Numerous factors are thought to be advantageous for non-native language learning although they are typically investigated in isolation, and the interaction between them is not understood. Firstly, bilinguals are claimed to acquire a third language easier than monolinguals acquire a second. Secondly, closely related languages may be easier to learn. Thirdly, certain phonetic features could be universally more difficult to acquire. We tested these hypotheses used as explanations by having adults learn vocabularies that differentiated words using foreign phonetic contrasts. In Experiment 1, Mandarin–English bilinguals outlearned English monolinguals, and the Mandarin-like (retroflex) artificial language was better learned than the English-like (fricative voicing). In Experiment 2, bilinguals again outlearned English monolinguals for the Mandarin-like artificial language. However, only Korean–English bilinguals showed an advantage for the more difficult Korean-like (lenition) language. Bilinguals, relative to monolinguals, show a general advantage when learning ‘easy’ contrasts, but phonetic similarity to the native language is useful for learning universally ‘difficult’ contrasts.

Keywords: bilingualism, phonetic learning, cognitive advantage, language universals, language-specific experience

Successful learning depends on a complex interaction between numerous internal and external factors (Wong & Ettlinger, 2011). Internal factors refer to characteristics of the learner, such as intelligence, working memory, and motivation, whereas external factors describe the learning situation, including the material or skills to be learned, the structure of the learning task, and whether feedback is provided (Perrachione, Lee, Ha & Wong, 2011). Language is one of the most complex skills that humans are capable of learning because it involves the mastery of numerous aspects including syntax, grammar, vocabulary, semantics, pragmatics, and phonetics. The focus of the present investigation is on the learning of non-native phonetic distinctions that are used to signal word meaning. Although learning to successfully differentiate the sounds of a non-native language can be challenging, several factors have been identified that are thought to be advantageous for non-native language learning. Firstly, it has been claimed that bilinguals acquire a third language (L3) easier than monolinguals acquire a second
language (L2) (Abu-Rabia & Sanitsky, 2010). Secondly, it is often suggested that closely related languages might be easier to learn than others, possibly due to the effects of native language (L1) transfer (Lado, 1957; Major, 2008; Odlin, 1989; Weinreich, 1953) or cross-language phonetic similarity (Best & Tyler, 2007; Flege, 1995). Thirdly, certain linguistic features may be universally easier to acquire than others (Eckman, 2008). In this study, we test these possibilities and the potential interaction between them as a first step to determining whether it is factors that are internal to the learner or the phonetic features of the language (external factors) that matter most in successful phonetic learning. The ultimate goal of this line of research is to explain the cognitive processes underlying subsequent language acquisition by establishing a rank ordering of the factors that determine learning success, and thus contribute to the larger understanding of human cognition. The series of experiments described here serves as a starting point in this quest.

A commonly accepted view, with some support from the scientific literature is that bilingual individuals learn subsequent languages easier than monolinguals (e.g., Albert & Obler, 1978; Jacobsen & Inhoof, 1974; Lerea & Kohut, 1961; Odlin, 1989). One potential explanation for this superior learning ability is the cognitive advantage associated with bilingualism. In their seminal study, Peal and Lambert (1962) demonstrated that bilingual children scored higher on verbal and nonverbal tests of cognitive ability compared to monolinguals. More recent studies have gone on to link bilingualism with improvements in divergent thinking (Baker, 2001; Ricciardelli, 1992), increased metalinguistic awareness (Ben-Zeev, 1977; Bialystok, 1991, 2001; Cummins, 1978; Ianco-Worrall, 1972; Jessner, 1999; Thomas, 1992), and better inhibitory control (Bialystok, Majumder & Martin, 2003; Bialystok & Viswanathan, 2009; Bruck & Genesee, 1995; Galambos & Goldin-Meadow, 1990; Green, 1998; Michael & Gollan, 2005). These cognitive advantages are thought to drive the additive effect of bilingualism on L3 acquisition (for a review see Cenoz, 2003). A number of incidental studies in formal educational settings have found that bilingualism does indeed have a positive effect on subsequent language learning (Cenoz & Valencia, 1994; Clyne, Hunt & Isaakidis, 2004; Keshavarz & Astaneh, 2004; Swain, Lapkin, Rowen & Hart, 1990; Thomas, 1988), even when demographic and socioeconomic variables and education were controlled for (Sanz, 2000).

Recently, some authors have questioned the reliability of the reported cognitive effects of bilingualism, particularly concerning which specific cognitive processes generate these advantages (Duñabeitia, Hernández, Antón, Macizo, Estévez, Fuentes & Carreiras, 2014; Paap & Greenberg, 2013). Several studies have also failed to find a bilingual advantage in L3 learning (Jaspaert & Lemmens, 1990; Sanders & Meijers, 1995; Schoonen, van Gelderen, de Groot, Hulstijn, Simis, Snellings & Stevenson, 2003; van Gelderen, Schoonen, de Groot, Hulstijn, Snellings, Simis & Stevenson, 2003). The findings of these studies suggest that the bilingual advantage that has been reported elsewhere may not apply to all bilinguals (learner-internal characteristics), or all language pairings (learner-external characteristics).

