnative phones to the same set of native categories, or to some overlap between the two sets of categories, or to entirely different categories for each of the uncategorized nonnative phones, which in turn leads to a range of possible discrimination patterns for assimilations involving uncategorized non-native phones. Discrimination may range from poor for completely overlapping categories (functionally akin to an SC assimilation type) to excellent for completely non-overlapping categories (functionally akin to TC assimilation).

PAM has been influential in accounting for non-native perception of naive, monolingual listeners, but was not designed in its original form to explain bilinguals' perception of non-native contrasts, that is, whether the L1 and L2 interact to enhance or inhibit discrimination of non-native contrasts. Thus, PAM-L2 (Best and Tyler, 2007) was created as an extension of PAM to L2 perceptual learning, specifically to predict the success with which L2 contrasts should be learned, and which new L2 phonetic and phonological categories are most likely to become established. When a learner begins acquiring an L2, L2 phones are first assimilated into already existing L1 categories or dissimilated from existing categories and established as new categories firstly on a phonetic level. As the learner's L2 vocabulary increases they attend increasingly 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. According to PAM-L2, a common L1/L2 phonological category can include distinct phonetic categories for each language, and those language-specific phonetic categories may evolve without necessarily influencing one another. Studies to date have tested PAM-L2 predictions by investigating bilinguals' perception of L1 and L2 contrasts (e.g., Antoniou et al., 2012), but have not yet investigated bilinguals' perception of unfamiliar non-native contrasts. If monolinguals are constrained in their non-native perception by native phonological categories, then it follows that bilinguals will be constrained by the categories that have formed as a result of their L1, and also L2, experience.

It is important to consider that they may also be constrained by other factors affecting their L1 and L2 experience and use. The contribution of age of L2 learning is known to be modulated by the relative quantity of input from native L2 speakers (Flege and Liu, 2001; Jia et al., 2006), length of residence in an L2-speaking environment,

and relative L1:L2 usage (Flege and MacKay, 2004). Based on these findings, it follows that bilinguals who have used their L2 so extensively from an early age 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. L2-dominant bilinguals are common, particularly among migrant populations in the US, UK, and Australia. However, little research has focused on speech perception in this type of bilingual. For these reasons, investigating L2-dominant bilingualism could be highly informative to our understanding of L2 speech category formation. Therefore, in this research project, we investigated the pattern of L1-L2 interaction in L2-dominant bilinguals' perception of nonnative contrasts.

Given current theoretical considerations, a bilingual's L1 and L2 may interact and affect perception of non-native contrasts in one of four ways. One possible pattern that we will refer to as a merged L1/L2 lens is that bilinguals, like monolinguals, are constrained by their L1, but unlike monolinguals, are also affected by learning their L2, and by the resultant interaction that may occur between the two languages (for a discussion see Best and Tyler, 2007). If so, early L2 acquisition may alter boundaries of L1 categories such that they are not the same as those of either group of monolingual listeners. Thus, L1-L2 effects on perception are likely to be bidirectional, although not necessarily symmetrical. The position taken in the speech learning model (SLM) (Flege, 1995) is largely consistent with this account. Specifically, SLM posits that L2-learning results in the systematic shifting of L1 and L2 phonetic categories as (a) L2 phones that are *equivalence classified* into existing L1 categories, resulting in L1/L2 merged categories, or (b) new L2 categories are formed that *dissimilate*, or deflect away from, the nearest L1 categories so that an L1-L2 phonetic contrast is maintained between them. Category dissimilation is more likely to occur in early L2 learners whose L1 categories are not yet fully formed, than in late learners (Flege *et al.*, 2003). The processes of equivalence classification or category dissimilation change both the L1 and L2 categories so that they no longer match either monolingual group. From the SLM view then, bilinguals do not have two lenses shaped by experience, but rather have one integrated L1/L2 lens that should result in discrimination that differs from both monolingual listeners who speak the bilinguals' L1 and those who speak the bilinguals' L2.

A second theoretical possibility, which we will refer to as a persistent L1 lens, is that even though early bilinguals learn and use an L2 fluently or even dominantly, the L1 may still exert persistent effects on perception. A number of studies of early Spanish–Catalan bilinguals have presented evidence for just this type of maintained L1 bias in perception (Pallier *et al.*, 2001; Sebastián-Gallés and Soto Faraco, 1999). For example, one such study revealed poor discrimination of L2 contrasts that fit an SC pattern with respect to monolinguals of their L1, suggesting long-term effects of L1 phonological organization on perception of L2 contrasts, even in bilingual listeners who have been fluent in the L2 from a young age (Pallier *et al.*, 1997). Similar results have been found across a diverse array of experimental tasks and contrasts (Caramazza *et al.*, 1973; Flege and Eefting, 1987a; Flege and Liu, 2001). These results suggest that even early bilinguals have one language-tuned lens and it is that of the L1. That is, even if these early bilinguals use each of their languages in everyday life, in perception they remain like L1 monolinguals. If this account is correct, then discrimination of non-native contrasts should show the same pattern in bilinguals as in monolingual listeners of their L1.

But there are still two other possible outcomes for non-native speech perception by early sequential bilinguals. The third possibility is that bilinguals' perception will show the effects of their language dominance rather than simply an L1 bias, that is, bilinguals may show an L2 bias when they are dominant in the L2. We refer to this possibility as a dominant language lens. Flege *et al.* (2002) were the first to suggest that L2-dominant bilinguals, because of their fluency in the L2, may be the most likely to suppress interference from the L1 on the L2. From this viewpoint, language dominance is hypothesized to determine which language (L1 or L2) will influence speech perception. It is possible to reinterpret past findings as indicating that it is not the "L1 filter" *per se* that exerts an influence on L2 perception, but rather that the bias is due is the bilinguals' dominance in the L1. Indeed, in the vast majority of past work, L2-dominant bilinguals have been ignored, and consequently, L1-dominance and order of acquisition have been confounded. Interestingly though, some studies have found larger perceptual switching effects on perception of L1 and L2 contrasts in bilinguals who are more proficient in the L2 than in those who are less proficient L2 users (Elman *et al.*, 1977; Hazan and Boulakia, 1993). García-Sierra *et al.* (2009) reported that Spanish–English bilinguals had an English-like (i.e., L2-like) phonemic boundary for /ga-/-/ka/ categorization, seemingly free of the persistent L1 effect reported in much past research. Importantly, in their study, 12 of the 15 bilinguals were more proficient in English than in Spanish, and used English more than Spanish on a daily basis, suggesting that they were L2-dominant. Our own past work presents some support for the language dominance account. We demonstrated that on a discrimination task, Greek–English bilinguals were indistinguishable from English monolinguals, reflecting their language dominance (L2-English) at the time of testing (Antoniou *et al.*, 2012). If this dominant language account is correct, then we would expect that after many years of predominant L2 use, L2-dominant bilinguals are no longer unilaterally biased by their L1 but rather by their L2, and this should be the case even when perceiving non-native speech contrasts.

A fourth possibility is that, as a result of their L1 and L2 learning, bilinguals have developed separate lenses for the L1 and L2 and use one or the other according to their current language mode. Consequently, speech perception performance will be more L1- or L2-like depending on which language is activated at the time of testing. In order to test this hypothesis, it is necessary to activate and test each language separately in experimental tasks. That is, the testing situation must be the same for bilinguals in each language mode as it is for the corresponding monolingual participants of their L1 and L2. The language mode framework (Grosjean, 1998,

J. Acoust. Soc. Am., Vol. 133, No. 4, April 2013

Antoniou *et al.*: Focusing the lens of language experience   2399

2001) provides a compatible model, as it assumes that in communicative contexts bilinguals function somewhere along a language-activation continuum ranging from monolingual to bilingual. They can effectively function as monolinguals in one of their languages, when only that one language is activated (though Grosjean also argues that the other language is never completely deactivated). Monolingual mode is achieved experimentally by having all contact, instructions, carrier sentences, and feedback occur in only one language throughout a study. Language mode is posited to affect all levels of language processing (Grosjean, 1998, 2001), although there have been only a handful of studies of its impact on perception of the bilinguals' L1 and/or L2 (Antoniou et al., 2012; Bohn and Flege, 1993; Caramazza et al., 1973; Elman et al., 1977; Flege and Eefting, 1987a,b; García-Sierra et al., 2009; Hazan and Boulakia, 1993; Williams, 1977). Recently, García-Sierra et al. (2012) observed a voicing boundary shift in Spanish–English bilinguals in an electrophysiological study when bilinguals were in a Spanish versus an English monolingual mode, maintained by having subjects read a magazine in the language consistent with the designated language mode during the experimental task (although see Winkler et al., 2003). Some of the sparse evidence of language mode effects on speech perception suggests that language mode affects discrimination of not only L1 and L2, but also of unfamiliar non-native contrasts. The clearest example comes from Calderón and Best (1996), who tested Spanish–English bilinguals under Spanish (L1) versus English (L2) language mode conditions, as compared to monolingual English listeners, on discrimination of three unfamiliar Xhosa bilabial stop-voicing contrasts. Spanish has a prevoiced versus short-lag voice onset time (VOT) distinction for stop-voicing contrasts, whereas English has a short- versus long-lag VOT distinction. Listeners were presented with three Xhosa contrasts: prevoiced implosive versus prenasalized /ɓ/-/ᵐb/, prevoiced implosive versus short-lag plosive /ɓ/-/b/, and voiceless ejective versus long-lag aspirated /p'/-/pʰ/. The English monolinguals outperformed the Spanish- but not the English-mode bilinguals on /ɓ/-/mb/ where both items are prevoiced (assimilated to English /b/ and /mb/ as in bubble vs bumble), and on the more English-like /p'/-/pʰ/, whereas the Spanish-mode bilinguals greatly outperformed the English monolinguals on the Spanish-like prevoiced implosive versus short-lag plosive /ɓ/-/b/. These results suggest that bilingualism can both enhance and inhibit non-native discrimination performance, depending on both the non-native contrast and the language mode of the bilinguals. This is consistent with Grosjean's claim that language mode affects all levels of language processing including phonetic and phonological levels, at least for these tasks.

