

Age of acquisition and proficiency in a second language independently influence the perception of non-native speech*

PILAR ARCHILA-SUERTE

University of Houston

JASON ZEVIN

Sackler Institute for Developmental

Psychobiology—Weill Medical College of Cornell

University

FERENC BUNTA

University of Houston

ARTURO E. HERNANDEZ

University of Houston

(Received: June 8, 2010; final revision received: January 25, 2011; accepted: February 15, 2011; First published online 7 September 2011)

Sensorimotor processing in children and higher-cognitive processing in adults could determine how non-native phonemes are acquired. This study investigates how age-of-acquisition (AOA) and proficiency-level (PL) predict native-like perception of statistically dissociated L2 categories, i.e., within-category and between-category. In a similarity task, participants rated the level of similarity between pairs of English syllables from 1 (similar) to 4 (dissimilar). Early L2 acquisition predicts accurate within-categorization and high proficiency in late L2 acquisition predicts improved between-categorization. Our results suggest that the manner in which bilinguals learn to categorize non-native sounds depends on the cognitive processes available at the age of L2 exposure.

Keywords: bilingualism, speech, age-of-acquisition, proficiency, sensorimotor

Introduction

The perceptual accuracy of non-native phonemes may rest on the learning mechanisms used to process speech. An account of general skill learning, the sensorimotor hypothesis (Hernandez & Li, 2007), highlights that the level of performance in a given sensorimotor skill (e.g., foreign language, music, sports) depends greatly on the person's age of acquisition. Thus, an early acquired skill leads to better performance than a late acquired skill. Some learners of L2, for example, demonstrate great difficulty acquiring non-native phonemes. Typically, early L2 learners have native-like production of the second language and late L2 learners struggle with foreign accents. The sensorimotor activity of coordinating sound perception to articulation, known as the motor theory of speech perception (Liberman & Mattingly, 1985), also suggests that the manner in which L1 and L2 phonemes are learned is significantly impacted by age. However, this early–late age of acquisition (AOA) demarcation is not precise. There is a broad range of

performance in late bilinguals in the perception of non-native sounds. While it appears that natural immersion and explicit instruction aids late bilinguals in discriminating L2 contrasts (Aoyama, Flege, Guion, Akahane-Yamada & Yamada, 2004), an early acquisition of the second language does not guarantee native-like perception. Early bilinguals who rely predominantly on their L1, for example, can have detectable foreign accents in L2 (Flege, Frieda & Nozawa, 1997).

Studies in cognitive development indicate that, due to biological constraints, implicit knowledge develops before explicit knowledge. That is, implicit knowledge develops in infancy and explicit knowledge develops in the preschool years through levels of increasing explication (Clements & Perner, 1994; Karmiloff-Smith, 1991). Implicit knowledge, as an early developing process, intertwines with the sensorimotor stage in which children acquire knowledge upon direct interaction with their world (Westermann, Mareschal, Johnson, Sirois, Spratling & Thomas, 2007). In the early acquisition of native or non-native speech, infants and young children have the opportunity to imitate articulatory gestures and practice language sounds by cooing and babbling; thus, rendering the learning of language sounds implicit and procedural. On the other hand, older children and adults rely on advanced stages of cognitive development, including concrete and formal operations that intertwine with later developing mechanisms of explicit knowledge. Therefore, while younger children use

* We would like to extend our gratitude to Lee Branum-Martin for his guidance with the analysis and to our research assistant Gabriela Ochoa for her help collecting data for this project. This research was supported by R21HD059103-01 Neural correlates of lexical processing in child L2 learners and by the Institute for Biomedical Imaging Science (IBIS) for Plasticity in Speech Perception in Early Bilingual Children.

Address for correspondence:

Pilar Archila, Department of Psychology, University of Houston, 126 Heyne Bldg., Houston, Texas 77204-5022, USA
mparchil@mail.uh.edu

an implicit/procedural process to categorize L1 and L2 phonemes, older children and adult learners use explicit rules to learn L2 phonemes. Neuroimaging studies have also found that in the performance of an explicit–implicit motor sequence learning task, children recruit subcortical regions that reach maturity earlier in life (i.e., basal ganglia), whereas adults recruit motor and parietal cortical regions that reach maturity later (Gao, Parsons, Bower, Xiong, Li & Fox, 1996). The distinction made between implicit/sensorimotor and explicit/higher-cognitive processes provides us with a new way of conceptualizing how early and late bilinguals learn non-native sounds.

In the categorical perception of speech, a phoneme that physically changes with respect to one or more parameters (e.g., VOT; Voice Onset Time, the period of time between the initial release of a plosive consonant and the production of a vowel sound) along a continuum is perceptually judged to belong to one or more categories depending on how similar or dissimilar each exemplar of the phoneme sounds to another (Kent, 1997). When various phonemic exemplars are perceived as similar or identical, the person creates a perceptual category for the given phoneme (e.g., /b/). Here, we treat this acquired perceptual similarity of phonemes as *within-category*. On the other hand, when various phonemic exemplars sound noticeably different from one another, a perceptual boundary develops between categories (e.g., /b/–/p/). This acquired perceptual contrast between clusters of phonemes is treated as *between-category*.

Although the majority of studies that have investigated the dichotomy of implicit–explicit learning in L2 have focused on the acquisition of L2 grammatical structures (Green & Hecht, 1992), the overall consensus in the literature is that explicit instruction is beneficial if L2 material is salient, regularly structured, limited in its number of variables and intelligible. On the other hand, implicit learning appears to be beneficial if L2 material is random and a specific rule cannot be generated (Reber, Kassin, Lewis & Cantor, 1980). Analogous to grammatical structures, speech categories dissociated as between- and within- could be interpreted as discriminable and salient or indiscriminable and random, respectively. Therefore, by means of explicit strategies, proficient late learners may learn to discriminate novel L2 contrasts if attention to the relevant cues in the phonemic boundaries is enhanced. On the other hand, accurate categorization of within-categories may require a sensorimotor and implicit process to deal with the randomness and fine-grained arrangement of similar-sounding phonemes – in which case only monolinguals and early bilinguals would be able to learn as children.