The specific mechanisms of how being bilingual might aid subsequent language learning are poorly understood. Bartolotti and Marian (2012) propose that one mechanism might be that bilinguals manage cross-language interference more effectively than monolinguals, and this may facilitate access to a newly learned language. Consistent with this view, Kaushanskaya and Marian (2009a) found that bilinguals have an advantage in vocabulary-learning of unfamiliar non-native words, and in a follow-up study demonstrated that this persists regardless of phonetic similarity to their known languages (Kaushanskaya & Marian, 2009b).

Phonetic learning differs greatly from other aspects of language learning, such as the learning of vocabulary, because even after years of continued L2 use and having mastered some aspects of the L2, many speakers have trouble discerning non-native phonetic contrasts (Palier, Colombé & Sébastián-Gallés, 2001). As such, phonetic learning is thought to be particularly difficult for non-native learners. Studies that have investigated the effect of bilingualism on phonetic discrimination (which is likely to be related to initial learning success, see Best & Tyler, 2007) have yielded mixed results. On one hand, there is evidence suggesting that bilinguals have an advantage when discriminating non-native Russian contrasts (Rabinovitch & Parver, 1966), Japanese single versus geminate stops (e.g., iken vs. ikken) (Enomoto, 1994), and the production of non-native initial phoneme sequences (Cohen, Tucker & Lambert, 1967). There is also evidence that the bilingual advantage for phonetic discrimination persists irrespective of degree of bilingualism. Gallardo del Puerto (2007) did not observe any difference between more or less balanced bilinguals on a vocabulary learning task although monolinguals were not examined. On the other hand, there are conflicting reports that bilinguals do not show an advantage relative to monolinguals in the phonetic discrimination of non-native phoneme sequences (Davine, Tucker & Lambert, 1971; Gonzalez-Ardeo, 2001; Lambert & MacNamara, 1969; Werker, 1986). In summary, the evidence for a bilingual advantage in phonetic discrimination is mixed. A possible explanation for these conflicting findings may be that the phonetic similarity between the subjects’ languages and the target language may attenuate or enhance a more general bilingual advantage. Consequently, the implications of bilingualism for phonetic learning are not yet understood and warrant systematic investigation with
a broad range of learner-internal and -external factors considered.

Apart from a general bilingual cognitive advantage to learn subsequent languages, it is often suggested that closely related languages are easier to learn than others. The locations of phonetic boundaries differ across languages and exposure to a specific language sharpens perceptual sensitivity near phonetic boundaries in that language, and results in some features being weighted optimally (Chandrasekaran, Sampath & Wong, 2010; Holt & Lotto, 2006; Iverson, Hazan & Bannister, 2005). A large body of research has demonstrated that the native language has a profound influence on non-native perception, that is, non-native sounds are perceived in relation to native categories, termed assimilation. Discrimination of contrasts that do not have phonemic status in the native language may vary based on acoustic-phonetic factors (Polka, 1991), as well as the existing contrasts and phonemes in the native language (Ettlinger & Johnson, 2009). Specific predictions concerning phonetic learning come from models of non-native speech perception that describe how past language experience shapes perception of novel speech sounds. The leading models of speech perception—such as the Perceptual Assimilation Model (Best, 1995), the Speech Learning Model (Flege, 1995), and the Native Language Magnet model (Kuhl, 1993)—agree that the native language has a profound effect on perception of non-native languages. The Perceptual Assimilation Model in particular makes a range of predictions of discrimination success in naïve monolingual listeners on the basis of category assimilation and phonetic goodness-of-fit. When two non-native phones are assimilated as equally good (or poor) exemplars of a single native category, discrimination will be poor, whereas when two non-native phones are assimilated to separate native categories, discrimination will be excellent (for a full list of assimilation patterns see Best, 1995). Similarly, the Speech Learning Model predicts that non-native phones are perceived relative to native phonetic categories, and this cross-linguistic similarity will determine whether new categories are established. Although they do not address subsequent phonetic learning in a target L3 by bilinguals, it may be possible to adapt principles of these models here.

One way of extending these models to L3 phonetic learning is by positing that bilingualism and phonetic similarity might interact in phonetic learning. Indeed, there is evidence that typological distance, especially when comparing L1 and L2 to L3, has been found to affect L3 acquisition (Cenoz, 2001; Hammarberg, 1997, 2001). Learning a typologically close language may allow the learner to benefit from their existing L1 (and in the case of bilinguals, also L2) knowledge and facilitate learning. Bilinguals who speak a language typologically similar to the target language tend to achieve a significantly better acquisition of the L3 than bilinguals who do not have a language typologically close to the L3 in their linguistic background (Balke-Aurell & Linblad, 1982; Swain et al., 1990). From the results of these studies, it stands to reason that whether bilinguals will exhibit a phonetic learning advantage over monolinguals may be determined by the phonetic similarity of the target language to their native languages.