Still, overall the findings on language mode effects in perceptual tasks have been mixed. In Antoniou et al. (2012), language mode affected Greek–English bilinguals' categorization of L1 and L2 stop-voicing contrasts, but had no effect on their discrimination of those same contrasts. We argued in that report that language mode affects *phonological* level judgments on L1 and L2 contrasts when the task specifies which language is to be used (e.g., phoneme categorization

and category-goodness ratings). However, because discrimination accesses a common L1-L2 *phonetic* space and does not require phonological judgments, bilinguals are free to use either of their L1 and L2 phonetic categories, and under these conditions they may be more likely to use those of their dominant language regardless of their current language mode. However, that study involved two languages that were both familiar to the bilinguals (their L1 and L2). A different pattern may be observed for unfamiliar non-native contrasts.

We therefore investigated how language mode affects bilinguals' categorization and goodness ratings, as well as discrimination, of non-native stop-voicing contrasts. We compared Greek–English bilinguals to monolingual listeners of English and Greek. English initial voiceless stops are typically long-lag aspirated, whereas voiced stops are unaspirated. Each stop is represented by a unique orthographic character in English (p, t, k and b, d, g, respectively). Greek voiceless stops are unaspirated and represented by a unique character (/p/ = π, /t/ = τ), whereas voiced stops are prevoiced and represented by digraphs (e.g., /b/ = μπ, /d/ = ντ), reflecting their Classical Greek origin as sequences of nasal + voiceless stop (see Arvaniti and Joseph, 2000 for a discussion of nasalization in Greek voiced stops). Due to this historical origin and current variations in presence of nasalization in voiced stops, Greek listeners have difficulty discriminating prevoiced from prenasalized stops (Antoniou et al., 2012; Malikouti-Drachman, 2001; but recall that Calderón and Best, 1996, found evidence of similar difficulties in Spanish-English bilinguals, whose L1 does not involve nasalization of voiced stops either historically or in current usage).

We took advantage of these language-specific differences and tested Greek-English bilinguals as compared to Greek and English monolinguals using Ma'di, a language containing a set of voicing (i.e., laryngeal) contrasts to which they had no prior linguistic exposure. Ma'di is a Nilo-Saharan language, spoken in the south of Sudan and the north of Uganda. It has a rich consonant inventory and contrasts four different kinds of stop-consonant voicing. For bilabials and coronals, Ma'di contrasts voiceless unaspirated /p, t/, prevoiced /b, d/, prenasalized /ᵐb, ⁿd/, and prevoiced implosive stops /ɓ, ɗ/ (Andersen, 1986; Blackings and Fabb, 2003; Kilpatrick, 1985).

Based on both articulatory and acoustic similarities, according to PAM, English monolinguals should assimilate the Ma'di prevoiced stops /b, d/ as very good exemplars of English /b, d/, and the Ma'di voiceless unaspirated stops /p, t/ as moderate-to-good exemplars of English /b, d/, which should result in CG assimilation and moderate levels of discrimination. This perceptual assimilation is expected because for English listeners, prevoicing is not in a crowded part of the VOT continuum, whereas voicing lag is. That is, prevoicing is hyper-voiced and should result in consistent categorization. Compatible with this reasoning, Bohn and Flege (1993) found that Spanish monolinguals consistently categorized English long-lag /t/ as "t," which is short-lag unaspirated in their L1. For Spanish listeners, long lag is not in a crowded part of the VOT continuum, and thus for them it appears that long lag is hyper-voiceless, and results in

Antoniou et al.: Focusing the lens of language experience

consistent categorization. The Ma'di prevoiced versus prenasalized contrasts /b/-/ᵐb/ and /d/-/ⁿd/ should instead yield UC assimilations because English does not contrast /b, d/ versus /mb, nd/ in word-initial position but does contrast them in medial position in disyllables, and because it does not require systematic prevoicing of /b/ in initial position but does require it intervocalically. As such, /ᵐb/ and /ⁿd/ should result in partial assimilations and L1-English category overlap is expected with /b/ and /d/, respectively, and moderate discrimination should result. Finally, for English monolinguals the prevoiced plosive versus implosive contrasts /b/-/ɓ/ and /d/-/ɗ/ should result in SC assimilations to English /b/ and /d/, respectively, and poor discrimination performance. This prediction is consistent with prior work on English monolingual listeners who assimilated the Zulu plosive-implosive contrast to a single native category and discriminated it relatively poorly (Best *et al.*, 2001). The Greek monolinguals, on the other hand, should assimilate the Ma'di prevoiced stops /b, d/ and voiceless unaspirated stops /p, t/ as excellent exemplars of Greek /b, d, p, t/, respectively, resulting in TC assimilations and excellent discrimination. But they should categorize the Ma'di prenasalized /ᵐb, ⁿd/ and voiced implosive stops /ɓ, ɗ/ as moderate-to-good exemplars of Greek prevoiced /b, d/, resulting in SC assimilations, and poor discrimination. Predicted assimilation patterns for the English and Greek monolingual groups are shown in Table I.

For the bilinguals, there are four possible ways in which they may perform relative to the monolinguals.

(1) Merged L1/L2 lens. Consistent with the SLM, bilinguals' discrimination performance on completely non-native voicing (laryngeal) stop contrasts may be intermediate to the two monolingual groups, reflecting the cumulative effects of the merged L1 and L2 phonetic systems.
(2) Persistent L1 lens. Bilinguals may discriminate non-native stop voicing in a manner similar to L1 (Greek) monolinguals, regardless of their language mode, consistent with the account of persistent L1-effects observed in early bilinguals despite many years of continued L2 use of Pallier *et al.* (1997).
(3) Dominant language lens. L2-dominant bilingual participants may instead discriminate non-native stop-voicing contrasts in a manner similar to L2 (English) monolinguals regardless of their language mode, reflecting their L2 dominance, consistent with the findings of Antoniou *et al.* (2012) on perception of stop-voicing contrasts in their own two languages.

(4) Separate L1 and L2 lenses. If, as Grosjean (1998, 2001) posits, language mode affects all levels of language processing, then bilinguals should show language-specific sensitivity shifts with language mode, by categorizing and assigning goodness ratings relative to the categories of the L1 or L2. That is, there should be an advantage in discrimination for the Greek mode bilinguals for Ma'di /b/-/p/ and /d/-/t/, and for the English mode bilinguals for /b/-/ᵐb/ and /d/-/ⁿd/. This would be consistent with findings that non-native stop-voicing discrimination can be enhanced or inhibited depending on language mode (Calderón and Best, 1996).

II. METHOD

A. Participants

Forty bilinguals were recruited from the Greek-Australian community in Sydney and were assigned to two groups based on language mode. All bilinguals came from the same population, and strict selection criteria were employed to ensure that they did not differ in their pattern of language acquisition, dominance, and use. The English mode (EM) and Greek mode (GM) groups were each comprised of ten males and ten females, all of whom were born in Sydney, had been exposed to Greek since birth, learned English as an L2 but no later than by age 5, and used English more than Greek to such an extent that they had become dominant in English, their L2. Their language histories had been acquired via questionnaire. The EM and GM bilingual groups were matched in age, the ages at which they began acquiring Greek and English, their mean self-ratings for their mastery (understanding, speaking, reading, writing) of Greek and English, and their self-reported daily use[1] of Greek and English (see Table II). All continued to use both Greek and English in their everyday lives, and were literate in Greek, although English was used more frequently, and in a wider variety of social situations.

Twenty-five English monolinguals ($M_{age} = 24.2$ years; thirteen males and twelve females) were recruited from the undergraduate student population at the University of Western Sydney. Twenty Greek monolingual ($M_{age} = 23.7$ years; ten males and ten females) residents of Athens, Greece, were recruited. Some had limited knowledge of English from school but none had spent extended time in an English-speaking country. Participants were paid the standard local rates for their participation. One Greek monolingual and one GM bilingual did not complete the tests and were removed from statistical analyses.