Age of acquisition in non-native speech perception

The ability to discriminate non-native phonemes gradually diminishes in the first year of life and early childhood

(Werker & Tees, 1984). As native language input increases, children become neurally committed to the phonetic system to which they are exposed – this is known as the native language magnet (NLM; Kuhl, 2000). In time, adult monolingual listeners become mainly receptive to the boundary that demarcates the change between phonemes (between-category). In bilinguals, the rapid organization of the L1 sound system may interfere with the learning of L2 phonemes because the new L2 sounds must be filtered through the L1 (Kuhl, 2000). An implication of the NLM is that individualized training for the improvement of non-native perception in late sequential bilinguals is insufficient because the early distortions that occurred in the phonetic space as a result of L1 acquisition cannot be reversed.

Proficiency level (PL) in non-native speech perception

Some late bilinguals can improve their perception of second language sounds. For example, proficient Spanish–English bilinguals have a perceptual vowel space that resembles that of English monolinguals (Flege, Munro & Fox, 1994). Experience in L2 can lead to the reorganization of perceptual assimilation patterns (Best & Strange, 1992), especially in environments of full immersion where new memory traces and new category boundaries can be formed (Peltola, Kuntola, Tamminen, Hämäläinen & Aaltonen, 2005). The Speech Learning Model (SLM) developed by Flege (1995) proposes that the capacity to perceive and categorize new sounds remains intact throughout life. Hence, as the L2 network becomes denser with new phonetic exemplars and categories, the less filtering of L2 speech through the L1 network will occur. It is important to note that category formation may be facilitated or impaired by the amount of perceptual distance that exists between L1 and L2 sounds. A long acoustic distance between two phonemes may improve discrimination while a short distance may lead to perceptual confusability. Therefore, introducing L2 learners to natural linguistic environments that emphasize the acoustic distance between L1 and L2 can help in the formation of new L2 speech categories.

Predicted effects of AOA and PL on the perception of within- and between-categories

The present study investigates how early and late bilinguals with varying proficiency levels perceive non-native speech syllables, exemplified here as within- and between-categories. First, we expect early bilinguals to cluster similar-sounding syllables as solo within-categories, thus ignoring the acoustic irrelevancies of L2 sounds that are alike. Likewise, early bilinguals should be sensitive to the linguistic information provided by the phonemic boundary, resulting in the discrete

categorization of distinct categories. Early bilinguals, in general, should resemble monolinguals since they were exposed to the L2 before any presumed perceptual distortions to their sound map occurred. By the same token, early bilinguals should resemble monolinguals because the manner in which they learned L2 phonemes relied in sensorimotor/implicit processes. No effects of L2 proficiency are expected to affect within- or between-categorization in this cohort of bilinguals. On the other hand, we expect later bilinguals' perception to differ depending on their PL. Given the early perceptual warping that results from L1 acquisition, we expect late bilinguals to perceive similarity in prototypical exemplars that are acoustically different, and perceive dissimilarity in prototypical exemplars that are acoustically alike. In other words, the clusters of each within-category should be broad and undefined as late bilinguals take into account acoustic cues that are irrelevant to native-like categorization. However, a high PL in late bilinguals is expected to improve discrimination of between-categories. Since the contrast between syllables from different categories is more noticeable, it is possible that these exaggerated cues obtained from the phonemic boundary can be learned explicitly, even in adulthood. As it has been suggested in the literature, high proficiency – as it correlates with more L2 use and attention – helps late bilinguals develop larger L2 speech networks, which in turn helps omit the filtering of L2 sounds via L1.

Considering the stimuli under study, we think that late bilinguals will more readily discriminate the vowels /æ/ and /ɛ/ (e.g., *bat* and *bet*) regardless of their PL in English, as these sounds may be assimilated to two distinct L1 categories /a/ and /e/ (e.g., *casa* and *leche*). On the contrary, the English vowels /ɑ/ and /ʌ/ (e.g., *hot* and *hut*) may be perceived as instances of the Spanish /a/, similar to the findings of English–Italian bilinguals (Guion, Flege, Akahane-Yamada & Pruitt, 2000). To sum up, we expect early AOA to result in native-like within- and between-categorization independent of their level of proficiency, thus suggesting sensorimotor/implicit processing of L2 sounds akin to monolinguals. High-proficient late AOA should result in improved between-categorization due to increased attention to the phonemic boundaries of novel L2 phonemes.

Method

Participants

Ninety-eight college students from the University of Houston (18 male and 80 female), between the ages of 18 and 36 years, participated in this study. The control group of monolinguals was composed of 28 native English speakers whose ethnic background comprised 12 Caucasians, 1 Asian American, 1 Latino, 11 African

Americans and 3 “Other”. Belonging to a particular ethnicity did not bias perception of English syllables for within-category ($F(4,23) = .18, p = .94$) or between-category discrimination ($F(4,23) = 1.56, p = .21$). The mean age in monolinguals was 22 years ($SD = 3.2$, range 18 to 32). Bilinguals were classified as early, intermediate or late depending on the age of exposure to English. All participants in the bilingual groups ($n = 70$) were of Hispanic descent and spoke Spanish as their first language and English as their second.

Early bilinguals

These were the participants exposed to English before five years of age. The group was composed of 31 participants with a mean age of 23.7 ($SD = 4.3$, range 18 to 33). Most early bilinguals were born in the US and began learning English in kindergarten or elementary school. Eight early bilinguals immigrated to the US before the age of five. These eight bilinguals had all resided in the US for at least fifteen years at the time of testing. The average number of years of formal education in English for the early bilingual group was 16.25 ($SD = 3.02$, range 12 to 23).

Intermediate bilinguals

These were the participants exposed to English between the ages of six and nine. Sixteen participants with a mean age of 20.8 ($SD = 2.07$, range 18 to 26) composed this group. The average number of years residing in the US was 11.75 ($SD = 8.7$, range 0 to 25) and the average number of years of formal education in English was 13.12 ($SD = 4.04$, range 6 to 23). Bilinguals in this group included those who were born in the US but only received full exposure to the language in elementary school and those who moved to the US around the age period that corresponds with this AOA group.

Late bilinguals

These were the participants exposed to English after ten years of age. In this study, the “latest” AOA was twenty-four years of age. Twenty-three participants with a mean age of 25.39 ($SD = 4.78$, range 19 to 36) composed this group. The average number of years residing in the US was 10.3 ($SD = 3.52$, range 7 to 17) and the average number of years receiving a formal education in English was 9.39 ($SD = 2.93$, range 4 to 15).