Differing levels of phonetic discrimination success have also been observed in bilingual speakers from the same population when presented with different languages. L2-dominant Greek-English early bilinguals (exposed to Greek from birth, but are dominant in their L2 English which they acquired at 3–4 years of age) discriminated native Greek stop-voicing contrasts less well than did Greek monolinguals, but were indistinguishable from English monolinguals in their discrimination of English (L2) contrasts (Antoniou, Tyler & Best, 2012). In a follow-up investigation, Greek-English bilinguals from the same population discriminated non-native Ma’di laryngeal stop contrasts at a level intermediate to English and Greek monolinguals (Antoniou, Best & Tyler, 2013). These differing patterns of results across languages suggest that phonetic discrimination (and potentially phonetic learning) may rely on a bilingual’s prior phonetic experience.

In addition to the bilingual cognitive advantage and the phonetic similarity between the native and target languages that are thought to contribute to subsequent learning success, there is also evidence that some language features are universally easier to learn than others, regardless of the learner’s linguistic background. Jakobson (1968) proposed a theory of universals in phonetic development and that every natural language contains ‘unmarked’ phonetic contrasts (sound patterns based on only a few relatively easy phonetic features). Each successive stage of phonetic development will follow a universal markedness hierarchy, with the most complex sounds (which are relatively rare across natural languages) coming in last. Research on the acquisition of spoken language in children has demonstrated that some speech sounds are produced earlier than others (Grunwell, 1981; Sander, 1972; Smit, Hand, Freilinger, Bernthal & Bird, 1990). For instance, English-speaking children first produce the consonants /p/, /m/, /b/, /n/, /w/ and /b/, whereas /dʒ/, /w/, /θ/, /ð/, /ʃ/ are among the last to be mastered (Rutherford, 1982). The implication is that certain articulatory sequences are easier to coordinate than others. The frequency with which speech sounds occur across the phonemic inventories of the world’s languages is thought to reflect this articulatory difficulty (e.g., voiced bilabial stops are more common than voiced velar stops due to the aerodynamics required to sustain voicing). This position is consistent with the argument that cognitive biases have favored the acquisition of phonetically natural
regularities (Jusczyk, 1998; Wilson, 2006). Similarly, sound changes that take place over long timeframes have also been argued to favor phonetically natural regularities over unnatural ones (Blevins, 2004; Ohala, 1992). Artificial language learning paradigms provide converging experimental support for the notion that phonetic naturalness affects the learning of phonological rules (Peperkamp & Dupoux, 2007; Peperkamp, Le Calvez, Nadal & Dupoux, 2006). In regards to phonetic learning, Eckman (1977, 1985) proposed the Markedness Differential Hypothesis which states that some phonetic features are relatively more difficult to learn than others, and this may interact with the learners’ native language to determine L2 acquisition success. Here, we extend this hypothesis to bilinguals learning an L3.

The present study sought to systematically test the questions of whether bilingualism, cross-language phonetic similarity, and universal phonetic difficulty interact in phonetic learning (rather than simply auditory discrimination) by testing novel word learning in bilinguals and monolinguals. Monolinguals and bilinguals were tested in their ability to learn novel words from several artificial languages that were comprised of a small set of words that differed on a single, crucial phonetic feature.

Based on the review of the literature thus far, several hypotheses were proposed to account for the bilingual advantage in phonetic learning. First, the bilingual cognitive hypothesis posits that simply being bilingual facilitates L3 acquisition due to general cognitive advantages that are associated with bilingualism, such as improved working memory and executive functioning, consistent with the studies of Kaushanskaya and Marian (2009a, 2009b). These cognitive advantages include enhanced inhibitory control and selective attention in nonverbal conflict resolution tasks, such as the Stroop and Simon tasks (Bialystok & Craik, 2010; Bialystok & Viswanathan, 2009; Carlson & Meltzoff, 2008). Bilinguals may have an increased ability to focus on relevant stimuli (in the case of phonetic learning, the relevant phonetic features) while ignoring conflicting information from their L1 or L2, which may ultimately facilitate L3 acquisition (consistent with Bartolotti & Marian, 2012). From this bilingual cognitive perspective, a bilingual’s specific linguistic background is not of crucial importance, but rather, the experience of having acquired multiple languages should provide an advantage for subsequent language acquisition. For example, Spanish-English and Mandarin–English bilinguals would both be expected to show an advantage in phonetic learning regardless of which phonetic feature is to be learned compared to monolinguals.