TABLE I. Predicted assimilation patterns for the English and Greek monolinguals for the Ma'di prevoiced-voiceless, prevoiced-prenasalized, and prevoiced plosive-implosive contrasts. Discrimination performance is predicted to follow the gradient TC > CG > SC, and UC will depend on the amount of overlap in the assigned category labels for the two Ma'di phones comprising the contrast.

Monolingual group	prevoiced-voiceless /ba/-/pa/, /da/-/ta/	prevoiced-prenasalized /ba/-/ᵐba/, /da/-/ⁿda/	prevoiced plosive-implosive /ba/-/ɓa/, /da/-/ɗa/
English monolinguals	CG	UC	SC
Greek monolinguals	TC	SC	SC

Bilingual group	Age (years)	Age learned (years)		Self-ratings (1–5)		% daily use	
		Greek	English	Greek	English	Greek	English
English monolinguals	23.7		0		5		100%
Greek monolinguals	24.2	0		5		100%	
EM	25.7	0	2.9	3.6	5.0	31%	86%
GM	24.6	0	3.0	3.5	5.0	33%	88%

B. Stimulus materials

The stimuli were recorded from two male Ma'di native talkers of the Lokai dialect. Both were highly literate and phonetically trained. Speech was elicited from printed targets that were presented in quasi-random order on a computer monitor.

The target pseudo-words /ba, pa, ᵐba, ɓa, da, ta, ⁿda, ɗa/ were embedded in a Ma'di carrier phrase and were later excised from the speech recordings using PRAAT acoustic analysis software (Boersma and Weenink, 2001). For the purposes of measuring VOTs, markers were placed at the beginning of the closure phase of the target stop, at the moment of consonantal release, and at the first periodic pitch pulse at the onset of the vowel (see illustrative oscillograms and spectrograms of Ma'di voiceless and voiced bilabial stops in Figs. 1 and 2).

For prenasalized stops, it can be difficult to identify a boundary between the nasal and stop prevoicing. Nasals are produced with closure of the oral cavity, open velum and radiation through the nasal cavity. They are characterized by the presence of the nasal formant (a high-intensity low-frequency F1), another peak around 1000 Hz, and antiresonances that dampen the higher frequencies (Stevens, 1999). Therefore, we separated the nasal from the stop prevoicing at the point where sudden energy loss occurred in the frequency components above approximately 250 Hz, caused by the closure of the velum to produce the oral stop, and measured the prevoicing from that point up to the release burst (see Fig. 3).

Prevoiced implosives were measured using the same procedures as for oral stops. Implosives are produced by lowering the larynx in a piston-like manner with vibrating vocal folds during oral closure. The downward sliding of the larynx reduces the air pressure in the oral cavity, and results in a momentary influx of air into the mouth. Markers could therefore be placed at the beginning of the closure phase of the stop, at the consonantal release and at the vowel onset (see Fig. 4).

Two tokens of each target were selected from each talker, for a total of 32 tokens. The contrasting tokens were selected to match as closely as possible on duration, fundamental frequency and amplitude measures (see Table III).

For categorization and goodness rating, the Ma'di bilabial /b, p, ᵐb, ɓ/ and coronal /d, t, ⁿd, ɗ/ stops were presented in syllable-initial /Ca/ position. For discrimination, three types of syllable-initial contrasts were used: prevoiced versus voiceless unaspirated stops /ba/-/pa/, /da-ta/; prevoiced versus prenasalized stops /ba/-/ᵐba/, /da/-/ⁿda/; and prevoiced plosive versus implosive stops /ba/-/ɓa/, /da/-/ɗa/.

C. Procedure

All participants, both monolingual and bilingual, were tested by the same simultaneous Greek–English bilingual experimenter (MA). The bilinguals were effectively treated as monolinguals in that all communication and linguistic content was in only one language, Greek for half of the bilinguals, and English for the other half. The English monolinguals and EM bilinguals were only spoken to in English, whereas the Greek monolinguals and GM bilinguals were spoken to in Greek. All written and onscreen instructions and interactions with the experimenter were
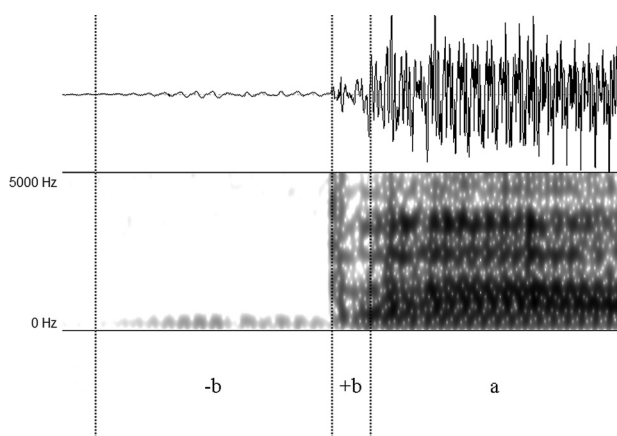

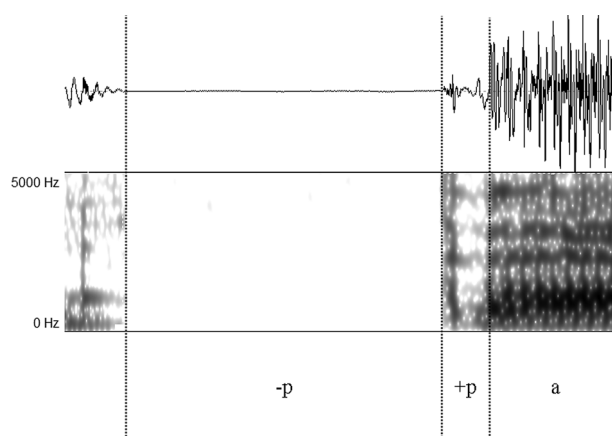FIG. 1. Ma'di prevoiced bilabial plosive /b/.

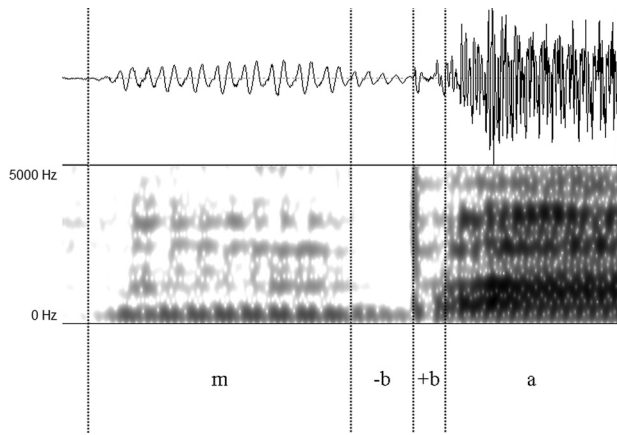FIG. 2. Ma'di voiceless unaspirated bilabial plosive /p/.

FIG. 3. Ma'di prenasalized bilabial plosive /^{m}b/.

consistent with that language. The designated language mode was maintained for each participant throughout the entire experiment session, including both verbal and written materials and discussion during experiment breaks. At the commencement of the experiment, participants were greeted and conversed with the experimenter about language appropriate topics. Great care was taken to avoid topics of discussion that might activate the other language. For example, EM bilinguals were usually asked about their occupation, current Australian news headlines and events, whereas those in GM were asked about their culture, family, church life and visits to Greece. The consent form and language background questionnaire were completed, and the participant received instructions and was familiarized with the AXB procedure. Half of the AXB task was completed and then a break was given. During this time, the participant chatted with the experimenter. The second half of the AXB task was then completed. A second break was then given, during which the participant was instructed regarding the categorization procedure and was familiarized with the onscreen choice category labels they should use to indicate what the consonant in each Ma'di syllable played to them sounded liked with respect to Greek or English consonants, depending on their language mode. The labels were either in Greek (for Greek and GM participants) or in English orthography (for English and EM
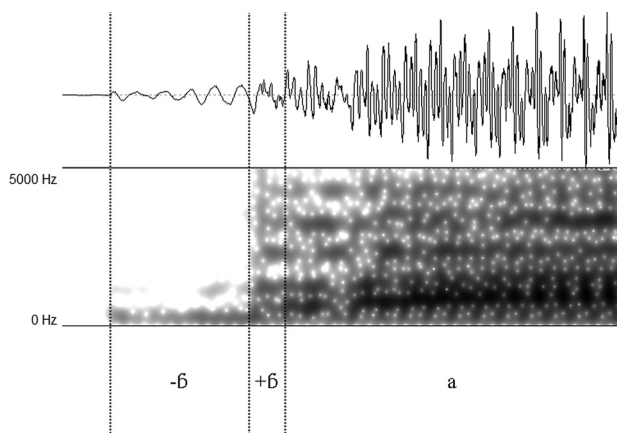

FIG. 4. Ma'di prevoiced bilabial implosive /ɓ/.

participants). Each label was paired with sample words from Greek or English, depending on the language mode for the participant, as examples of the labeled Greek or English consonant(s). It was made clear to participants that the Ma'di phonetic realization they heard might not match their own (e.g., not many people produce /^{m}ba/ in initial position) but what was important was that they understood the difference between the labels. So as not to interfere with this single language procedure, upon completion of the experiment, bilingual participants completed self-ratings and language use questionnaire items for their other language after completing the test, i.e., during the debriefing.