The bilingual groups differed significantly in their length of US residency (LOR) ($F(2,67) = 3.11, p = .05$), the quantity of L2 education ($F(2,67) = 29.44, p < .0001$) and L2 use ($F(2,67) = 7.06, p < .0017$) (see Table 1 for detailed characteristics of all groups). All participants reported normal hearing and no history of language disorders.

Table 1. Group summary descriptive statistics.

Variables measured	Group			
	Monolingual	Early	Intermediate	Late
Chronological age	22(3.2)	23.7(4.3)	20.8(2.0)	25.3(4.7)
Age of L2 learning	–	3.6(1.5)	6.9(0.9)	14.5(3.8)
Length of residence in US	–	20.25(3.63)*	11.7(8.7)	10.3(3.5)
L2 education	–	16.25(3.0)	13.1(4.0)	9.60(3.0)
L2 proficiency	73.9(6.0)	69.7(7.1)	66.7(7.4)	64.1(7.0)
L1 proficiency	–	61.9(10.3)	66.1(5.1)	75.9(4.2)

NOTES: Means (with standard deviations in parentheses) are reported in years for the following variables: chronological age, age of L2 learning, length of residency and L2 education. * The value entered for length of residency for the early bilingual group only applies to the eight subjects who immigrated to the US before five years of age. The rest of the early bilinguals were born in the US.

Stimuli

Recordings of natural speech of the English syllables *saf*, *sef*, *sof* and *suf* were obtained from an adult male monolingual English speaker in a sound-attenuated room. The 160 unique tokens (40 tokens for each syllable) were digitized at 44,100 Hz using a Sony MZ-NH800 mini-disc recorder and a Sony ECM-CS10 stereo microphone. Once recorded, the computer program Praat was used to normalize peak amplitude and ensure audibility of all stimuli. The exemplars for each syllable type varied in intonation, timbre and duration. This manipulation of the stimuli was implemented to investigate within-categorization. All four vowels used in this task differ enough in their spectral quality to be characterized as separate categories. Previous research has indicated that English vowels in words like *bat* and *bet* can be perceived as instances of the Spanish vowels /a/ and /e/ in inexperienced Spanish–English bilinguals (Flege, 1991). On the other hand, English vowels in words like *hot* and *hut* can be perceived as instances of /a/ in Italian–English bilinguals (Flege, 2003). Given the phonetic similarity between Italian and Spanish, it is likely that our subjects will confuse the vowels in *sof* and *suf* as instances of the Spanish /a/. The English voiceless fricatives /s/ and /f/, with the random acoustic noise that characterizes them, were used to benefit the recognition of the vowel within the syllable.

Linguistic assessments

Picture vocabulary (PV)

A test selected from the Woodcock Language Proficiency Battery-Revised. It evaluated expressive competence in English and Spanish. Participants overtly named pictures that gradually ranged from easy to difficult. Monolinguals were tested in English and bilinguals were tested in English and Spanish.

Listening comprehension (LC)

A test also extracted from the Woodcock Language Proficiency Battery-Revised. This sentence completion task assessed receptive competence in each language. Participants listened to the beginning of a sentence and were asked to complete the end of it. For example: “Radios play pretty _____ (music).”

In the analysis, the tests of picture vocabulary and listening comprehension were combined for each language version (i.e., PV–LC English, PV–LC Spanish) to account for the level of proficiency in L1 and L2.

Similarity-judgment task

In a soundproof booth, participants completed a similarity-judgment task using Paradigm software (Perception Research Systems Inc.). The task presented 780 pairs of syllable sounds via two computer speakers while simultaneously displaying a 4-point likert scale on the monitor screen. The participants were asked to rate the level of similarity between the two consecutively presented syllables (1 = very similar, 4 = not similar). The duration of the experiment was approximately 45 minutes.

Results

Proficiency in English (L2)

We analyzed the proficiency levels of English and Spanish to understand how they played a role in the perceptual categorization of L2 English syllables (i.e., *saf*, *sef*, *sof* and *suf*). Note that English is the first language in the monolingual group. First, a one-way ANOVA was carried out to examine differences in English proficiency, with AOA (monolinguals vs. early vs. intermediate vs. late) serving as the between-subjects factor. The tests composing the variable “English proficiency” were picture

vocabulary and listening comprehension. We calculated this variable by adding the percent correct of each test and then dividing the total by 2, thus obtaining a global average measure of language proficiency. A moderately significant correlation between the tests indicated that picture vocabulary and listening comprehension could be suitably combined and used as a single measure of proficiency ($r = .47, p < .0001$). These two tests have also been demonstrated to be in the oral proficiency end of the standardized Woodcock scale (Woodcock & Muñoz-Johnson, 2005). The analysis revealed that the means in English proficiency obtained for monolinguals ($M = 73.9, SD = 6.08$), early bilinguals ($M = 69.7, SD = 7.17$), intermediate bilinguals ($M = 66.7, SD = 7.44$) and late bilinguals ($M = 64.1, SD = 7.05$) differed significantly; omnibus ($F(3,91) = 8.80, p < .0001$). However, further post-hoc pairwise comparisons indicated that the intermediate bilinguals did not significantly differ from early bilinguals ($F(1,91) = 2.04, p > .15$) or from late bilinguals ($F(1,91) = 1.24, p > .26$) in English proficiency. A trend is observed as the higher means obtained in monolinguals are followed by lower means in early, intermediate and late bilinguals. In addition, we found that English proficiency significantly correlated with English use ($r = .434, p < .02$).

Separate one sample *t*-tests were run to investigate the amount of variability in English proficiency *within each group*. The results showed that English proficiency did not significantly differ within monolinguals ($t(27) = .08, p = .93$) or early ($t(30) = .06, p = .95$), intermediate ($t(15) = .03, p = .96$) or late bilinguals ($t(19) = .05, p = .96$).