Second, the language specific hypothesis states that knowledge of a wider range of phonetic and linguistic features facilitates the learning of new, but related, features. According to this view, it should be easier to learn a language whose phonetic features are close to those of the learner’s native language. For example, the stop voicing contrast /b/-/p/ is realized in both French and Italian as a voicing lead versus short-lag voice onset time distinction. A likely and straightforward prediction by the leading models of speech perception (Best, 1995; Flege, 1995; Kuhl, 1993) would be that knowledge of Italian would lead to accurate perception and production of the similarly realized French /b/-/p/ contrast. However, the predictions put forth by the language specific hypothesis go beyond such simple cases of perceptual assimilation. According to the language specific hypothesis, language experience is thought to facilitate the learning of new but phonetically related features. Language experience with a given phonetic feature will aid learning of a common feature in a foreign language, even if realized at a different place of articulation or with a different manner. For example, English uses voicing as a contrastive feature for labiodental fricatives /f/-/v/ (e.g., fan vs. van), but does not contain bilabial fricatives /ɸ/-/β/. Nevertheless, according to the language specific account, due to their experience with a fricative voicing distinction (albeit in a different place of articulation) English listeners (be they monolingual or bilingual) should learn to distinguish the bilabial fricative voicing distinction better than speakers of a language that does not use fricative voicing contrastively (e.g., Mandarin). In contrast, Mandarin uses the retroflex tongue posture as a contrastive feature (e.g., /s/-/ʃ/) whereas English does not. Thus, it follows that a Mandarin speaker may have an advantage in learning a novel language that also uses retroflex postures contrastively. The central tenet of the language specific account is that past language experience, rather than bilingualism, is the most important factor in phonetic learning.

A possible alternative to the language specific hypothesis is the language universal hypothesis which specifies that certain phonetic features are easier to learn than others, independent of language experience, perhaps due to a ranking in terms of perceptual salience, articulatory complexity, or markedness. According to the language universal hypothesis, all learners should find the ‘marked’ contrasts the most difficult to learn, regardless of language experience.

These three hypotheses were tested in a series of two experiments. In both experiments, adult subjects learned a vocabulary that required use of foreign phonetic contrasts to signal word meaning. Successful learning of the vocabulary necessarily entailed learning to use a critical phonetic dimension to tell apart the minimal pair words. In Experiment 1, English monolinguals and Mandarin–English bilinguals were compared in their ability to learn English-like (fricative voicing) and Mandarin-like (retroflex consonants) phonetic contrasts. In Experiment
2, we compared English monolinguals and Mandarin–English and Korean–English bilinguals on Mandarin-like (retroflex) and Korean-like (lenition) phonetic contrasts. Specific predictions are provided in the prelude to each experiment.

**Experiment 1: English- and Mandarin-like artificial languages**

i. If the bilingual cognitive hypothesis is correct, we would expect the Mandarin–English bilinguals to outperform the English monolinguals on both the English-like and Mandarin-like artificial languages.

ii. If the language specific hypothesis is correct, we would expect the Mandarin–English bilinguals to show an advantage for the Mandarin-like artificial language only.

iii. Another possibility is that there might be some interaction between these two hypotheses, where the bilinguals show an overall phonetic learning advantage but it is exaggerated for the Mandarin-like artificial language.

iv. If the language universal hypothesis is correct, both groups will find the same phonetic contrast more difficult to learn.

**Method**

**Participants**

Experiment 1 involved 12 English monolinguals ($M_{age} = 21$, age range = 19–23, 9 females, 3 males) and 12 Mandarin–English bilinguals ($M_{age} = 21$, age range = 20–23, 9 females, 3 males). All subjects were students at Northwestern University, had been born in the United States, and completed a detailed demographics questionnaire to ensure that the groups were matched for age, sex, and musical training (English monolinguals: onset = 7.9, experience = 7.4; Mandarin–English bilinguals: onset = 8.0, experience = 8.1). To be included in the monolingual group, participants had to be proficient in only English, although they may have had some foreign language experience, but that language was not used on a daily basis (less than 3 years of foreign language experience, usually in high school). Importantly, none of the monolingual subjects had any prior Mandarin experience (four had studied Spanish and two had studied French in high school). The bilinguals had been exposed to both Mandarin and English from birth, and were not proficient in an L3 (one subject reported having studied Spanish). All Mandarin–English subjects spoke a dialect of Mandarin in which retroflexion is a contrastive feature. None of the subjects reported any history of audiologic or neurological deficits. All passed a pure tone screening at 25 dB at 500, 1,000, 2,000 and 4,000 Hz.