On the day of testing, participants first completed an AXB discrimination task, followed by a categorization and goodness rating task, using a laptop computer, Roland UA-25 audio interface and Sennheiser HD 650 headphones.

On each trial of the AXB discrimination test, three stimuli were presented. Participants chose whether the first (A) or last item (B) matched the middle item (X), and indicated their response by pressing one of two keys (covered with a label that read in English, FIRST or LAST, or in Greek, ΠΡΩΤΗ or ΤΕΛΕΥΤΑΙΑ). Tokens recorded from talker 1 were used for the A and B tokens and those recorded from talker 2 were used for the X tokens. This multiple-talker presentation made the task more difficult because participants could not rely on within-talker phonetic cues for discrimination. A block of sixteen triads was presented for each contrast in random order. The order of the contrast blocks was also randomized across participants. The interstimulus interval was 1 s and intertrial interval was 2 s. Participants were required to make their response within a 3.5 s time limit. If the participant exceeded that limit, the next AXB trial was presented and the "missed" trial was presented again at the end of the contrast block. Less than 3% of all trials were missed.

For the categorization test, eight labels were presented onscreen for each language condition (English: b, d, p, t, m + b, n + d, m + p, n + t; Greek: μπ, ντ, π, τ, μ + μπ, ν + ντ, μ + π, ν + τ), offering both voiced and voiceless stops as well as nasal + stop combinations for onscreen selection, as has been used in our previous research on Greek–English bilinguals (Antoniou *et al.*, 2012). Participants were familiarized with the responses before the experiment began. After the participant assigned a category label by clicking on one of the onscreen options using a mouse, they heard the same token again and had to assign a goodness rating ranging from 1 (very strange) to 7 (perfect) relative to the selected native category. The trial pairs (i.e., categorization trial followed by goodness-rating trial) were presented in random order.

The English monolinguals and the two groups of bilinguals were either tested in a quiet testing booth at the MARCS Institute, or in a quiet room in their home. The Greek monolinguals were tested in a quiet room in their home in Athens. Headphone volume was set using a sound-level meter to equalize the signal-to-noise ratio across participants (+35 dB), as the ambient noise level varied across the different testing environments.

TABLE III. Acoustic attributes of the Ma'di bilabial and coronal stimulus tokens. Descriptions of measurement procedures for stop closure, burst, and VOT durations are presented in text.

Stimulus syllable	Annotation	Duration (ms)	VOT (ms) M (SD)	Overall dB (RMS ampl.)	Overall F0 (Hz)	Mean dB (RMS ampl.)
Bilabials						
/ba/	b (closure)	96.3	−96.3 (30.8)	44.8		
	b (burst)	15.7		68.6		
	a	296.1		73.2	165.9	75.9
	Total	408.1				
/pa/	p (closure)	137.3		38.2		
	p (burst)	19.6	19.6 (3.8)	66.0		
	a	323.0		71.1	165.7	71.1
	Total	479.9				
/ᵐba/	m	102.5		61.5		
	b (closure)	38.1	−38.1 (13.3)	58.0		
	b (burst)	12.0		67.7		
	a	351.8		72.6	161.8	72.6
	Total	504.4				
/ɓa/	ɓ (closure)	32.3	−32.3 (5.6)	59.4		
	ɓ (burst)	6.2		70.1		
	a	340.8		73.1	166.5	73.1
	Total	379.3				
Coronals						
/da/	d (closure)	111.3	−111.3 (19.4)	51.4		
	d (burst)	13.6		68.8		
	a	331.5		73.3	165.9	73.3
	Total	456.4				
/ta/	t (closure)	136.8		41.6		
	t (burst)	10.6	10.6 (3.6)	68.6		
	a	294.6		71.1	162.6	71.1
	Total	442.0				
/ⁿda/	n	107.5		62.4		
	d (closure)	25.4	−25.4 (7.2)	60.5		
	d (burst)	11.2		70.5		
	a	331.2		72.1	173.3	72.1
	Total	475.3				
/ɗa/	ɗ (closure)	59.6	−59.6 (10.3)	53.8		
	ɗ (burst)	8.2		70.3		
	a	350.8		72.2	159.6	72.2
	Total	418.6				

III. RESULTS

A. Categorization and goodness ratings

We report the categorization results first for theoretical reasons, because PAM predicts discrimination success from assimilation patterns, which we derived from the categorization data. We adopted the identification convention used in past research, which specifies that if one label is applied for 70% or more of all responses for a given non-native phone, then that phone is considered to be categorized (Antoniou et al., 2012; Bundgaard-Nielsen et al., 2011). Category goodness-of-fit ratings ranged from 1 to 7, where average ratings between 1.0–3.0 were considered poor, 3.1–5.0 were considered moderately good, and 5.1–7.0 were considered very good/excellent. When two non-native phones were categorized to the same native category, a t-test was conducted to determine whether the goodness-of-fit ratings differed (in which case the assimilation pattern is a CG type) or if they did not significantly differ (SC assimilation type). The categorization and goodness ratings for the Ma'di syllable-initial bilabial and coronal stop consonants provided by the English monolinguals, Greek monolinguals, and bilinguals in English and Greek modes are shown in Tables IV and V, respectively.

1. Categorization of Ma'di bilabial stops

We will first review the data for the categorization and goodness-of-fit ratings for the bilabials and then turn to the coronal stops. A summary of the assimilation types for the Ma'di stop contrasts derived from the categorization and goodness rating data is shown in Table VI. For the bilabials,

TABLE IV. Mean percent identification (and goodness ratings 1–7) of Ma'di syllable-initial bilabial stop consonants by monolinguals and bilinguals in English and Greek language modes.

Group	Consonant stimuli	b/μπ	p/π	m + b/μ + μπ	m + p/μ + π	Criterion met[a]	Categorized as
Monolingual English	/ba/	**96.2**	1.3	2.5		C	/b/
		(4.6)	(3.0)	(3.5)			
	/pa/	17.7	**70.9**	1.3	10.1	C	/p/
		(4.2)	**(3.6)**	(3.0)	(3.9)		
	/ᵐba/	**46.3**	1.3	**45.0**	7.5	U	
		(3.9)	(6.0)	**(4.5)**	(3.2)		
	/ɓa/	**94.9**	1.3	3.8		C	/b/
		(4.2)	(4.0)	(1.7)			
Monolingual Greek	/ba/	**63.8**	**22.5**	5.0	8.8	U	
		(5.4)	**(3.6)**	(4.3)	(3.1)		
	/pa/		**100**			C	/p/
			(6.0)				
	/ᵐba/	**50.0**	1.3	**37.5**	11.3	U	
		(5.3)	(1.0)	**(5.1)**	(4.7)		
	/ɓa/	**97.5**	1.3		1.3	C	/b/
		(5.5)	(1.0)		(5.0)		
English mode	/ba/	**87.3**	3.8	7.6	1.3	C	/b/
		(5.4)	(4.0)	(5.6)	(6.0)		
	/pa/	**19.2**	**67.9**	2.6	10.3	U	
		(4.2)	**(4.6)**	(5.5)	(4.0)		
	/ᵐba/	**36.3**	3.8	**51.3**	8.8	U	
		(4.2)	(4.0)	**(5.3)**	(4.1)		
	/ɓa/	**86.7**	5.3	8.0		C	/b/
		(5.0)	(3.0)	(5.5)			
Greek mode	/ba/	**71.3**	16.3	5.0	7.5	C	/b/
		(4.6)	(4.2)	(3.5)	(3.2)		
	/pa/	3.8	**91.0**	1.3	3.8	C	/p/
		(3.5)	**(5.2)**	(3.0)	(4.3)		
	/ᵐba/	**50.0**	1.3	**40.0**	8.8	U	
		(3.9)	(4.0)	**(4.6)**	(2.7)		
	/ɓa/	**89.9**	3.8	6.3		C	/b/
		(4.5)	(1.3)	(4.6)			

[a]< 70% was considered categorized (C), whereas < 70% criterion was uncategorized (U). Boldface indicates most chosen category for C and top two categories for U.

monolingual English participants consistently categorized Ma'di /ba/ and /ɓa/ (both > 90%) as moderately good exemplars of English /b/ (4.6 and 4.2 out of 7, respectively, which did not differ significantly by t-test, p = 0.075), yielding SC assimilation, which should result in poor discrimination. The English monolinguals just reached > 70% for categorizing Ma'di /pa/ as moderately good English /p/ (3.6), yielding TC assimilation for Ma'di /ba/-/pa/, for which PAM predicts excellent discrimination. They did not categorize /ᵐba/ reliably, spreading their responses across English /b/ and /m/ + /b/ (46.3% and 45%), respectively, resulting in UC assimilation for /ba/-/ᵐba/, which we would expect to result in moderate-to-good discrimination.

