Child–Adult Differences in Second-Language Phonological Learning: The Role of Cross-Language Similarity

Wendy Baker[1], **Pavel Trofimovich**[2], **James E. Flege**[3], **Molly Mack**[4], **Randall Halter**[2]

[1] *Brigham Young University, Provo, Utah, U.S.A.*
[2] *Concordia University, Centre for the Study of Learning and Performance, Montréal, Canada*
[3] *Istituto di scienze e tecnologie della cognizione (Institute of Cognitive Sciences and Technologies), National Research Council, Rome, Italy*
[4] *University of Illinois, Urbana-Champaign, U.S.A.*

Key words

child–adult differences

cross-language similarity

second language

speech perception

speech production

Abstract

This study evaluated whether age effects on second language (L2) speech learning derive from changes in how the native language (L1) and L2 sound systems interact. According to the "interaction hypothesis" (IH), the older the L2 learner, the less likely the learner is able to establish new vowel categories needed for accurate L2 vowel production and perception because, with age, L1 vowel categories become more likely to perceptually encompass neighboring L2 vowels. These IH predictions were evaluated in two experiments involving 64 native Korean- and English-speaking children and adults. Experiment 1 determined, as predicted, that the Korean children were less likely than the Korean adults to perceive L2 vowels as instances of a single L1 vowel category. Experiment 2 showed that the Korean children surpassed the Korean adults in production of certain vowels but equaled them in vowel perception. These findings, which partially support the IH, are discussed in relation to L2 speech learning.

1 Introduction

Over the last several decades, speech perception and production research has yielded considerable evidence that, overall, children are more successful at learning a second

Acknowledgments: This research was partially supported by NIH grant 501 DC02892–05; CFDA #93.173 to James E. Flege and research grants to Wendy Baker from Brigham Young University, and to Pavel Trofimovich from the University of Illinois at Urbana-Champaign. Many thanks are extended to Gabseon (Sunny) Lee for her help in recruiting the Korean participants. An earlier version of this article was presented at the 26th Boston University Conference on Language Development. The authors thank Ocke-Schwen Bohn, Christophe Pallier, and four anonymous reviewers for their helpful suggestions on earlier drafts of this manuscript.

This article is dedicated in loving memory to Molly Mack (1950–2008) by her grateful students and friends.

Address for correspondence. Wendy Baker, PhD, Department of Linguistics and English Language, Brigham Young University, 4057 JFSB, Provo, UT 84601, U.S.A.; e-mail <wendy_baker@byu.edu>

Language and Speech; 0023-8309; Vol 51(4): 317–342; 099067; DOI:10.1177/0023830908099068
http://las.sagepub.com

language (L2) than are adults (Birdsong & Molis, 2001; Flege, Yeni-Komshian, & Liu, 1999; Johnson & Newport, 1989; Stevens, 1999), especially in the realm of phonetics and phonology (Flege, MacKay, & Meador, 1999; Guion, Flege, Lin, & Yeni-Komshian, 2000; Munro, Flege, & MacKay, 1996; Oyama, 1976). To date, most evidence for child–adult differences in L2 phonological learning has come from studies examining children's and adults' L2 perception and production after an extensive amount of L2 experience or exposure (e.g., Flege, MacKay et al., 1999). However, besides child–adult differences per se, any evidence of child superiority observed in such cases might also be attributed to differences in L2 input that children and adults receive throughout L2 development (see, e.g., Flege, Yeni-Komshian et al., 1999; Jia & Aaronson, 2003). The present study sidestepped this criticism in making child–adult comparisons by testing children and adults at the onset of L2 phonological learning in a context where the target language is spoken natively. The overall aim of the study was to provide new insights into child–adult differences in L2 speech learning.

1.1 Child–adult differences in L2 learning

One common explanation for child–adult differences in L2 learning is the existence of a critical period for the acquisition of an L2. The putative critical period is usually associated with time-sensitive neurobiological maturation processes, resulting in a decrease in the ability to fully acquire an L2 (e.g., Lenneberg, 1967). This approach claims that, in order to reach native-speaker proficiency in an L2, learners need to be exposed to the L2 within a narrow, neurologically-determined "window." Other researchers argue for a sensitive rather than critical period, stating that early childhood represents an ideal, but not the only, time when a language can be acquired with native-like accuracy (e.g., Abu-Rabia & Kehat, 2004; Hakuta, Bialystok, & Wiley, 2003). According to this view, early childhood represents a time where optimal input and interaction with the L2 can take place (Mack, 2003; Pallier, Bosch, & Sebastián-Gallés, 1997). (See Mack, 2003, for one conceptualization of the relationship between critical and sensitive periods in L2 development.) Support for a critical or sensitive period for L2 learning (e.g., Johnson & Newport, 1989; Newport, 1990) has parallels in investigations of diverse aspects of sensory motor development in animals, for example, in the ability of frogs to learn to swim (Moorman, Cordova, & Davies, 2002), cats to see (Blausdel, Mitchell, Muir, & Pettigrew, 1977), birds to sing (Rozin, 1976), or in the development of limb–eye coordination in rats (Fifkova, 1968).

While biologically-determined explanations for child–adult differences in L2 learning are appealing, testing the assumptions underlying the critical-sensitive period hypothesis has proven quite difficult. In fact, adequately testing this hypothesis would require attempting to ensure, for example, that children and adults receive the same amount and type of L2 input and that they do not differ in their motivation to learn an L2, in their extent of L2 use, nor in their degree of affiliation to an ethnic group (Flege, Yeni-Komshian et al., 1999; Gatbonton & Trofimovich, 2008; Jia & Aaronson, 2003). After these conditions have been met, a valid test of this hypothesis would also require *direct* measures of the neurobiological changes thought responsible for changes in learning, with these changes measured throughout the learning process. To our knowledge, none of these tests have ever been done.