**Stimuli**

Two artificial languages served as stimuli for this experiment. A female native speaker of Gujarati produced the Mandarin-like stimuli, and a phonetically-trained male native speaker of English produced the English-like stimuli. Each language was comprised of four minimal pairs that differed on a single, critical phonetic feature (eight words per language). All of the words were monosyllabic /CV/ tokens and ended with the vowels /e/, /i/, /o/, and /u/. One of the artificial languages was Mandarin-like in that the retroflex versus non-retroflex contrast (/t/-/t/) differentiated the /CV/ monosyllabic words, which is a native contrast in Gujarati. The other was English-like in that fricative voicing was used to differentiate the /CV/ monosyllabic words (/f/-/f/). Crucially, none of these phonemes are actually present in Mandarin and English, respectively, meaning that all subjects were required to learn new phonemes.

The stimulus recordings were conducted inside a sound attenuated booth using a Shure SM58 cardioid vocal microphone attached to a desktop boom stand. The talker was positioned approximately 10 cm from the microphone. Speech was digitized to computer with a sampling rate of 44.1 KHz and a 16-bit sampling depth. The intensity of all stimuli was normalized.

**Procedure**

The experiment began with a passive exposure phase, during which a picture and a sound were paired (8 words × 12 repetitions = 96 trials total). Words were presented at the rate of every 3.5 s, and thus the exposure phase lasted for approximately 6 minutes. No response was required from the subject to advance to the next word. This was then immediately followed by a test phase (8 words × 8 repetitions = 64 trials total), during which each word was auditorily presented and all eight pictures were displayed on screen. The participant was required to match the heard word to the correct picture by pressing keys 1–8 on the keyboard. The test phase was self-paced. The experiment was presented using Sennheiser HD 280 Pro headphones and a computer running E-Prime software. Stimulus output level was calibrated to 72 dB SPL. The order in which the two artificial languages were learned was counterbalanced across subjects.

**Results**

Figure 1 depicts the English monolinguals’ and the Mandarin–English bilinguals’ accuracy in language-learning. Inspection of the bars shows that both groups...
were well above chance for both languages (both over 12.5%). It also appears that the Mandarin–English bilinguals show an advantage relative to the monolinguals.

Assumptions of Analysis of Variance (ANOVA) were satisfactory, and a mixed factorial $2 \times 2$ ANOVA was conducted with a between-subjects factor of group (monolingual vs. bilingual) and a within-subjects factor of artificial language (Mandarin-like vs. English-like). A significant main effect of group revealed that the Mandarin bilinguals performed better than the English monolinguals overall across both artificial languages, $F(1, 22) = 5.04, p = .035, \eta_p^2 = .187$. There was also a significant main effect of artificial language, $F(1, 22) = 5.04, p = .035, \eta_p^2 = .186$, indicating that both groups learned the Mandarin-like artificial language better than the English-like artificial language. There was no significant interaction.

Discussion

The results of Experiment 1 suggest that bilinguals learn phonetic distinctions better than monolinguals, irrespective of phonetic similarity to the native language(s), lending support to the bilingual cognitive hypothesis. The Mandarin–English bilinguals showed an advantage in learning both the Mandarin-like and English-like artificial languages. These results are consistent with past work that has demonstrated that bilingualism aids the acquisition of non-native language features, possibly because of cognitive advantages that have been attributed to bilingualism (Bartolotti & Marian, 2012; Kaushansky & Marian, 2009a, 2009b).

Support was also found for the language universal hypothesis, which specifies that some phonetic features might be easier to learn than others, irrespective of the learner’s language background. Both groups of learners performed better on the Mandarin-like artificial language than on the English-like. Possible reasons for this difference in learning difficulty might be that the Mandarin-like fricative retroflex contrast is more acoustically salient, marked, or gesturally distinct than the English-like bilabial fricative voicing contrast. At this point, such underlying explanations are speculative, and more phonetic contrasts and listeners from differing language backgrounds need to be investigated before the influence of a universal phonetic difficulty is more fully understood.

Although the results of Experiment 1 provide a valuable starting point, they are limited because only one bilingual group was tested, and only two languages were examined. This makes it difficult to determine if bilingualism or the phonetic feature is the most important factor in determining phonetic learning success. Now that bilingual, as well as language-based, differences have been found in phonetic learning, the process of systematic investigation of these two factors may continue by comparing the monolinguals to an additional group of bilinguals who speak a language other than Mandarin. Therefore, an experiment was conducted in which monolinguals were compared to not one, but two groups of bilinguals, in their ability to learn phonetic features in two artificial languages.