Monolingual Greeks unanimously categorized Ma'di /pa/ as an excellent version of Greek /p/ (6.0) and /ɓa/ as a very good Greek /b/ (5.5), a TC assimilation pattern. As expected, they had difficulty categorizing /ᵐba/ (50% of responses assigned to /b/ and 37.5% to /m/ + /b/). Surprisingly, the Greeks did not categorize Ma'di /ba/, falling short

of the categorization criterion (63.8% as Greek /b/, but with a high rating of 5.4), resulting in UC assimilations for /ba/-/pa/ and /ba/-/ɓa/, and UU assimilation for /ba/-/ᵐba/. Despite the latter seemingly similar assimilation types, recent developments of PAM would predict differences in discrimination success according to the degree of category overlap. Because /ba/ was mostly categorized as Greek /b/, we expect good discrimination of /ba/-/pa/ because there is relatively little overlap in the Greek categories assigned to the two Ma'di phones, whereas /ba/-/ɓa/ is likely to result in poor discrimination because of the overlap in categorization of both of these phones as good versions of Greek /b/.

EM bilinguals categorized both /ba/ (87.3%) and /ɓa/ (86.7%) as good exemplars of English /b/ (5.4 and 5.0 out of 7, respectively), yielding SC assimilation for /ba/-/ɓa/ (because goodness-of-fit ratings did not differ, p = 0.393) which should result in poor discrimination. They did not categorize /pa/ (67.9% of responses assigned to /p/, 19.2% to /b/, and 10.3% to /m/ + /p/) or /ᵐba/ (51.3% to /m/ + /b/ and

TABLE V. Mean percent identification (and goodness ratings 1–7) of Ma'di syllable-initial coronal stop consonants by monolinguals and bilinguals in English and Greek language modes.

Group	Consonant stimuli	Category label				Criterion met[a]	Categorized as
		d/ντ	t/τ	n+d/ν+ντ	nt/ν+τ		
Monolingual English	/da/	**97.5**		2.5		C	/d/
		(4.2)		(3.0)			
	/ta/	**72.3**	21.5	1.5	4.6	C	/d/
		(3.9)	(3.2)	(1.0)	(3.0)		
	/ⁿda/	**46.8**	2.5	**48.1**	2.5	U	
		(3.7)	(3.0)	**(4.1)**	(2.5)		
	/ɗa/	**96.6**	1.7	1.7		C	/d/
		(3.6)	(1.0)	(2.0)			
Monolingual Greek	/da/	**77.5**		10.0	12.5	C	/d/
		(5.3)		(3.8)	(3.3)		
	/ta/	9.8	**83.6**		6.6	C	/t/
		(4.5)	**(5.8)**		(3.8)		
	/ⁿda/	**47.5**		**42.5**	10.0	U	
		(4.8)		**(4.8)**	(4.9)		
	/ɗa/	**82.1**		1.8	16.1	U	
		(4.2)		(4.0)	(3.8)		
English Mode	/da/	**92.4**		7.6		C	/d/
		(5.3)		(5.7)			
	/ta/	**51.3**	**42.1**		6.6	U	
		(4.5)	**(4.1)**		(5.5)		
	/ⁿda/	**31.2**	1.3	**54.5**	13.0	U	
		(4.5)	(3.0)	**(5.1)**	(4.7)		
	/ɗa/	**82.1**		16.4	1.5	C	/d/
		(3.5)		(5.0)	(4.0)		
Greek Mode	/da/	**80.0**	2.5	3.8	13.8	C	/d/
		(4.6)	(2.5)	(5.7)	(2.5)		
	/ta/	12.3	**78.5**	3.1	6.2	C	/t/
		(4.3)	**(5.0)**	(4.0)	(3.3)		
	/ⁿda/	**36.7**		**39.2**	24.1	U	
		(3.5)		**(4.7)**	(3.2)		
	/ɗa/	**66.7**	4.8	9.5	**19.0**	U	
		(3.1)	(3.3)	(3.2)	**(2.9)**		

[a]< 70% was considered categorized (C), whereas < 70% criterion was uncategorized (U). Boldface indicates most chosen category for C and top two categories for U.

36.3% to /b/), and therefore both /ba/-/pa/ and /ba/-/ᵐba/ resulted in UC assimilations. Because most responses for /pa/ were /p/ and for /ᵐba/ were /m/+/b/ (albeit uncategorized), moderate-to-good discrimination should result for both contrasts.

GM bilinguals categorized both /ba/ (87.3%) and /ɓa/ (89.9%) as moderate exemplars of Greek /b/ (4.6 and 4.5,

respectively), yielding SC assimilation (goodness-of-fit ratings did not differ, $p = 0.926$) from which poor discrimination should result. They categorized /pa/ (91%) as good Greek /p/ (5.2), resulting in TC assimilation for /ba/-/pa/, and thus discrimination should be excellent. The GM bilinguals did not categorize /ᵐba/ (50% of responses assigned to /b/, 40% to /m/+/b/), resulting in UC assimilation for

TABLE VI. Monolinguals' and bilinguals' assimilation patterns for Ma'di syllable-initial stop contrasts.

	Contrast					
Group	/ba/-/pa/	/da/-/ta/	/ba/-/ᵐba/	/da/-/ⁿda/	/ba/-/ɓa/	/da/-/ɗa/
English monolinguals	TC	SC	UC	UC	SC	SC
Greek monolinguals	UC	TC	UU	UC	UC	CG
English Mode bilinguals	UC	UC	UC	UC	SC	CG
Greek Mode bilinguals	TC	TC	UC	UC	SC	UC

/b/-/^{m}ba/, but from the slight difference in goodness-of-fit ratings for the overlapping category (Greek /b/) we predict moderate discrimination.

2. Categorization of Ma'di coronal stops

For the coronals, monolingual English listeners categorized prevoiced /da/ (97.5%), voiced implosive /ɗa/ (96.6%), and voiceless unaspirated /ta/ (72.3%) all as moderately good exemplars of voiced English /d/ (4.2, 3.6, and 3.9 out of 7, respectively). This differs from their categorization for Ma'di short-lag /pa/ reported above which met the categorization criterion for the English voiceless bilabial stop /p/. One possible reason for this different pattern of categorization between the bilabial and coronal voiceless stops is that the VOT for Ma'di /da/ was 10.6 ms, close to that of English voiced /d/ (i.e., voiceless unaspirated [t]), whereas the mean VOT of Ma'di /pa/ was 19.6 ms, which is closer to the English voiceless aspirated VOT for /p/ ([p^h], typically > 40 ms). Thus, /da/-/ta/ and /da/-/ɗa/ yielded SC assimilations (goodness-of-fit ratings did not differ for either contrast, $p = 0.09$ and $p = 0.12$, respectively) and both should result in poor discrimination. The English monolinguals did not meet the categorization criterion for Ma'di /^{n}da/, for which responses were spread across English /d/ (46.8%) and /n/ + /d/ (48.1%), both rated as moderately good exemplars (3.7 and 4.1, respectively). Thus, /da/-/^{n}da/ yielded UC assimilation, and because of the partial overlap with /d/, discrimination should be moderate, though higher than for the SC contrasts.

Monolingual Greek participants reliably categorized all except for the prenasalized stop /^{n}da/, which Greek listeners typically have difficulty with. The Greek monolinguals categorized /da/ (77.5%) and /ɗa/ (82.1%), respectively, as very good and moderately good exemplars of Greek /d/ (5.3 for /da/, 4.2 for /ɗa/). This difference in goodness-of-fit ratings, $t(18) = 4.5$, $p < 0.001$, resulted in CG assimilation and from this we expect moderate discrimination. The Greeks categorized Ma'di /ta/ (83.6%) as a very good exemplar of Greek /t/ (5.6), yielding TC assimilation for /da/-/ta/, for which we expect excellent discrimination. The Greeks did not meet the categorization criterion for Ma'di /^{n}da/. Their responses were spread across /d/ (47.5%) and /n/ + /d/ (42.5%), resulting in UC assimilation for /da/-/^{n}da/. We would expect discrimination of /da/-/^{n}da/ to be better than for CG /da/-/ɗa/ as there was less category overlap in the assimilation pattern, but neither of these contrasts should be discriminated as well as TC /da/-/ta/.

The EM bilinguals categorized Ma'di /da/ (92.4%) and /ɗa/ (82.1%), respectively, as very good and moderately good exemplars of English /d/ (5.3 and 3.5 out of 7). This difference in goodness ratings, $t(19) = 5.3$, $p < 0.001$, resulted in CG assimilation of /da/-/ɗa/ from which we expect moderate discrimination. Neither /ta/ nor /^{n}da/ met the 70% categorization criterion. Responses for /ta/ were shared across English /d/ (51.3%) and /t/ (42.1%), both as moderately good exemplars (4.5 and 4.1, respectively), and responses for /^{n}da/ were shared across /n/ + /d/ (54.5%) and /d/ (31.2%), as very good and good exemplars (5.1 and 4.5, respectively). Both /da/-/ta/ and /da/-/^{n}da/ yielded

UC assimilations, but due to overlap with English /d/, /da/-/ta/ should result in moderate discrimination, whereas /da/-/^{n}da/ showed less overlap, and thus discrimination should be better. Importantly, the EM bilinguals did not categorize the voiceless unaspirated Ma'di /ta/ as English /d/, as had the English monolinguals, suggesting that bilinguals may have narrower VOT categories than monolinguals.