At the same time, some researchers have called into question the *very existence* of a critical or a sensitive period for L2 learning (e.g., Hakuta et al., 2003) due to the abundance of conflicting evidence. For example, Abu-Rabia and Kehat (2004) reported that older learners were more native-like in their L2 pronunciation than were younger learners. Other studies have reported age-of-acquisition effects for learners who were clearly beyond a critical or a sensitive period (e.g., Hirsh, Morrison, Gaset, & Carnicer, 2003; Trofimovich & Baker, 2006). Such inconsistent findings suggest that other explanations of age-of-acquisition effects are needed to account for the results of those studies that do not support the critical-sensitive period hypothesis (e.g., Birdsong & Molis, 2001). In principle, such explanations can be compatible with the critical-sensitive period hypothesis but they cannot draw *solely* on neuropsychological evidence, at least until the current and future methodologies of neurophysiological assessment can allow researchers to link language development to neurobiological changes in the brain (see Pallier et al., 2003, for further discussion). Our goal here was therefore to describe and test one such explanation of child–adult differences in L2 phonological learning.

1.2
The interaction hypothesis

According to Flege's Speech Learning Model (Flege, 1995), one possible way in which child–adult differences manifest themselves, at least in the realm of L2 phonological learning, is through the interaction (bidirectional influence) of learners' native language (L1) and their L2. More specifically, learners' L1 and L2 are believed to interact in a different manner depending on the age at which the L2 is learned. This tenet of the Speech Learning Model has been termed the "interaction hypothesis," or IH (Flege, 1992, 1999; Walley & Flege, 1999). It is likely that the L1 and L2 always influence each other to some degree in all learners (Baker & Trofimovich, 2005; Mack, 1990). The IH holds, however, that the nature of the L1–L2 interaction changes as a function of the state of development of the L1 phonetic system at the time L2 learning begins.

As discussed below, there is evidence that the long-term memory representations (categories) for vowels and consonants develop slowly through childhood and into adolescence (Hazan & Barrett, 2000; Walley & Flege, 1999). As the L1 categories develop, they are said to become stronger "attractors" for L2 sounds. Put another way, L2 vowels and consonants become increasingly likely to be treated as variants of an L1 category as the L1 phonetic system develops, even when L1–L2 differences can be auditorily detected (see Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992, for a discussion of L1 categories as perceptual magnets, and Best, 1995, for a discussion of perceptual assimilation).

Numerous studies to date have demonstrated that the L1 indeed exerts a powerful influence on adults' ability to learn L2 sounds, and that adults' variation from native L2 speaker perception and production is often traceable to their L1 (e.g., Flege, Bohn, & Jang, 1997; see Best, 1995, and Flege, 1995, for reviews). By contrast, the influence of the L1 on children's ability to learn L2 sounds is often less apparent. Children overall appear more likely than adults to overcome L1 influences and to approximate native L2 speaker perception and production more closely (Baker & Trofimovich, 2005; Flege, MacKay et al., 1999; Flege, Yeni-Komshian et al., 1999). Thus, the IH proposes that child–adult differences in L2 phonological learning depend on the degree to which the L1 and L2 interact: the degree of interaction is weaker in younger learners and is more pronounced in older ones. Conceptualized as the extent to which L1 and L2 phonetic systems interact in learners, child–adult differences in L2 phonological learning can then be understood as a consequence of prior learning itself (Bever, 1981), a point we revisit below.

1.3
Tenets of the IH

1.3.1
L1 categories develop slowly

The IH hypothesis is founded on the observation that L1 categories develop slowly. Previous research has shown that L1 phonological learning is gradual, characterized by a slow attunement of L1 sound categories. Children learning their L1 display a strong preference for the sounds of their L1 early in life, demonstrating a change from language-general to language-specific processing of vowels at around six months (Kuhl et al., 1992) and consonants at around 10 months of age (e.g., Werker & Tees, 1984). However, it appears that children's categories for specific L1 sounds and their allophonic variants continue to become refined throughout childhood and perhaps into adolescence (Hazan & Barrett, 2000; Walley & Flege, 1999; but see Sussman & Carney, 1989). For example, native English children become increasingly more adept between the ages of 6 and 12 years at discriminating English consonants, although they do so less accurately than adults (Hazan & Barrett, 2000). When compared with adults, children also have been shown to display less sensitivity in their discrimination of L1 sounds (Flege & Eefting, 1986; Walley & Flege, 1999). Moreover, recent neurophysiological investigations exploring brainwave activity in response to speech stimuli have demonstrated a progressive, developmental change in how speech is processed across participants aged five years to adulthood (Cunningham, Nicol, Zecker, & Kraus, 2000; Sharma, Kraus, McGee, & Nicol, 1997). When taken together, these findings suggest that children's and even adolescents' L1 phonetic representations are often not fully adult-like.

1.3.2
Mature L1 categories mask L1–L2 perceptual differences

As L1 categories develop, they become more robust, leading to more accurate performance in L1 speech processing tasks by older adolescents and adults. The structure of the developing L1 sound categories, believed to be altered by the statistical distribution of speech sounds in the L1 (Kuhl, 1998; Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002), have been shown to lead to changes in phonological encoding (Best, 1994), selective attention to speech (Pisoni, Lively, & Logan, 1994), and auditory-perceptual sensitivities within L1 sound categories (Kuhl, 1998; Kuhl et al., 1992; Werker, 1994). One consequence of L1 category development, therefore, is that adults possessing fully-developed L1 categories may be less likely to detect between-language phonetic differences than children, and thus are more likely to perceive L2 sounds as instances of L1 sound categories (Guion, Flege, Akahane-Yamada, & Pruitt, 2000; Iverson et al., 2003). Moreover, a reason to suppose that children will be more likely to detect between-language phonetic differences is that their L1 sound system is still developing (Kuhl, 1991, 1998; Werker, 1994). Take, for example, the acquisition of English /ɹ/ and /l/ by Japanese adults. Japanese has just one liquid /ɾ/, whereas English has two, /ɹ/ and /l/. Japanese adults tend to identify both English liquids as instances of the one Japanese liquid, attending to acoustic properties that are useful in distinguishing Japanese consonants but not English /ɹ/ and /l/ (Best, 1995; Guion et al., 2000; Iverson et al., 2003). The perceived phonetic difference between Japanese /ɾ/ and English /ɹ/ is somewhat greater than that between Japanese /ɾ/ and English /l/, leading to somewhat better learning of English /ɹ/ (Guion et al., 2000). However, the overall perceived similarity of /ɹ/ to /ɾ/ appears to block category formation in most Japanese adults (Guion et al., 2000) but not children (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004).