Proficiency in Spanish (L1)

Similarly to the way we calculated English proficiency, adding the percent correct of Spanish listening comprehension and picture vocabulary and dividing the score by 2 quantified overall Spanish proficiency. Our calculation here was supported by a strong correlation between L1 picture vocabulary and listening comprehension ($r = .74, p < .0001$). A one-way ANOVA was used to examine differences in L1 proficiency in the three bilingual groups. The results showed that the early ($M = 61.98, SD = 10.33$), intermediate ($M = 66.19, SD = 5.19$) and late bilinguals ($M = 75.98, SD = 4.22$) significantly differed in their Spanish proficiency; omnibus ($F(3,91) = 19.49, p < .0001$). The early bilinguals had marginally lower proficiency in Spanish than later bilinguals ($F(1,91) = 3.04, p > .08$). A trend, in the opposite direction from that of English proficiency, indicated lower means of Spanish proficiency in early bilinguals followed by higher means in intermediate and late bilinguals. Similarly to what we found in L2 proficiency, L1 proficiency was significantly correlated with L1 use ($r = .415, p < .05$).

Just like English proficiency, separate one sample *t*-tests were run to investigate the amount of variability in Spanish proficiency *within each group*. Here we found that Spanish proficiency did not significantly differ within early ($t(30) = .04, p = .96$), intermediate ($t(15) = .08, p = 0.94$) or late bilinguals ($t(19) = .09, p = .92$).

Interestingly, there was a significant difference between the level of proficiency in L1 and L2 in early ($t(30) = 4.87, p < .0001$) and late bilinguals ($t(19) = -7.34, p < .0001$), but not in intermediate bilinguals ($t(15) = .28, p = .784$). These results indicate that early bilinguals who are highly proficient in English are not as proficient in Spanish and late bilinguals who are highly proficient in Spanish are not as proficient in English. The level of dominance in one language appears to occur at the expense of the other. Intermediate bilinguals, however, have comparable proficiencies in both languages despite the fact that their scores are not as high as those of monolinguals in English or late bilinguals in Spanish. These results were further corroborated with two significant correlations between the predictor variables. AOA negatively correlated with English proficiency ($r = -.29, p = .017$) but positively correlated with Spanish proficiency ($r = .58, p < .0001$), meaning that late acquisition of the L2 results in lower proficiency in English, but at the same time, higher proficiency in L1 Spanish. It is important to highlight these correlations, as the relationship between AOA and proficiency can be difficult to tease apart.

Multidimensional scaling and one-way ANOVA

We obtained the perceptual distances according to the participants' responses on the similarity judgment task via multidimensional scaling (MDS) analysis. The mean distances between exemplars belonging to the same category (i.e., within-category) were obtained from a 40×40 matrix that contained four 10×10 submatrices belonging to the *saf*, *sef*, *sof* and *suf* categories, respectively. The median or center of each category was used to calculate the perceptual distances between different categories (i.e., between-category) (see Figure 1). The distances obtained for within- and between-categories in each participant were used in one-way ANOVAs to examine group differences in the perception of non-native syllables and in multiple regression analyses to investigate how the range of AOA and varying PL predicted the categorization of such. Note that in our MDS analysis, the direction of the axes is trivial, as we are mainly concerned with calculating the respective perceptual distances among exemplars.

A composite variable named "within-category" averaged the perceptual distances among the exemplars of single categories. Within-category perception of all syllable types (i.e., *saf*, *sef*, *sof*, *suf*) significantly differed across all four groups; omnibus ($F(3,94) = 5.14, p = .0025$).

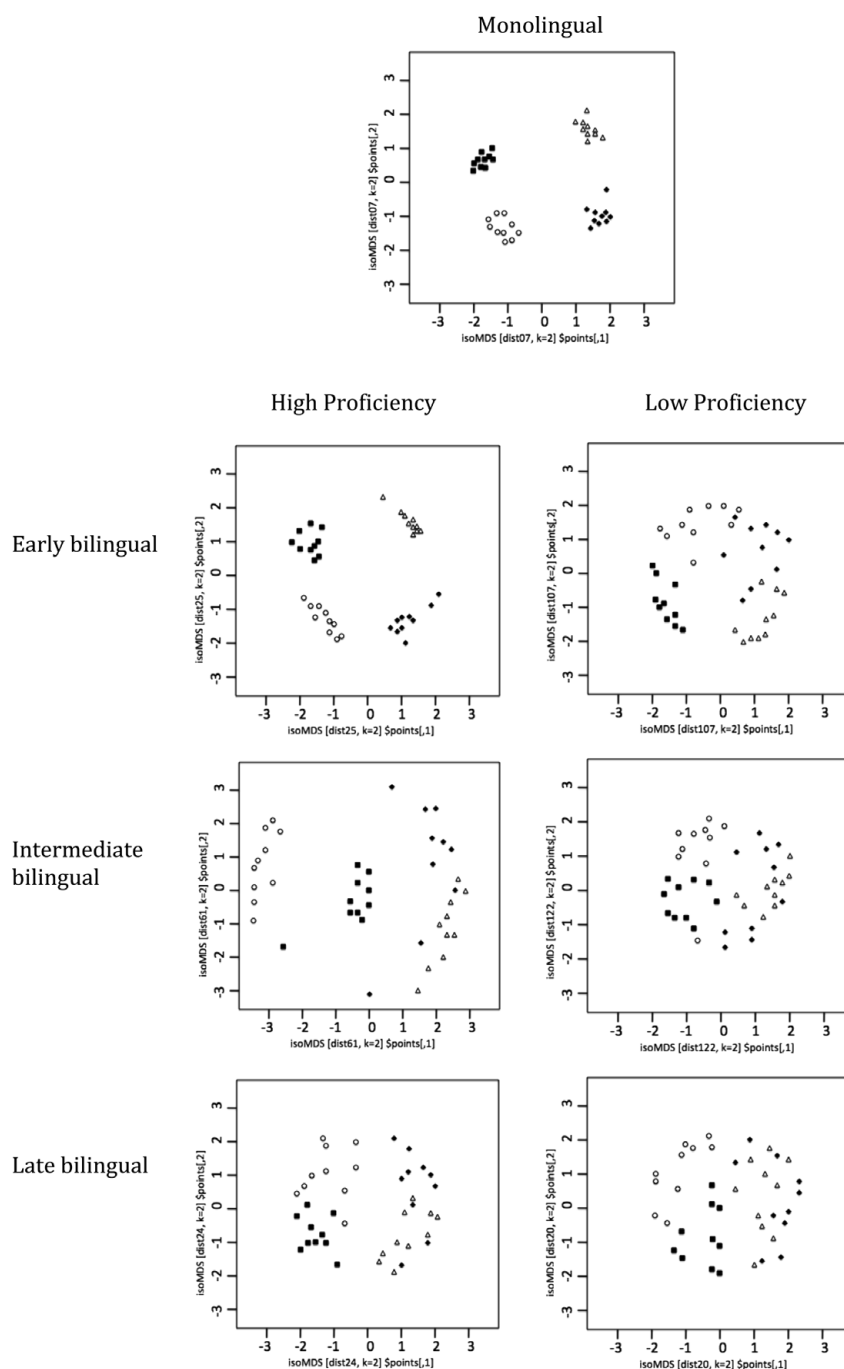


Figure 1. Multidimensional scaling plots – graphic representations of the perceptual acoustic distances of within- and between-categories subdivided by group type and proficiency level. Each symbol represents a phonemic category: black squares for “saf”, black diamonds for “suf”, white circles for “sef”, and white triangles for “sof”.