Experiment 2: Korean- and Mandarin-like artificial languages

Experiment 2 examined whether the bilingual advantage for phonetic learning observed in Experiment 1 is due to a general bilingual cognitive advantage. As in Experiment 1, English monolinguals were compared to Mandarin–English bilinguals, but in order to evaluate the claim that bilinguals possess a general advantage in phonetic learning, a second group of bilinguals (i.e., Korean–English) was also recruited. These three groups of subjects were asked to learn two artificial languages: the Mandarin-like artificial language from Experiment 1, and a new Korean-like artificial language. Korean stops are typologically unusual in that the three-way contrast uses the same airstream mechanism (i.e., pulmonic egressive).
and the three different categories are characterized as lenis, fortis, and aspirated (Cho, Jun & Ladefoged, 2002). The Korean-like artificial language contained a lenition contrast similar to that found in Korean stops, but was realized in voiceless interdental fricatives. The crossed design of Experiment 2 provides a critical test and allows for the examination of whether bilinguals possess an overall advantage (consistent with the bilingual cognitive hypothesis) or if the bilingual advantage is greatest for phonetically similar languages (i.e., whether the bilingual cognitive and language specific hypotheses interact).

i. If the bilingual cognitive hypothesis is correct, we would expect both the Mandarin–English and Korean–English bilinguals to outperform the English monolinguals on both artificial languages.

ii. If the language-specific hypothesis is correct, we would expect the Mandarin–English bilinguals to show an advantage relative to monolinguals for the Mandarin-like artificial language only, and for the Korean–English bilinguals to show an advantage for the Korean-like artificial language only. This seems unlikely because the Mandarin–English bilinguals outperformed the English monolinguals on both artificial languages in Experiment 1.

iii. A third possibility is that these two possibilities may interact, where the bilinguals have an overall advantage but it is exaggerated for the artificial language that is phonetically similar to their L1. If this interaction hypothesis is correct, both groups of bilinguals should outperform the monolinguals, and the Mandarin–English bilinguals should possess an exaggerated advantage for the Mandarin-like artificial language and Korean–English bilinguals for the Korean-like artificial language.

iv. Finally, if the language universal hypothesis is correct, all groups should find the same phonetic features easier to learn than others, independent of language experience.

Method

Participants

Three groups of participants were recruited for Experiment 2. A new group of 24 English monolinguals (M_{age} = 20.4, age range = 18–24, 14 females, 10 males), a new group of 23 Mandarin–English bilinguals (M_{age} = 21.5, age range = 18–28, 15 females, 8 males), and a third group of 21 Korean–English bilinguals (M_{age} = 19.9, age range = 18–24, 15f, 6m). All subjects had been born in the United States and were students at Northwestern University. Attempts were made to match the groups in terms of their ages, sex and musical training (English monolinguals: onset = 7.8, experience = 6.4; Korean–English bilinguals: onset = 6.7, experience = 10.7; Mandarin–English bilinguals: onset = 7.0, experience = 6.9) although the Korean–English bilingual group possessed slightly greater musical experience than the other groups.\(^1\) English monolingual subjects did not possess any Mandarin or Korean language experience (eight had studied Spanish, three had studied French, and two had studied German in high school). The bilinguals had been exposed to both of their languages from birth (i.e., Mandarin and English or Korean and English) and were not fluent in an L3 (two Mandarin–English bilingual subjects had studied Italian or Spanish, and five Korean–English bilingual subjects had studied Spanish or French in high school). All Mandarin–English subjects spoke a dialect of Mandarin in which retroflexion is present, and all Korean–English subjects spoke a Korean dialect in which lenition is used contrastively. None of the subjects reported any history of audiologic or neurological deficits. All passed a pure tone screening at 25 dB at 500, 1,000, 2,000 and 4,000 Hz.

Stimuli

Two artificial languages were presented in Experiment 2. All of the words were monosyllabic /CV/ tokens. One of the artificial languages was the Mandarin-like language used in Experiment 1 that employed a retroflex versus non-retroflex contrast to differentiate the /CV/ monosyllabic words. The other artificial language was Korean-like in that a lenition contrast (\(/θ/-/θ\)/) was used to differentiate the /CV/ monosyllabic words, but was realized in interdental fricatives (and not in stops as in Korean). A female native speaker of Korean with phonetic training produced the Korean-like stimuli. As in Experiment 1, each language contained four /CV/ minimal pairs that differed on a single, critical phonetic feature (2 consonants × 4 vowels = 8 words per language). The stimulus recording procedure was identical to that employed for the artificial languages in Experiment 1.

Procedure

The artificial language learning procedure was identical to that employed in Experiment 1.

Results

Figure 2 displays the learning success for the Mandarin-like and Korean-like artificial languages by the English monolinguals and the Mandarin–English and Korean–English bilinguals. Inspection of the bars seems to indicate

\(^1\) This slight difference in musical experience did not make a unique contribution to the results and thus will not be discussed further.
Figure 2. English monolinguals’ and Mandarin–English and Korean–English bilinguals’ accuracy in learning Mandarin-like and Korean-like artificial languages.

that the Korean–English bilinguals outperformed the other 2 groups. It also appears that the Korean-like artificial language was more difficult to learn for all three groups.