1.3.3
Testing the IH

How might the IH be tested in an L2 study? According to the IH, children are less likely than adults to perceive L2 sounds as instances of L1 sound categories and, because of this, more likely to perceive and produce L2 sounds more accurately. Trofimovich, Baker, and Mack (2001) showed that Korean adults typically identify tokens of English /ʊ/ as instances of Korean /u/. If the IH is correct, then the Korean children who have received a similar amount of L2 input as the Korean adults should be less likely than the Korean adults to perceive English /ʊ/ as an instance of Korean /u/. It should also hold true that these same Korean children should be more accurate than the Korean adults in their perception and production of English /ʊ/.

1.4
The present study

Although the IH is well motivated on theoretical grounds, its predictions have not been extensively investigated. In fact, only a few known studies have demonstrated that children are less likely than adults to treat L2 sounds as unambiguous instances of a single L1 sound category. For example, Bond and Adamescu (1979; see also Butcher, 1976) showed that native English-speaking four-year-olds were more accurate than adolescents (11–13 years old) and adults in perceiving a difference between English plosive consonants and Hausa implosive consonants (unfamiliar to them). In particular, the four-year-olds but not the adolescents or the adults distinguished the foreign consonants from L1 consonants at significantly above-chance rates (72% accuracy). Suggestive as these findings are, they do not clearly demonstrate that these same children are also more likely to perceive and produce non-native sounds more accurately than adults, either initially or after several years of exposure to them. In other words, it is still unknown whether children's ability to perceive L1–L2 differences confers upon them an advantage over adults at perceiving and producing L2 sounds.

In the present study, our overall goal was, therefore, to link children's perception of the relationship between sounds in the L1 and L2 to the accuracy with which they produce and perceive L2 sounds. In Experiment 1, we tested whether children are indeed less likely than adults to perceive L2 sounds as instances of L1 sound categories. Such a finding would indicate that children's L1 sound categories do not interact with L2 sounds to the same degree as adults' L1 sound categories do. In particular, we asked native Korean children and adults to identify English vowels in terms of Korean vowel categories in a forced-choice identification task and to rate the English vowels for "goodness of fit" to the same Korean vowel categories. In Experiment 2, we examined how the interaction of children's and adults' L1 and L2 phonetic systems relate to their perception and production of English vowels. To this end, we tested the same Korean adults and children in perception and production tasks of the same English vowels.

2 Experiment 1

The purpose of Experiment 1 was to test the hypothesis that children are less likely than adults to perceive L2 sounds as instances of L1 sound categories in the initial stages of L2 learning. To this end, native Korean children and adults were asked to classify English vowels in terms of Korean vowel categories in a cross-language perceptual identification task. The stimuli were multiple natural tokens of English vowels. Upon hearing each token, the participants were required to classify it as an instance of one of the 10 vowels of standard Korean, and then to rate its goodness of fit to the selected category. An English vowel was considered to have been an instance of a Korean vowel category if most of its tokens were classified in terms of just one Korean vowel category. The hypothesis tested here was that the Korean adults would perceive the instances of each of the English vowels as members of a single Korean vowel category to a greater extent than would the Korean children.

2.1 Method

2.1.1 Participants

The 32 participants in Experiment 1 were native speakers of Korean who resided in Illinois (Urbana-Champaign or Chicago) at the time of the study. Half were children and half were adults. The children (9 males, 7 females), who had arrived in the U.S.A. at an average age of 10;0 (6;0–13;5) and had lived there for an average of 9 months (5–13), had a mean age of 10;10 years (7;5–14;0). Two of these children (one 12- and one 14-year-old) were slightly older than the rest. Although these two participants might be considered adolescents, we grouped them with children for two reasons. First, pre-adolescents and adolescents of this age, like younger children, process speech differently than adults do (e.g., Cunningham et al., 2000; Hazan & Barrett, 2000). Second, these two participants' response patterns were similar to those shown by the rest of the children. The adults (6 males, 10 females), who had arrived in the U.S.A. at an average age of 25;1 years (19;8–31;0) and had lived there for an average of 6 months (1–22), had an average age of 25;7 years (21;2–31;2).

The participants were selected on the basis of English-language experience and language proficiency. The goal was to recruit children and adults having as similar and as minimal an amount of exposure to English in the U.S.A. as possible. The participants, whose length of U.S. residence ranged between 1 and 22 months, thus represented L2 learners with relatively minimal amount of exposure to the L2 in the target country (see Tsukada et al., 2005, for a comparison of Korean children and adults with 3–5 years of U.S. residence). Using a 10-point scale (1 = *I don't speak any English*, 10 = *I am a native speaker of English*), the participants rated their ability to speak, read, write, and comprehend English. (At least one of the children's parents was present to help verify that the child's responses to the questions were accurate.) The obtained mean self-ratings were 4.3 (2–7) for the adults and 4.5 (2–7) for the children. At least two children and one adult rated themselves relatively high in English (i.e., a 7 on a 10-point scale). It is likely that these participants overestimated their English ability. An informal pre-experiment interview revealed that all participants were poor at listening and speaking. In fact, none could easily carry on a simple conversation in English, suggesting that all participants were beginning learners of English, at least with respect to their speaking and listening ability. The participants also estimated their percent of daily Korean use. The obtained mean ratings were 66% (30–80%) for the children and 61% (40–90%) for the adults. The children and the adults did not differ significantly in terms of self-rated L2 proficiency or use, $p > .05$.