As expected, early and late bilinguals significantly differed in the clustering of within-category stimuli ($F(1,94) = 10.66, p = .0015$). Also in line with our predictions, monolinguals and early bilinguals had similar patterns of within-categorization ($F(1,94) = 1.15, p = .28$), indicating that an early acquisition of L2 allows the proper grouping of L2 within-categories; presumably because these phonemes were acquired by means of

sensorimotor and implicit processes early in life. Notably, intermediate and late bilinguals did not differ in the pattern of *altered* within-categorization ($F(1,94) = .01, p = .92$), suggesting that native-like phonemic mapping of L2 exemplars of the same category begins to warp around six years of age when preschool children begin to use explicit rules to understand the world. It may be possible that higher-cognitive systems are interfering with the

implicit processing of within-categories. Irrespective of proficiency, the largest variation of stimuli categorization was observed within the new *suf* category in intermediate ($M = 1.19$, $SD = .46$) and late bilinguals ($M = 1.07$, $SD = .28$).

Whether as a combination of syllable pairings (i.e., *saf-sef*, *saf-sof*, *saf-suf*, *sef-sof*, *sef-suf*, *sof-suf*) or as individual pairs, all groups were able to distinguish between-categories (omnibus ($F(3,94) = 1.90$, $p = .13$)), with the exception of the *sof-suf* pair. The discrimination of *sof-suf* differed across monolinguals and bilinguals (omnibus ($F(3,94) = 5.97$, $p = .0009$)), with a trend showing monolinguals to be better than any other group in correctly discriminating *sof* from *suf* ($F(1,94) = 8.16$, $p = .0053$). The likelihood that a bilingual group properly distinguished the two novel L2 categories – *sof* and *suf* – gradually changed as a function of AOA. A careful look at the means – early ($M = 2.23$, $SD = .82$), intermediate ($M = 2.09$, $SD = 1.14$) and late bilinguals ($M = 1.38$, $SD = .75$) – reveals this trend. Furthermore, we found that late bilinguals' categorization of L2 syllables that have comparable categories in L1 (i.e., *saf-sef*) was significantly different from the categorization of novel L2 syllables (i.e., *sof-suf*) ($t(22) = 4.36$, $p < .0003$), as illustrated in the tighter clusters for *saf* and *sef* and broader clusters for *sof* and *suf* categories in late bilinguals. This finding indicates that L1 phonemic categorization affects L2 categorization, as shown in previous studies. Finally, it is important to note that even though our group of subjects was considerably unbalanced in regards to gender (80 females, 18 males), there were no significant differences in the perception of within- or between- L2 categories across groups ($F(1,96) = .04$; and $F(1,96) = .74$, respectively).

Multiple regressions

We were interested in predicting native-like perception of L2 speech sounds in bilinguals given the allocation of syllabic exemplars to within-category or between-category. In the literature, the predictor variables of AOA, L1 and L2 proficiency are thought to contribute to the categorization of non-native speech sounds. Therefore we used these three variables to conduct two separate multiple regressions on the metrics obtained from MDS (i.e., perceptual distances); one multiple regression for within-category and the other for between-category. It is worth noting that the similarity scores from MDS represented greater similarity if the numbers were small, and greater dissimilarity if the numbers were large. In the first multiple regression, we found that AOA significantly predicted within-category discrimination ($r = .4416$, $p = .0062$ ($pr^2 = .11$, $sr^2 = 0.10$)), while proficiency in English ($r = -.07$, $p = .55$ ($pr^2 = .005$, $sr^2 = .004$)) and proficiency in Spanish ($r = -.05$, $p = .72$ ($pr^2 = .001$, $sr^2 = .001$)) did not. No interactions between Spanish proficiency, English

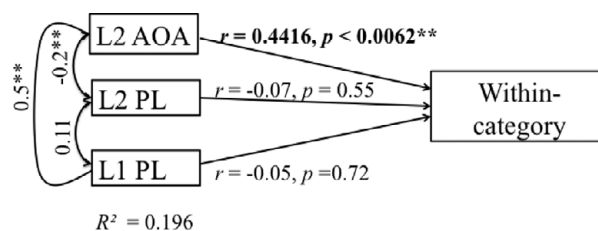


Figure 2. Multiple regression on within-category. AOA is a significant predictor.

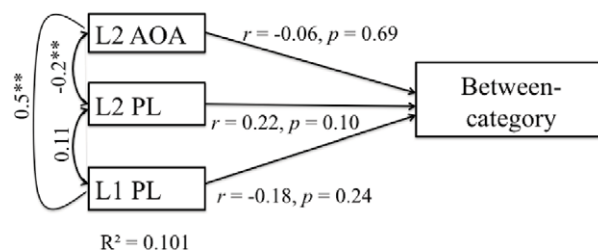
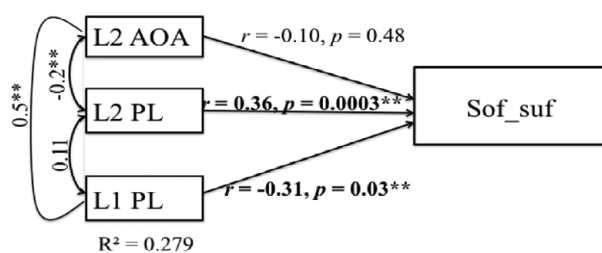


Figure 3. Multiple regression on between-category. No significant predictors are found.

proficiency and AOA were found when predicting within-category stimuli. We obtained a small to medium effect size for this regression ($f^2 = .243$, with a high power of .92). These findings are in accord with our hypotheses, while early acquisition of L2 results in the tight clustering of within-category exemplars, later acquisition results in dispersed discrimination of exemplars. Like native speakers, early bilinguals correctly group a range of exemplars as members of the same category, probably due to the sensorimotor/implicit manner in which these sounds were acquired. In this analysis, 20% of the variance was accounted for by all three predictor variables (see Figure 2).