The GM bilinguals categorized /da/ as /d/ (80%) and as a moderately good exemplar (4.6). They categorized /ta/ (78.5%) as a good exemplar of Greek /t/ (5.0), and thus /da/-/ta/ yielded TC assimilation. The GM bilinguals did not categorize Ma'di /^{n}da (36.7% of responses assigned to /d/, 39.2% to /n/ + /d/, and 24.1% to /n/ + /t/) or /ɗa/ reliably (66.7% assigned to /d/, 19% to /n/ + /t/), and thus /da/-/^{n}da/ and /da/-/ɗa/ both resulted in UC assimilations. Due to differences in the overlap of the assigned categories, discrimination of /da/-/^{n}da/ should be good, but not as good as for TC assimilations, whereas discrimination of /da/-/ɗa/ should be poor due to the high degree of overlap in the assigned category labels.

It is clear that the two bilingual groups' categorization was affected by language mode and this is reflected in their assimilation types. For example, the GM bilinguals successfully categorized both Ma'di /ba/-/pa/ and /da/-/ta/ as TC contrasts, whereas the EM bilinguals fell short of the categorization criterion for /pa/ and /ta/, thus showing UC rather than TC assimilation for the Ma'di /ba/-/pa/ and /da/-/ta/ contrasts. However, it is also clear from Table VI that the assimilation patterns of each bilingual group also differed in some ways from those of the corresponding monolingual groups.

According to PAM, TC assimilations should result in the best levels of discrimination, followed by CG and then SC assimilation types. Assimilation types involving an uncategorized phone (UC and UU) may vary in how well they are discriminated depending on the phonetic similarity and amount of overlap of the partial assimilations of the uncategorized phone and the categorized one (or in the case of UU assimilation, the overlap between the two uncategorized phones).

B. AXB discrimination

In contrast to the excellent discrimination PAM would predict from their TC assimilation for /ba/-/pa/, but consistent with the poor discrimination predicted for /da/-/ta/, the English monolinguals were constrained by the English short- versus long-lag stop-voicing distinction and discriminated the Ma'di prevoiced versus voiceless unaspirated contrasts poorly. They showed notably better but still far from excellent discrimination of the prevoiced versus prenasalized contrasts and poor discrimination of the prevoiced plosive versus implosive contrasts (see mean discrimination performance levels in Table VII).

Consistent with PAM, on the other hand, the Greek listeners' mean discrimination levels were very good for prevoiced versus voiceless unaspirated contrasts on which they had shown TC (/da/-/ta/) and UC (/ba/-/pa/) assimilation. Still, they were not at ceiling, suggesting that the Ma'di

TABLE VII. Monolinguals' and bilinguals' mean % correct discrimination of Ma'di initial position stop-voicing contrasts (standard error in parentheses).

Group	Contrast					
	/ba/-/pa/	/da/-/ta/	/ba/-/ᵐba/	/da/-/ⁿda/	/ba/-/ɓa/	/da/-/ɗa/
English	63.3	56.0	72.3	77.5	53.5	58.3
	(2.0)	(2.0)	(2.4)	(2.4)	(2.0)	(2.0)
Greek	80.6	90.5	69.4	64.5	59.2	57.9
	(2.0)	(2.0)	(2.9)	(2.9)	(3.4)	(3.4)
EnMode	72.5	67.8	71.9	71.9	56.6	60.9
	(3.1)	(3.1)	(3.3)	(3.3)	(1.6)	(1.6)
GrMode	71.1	71.7	70.4	69.1	53.9	57.9
	(2.8)	(2.8)	(3.1)	(3.1)	(2.3)	(2.3)

prevoiced versus voiceless unaspirated distinction was not a 'perfect fit' to the Greek stop-voicing distinction, especially for /ba/-/pa/. As expected, they showed relatively poor discrimination of the prevoiced versus prenasalized contrasts, and even poorer discrimination of the prevoiced versus implosive contrasts.

The two groups of bilinguals did not appear to differ in their discrimination as a function of language mode. Both bilingual groups appeared to be intermediate to the two monolingual groups for the prevoiced-voiceless unaspirated and prevoiced-prenasalized contrasts, and their discrimination of the prevoiced plosive-implosive contrasts was as poor as that of the two monolingual groups.

To evaluate the reliability of those observations, a mixed $2 \times 2 \times (2 \times 3)$ analysis of variance was conducted with between-subjects factors of language context (English vs Greek) and lingualism (mono- vs bilingual), and within-subjects factors of place (bilabial vs coronal) and contrast (prevoiced-voiceless unaspirated vs prevoiced-prenasalized vs prevoiced plosive-implosive). A main effect of contrast revealed that not all contrasts were equally difficult to discriminate, $F(2, 78) = 47.3$, $p < 0.001$, $\eta^2_p = 0.548$. Post hoc tests confirmed that overall, the prevoiced plosive versus implosive contrasts were more difficult to discriminate than prevoiced versus voiceless unaspirated ($M_{diff} = 13.6$), $F(1, 79) = 77.8$, $p < 0.001$, and prevoiced versus prenasalized contrasts ($M_{diff} = 14.0$), $F(1, 79) = 68.2$, p < 0.001. Discrimination performance did not differ between prevoiced versus voiceless unaspirated and prevoiced versus prenasalized contrasts.

A significant two-way language × contrast interaction revealed that the language groups differed in their discrimination of the different contrast types, $F(2, 78) = 18.0$, $p < 0.001$, $\eta^2_p = 0.315$, and a higher order three-way language × lingualism × contrast interaction suggested that the language group differences for the contrast types were mediated by whether the listeners were mono- or bilingual, $F(2, 78) = 11.5$, $p < 0.001$, $\eta^2_p = 0.228$. Simple interactions revealed that the language × lingualism interaction held for prevoiced versus voiceless unaspirated contrasts, $F(1, 79) = 24.9$, $p < 0.001$, but not for prevoiced versus prenasalized or prevoiced plosive versus implosive contrasts. Further breakdown simple effects analyses revealed that the effect was driven by the difference between the Greek monolinguals'

excellent discrimination (85.5%) and the English monolinguals' poor discrimination (59.6%) of the prevoiced versus voiceless unaspirated contrasts, $F(1, 79) = 57.8$, $p < 0.001$. The bilinguals' discrimination of prevoiced versus voiceless unaspirated contrasts, however, was intermediate to the two monolingual groups (EM: 70.2%, GM: 71.4%). Importantly, the two bilingual groups did not differ in their discrimination for any of the six contrasts as a function of language mode. Their performance on both the prevoiced versus voiceless unaspirated and prevoiced versus prenasalized contrasts was intermediate to the two monolingual groups, regardless of language mode. All four groups performed poorly on the prevoiced plosive-implosive contrasts. All other effects not reported here were not statistically significant.

IV. DISCUSSION

The present findings demonstrate that when categorizing (assimilating) non-native stop consonant voicing distinctions to one of their two languages, bilingual listeners are sensitive to the phonetic settings of both of their languages. The bilinguals categorized and assigned goodness-of-fit ratings of non-native Ma'di stops relative to the stop-voicing categories of the L1 or L2 consistent with the language mode they were operating in. However, when discriminating the Ma'di stop-voicing contrasts, bilinguals were unaffected by language mode. Instead, they seemed to discriminate the non-native prevoiced versus voiceless unaspirated and prevoiced versus prenasalized contrasts at levels intermediate to the English and Greek monolinguals. The Greek monolinguals discriminated the prevoiced versus voiceless unaspirated contrasts best, English listeners performed worst, and bilinguals were in the middle. For the prevoiced versus prenasalized contrasts, conversely, English monolinguals performed best, Greeks were worst (although this difference did not reach significance), and bilinguals were once again in the middle. All four groups of listeners poorly discriminated the prevoiced plosive versus implosive contrasts.

The implications of these findings are that bilinguals do not possess a cumulative advantage for discriminating unfamiliar, non-native contrasts. Rather, they appear to integrate the phonetic properties of their L1 and L2, resulting in intermediate discrimination performance relative to monolinguals of each language, as was most clearly observed for prevoiced versus voiceless contrasts. Discrimination is the ability to judge whether or not two phones differ, whereas categorization also requires that the listener apply a label to the phone. For these reasons, in the fields of audiology and psychophysics, categorization is considered to be a higher level skill than discrimination (Macmillan, 1987). Discrimination, as a lower level skill, seems to access a common phonetic space that reflects the influence of both of the bilingual's languages, compatible with merged lens accounts of L1-L2 interaction, such as that of the SLM (Flege, 1995). Perhaps it is for this reason that discrimination is unaffected by situational factors, such as language mode.

The lack of a language mode effect on discrimination stands in contrast with the categorization results, however, for which a clear effect of language mode was observed.