Most of the adults were enrolled in a beginning ESL program at the University of Illinois; most of the children were enrolled in ESL classes at their public school. Prior to arriving in the U.S.A., the adults had studied English in Korea as part of regular classroom instruction for an average of 9.2 years (9–13) and had been first exposed to English at a mean age of 12;6 years (9;0–14;0). However, they had little exposure to native speaker input before arriving in the U.S.A. and, despite several years of language instruction, had little experience in listening or speaking English. Because of the participants' low proficiency in English, all instructions were either read by participants in Korean or given to them by a native Korean speaker.

2.1.2
Materials

This experiment focused on the contrast between English /i/–/ɪ/ (as in "beat" and "bit") and /u/–/ʊ/ (as in "boot" and "book"). Both pairs of vowels tend to be confused, in production and perception, by Korean learners of English (Flege et al., 1997). Three native English speakers (see below) produced all four target vowels in three phonetic contexts (b_t/k, n_t/k, h_d), yielding a series of monosyllabic English words (/i/: *beat*, *neat*, *he'd*; /ɪ/: *bit*, *knit*, *hid*; /u/: *boot*, *nuke*, *who'd*; /ʊ/: *book*, *nook*, *hood*). It was hoped that the multiple contexts would minimize the idiosyncratic effects of any single context on the results obtained here (see Strange, Akahane-Yamada, Kubo, Trent, & Nishi, 2001). English words containing four additional English vowels (/æ/: *bat*, *gnat*, *hat*; /ɛ/: *bet*, *net*, *head*; /ɑ/: *bought*, *not*, *hot*; /ʌ/: *but*, *nut*, *hut)* were also recorded as distractors in this and the following experiment. (For data pertaining to these vowels, see Baker, Trofimovich, Mack, & Flege, 2002 and Trofimovich, Baker, & Mack, 2001.)

Given that children's and adults' familiarity with individual lexical items influences measures of L2 perception and production (e.g., Walley & Flege, 1999), the participants were asked to complete a vocabulary knowledge task after the test session. Following Bradlow and Pisoni (1999), the participants rated their familiarity with each stimulus word, presented in written form, using a 7-point scale (1 = *I don't know the word*, 7 = *I know the word and can use it in a sentence*). The experimenter gave the participants an

Table 1

Native Korean children's and adults' mean word familiarity ratings (1–7 scale) for the stimulus words used in Experiments 1 and 2 (standard deviations in parentheses)

	English vowel							
Group	/i/		/ɪ/		/u/		/ʊ/	
Children	beat	5.6 (2.1)	bit	5.0 (2.4)	boot	6.1 (1.8)	book	6.8 (0.8)
	neat	4.4 (2.6)	knit	3.4 (1.8)	nuke	2.6 (2.0)	nook	2.8 (2.1)
	he'd	5.4 (2.2)	hid	5.9 (1.9)	who'd	5.2 (2.4)	hood	4.7 (2.1)
	bead*	4.4 (1.5)	big*	6.9 (0.3)	booed*	2.6 (1.9)	good*	6.6 (1.0)
	heat*	5.4 (2.2)			hoop*	4.4 (2.4)		
Adults	beat	7.0 (0.0)	bit	6.8 (0.6)	boot	6.4 (1.4)	book	7.0 (0.0)
	neat	6.8 (0.6)	knit	3.7 (2.5)	nuke	2.7 (1.7)	nook	2.5 (1.6)
	he'd	6.6 (0.8)	hid	6.2 (1.9)	who'd	6.5 (1.0)	hood	5.7 (1.3)
	bead*	4.3 (2.5)	big*	7.0 (0.0)	booed*	2.5 (1.9)	good*	7.0 (0.0)
	heat*	7.0 (0.0)			hoop*	3.2 (1.8)		

Note: Asterisks designate words used in the vowel production task in Experiment 2.

example of how to rate the words, offering the Korean translation of the scale labels. If reading seemed to be a problem (which rarely happened, but was more often true for children than adults), the experimenter read the word aloud. As shown in Table 1, some words (*book*, *boot*, *beat*) were more familiar than others (*nook*, *nuke*, *knit*) for both groups. *T*-tests, however, revealed a significant child–adult difference in familiarity ratings for just two words: *bit*, $t(30) = 2.45$, $p = .02$, and *neat*, $t(30) = 3.04$, $p = .01$.

The native English speakers who produced the stimulus words were all adult males from the Pacific Northwest region of the U.S.A. (mean age: 24.2) who spoke a variety of General American English. None were bilingual, and none had professional training in speech or linguistics. The test words were read three times each in a carrier phrase (*I say __ for you*) at a self-selected normal speaking rate and were recorded using a DAT recorder (Sony TCD-D8) and a Shure (SM10A) microphone. The utterances were sampled at 22.05 kHz, and the test words edited out. The first two authors selected what they judged to be the best two tokens of each test word produced by each talker. They did so by examining each of the tokens both auditorilly and acoustically. The resulting 36 test words (12 words × 3 speakers) were later normalized for peak intensity to reduce differences in perceived loudness.