In the second multiple regression, none of the predictor variables – AOA, L1 and L2 proficiency – were found to significantly predict between-categorization ($r = -.06$, $p = .69$, $r = -.18$, $p = .24$, and $r = .22$, $p = .10$, respectively). Similarly to within-category regression, no interactions between Spanish proficiency, English proficiency and AOA were found when predicting the perception of between-category stimuli. Also, less variance was contributed by the predictor variables in the between-category regression analysis ($R^2 = .101$) (see Figure 3).

However, further analyses revealed that discrimination of the *sof-suf* pairing was significantly predicted by English proficiency ($r = .36$, $p = .003$). In other words, bilinguals with high L2 proficiency better discriminated the tokens of *sof* and *suf* than bilinguals with low L2 proficiency. This effect of proficiency was primarily driven by the late bilingual group ($r = .43$, $p = .03$) (see Figure 4), suggesting that increased attention to



Late-bilinguals only:

L2 AOA: $r = 0.04, p = 0.80$

L2 PL: $r = 0.43, p = 0.03**$

L1 PL: $r = -0.66, p = 0.003**$

Figure 4. Multiple regression on *sof*–*suf* pairing. Proficiency in the first (Spanish) and second (English) languages are significant predictors; the effect is primarily driven by late bilinguals.

the phonemic boundary of novel L2 sounds can help a late bilingual learn L2 categorization explicitly. For this regression, Cohen's f showed a medium effect size ($f^2 = .369$) and a low power (.17). An unexpected result showed that *low* Spanish proficiency also predicted *sof*–*suf* categorization ($r = -.31, p = .03$), and the effect was also driven by late bilinguals ($r = -.66, p = .0032$). This surprising finding may be the result of L1 language attrition that in late bilinguals renders perception of L2 sounds more transparent by circumventing the L1 sound network. Another possibility is that some late bilinguals with low PL scores in Spanish did not develop well-defined phonemic boundaries in the L1. This would translate into a sustained auditory malleability that would in turn help them acquire the new L2 sounds more efficiently. In summary, our results showed that early L2 acquisition is an important factor in accurate within-category discrimination and high L2 proficiency in late bilinguals is fundamental for the perception of boundaries between categories of novel L2 sounds.

Discussion

As predicted in the 'Introduction', our results showed that the perception of within- and between- categories is differentially affected by age of acquisition and proficiency level in the second language. Specifically, early bilinguals demonstrated reduced sensitivity to the small acoustic variations that exist among exemplars of the same category and increased sensitivity to the phonemic boundaries between categories. On the other hand, late bilinguals attended to irrelevant acoustic cues of the same phoneme resulting in the perceptual distortion of within-categories, but recognized meaningful phonemic

changes that indicated a switch in L2 categories. In line with the literature on brain and cognitive development of implicit/explicit processes, it appears that children and adults respectively rely on the neural–cognitive mechanisms available at the time of learning. The fine-grained and random aspect of within-category phonemes compels the use of implicit learning strategies that only early bilinguals can readily utilize early in life via sensorimotor/perceptual activities of subcortical areas. Unlike within-category perception, the distinctive arrangement of separate phonemic clusters (i.e., between-category) allows highly proficient late bilinguals to explicitly learn the acoustic cues present at the phonemic boundaries via high-level cognitive processes sustained by cortical areas in adulthood. A neuroimaging study conducted by Joanisse, Zevin and McCandliss (2007), for example, investigated the categorization of speech sounds in monolingual adults and found that within-category stimuli activated subcortical regions, while between-category stimuli activated cortical regions like the left superior sulcus, middle temporal gyrus and inferior parietal cortex. It may be the case then that early bilinguals discriminate within- and between-categories by means of perceptual processes supported by early developing subcortical activity, whereas late high-proficient bilinguals discriminate between-category speech sounds by means of higher-order cognition sustained by later-developing cortical activity.

Akin to the behavioral results presented here, recent fMRI findings in our laboratory with 66 Spanish–English bilinguals ($N = 83$, 17 monolinguals) varying in L2 AOA and PL revealed that passively listening to pairs of “same” (within-category; e.g., *saf*–*saf*) or “different” (between-category; e.g., *saf*–*suf*) non-native sounds led to the recruitment of the superior temporal gyrus (STG) and inferior frontal gyrus (IFG) bilaterally in early bilinguals (Archila, Ramos, Zevin & Hernandez, 2010). These two brain regions are known for their involvement in early auditory processing and speech categorization (Binder et al., 2000) and have been reported in monolingual adults and monolingual infants processing L1 speech (Dehaene-Lambertz, Dehaene & Hertz-Pannier, 2002; Imada, Zhang, Cheour, Taulu, Ahonen & Kuhl, 2006). On the other hand, various bilateral regions of the frontal lobe and parietal lobe, including the superior frontal gyrus, precuneus, supramarginal gyrus and precentral gyrus, showed increased activity in highly proficient bilinguals. The latter areas are known for their role in attention (Hugdahl, Wester & Asbjørnsen, 1991; Rueckert & Grafman, 1996; Sabri, Binder, Desai, Medler, Leitl & Liebenthal, 2007), active phonetic change judgments (Zevin & McCandliss, 2005), and mappings of articulatory features (Pulvermüller, Huss, Kherif, Martin, Hauk & Shtyrov, 2006). It appears then that early acquisition and high proficiency in a second

language independently activates regions associated with perceptual or higher-order cognition. Therefore, early bilinguals draw on similar neural substrates to discriminate speech tokens that are either “same” or “different” because processing these sounds depends on early developing brain areas, such as primary auditory areas. On the other hand, high proficiency in a second language, which is only acquired with years of experience, draws on later-developing brain regions in frontal and parietal cortices to process non-native speech.