Specifically, the categorization results indicate that relatively higher-level, that is, *phonological* judgments maintain a clear separation between early bilinguals' two languages, which is mediated by language mode. The categorization results thus partially complement studies on Spanish–Catalan bilinguals in Barcelona that have suggested a persistent L1 influence in early bilinguals across a variety of tasks (e.g., Pallier *et al.*, 2001; Sebastián-Gallés and Soto Faraco, 1999). Note, however, that in latter studies the bilinguals were not L2-dominant like our participants were, and the language mode was not manipulated, as it was in the present study. In fact, it is not explicitly stated in those papers which language was used by the experimenter during test sessions. These methodological differences, language environment differences, and/or differences in the phonetic similarities between the bilinguals' languages (Spanish and Catalan are both Romance languages and more closely related than are English, a Germanic language, and Greek, a Hellenic language), may explain why we did not replicate the persistent L1-influence described for Spanish–Catalan bilinguals. Further research would be needed to narrow down the source(s) of the different outcomes for the Spanish-Catalan studies versus our own.

The varying language mode effects observed here suggest the importance of task effects, as described in the recently developed automatic selective perception (ASP) model (Strange, 2011). It may be the case that attention to different levels of information in the stimuli (attentional focus in ASP terms) are needed by bilinguals for discrimination versus assimilation tasks. The language-specific differences between the two monolingual groups suggest that discrimination relies on non-contrastive phonetic information (e.g., gradient category-goodness information), as in our task discrimination could not be based on simply low level acoustic information (e.g., idiosyncratic differences between utterances from a single talker), as the A and B comparators were produced by a different speaker than the target item X. This interpretation is further supported by the intermediate discrimination performance of the bilingual groups relative to the two monolingual groups for the prevoiced versus voiceless contrasts. In contrast, the categorization task requires phonological-level judgments, that is, assimilating a non-native phone to a native category and judging its goodness-of-fit to that category. We argue that it is for this reason that language mode effects were observed between the bilingual groups on the perceptual assimilation tasks. Across both tasks, then, our interpretation is not fully consistent with Grosjean's (1998, 2001) statements that language mode affects *all* levels of language processing. Specifically, we found language mode effects for phonological-level judgments, but *not* for lower-level phonetic discrimination performance. A speculative conclusion at this point is that this disjoint between the phonetic and phonological levels may, in the case of bilinguals, have led to the discrepancies between their assimilation patterns in the different language modes and their discrimination performance.

Our findings from the monolingual listeners are largely consistent with the predictions of PAM, with the caveat that the excellent discrimination predicted for Ma'di /ba/-/pa/ by the English monolinguals did not occur. This may have been due to a categorization artifact due to the 70% categorization criterion that was employed (Ma'di /pa/ just reached the threshold at 70.9%). A stricter categorization criterion (such as 75%) would have resulted in UC assimilation for /ba/-/pa/, and when the partial overlap of assignments to English /b/ are taken into account, the data appear to be more consistent with the monolinguals' moderate (not excellent) discrimination performance. An alternative possibility is that because the discrimination task was administered first in order to minimize effects of categorization on discrimination performance, it is possible that the listeners may not have been relying consistently on assimilations to native stop-voicing categories during the discrimination task.

Although PAM predictions were largely accurate with regards to the monolinguals' discrimination, it appears that PAM predictions may not be simply transferred and applied to bilinguals, particularly as the effect of language mode clearly depends on which perceptual task is involved. We have shown that when asked, bilinguals categorize non-native stops relative to categories of the language mode, that is, their L1 or their L2. Although the effects of language mode on categorization are quite clear between the bilingual groups, neither group performed identically to either group of monolinguals in categorization. Additionally, these language mode effects do not influence their discrimination. Thus, both groups of bilinguals' assimilation patterns did reflect the effect of language mode on phonological categorization and rating judgments of the non-native stops, but for the GM bilinguals, these were not accurate predictors of discrimination performance, which is not consistent with what would be predicted by either PAM, PAM-L2, or SLM.

For over 50 years, research on bilingualism has tried to answer the question of whether a bilingual's two languages are integrated or kept separate. Traditionally, performance that is equal to that of monolingual native speakers (on whatever measure examined) of the two languages has been interpreted as evidence for the maintenance of separate L1 and L2 phonological systems. In a series of tightly controlled studies on the same population of L2-dominant Greek–English bilinguals, we have investigated speech production (Antoniou *et al.*, 2010), code-switching in production (Antoniou *et al.*, 2011), perception of native (Greek and English) stop-voicing distinctions (Antoniou *et al.*, 2012), and here, perception of non-native stop-voicing contrasts. Across these different investigations, we have observed a variety of patterns of L1-L2 interaction that have differed by task. In production of native stop-voicing contrasts in word-initial (CV) and word-medial post-nasal (VNCV) contexts, bilinguals in each language mode matched the VOTs of the corresponding English and Greek monolingual speakers, findings compatible with the view that the L1 and L2 are maintained separately (separate L1 and L2 lenses). However, and importantly, the findings from production of stops in medial intervocalic position (VCV) demonstrated that such interpretations of separate L1 and L2 systems that are free of L1-L2 interaction are too simplistic. Specifically, the EM bilinguals produced English voiced stops in intervocalic VCV position with significantly longer prevoicing than did native English speakers, reflecting an influence of their Greek L1 on their English L2 (albeit

modest). Such L1-influence could not emerge in our carefully controlled language mode conditions if the L1 and L2 phonological systems were completely separate. Therefore, we may reject bilingual theories that posit completely separate and non-interacting L1 and L2 systems. Indeed, in the study of Antoniou *et al*. (2011), when bilinguals were asked to use both languages in a code-switching task, that is, to produce L1 targets within an L2 context or vice versa, this L1 interference on L2 stop voicing productions increased (consistent with a persistent L1 lens). The fact that a VOT shift occurred in the code-switching conditions indicates that the phonetic categories of each language must be linked at a higher (and according to PAM-L2, phonological) level. However, a different pattern of L1-L2 interaction was observed in *perception* of native (Greek and English) stop-voicing contrasts (Antoniou *et al*., 2012). Bilinguals showed a clear language mode effect for categorization, but discrimination was unaffected—a finding that has been replicated here. In discrimination of their L1 and L2 stop voicing contrasts, both groups of bilinguals approximated the English monolinguals, suggesting a reverse influence of the later acquired, but dominant, L2 on the L1 (dominant language lens). The present investigation on nonnative perception of Ma'di stop-voicing contrasts has shown a fourth pattern of L1-L2 interaction: monolingual-like effects of language mode on assimilation performance (categorization and goodness-of-fit ratings), but *bidirectional L1-L2 interference* in discrimination, resulting in intermediate performance relative to L1 and L2 monolingual baselines for the prevoiced versus voiceless contrasts (consistent with a merged L1-L2 lens). The present results are strong evidence that although bilinguals attune perceptually to the L2 (dominant language), L1-L2 interaction effects persist, even after many years of continued use and dominance of the L2. The findings across this series of experiments suggest that the relationship between speech perception and production, at least in early sequential bilinguals, is more complicated than has been previously imagined. At the very least, it does not appear that perception and production are different sides of the same coin as was once thought (Liberman and Mattingly, 1985). Rather, our findings from this series of studies on Greek–English bilinguals demonstrate that the observed pattern of L1-L2 interaction is dependent on the communicative task, including whether it involves listening or speaking.

While these combined findings provide a range of new insights in the fields of speech production and perception, both within and across language boundaries, a number of questions remain, and new questions beckon. Current theories of cross-language speech perception (PAM), and of accented L1 and L2 production (SLM) and perception (PAM-L2), account for the perception abilities of naive monolingual listeners or inexperienced L2 learners, but have failed to account for the performance of fluent, stable bilinguals. The important theoretical contribution of the present work is that it lays the groundwork for the development of a framework that addresses the segmental production and perception of fluent bilingual speakers, and how it is swayed by language dominance and situational factors, such as language context and the communicative task. We posit, based on our own and others' findings, that language dominance interacts with bilingual phonetic and phonological organization, such that (1) L2-dominant bilinguals possess common interlanguage L1/L2 phonological categories that respond to separate L1 and L2 phonetic realizations within a multidimensional, multilingual phonetic landscape, consistent with the phonetic and phonological architecture proposed by PAM-L2, (2) the existence of these L1 and L2 phonetic categories allow for short-term perceptual reattunement to occur on the basis of, and in relation to, the immediate surrounding language context/mode, and (3) for L2-dominant bilinguals, task demands influence performance according to the attentional processing required for the task at hand, that is, whether they are required to attend to information that is relevant to the L1 and/or L2.