2.1.3
Procedure

In this and the following experiment, participants were tested individually in a quiet room, using a personal computer and speech presentation software (Smith, 1997). The 144 stimulus words (36 test words × 2 repetitions, 36 distractors × 2 repetitions) were randomly presented over stereo headphones (Sennheiser HD 535). As in previous research, the participants provided two perceptual responses to each stimulus (Guion et al., 2000; Strange et al., 2001). They first identified each stimulus in terms of one of the 10 vowels of standard Korean (/i/, /o/, /ɨ/, /y/, /œ/, /ɛ/, /e/, /ʌ/, /a/, /u/). The Korean vowel response alternatives were displayed in Hangul characters on a computer screen. All participants were able to read Korean. Prior to the experiment, the participants were asked to say the names of the characters to ensure their familiarity with the characters and to minimize orthographic confusions among them. Although it might be argued that listeners in such a task compare L2 tokens to abstract L1 phonological representations or to specific exemplars of L1 sounds (see Pierrehumbert, 2003, for discussion), in either case they are likely relying on their existing "knowledge" of the L1 phonetic system (see Trofimovich et al., 2001).

The same stimulus was presented a second time 5 msec after the classification was provided. The participants' task on the second presentation was to rate the similarity of the token they had just heard in terms of its goodness of fit to the selected Korean category. They used a 7-point scale (1 = *sounded very dissimilar*, 7 = *sounded very similar*). Ratings of this kind are often obtained in conjunction with cross-language identification responses (Guion et al., 2000). They provide a gradient of similarity (or difference) for a perceptual match between an L2 sound token and the chosen L1 sound category. Typically estimated on an ordinal scale, goodness-of-fit judgments may in fact underlie cross-language identification responses, which suggests that listeners use the same information in each token for both perceptual identification and goodness-of-fit rating (Takagi, 1995). After the rating was given, the next stimulus was presented twice (once for a similarity judgment and once for a goodness-of-fit rating), and so on.

Before testing, the participants were given a 10-item practice session to familiarize them with the procedure. They were able to ask the researcher questions as they performed the practice. The researcher made sure that the participants clicked the buttons labeled with Korean vowels after the first presentation of each word and that they clicked the buttons labeled with goodness ratings (1 through 7) after the second presentation of each word. Because the researcher was familiar with the stimuli, the practice was guided at first. For example, the researcher would give the following directions, especially to the children, during the first two or three practice trials: "That word was *bead*. Which vowel in Korean does the vowel in *bead* sound like?" The researcher could tell from the responses chosen whether or not the child comprehended the purpose of the task.

2.2
Results

In the cross-language identification task, responses were scored by computing how many times each participant identified all instances of a particular target vowel category (English /i/, /ɪ/, /u/, /ʊ/) in terms of each of the 10 Korean vowels. For example, the number of times each participant identified English /i/ as an instance of Korean /i/ was tabulated, as well as the number of times each participant identified this same English vowel with the nine other Korean vowels. Then the primary and most frequent (modal) Korean vowel response alternative chosen in response to the nine tokens of each English vowel (3 contexts × 3 speakers) was determined for each participant. The dependent variable was the number of times (maximum = 9) with which this modal Korean response category was chosen for the nine tokens of each of the four English target vowels. The goodness-rating responses were scored by computing each participant's goodness rating for the match between each token of an English vowel

Table 2

Mean number of times (out of nine) native Korean children (C) and adults (A) identified each of the four English vowels with 10 Korean vowel response alternatives (modal responses in bold)

	Korean vowel response alternative									
English vowel	/i/	/y/	/u/	/ɨ/	/a/	/ʌ/	/o/	/e/	/ɛ/	/œ/
/i/	**C: 7.74**	C: .27	C: .18	C: .45		C: .36				
	A: 8.64	A: .09	A: .09	A: .09		A: .09				
/ɪ/	**C: 5.76**	C: 1.08	C: .18	C: .72				C: .63	C: .63	
	A: 7.11	A: .18		A: .54				A: .99	A: .18	
/u/	C: .54	C: .36	**C: 6.39**	C: 1.17					C: .54	
	A: .18	A: .18	**A: 7.74**	A: .90						
/ʊ/	C: .45	C: .27	**C: 4.68**	C: 1.35		C: 1.71	C: .54			
			A: 6.21	A: 1.17		A: 1.08	A: .54			

Note: Chance identification rate = .90.

and its Korean modal response alternative only. That is, goodness ratings were not calculated for instances when the modal response was not chosen.

The two dependent variables just described (number of modal responses, mean ratings) were examined in (2) Age × (4) Vowel analyses of variance (ANOVAs). Simple effects and interactions were explored using Fischer's LSD tests using an alpha level of .05. Effect sizes were estimated using partial eta squared, η_p^2, calculated by dividing the effect sum of squares by the effect sum of squares plus the error sum of squares. For *t*-tests, effect sizes are reported as *r*.

Mean identification frequencies for each English vowel are presented in Table 2. Both the Korean children and the adults selected a single Korean vowel as their primary and most frequent (modal) response alternative a majority of the time in response to all four English vowels (/i/, /ɪ/, /u/, /ʊ/). All participants without exception chose Korean /i/ more frequently than any other Korean vowel when classifying both English /i/ and English /ɪ/. Similarly, all participants, both children and adults, used Korean /u/ more frequently than any other Korean vowel to classify both English /u/ and English /ʊ/.

The ANOVA examining how often (maximum) each participant chose the modal Korean response alternative in response to each of the four English target vowels yielded a significant main effect of age, $F(1, 30) = 6.93$, $p = .013$, $\eta_p^2 = .19$, and vowel, $F(3, 90) = 31.21$, $p = .0001$, $\eta_p^2 = .51$, but not a significant two-way interaction, $F(3, 90) = .50$, $p > .50$, $\eta_p^2 = .02$. Tests of simple main effects revealed that the children differed statistically significantly from the adults in the number of times they perceptually identified three English vowels (/ɪ/, /u/, /ʊ/) with their modal Korean response alternatives (/i/, /u/, /u/, respectively). That is, the children mapped fewer tokens of each of these three English vowels onto their respective modal Korean response alternatives than the adults did, $ts(30) > 2.05$, $ps < .05$, $rs > .35$. However, the difference between the children's and the adults' responses was non-significant for English /i/, $t(30) = 1.71$, $p = .10$, $r = .30$, although there was a numerical trend in the same direction as for the other three English target vowels (see Table 2).