The same paradigm of passive listening with 38 Spanish–English bilingual children in our laboratory showed that younger children between the ages of six and eight activate more subcortical and primary auditory areas (i.e., bilateral STG, caudate nucleus, hippocampus and right putamen) than older children between the ages of nine and ten. Moreover, better proficiency in Spanish demonstrated more neural activity in the right parietal lobe and left cingulate gyrus (unpublished). This latter set of activations reflects the involvement of sensory integration and awareness in linguistically advanced children. Consistent with our behavioral and neuroimaging results with adults, these findings suggest that the brain areas recruited for the detection of non-native speech sounds are a function of brain maturation processes and language proficiency. As the brain develops and children become more advanced learners of two languages, new higher-order cortical areas emerge to process speech.

The behavioral findings presented here and the findings from our recently completed fMRI projects with bilingual adults and bilingual children do not appear to be in line with the results of Callan, Jones, Callan and Akahane-Yamada (2004), and Golestani and Zatorre (2004). The different findings across studies may be rooted in methodological details. For example, Callan et al. studied two groups of monolingual speakers who either identified between-category speech sounds from their native tongue (i.e., English speakers actively listening to syllables with initial-position /l/ and /r/ in English) or identified sounds from a second language (i.e., Japanese speakers with some English experience listening to initial-position English /l/ and /r/ syllables). Their results showed that Japanese monolinguals identifying English consonants activated the STG, IFG, anterior insula, planum temporale, supramarginal gyrus and cerebellum. The study concluded that Japanese monolinguals with some second language experience activated a greater number of articulatory–auditory and articulatory–orosensory areas than strictly monolingual subjects. Similarly, Golestani and Zatorre studied the neural changes observed in monolinguals after five sessions of phonetic training with an unfamiliar Hindi contrast. Their results showed activity in the STG, IFG and insula-frontal operculum after training. The study concluded that efficient processing

of non-native speech results from activation of the same brain areas recruited for native speech. A fundamental difference lies between the groups investigated in each study. While monolinguals with some second language experience and monolinguals with a short period of phonetic training recruited the same temporal areas seen in L1 speech processing, these groups reported no activity in parietal regions – with the exception of activity in the SMG for Japanese monolinguals with some English experience. This suggests that at least six years of classroom instruction in a foreign language (see Callan et al., 2004, for a complete review of the study) are necessary to learn to discriminate non-native between-categories, as demonstrated by activity in the SMG. As mentioned, the SMG has been associated with phonetic change detection (Zevin & McCandliss, 2005). It is likely that no other parietal regions were activated because the monolingual groups did not have an adequate level of proficiency in the second language. Bilinguals, on the other hand, have undergone drastic neural changes throughout years of experience perceptually manipulating two phonetic codes. It may take a considerable amount of time to observe neural reorganization in the form of fronto-parietal activations. As our behavioral and neuroimaging data shows, only *highly proficient* bilinguals perceptually discriminate difficult non-native between-categorical contrasts (i.e., *sof–suf*) and recruit additional brain areas to support this process (right frontal regions, supramarginal gyrus, precentral gyrus and precuneus in adults and right parietal lobe and left cingulate gyrus in children).

While the Native Language Magnet (NLM) explains that a neural commitment to the L1 around twelve months of age skews the perception of L2, the Speech Learning Model (SLM) insists that perception is malleable and can be corrected with abundant L2 exposure even in adulthood. Given our classification of early bilinguals, it appears that native-like discrimination of within- and between-category exemplars between infancy and five years of age is still possible. Furthermore, late proficient bilinguals are able to categorize novel L2 speech phonemes. Therefore, our findings indicate the maintenance of perceptual plasticity beyond the cut-off age stipulated by the NLM. Although our data support the SLM, this model does not make reference to the underlying neurocognitive learning mechanisms that play a potential role in L2 speech acquisition in children and adults. However, our findings are consistent with the view that early and late bilinguals differentially learn L2 phonemes due to the cognitive processes recruited.

The sensorimotor hypothesis provides a theoretical framework for the fundamental differences in skill learning – including L2 speech – that surface as AOA increases. Along with the sensorimotor approach, other developmental theories hint at the notion that

discrepancies in learning occur because there are notorious coupled biological–cognitive constraints that compel children and adults to process speech information differently. These differences in early vs. late processing have also been noted in visual perception (Bronson, 1974), and face perception (Morton & Johnson, 1991). In both research areas, early processing has been associated with subcortical activity and activity in primary areas of the modality under question, whereas later processing has been associated with cortical activity and activity in secondary and tertiary areas. The question is not whether early or late bilinguals can learn to categorize non-native speech – after all, an absolute ending of the critical period in language has not been reported (Birdsong, 1999) – but rather *How does a brain in the midst of developmental change process L2 sounds?* The theory of interactive specialization proposed by Johnson (2001, 2005) holds that general functions of the human brain become increasingly specialized as cortical regions respond to stimuli's recurrent properties. According to this account, a sensitive period gradually closes as the system becomes more finely tuned. Consequently, in non-native speech perception, early bilinguals go through the process of increased phonological specialization in both languages simultaneously, while late bilinguals pick up on the sound system of L2 after specialization to L1 sounds has already occurred. Similar approaches allude to the notion of general-to-specific functions in speech processing (Kuhl, 2000), by advocating that infants perceive speech sounds through the extraction of statistical regularities and adults perceive L2 sounds through an L1 filter. Since infants and young children process information generally, they require large amounts of input to implicitly extract such regularities from the speech signal (Hudson-Kam & Newport, 2005). On the other hand, adults can adopt explicit strategies to override the specificity or commitment of the neural system to the L1 and thereby extract the cues necessary for L2 speech learning (Hudson-Kam & Newport, 2005).