ANOVA assumptions were satisfactory and a $3 \times (2)$ ANOVA was conducted with a between-subjects factor of group (English monolinguals vs. Mandarin–English vs. Korean–English bilinguals) and a within-subjects factor of artificial language (Mandarin-like vs. Korean-like). As in Experiment 1, there was a significant main effect of group, $F(2, 65) = 12.56$, $p < .001$, $\eta_p^2 = .279$. Post-hoc Sidak tests revealed that, overall, both the Korean–English bilinguals, $p < .001$ (mean difference = 18.5%), and Mandarin–English bilinguals, $p = .011$ (mean difference = 11.0%), did better than the English monolinguals, and the two bilingual groups did not differ, $p = .146$. There was also a main effect of artificial language, $F(1, 65) = 435.82$, $p < .001$, $\eta_p^2 = .870$, driven by the fact that all groups performed better on the Mandarin-like artificial language than on the Korean-like.

Most importantly, there was a significant Group × Language interaction, $F(2, 65) = 3.64$, $p = .032$, $\eta_p^2 = .101$. Posthoc tests revealed that both the Mandarin–English, $F(1, 70) = 7.04$, $p = .010$, and Korean–English bilinguals outperformed the English monolinguals on the Mandarin-like artificial language, $F(1, 70) = 12.67$, $p = .001$, and the two bilingual groups did not differ, $F < 1$. For the Korean-like artificial language, the Korean–English bilinguals outperformed both the English monolinguals, $F(1, 70) = 12.59$, $p = .001$, and the Mandarin–English bilinguals, $F(1, 70) = 7.16$, $p = .009$. The English monolinguals ($M = 38\%$) and Mandarin–English bilinguals ($M = 40\%$) did not differ on the Korean-like artificial language, $F < 1$.

Discussion

The results of Experiment 2 suggest that bilinguals may have an advantage over monolinguals in subsequent phonetic learning. Both bilingual groups showed an advantage over the English monolinguals for the Mandarin-like retroflex contrast, similar to what was observed in Experiment 1; however, only the Korean–English bilinguals (not the Mandarin–English bilinguals) showed a learning advantage for the Korean-like lenition contrast. This may be interpreted as lending partial support to the bilingual cognitive hypothesis but also confirms that, as shown in previous studies, the bilingual advantage may not be sufficient to boost learning of all foreign phonetic contrasts.

The findings are also consistent with the language universal hypothesis because all groups performed better on the Mandarin-like than the Korean-like artificial language. All groups exhibited poorer learning of the Korean-like lenition contrast, suggesting that the Korean-like interdental fricative lenition contrast is more difficult to acoustically differentiate, and consequently learn, than the Mandarin-like retroflex contrast. Crucially, although the Korean-like artificial language was the most difficult to learn, the Korean–English bilinguals outperformed the other groups on this language, suggesting that the bilingual cognitive and language specific hypotheses may interact to determine language learning success.

General Discussion

The findings from this series of experiments suggest that bilinguals, relative to monolinguals, show a general advantage of learning foreign phonetic contrasts, but importantly, this is modulated first by whether the contrast in question is ‘easy’ or ‘difficult’, and then by how similar these contrasts are to those in the learners’ native language(s). This provides partial support for the bilingual cognitive hypothesis that bilinguals possess a general advantage in subsequent phonetic learning, potentially due to cognitive advantages that are associated with bilingualism (Kaushanskaya & Marian, 2009a, 2009b; Tremblay & Sabourin, 2012). However, there are two important caveats, namely that the bilingual advantage was not equal across all three artificial languages, and phonetic similarity to the native language aided learning
of the difficult Korean-like lenition contrast. Thus, partial support has been observed for all three hypotheses.

The major finding from the present work is that an interaction between the general bilingual advantage for phonetic learning (learner-internal factor), the universal difficulty of some contrasts over others (learner-external factor), and the advantage for learning contrasts that are phonetically similar to those of the native language (learner-external factor) has been identified. Specifically, the data suggest that this interaction adheres to a particular rank ordering, such that bilinguals enjoy an advantage over monolinguals in phonetic learning, and beyond this bilingual advantage, certain contrasts are easier for all learners to acquire, and, finally, for difficult contrasts, language specific experience may be called upon to aid learning. This assertion is based on our observations that bilinguals show a phonetic learning advantage over monolinguals for contrasts that are universally ‘easier’ (e.g., Mandarin-like retroflex), but language-specific experience may boost learning of universally ‘difficult’ contrasts (e.g., Korean-like lenition). This explanation may illuminate why past investigations of the bilingual advantage in phonetic discrimination have sometimes yielded contradictory results. On one hand, bilinguals have shown an advantage in discriminating non-native speech contrasts (Enomoto, 1994; Rabinovitch & Parver, 1966), but in other studies bilinguals have not outperformed monolinguals (Davine et al., 1971; Gonzalez-Ardeo, 2001; Werker, 1986). Our findings suggest that the bilingual advantage for subsequent language learning reported in past research (Cenoz & Valencia, 1994; Clyne et al., 2004; Keshavarz & Astaneh, 2004; Sanz, 2000; Swain et al., 1990; Thomas, 1988) interacts with universal difficulty, as well as similarity to the native language.