The average goodness ratings ranged from 4.1 to 5.5. The ANOVA examining these ratings yielded a significant main effect of vowel, $F(3, 90) = 7.86$, $p = .0001$, $\eta_p^2 = .21$. However, neither the main effect of age, $F(1, 30) = 1.45$, $p < .20$, $\eta_p^2 = .05$, nor the age × vowel interaction, $F(3, 90) = 1.20$, $p > .25$, $\eta_p^2 = .04$, reached significance. This suggested that the children, as a group, did not differ from the adults in their goodness ratings when evaluating the perceptual match between English vowels and their modal Korean response alternatives. However, a closer analysis of individual response patterns revealed that the children tended to rate all English vowels as being less similar to their modal Korean response alternatives than the adults did. For example, the goodness ratings associated with the 90th percentile (the value below which 90% of the ratings fall) were consistently lower for the children (/i/: 6.4; /ɪ/: 6.0; /u/: 5.9; /ʊ/: 5.6) than for the adults (/i/: 7.0; /ɪ/: 6.7; /u/: 6.7; /ʊ/: 6.5).

2.3 Discussion

The results of this experiment suggested that the children appear to perceive fewer tokens of English vowels as unique and unambiguous instances of a single Korean vowel category. These child–adult differences reached statistical significance for three of the four English vowels (/ɪ/, /u/, /ʊ/). The children also demonstrated a numerical trend to identify English /i/ with Korean /i/ less often than did the adults. This difference did not reach statistical significance perhaps because English and Korean /i/ are perceptually highly similar, making it more likely for these two sounds to be perceptually "confusable" (Trofimovich et al., 2001; Yang, 1996). However, this finding only obtained in the analyses of cross-language identification. Although there was a tendency for the children to give lower goodness ratings in response to all four English vowels, at the group level the children did not differ from the adults in how they rated the perceptual match between English and Korean vowels. One explanation for this discrepancy may be that the procedure of cross-language goodness rating used here may not have been sufficiently sensitive to detect differences in the perception of the relationship between English and Korean vowels. A rating scale that is more fine-grained than the one used in the present study could perhaps make it possible in future research to detect potential differences between children's and adults' perceptual goodness ratings (see Aoyama et al., 2004).

These findings would not provide support for the hypothesis that children are less likely than adults to perceive L2 sounds as instances of L1 sound categories in the initial stages of L2 learning if the obtained child–adult differences were developmental in nature; that is, if the differences were due to the children's inability to perform the experimental task as intended. To evaluate this possibility, three post-hoc analyses were performed.

We first examined the Korean response alternatives chosen by the children and the adults in response to each English vowel. If the children had been guessing or had been off task, they could have chosen, at least to a certain extent, implausible or inconsistent Korean response alternatives in responding to English vowels. Inconsistent or implausible response alternatives would include, for example, those Korean vowels that are both perceptually similar and those that are dissimilar to the English vowels. This analysis, however, revealed that both the children and the adults

Table 3

Mean number of Korean response alternatives chosen and mean response consistency indexes for each English vowel (standard deviations in parentheses)

	Number of response alternatives chosen English vowel				*Response consistency index English vowel*			
Group	/i/	/ɪ/	/u/	/ʊ/	/i/	/ɪ/	/u/	/ʊ/
Children	2.1 (1.6)	2.9 (.99)	3.0 (1.3)	3.6 (.97)	.99 (.01)	.97 (.02)	.97 (.02)	.96 (.01)
Adults	1.4 (.52)	2.6 (.84)	2.0 (.67)	2.9 (.88)	.99 (.01)	.98 (.01)	.98 (.01)	.97 (.01)

selected acoustically and perceptually viable Korean vowel categories in response to each English vowel. For example, in response to English /i/, both the children and the adults chose the same three response alternatives: Korean /i/ (the modal response alternative), Korean /ɨ/ (high unrounded central vowel), and Korean /y/ (high unrounded front vowel). Because these three Korean vowels are located in the same region of the vowel space as English /i/ (Yang, 1996), they were considered viable response alternatives (see Table 2).

To determine if the children were on task and performed the task as intended, we next calculated the number of Korean vowel response alternatives (out of 10) selected by each participant in response to each English vowel. If the children had been merely guessing or had been off task, they would have chosen numerically more types of vowel response alternatives than would the adults. For example, if a participant selected Korean /i/, Korean /ɨ/, and Korean /e/ in response to different tokens of English /i/, this participant's type frequency count was 3 in this analysis. A two-way ANOVA examining these values (see Table 3) yielded a significant main effect of vowel, $F(3, 90) = 12.57$, $p = .0001$, $\eta_p^2 = .41$, but neither the main effect of age, $F(1, 30) = 4.07$, $p = .06$, $\eta_p^2 = .18$, nor the two-way interaction, $F(1, 30) = .59$, $p > .25$, $\eta_p^2 = .03$, reached significance. This suggested that the children were not more likely than the adults to choose quantitatively more Korean response alternatives in response to any English vowel.

The final post-hoc analysis aimed to determine if response consistency was comparable for the children and the adults. Response consistency, RC, was estimated using the following formula: $RC = 1-((p \log_{10} (1/p)) + (1-p) \log_{10} (1/1-p))$, where p is the probability of response for a particular stimulus and the log base is the number of response alternatives (Attneave, 1959). This measure, which has been used in previous research examining children's speech processing (Hazan & Barrett, 2000; Ohde & Sharf, 1988), estimates intra-subject response reliability. If children had been merely guessing or had been off task, their responses would have been less consistent than those of the adults. The ANOVA examining the RC values (see Table 3) yielded a significant main effect of vowel, $F(3, 90) = 30.87$, $p = .0001$, $\eta_p^2 = .51$, but not a significant main effect of age, $F(1, 30) = 4.19$, $p = .06$, $\eta_p^2 = .12$, nor a two-way interaction, $F(3, 90) = .03$, $p > .50$, $\eta_p^2 = .01$. This suggested that the children and the adults demonstrated a comparable degree of consistency in assigning tokens of English vowels to Korean vowel categories.