If a sensorimotor way of processing L2 speech results in native-like perception, why do some early bilinguals have foreign accents? Research shows that these particular individuals continue to use their L1 on a regular basis (Flege et al., 1997); therefore, it is possible that the prolonged use of a native language (a) interferes with accurate L2 learning or (b) decreases the number of opportunities for input and output of L2. Computer simulation studies, for example, have shown that the likelihood of a model to learn a second object depends on the length of time the first and second object are exposed to the model (Bolhuis, 1999). Furthermore, Pallier et al. (2003) has found that adult Korean adoptees who moved to France as children show complete loss of a first language. This group of individuals who learned their L2 between the ages of three and eight show no difference

in either behavioral or neural activity when compared to monolingual French speakers. That is, extreme immersion in L2 can lead to complete loss of L1. Therefore, the amount of L1 and L2 use is a relevant factor in non-native speech perception that needs to be more carefully investigated. It is also important to consider the linguistic backgrounds employed in non-native speech perception studies, as vowel inventories across languages can range from very similar to very dissimilar (Best, 1995), thus affecting the generalizability of the results. Finally, there is continued debate in the field of speech perception regarding whether tokens of stimuli should be synthesized or not. For our purposes in the present study, natural speech stimuli gave us greater ecological validity and stronger support for our findings, as our main interest was to understand how bilinguals cognitively distinguish standard L2 speech by dissociating their perceptual judgments as within- and between-categories. We encourage researchers in the field to investigate how the effect of L2 proficiency alters neurofunctional processes in intermediate and late bilinguals as this topic becomes of ever-increasing importance in our rapidly growing population of Spanish–English bilinguals entering the education system in the United States.

In summary, our results demonstrate that early L2 exposure results in accurate within- and between-categorization, suggesting that children use sensorimotor and implicit mechanisms for learning L2 speech. On the other hand, late L2 exposure only results in correct between-categorization of novel L2 sounds if the adult learner is highly proficient, suggesting that adults can make use of high-level cognitive processes like attention and other explicit strategies to learn the acoustic cues that determine the phonemic boundaries of L2.

References

- Aoyama, K., Flege, J. E., Guion, S., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32, 233–250.
- Archila, P., Ramos, A. I., Zevin, J., & Hernandez, A. E. (2010). How age of acquisition and proficiency predicts sound detection in bilinguals. Poster presented at Armadillo Conference, College Station, TX.
- Best, C. (1995). A direct realist view of cross language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, pp. 171–206. Timonium, MD: York Press.
- Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305–331.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10 (5), 512–528.

- Birdsong, D. (1999). Introduction: Whys and why not of the Critical Period Hypothesis for second language acquisition. In D. Birdsong (ed.), *Second language acquisition and the Critical Period Hypothesis*, pp. 1–22. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bolhuis, J. J. (1999). The development of animal behaviour: From Lorenz to neural nets. *Naturwissenschaften*, *86*, 101–111.
- Bronson, G. (1974). The postnatal growth of visual capacity. *Child Development*, *45* (4), 873–890.
- Bronson, G. (1990). Changes in infants' visual scanning across the 2–14 week age period. *Journal of Experimental Child Psychology*, *49*, 101–125.
- Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory–auditory/orosensory internal models. *Neuroimage*, *22* (3), 1182–1194.
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech in infants. *Science*, *298* (5600), 2013–2015.
- Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Hueber & C. Ferguson (eds.), *Crosscurrents in second language acquisition and linguistic theories*, pp. 249–289. Amsterdam/Philadelphia: John Benjamins.
- Flege, J. E. (1995). Second-language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, pp. 229–273. Timonium, MD; York Press.
- Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In N. Schiller & A. Meyer (eds.), *Phonetics and phonology in language comprehension*, pp. 319–355. Berlin: Walter de Gruyter.
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169–186.
- Flege, J. E., Munro, M., & Fox, R. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*, 3623–3641.
- Gao, J. H., Parsons, L. M., Bower, J. M., Xiong, J., Li, J., & Fox, P. T. (1996). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science*, *272*, 545–547.
- Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: Reallocation of neural substrates. *Neuroimage*, *21* (2), 494–506.
- Green, P., & Hecht, K. (1992). Implicit and explicit grammar. An empirical study. *Applied Linguistics*, *13* (2), 168–184.
- Guion, S., Flege, J., Akahane-Yamada, R., & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107* (5), 2711–2724.
- Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133* (4), 638–650.
- Hudson-Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.
- Hugdahl, K., Wester, K., & Asbjørnsen, A. (1991). Auditory neglect after right frontal lobe and right pulvinar thalamic lesions. *Brain & Language*, *41* (3), 465–473.
- Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's area: A developmental magnetoencephalography study. *NeuroReport*, *17* (10), 957–962.
- Joanisse, M. F., Zevin, J. D., & McCandliss, B. D. (2007). Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex*, *17* (9), 2084–2093.
- Johnson, M. H. (2001). Functional brain development in humans. *Nature Reviews Neuroscience*, *2*, 475–483.
- Johnson, M. H. (2005). *Developmental cognitive neuroscience*, 2nd edn. Oxford: Blackwell.
- Karmiloff-Smith, A. (1991). Innate constraints and developmental change. In S. Carey and R. Gelman (eds.), *Epigenesis of the mind: Essays in biology and knowledge*, pp. 171–197. Mahwah, NJ: Erlbaum.
- Kent, R. (1997). *The Speech Sciences*. San Diego: Singular Publishing Group.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97* (22), 11850–11857.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
- Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological Review*, *98*, 164–181.
- Pallier, C., Dehaene, S., Poline, J. B., LeBihan, D., Argenti, A. M., Dupoux, E., et al. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex*, *13*, 155–161.
- Peltola, M., Kuntola, M., Tamminen, H., Hämäläinen, H., & Aaltonen, O. (2005). Early exposure to non-native language alters preattentive vowel discrimination. *Neuroscience Letters*, *388* (3), 121–125.
- Pulvermüller, F., Huss, M., Kherif, F., Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103* (20), 7865–7870.
- Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 492–502.
- Rueckert, L., & Grafman, J. (1996). Sustained attention deficits in patients with right frontal lesions. *Neuropsychologia*, *34* (10), 953–963.
- Sabri, M., Binder, J., Desai, R., Medler, D., Leitl, M., & Liebenthal, E. (2007). Attentional and linguistic

- interactions in speech perception. *Neuroimage*, 39 (3), 1444–1456.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.
- Westermann, G., Mareschal, D., Johnson, M., Sirois, S., Spratling, M., & Thomas, M. (2007). Neuro-constructivism. *Developmental Science*, 10 (1), 75–83.
- Woodcock, R., & Muñoz-Johnson, A. (2005). *Woodcock-Munoz Language Survey: Normative Update*. Itasca, IL: Riverside Publishing.
- Zevin, J. D., & McCandliss, B. D. (2005). Dishabituation of the BOLD response to speech sounds. *Behavioral and Brain Functions*, 1 (1), 1–4.