Our findings also suggest that perceptual flexibility, which may be seen as responsiveness or sensitivity to one's linguistic context, is a crucial element in language use, particularly by bilinguals but also by monolinguals. Findings of substantial and rapid perceptual flexibility, in terms of shifting phonetic boundaries to accommodate to a given speaker after even short periods of exposure, have also been reported for monolinguals (Norris *et al*., 2003). And this flexibility may indeed be important, and likely indispensable, for second/foreign language learning beyond early childhood. As well, it is central to language change over the lifespan, and to the patterns of speech perception in bilinguals. Indeed PAM and SLM have addressed speech perception (and production), but without explicitly considering such developmental adjustments, and what facilitates/drives them. Finally, given that the combined results across studies demonstrate different patterns of L1-L2 interaction, we also recommend that researchers should exercise caution when generalizing from observations on a single task that tests only one specific language ability (see also Strange, 2011).

Although this series of experiments has tested a previously overlooked, but large, population of bilinguals, the present research has several limitations that should be noted. First, a restricted set of consonant contrasts was employed. Second, the strict sampling is simultaneously a great strength of our investigation, but also limits the generalizability of our findings to other bilingual populations. The decision to investigate L2-dominant bilinguals was driven by prior findings (Flege *et al*., 2003; Antoniou *et al*., 2012), theory (PAM-L2), as well as the fact that L2-dominant bilinguals are a large subgroup that is under-represented in the literature. L2-dominant bilinguals are common, particularly among migrant populations in the US, UK, and Australia. Finally, a fuller understanding of the influence of language dominance would be obtained by testing bilinguals who speak the same two languages but differ in language dominance. Unfortunately, to our knowledge there exists no sizeable English-L1 group that has acquired Greek in early childhood. It may be possible to examine this in other populations (e.g., Mandarin bilinguals who are dominant in their L1 versus in their L2). We leave these issues to future research.

In conclusion, this paper provides novel and important contributions to the study of bilinguals' perception of nonnative spoken language. The results demonstrate that bilinguals are uniquely configured and flexible language users who integrate both languages in a common phonetic/phonological

space, and that different and asymmetrical patterns of L1-L2 interaction emerge in categorization versus discrimination of non-native phones. Importantly, the findings indicate that at least some aspects of perception (categorization) may be more malleable than others (discrimination), even after many years of continued L2 exposure, usage, and even dominance.

ACKNOWLEDGMENTS

[1]Participants estimated their daily L1 and L2 use separately so as not to interfere with the language mode procedure. For this reason, the percentages do not sum to 100%. Importantly, they clearly illustrate the bilinguals' L2 dominance.

Andersen, T. (**1986**). "The phonemic system of Madi," Afrika Übersee **69**, 193–207.

Antoniou, M., Best, C. T., Tyler, M. D., and Kroos, C. (**2010**). "Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2," J. Phonetics **38**, 640–653.

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

Antoniou, M., Tyler, M. D., and Best, C. T. (**2012**). "Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode?," J. Phonetics **40**, 582–594.

Arvaniti, A., and Joseph, B. D. (**2000**). "Variation in voiced stop prenasalization in Greek," Glossologia **11–12**, 131–166.

Best, C. T. (**1995**). "A direct realist view of cross-language speech perception," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, edited by W. Strange (York Press, Baltimore), pp. 171–204.

Best, C. T., McRoberts, G. W., and Goodell, E. (**2001**). "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," J. Acoust. Soc. Am. **109**, 775–794.

Best, C. T., and Tyler, M. D. (**2007**). "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, edited by O.-S. Bohn and M. J. Munro (John Benjamins Publishing Company, Amsterdam), pp. 13–34.

Blackings, M., and Fabb, N. (**2003**). *A Grammar of Ma'di* (Mouton de Gruyter, New York).

Boersma, P., and Weenink, D. (**2001**). "PRAAT, a system for doing phonetics by computer," Glot Int. **5**, 341–345.

Bohn, O.-S., Best, C. T., Avesani, C., and Vayra, M. (**2011**). "Perceiving through the lens of native phonetics: Italian and Danish listeners' perception of English consonant contrasts," in *Proceedings of the 17th International Congress of Phonetic Sciences*, edited by W.-S. Lee and E. Zee (Department of Chinese, Translation and Linguistics, City University of Hong Kong, Hong Kong), pp. 336–339.

Bohn, O.-S., and Flege, J. E. (**1993**). "Perceptual switching in Spanish/English bilinguals," J. Phonetics **21**, 267–290.

Bundgaard-Nielsen, R. L., Best, C. T., and Tyler, M. D. (**2011**). "Vocabulary size matters: The assimilation of L2 Australian English vowels to L1 Japanese vowel categories," Appl. Psycholinguist **32**, 51–67.

Calderón, J., and Best, C. T. (**1996**). "Effects of bilingualism on non native phonetic contrasts," J. Acoust. Soc. Am. **99**, 2602.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., and Carbone, E. (**1973**). "The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals," J. Acoust. Soc. Am. **54**, 421–428.

Elman, J. L., Diehl, R. L., and Buchwald, S. E. (**1977**). "Perceptual switching in bilinguals," J. Acoust. Soc. Am. **62**, 971–974.

Flege, J. E. (**1995**). "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, edited by W. Strange (York Press, Baltimore), pp. 233–277.

Flege, J. E., and Eefting, W. (**1987a**). "Production and perception of English stops by native Spanish speakers," J. Phonetics **15**, 67–83.

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

Flege, J. E., and Liu, S. (**2001**). "The effect of experience on adults' acquisition of a second language," Stud. Second Lang. Acquis. **23**, 527–552.

Flege, J. E., and Mackay, I. R. A. (**2004**). "Perceiving vowels in a second language," Stud. Second Lang. Acquis. **26**, 1–34.

Flege, J. E., MacKay, I. R. A., and Piske, T. (**2002**). "Assessing bilingual dominance," Appl. Psycholinguist **23**, 567–598.

Flege, J. E., Schirru, C., and MacKay, I. R. A. (**2003**). "Interaction between the native and second language phonetic subsystems," Speech Commun. **40**, 467–491.

García-Sierra, A., Diehl, R. L., and Champlin, C. (**2009**). "Testing the double phonemic boundary in bilinguals," Speech Commun. **51**, 369–378.

García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., and Champlin, C. A. (**2012**). "Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study," Brain Lang. **121**, 194–205.

Grosjean, F. (**1998**). "Studying bilinguals: Methodological and conceptual issues," Biling. Lang. Cogn. **1**, 131–149.

Grosjean, F. (**2001**). "The bilingual's language modes," in *One Mind, Two Languages: Bilingual Language Processing*, edited by J. Nicol (Blackwell Publishing, Oxford), pp. 1–22.

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

Jenkins, J. J., Strange, W., and Polka, L. (**1995**). "Not everyone can tell a 'rock' from a 'lock': Assessing individual differences in speech perception," in *Assessing Individual Differences in Human Behavior: New Concepts, Methods, and Findings*, edited by D. Lubinski and R. V. Dawis (Davies-Black Publishing, Palo Alto, CA), pp. 297–325.

Jia, G., Strange, W., Wu, Y., Collado, J., and Guan, Q. (**2006**). "Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure," J. Acoust. Soc. Am. **119**, 1118–1130.

Kilpatrick, E. (**1985**). "Preliminary notes on Ma'di phonology," Occas. Pap. Study Sudan Lang. **4**, 119–132.

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

Macmillan, N. A. (**1987**). "Beyond the categorical/continuous distinction: A psychophysical approach to processing modes," in *Categorical Perception: The Groundwork of Cognition*, edited by S. R. Harnad (Cambridge University Press, New York), pp. 53–85.

Malikouti-Drachman, A. (**2001**). "Greek phonology: A contemporary perspective," J. Greek Linguist **2**, 187–243.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (**1975**). "An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English," Percept. Psychophys. **18**, 331–340.

Norris, D., McQueen, J. M., and Cutler, A. (**2003**). "Perceptual learning in speech," Cogn. Psychol. **47**, 204–238.

Pallier, C., Bosch, L., and Sebastián-Gallés, N. (**1997**). "A limit on behavioral plasticity in speech perception," Cognition **64**, B9–B17.

Pallier, C., Colomé, A., and Sebastián-Gallés, N. (**2001**). "The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries," Psychol. Sci. **12**, 445–449.

Polivanov, E. (**1931**). "La perception des sons d'une langue étrangère (Perception of sounds of a foreign language)," Trav. Cercle Linguist. Prague **4**, 79–96.

Sebastián-Gallés, N., and Soto-Faraco, S. (**1999**). "Online processing of native and non-native phonemic contrasts in early bilinguals," Cognition **72**, 111–123.

Stevens, K. N. (**1999**). *Acoustic Phonetics* (MIT Press, Cambridge, MA).

Strange, W. (**2011**). "Automatic selective perception (ASP) of first and second language speech: A working model," J. Phonetics **39**, 456–466.

Trubetzkoy, N. S. (**1939**). *Principles of Phonology*, translated by C. A. Baltaxe (Vandenhoek and Ruprecht, Berkeley, 1969).

Williams, L. (**1977**). "The perception of stop consonant voicing by Spanish-English bilinguals," Percept. Psychophys. **21**, 289–297.

Winkler, I., Kujala, T., Alku, P., and Näätänen, R. (**2003**). "Language context and phonetic change detection," Cogn. Brain Res. **17**, 833–844.