Because the children were not more likely than the adults to choose inexplicable response alternatives, to select a different number of response alternatives, or to respond inconsistently, the findings reported earlier can be taken as support for the hypothesis that children are less likely than adults to perceive L2 sounds as instances of a single L1 sound category. If children are indeed less likely than are adults to perceive L2 sounds as instances of L1 sound categories, children might also perceive and produce L2 sounds more accurately than adults. The next experiment was conducted to test this possibility.

3 Experiment 2

The purpose of Experiment 2 was to determine whether children's and adults' judgments of cross-language similarity obtained in Experiment 1 for the four target English vowels (/i/, /ɪ/, /u/, /ʊ/) might predict the accuracy with which the Korean children and adults perceive and produce these same English vowels. Given that the children appeared to perceive the English vowels as more dissimilar from L1 vowels than the adults, it was hypothesized that children would also perceive and produce the English vowels (/i/, /ɪ/, /u/, /ʊ/) more accurately than would the Korean adults.

3.1 Method

3.1.1 Participants

Half of the Korean participants took part in this experiment immediately before Experiment 1, and the remaining half did so following Experiment 1. In addition, 16 adult (7 males, 9 females) and 16 child (9 males, 7 females) monolingual native speakers of American English, all residents of Illinois, participated. The native English children and adults had average ages of 10;8 years (8.0–13;10) and 24;0 years (20;0–32;0), respectively. The English children had not been exposed to any language other than English. The English adults had not studied another language beyond minimal training in high school or college language courses. Neither the Korean and English children, nor the Korean and English adults, differed significantly in chronological age, $p > .05$.

3.2 Procedure

3.2.1 Vowel perception

The target vowels used in the vowel perception task were the same as in Experiment 1 (/i/, /ɪ/, /u/, /ʊ/). These vowels were produced in a CVC context, yielding the English strings he'd, hid, who'd, and hood. The h_d phonetic context is frequently used in speech perception research (e.g., Hillenbrand, Getty, Clark, & Wheeler, 1995) because it tends to minimize contextual influences; namely, the fricative /h/ typically does not influence the acoustic properties of neighboring vowels in English (Olive, Greenwood, & Coleman, 1993). The three adult male monolingual native English speakers who were selected to record the words had all been born and raised in Illinois (mean age:

25;4 years), because the participants, both native and Korean, had been exposed to the variety of English spoken in Illinois. In selecting the specific tokens to be used as stimuli, we took care to identify tokens of the four vowel categories that did not differ by more than 10 msec in duration. Although duration does indeed provide a secondary perceptual cue for native English-speaking listeners to the contrasts between English /i/–/ɪ/ and /u/–/ʊ/, Flege et al. (1997) found that Korean adults tended to focus exclusively on vowel duration, largely ignoring the primary cue to these vowel distinctions, namely, spectral quality differences. Therefore, we chose vowel tokens that were as similar in duration as possible.

The selected stimuli (1 token per speaker) were presented in pairs to test the discrimination of /i/–/ɪ/ (*he'd–hid*), /u/–/ʊ/ (*who'd–hood*), and /i/–/u/ (*he'd–who'd*). It seemed likely that the Korean participants would have little difficulty discriminating the vowels in *he'd–who'd* because these words contained English vowels, /i/ and /u/, that were likely to be perceived as instances of two phonemically distinct Korean vowels, /i/ and /u/, respectively (Trofimovich et al., 2001). Thus, performance on this vowel contrast was intended to serve as a way to determine if the participants had understood and were performing the task as intended.

In the categorial discrimination task used here (Flege, MacKay et al., 1999; Gottfried, Jenkins, & Strange, 1985), the three stimuli presented in each triad were always physically different tokens, each spoken by a different native English speaker. A total of 24 triads were used to test each of the three vowel contrasts. Half of the trials were "change" trials in which one token was drawn from a different vowel category than the other two tokens. The remaining half of the trials were "no-change" trials containing three physically different instances of the same vowel category. In response to each, the participants were to click one of four buttons appearing on the computer screen. The correct response for change trials was the serial position of the odd item out (i.e., the button marked *1*, *2*, or *3*). The correct response to no-change trials was the fourth button, marked *none* (to indicate "no odd item out"). The inter-stimulus interval between the three tokens in each trial was 800 msec. The next trial (triad) was presented 1 second after a response was received for the preceding trial. The participants were permitted to replay the stimuli in a triad if they wanted, but could not change their response once it had been given. To ensure that participants understood the task, they performed a 10-item practice using non-test vowel tokens consisting of the same vowels in the test condition before beginning the vowel discrimination test.

The dependent variable for each vowel contrast was A', a measure used to reduce the effect of response bias. The A' scores were derived from the proportion of hits (correctly choosing the "odd" item on change trials) and false alarms (incorrectly choosing an "odd" item from among same items on catch trials) for each participant and each vowel contrast (Snodgrass, Levy-Berger, & Haydon, 1985). An A' score of 1.0 indicates perfect discrimination while an A' score of .5 or lower indicates discrimination at or below chance.