To our knowledge, only one study has investigated the phonetic learning (rather than simply discrimination) abilities of bi- and multilinguals relative to monolinguals. Tremblay and Sabourin (2012) trained English monolinguals, English–French bilinguals, and multilinguals (exposed to English and French since birth, and at least one additional language) to discriminate the non-native Hindi dental-retroflex contrast. After three days of training, the multilinguals outperformed the monolinguals at post-test, and both the bilinguals and multilinguals outperformed a monolingual control group that received no training (whereas the trained monolinguals did not differ from this control group). This finding suggests that bi- and multilinguals possess superior phonetic learning abilities. However, a limitation of this study is that only one phonetic dimension from a single language was investigated, and thus it is not clear if the bi- and multilingual advantage is modulated by, or is robust across, other phonetic contrasts. The present study has addressed these shortcomings.

The observation that certain phonetic features are easier to learn than others lends some support to the language universal hypothesis. The better learning of Mandarin-like retroflex consonants by all learner groups suggests that this phonetic contrast appears to be the easiest to learn irrespective of linguistic experience. In contrast, the Korean-like lenition contrast was the most difficult to learn for all groups. Evidence for the relative universal discriminability of these contrasts can be informed by their relative frequency among the languages of the world. Retroflexion occurs in 20.1% of the world’s languages (Maddieson, 2013a) and fricative voicing is present in 34.6% of languages (Maddieson, 2013b). This is juxtaposed with the presence of fortis consonants in only 1.4% of the world’s languages (Maddieson, 1984). This suggests that fricative voicing and retroflexion are readily perceptible whereas the lenition contrast is objectively more difficult to perceive than others. Thus, these findings are consistent with the notion of a universal component of perceptibility because all groups performed better on the Mandarin-like than the Korean-like artificial language. All groups exhibited poorer learning of the Korean-like lenition contrast, suggesting that the Korean-like interdental fricative lenition contrast is more difficult to acoustically differentiate, and consequently learn, than the Mandarin-like retroflex contrast. There are a number of possible ways to account for these observations: it might be the case that certain phonetic features are perceptually more salient than others, or that certain physical/acoustic (Chang, Plauché & Ohala, 2001) or articulatory configurations (Browman & Goldstein, 1992) are easier to produce or perceive, or perhaps the observed differences in learning reflect markedness effects and ranking of phonetic features (Hayes & Steriade, 2004). These explanations are not necessarily mutually exclusive or incompatible.

The observation that language-specific experience aided the learning of the Korean-like lenition distinction is consistent with past research from the cross-language speech perception literature. It is well-established that the L1 exerts a persistent influence on non-native phonetic perception, extending well into adulthood. Models of speech perception agree that non-native sounds are perceived in relation to native phonetic categories, a process termed perceptual assimilation (Best, 1995; Flege, 1995). The contribution of the present series of experiments is to specify that native phonetic similarity may boost learning for non-native contrasts that are ‘difficult’, making a unique contribution to learning outcomes after bilingualism and universal explanations (e.g., markedness, articulatory difficulty, or perceptual salience) have been taken into account.

The present findings lay the ground for an exciting and fruitful line of research. Future studies should systematically explore which phonetic features are easier
to learn. For instance, it may be possible to base predictions on linguistic (Ettlinger & Johnson, 2009; Jakobson, 1968; Jusczyk, 1998; Ohala, 1992; Wilson, 2006) and perceptual (Best, 1995; Flege, 1995; Kuhl, 1993) theories, and develop a rank ordering of learnability of phonetic features. In addition, it will be necessary to compare bilinguals from diverse language backgrounds so as to more stringently test the bilingual cognitive, language universal, and language specific hypotheses. In the present series of experiments, we did not conduct cognitive assessments, and thus recommend that future work assesses aspects of cognition using standardized tests in order to pinpoint the cognitive factors that may contribute to the bilingual cognitive advantage (e.g., executive function, and specifically inhibitory control). 

Another useful avenue for research may be to compare the phonetic learning performance of bilinguals differing in their order of L2 acquisition (e.g., simultaneous, early and late bilinguals). It may also be informative to compare bi- and multilinguals to determine if knowledge of multiple languages is positively correlated with phonetic learning success (consistent with Tremblay & Sabourin, 2012). Ultimately, it might be possible to address the question of whether it is the phonetic features themselves or factors internal to the learner that matter most in successful language learning.

In conclusion, bilinguals possess an advantage over monolinguals when it comes to phonetic learning. However, this is modulated by the universal difficulty of the phonetic contrast in question as well as the phonetic similarity between the target language and the learners’ native language. Theories of foreign phonetic learning need to take into account the interaction and rank ordering of these internal and external factors that have been demonstrated to determine learning success.

References