3.2.2
Vowel production

The vowel production task focused on the same four target English vowels: /i/, /ɪ/, /u/, /ʊ/. These vowels were produced in English CVC words (/i/: *beat*, *bead*, *heat*; /ɪ/: *bit*, *big*, *hid*;

/u/: *boot*, *booed*, *hoop*; /ʊ/: *book*, *good*, *hood*). These words were not identical to those used in Experiment 1, with some words (e.g., *who'd* and *neat*) being replaced by more picturable ones (e.g., *hoop* and *bead*, respectively) to facilitate elicitation in a picture-naming task. The stimulus words were recorded by a monolingual female native English speaker from Illinois (age: 31;1 years), and were prepared as in Experiment 1. As in Experiment 1, the Korean children and adults did not demonstrate systematic differences in their familiarity with the words. An examination of familiarity ratings (see Table 1) indicated that both the Korean children and adults were more familiar with some words (e.g., *book*, *big*, *good*) than others (*booed*, *bead*, *hoop*), but the ratings from the adults and children differed significantly for only one of the 12 words: *heat*, rated as more familiar by adults than children, $t(30) = 2.39$, $p = .025$.

The participants were asked to name each word (and those of 12 additional distractors) as simple line drawings of each were displayed. The 24 words were elicited in a random order three times each. During the first elicitation, the participants heard the name of the picture via a loudspeaker (Harman/Kardon) as the picture was displayed. For the final two elicitations, the participants were asked to remember the name of the picture and to say it upon seeing the picture. The participants were thus not merely shadowing (imitating) the female speaker but were attempting to retrieve their own phonological representation for each word. In the few (< 2%) instances in which participants were unable to name the picture, the aural model was presented. The participants' productions were recorded using high quality equipment (Shure SM10A microphone, Sony TCD-D8 DAT recorder) for later analysis.

The last spontaneously produced token of each stimulus word was digitized. The 768 stimulus words (12 words × 64 participants) were then normalized for peak intensity, and presented to 10 adult listeners (aged 18–25 years) who were monolingual native speakers of American English and had taken at least two courses in phonetics in which they had received practice in transcribing English speech materials using IPA symbols. The stimulus words were presented to the listeners in a randomized list one at a time over stereo headphones (Sennheiser HD 535) with an inter-trial interval of 1 second. The listeners were instructed to listen to each stimulus word and to select one of the 14 vowels of standard American English (/i/, /ɪ/, /u/, /ʊ/, /æ/, /ɛ/, /e/, /o/, /ɑ/, /ʌ/, /ɔ/, /aʊ/, /aɪ/, /ɔɪ/) that most closely corresponded to the vowel in the stimulus word. The listeners were not told which vowel the participants had produced nor what vowel was intended. The listeners were instructed to base their judgments upon the vowels that were actually spoken, not the vowels that they thought were intended in the stimulus words. The dependent variable was the number of listeners (maximum = 10) who transcribed each vowel token as its intended target.

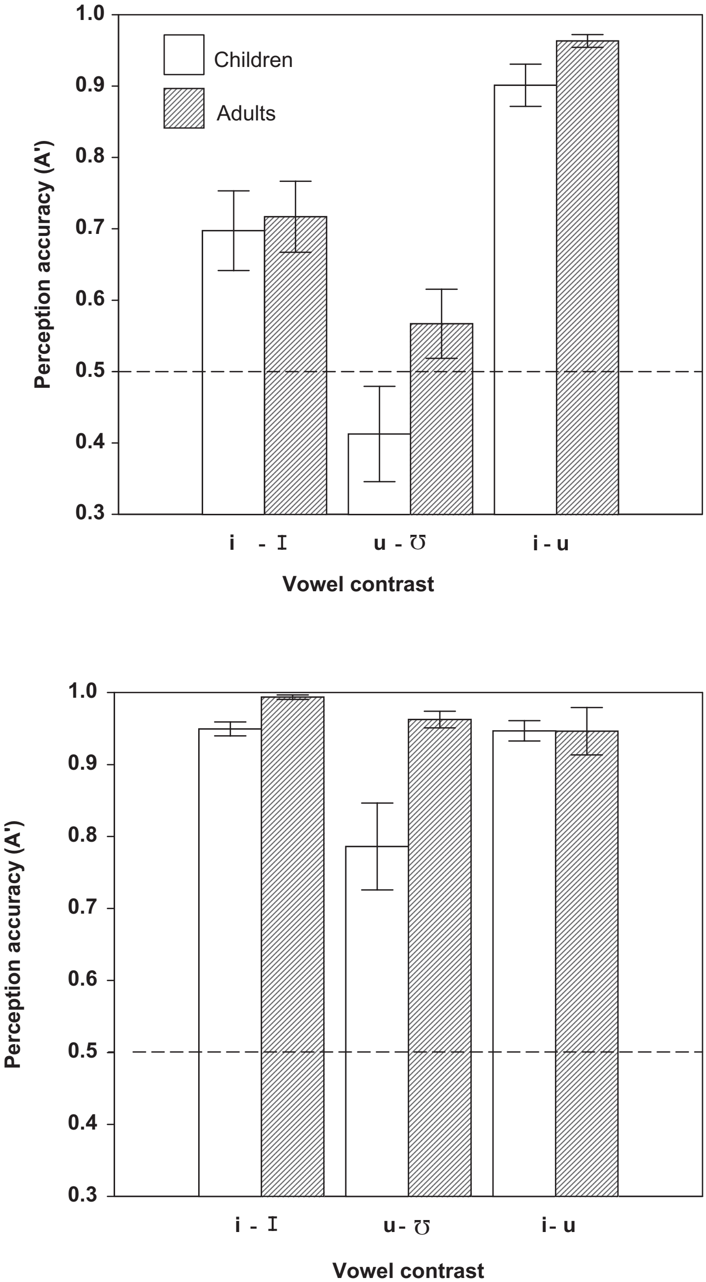
3.3 Results

3.3.1 Vowel perception

Figure 1 shows the average discrimination scores obtained for the three English vowel contrasts by the four groups of participants. These scores ranged between a high of .99 for the English adults' discrimination of /i/–/ɪ/ (mean hit rate: 10.6; mean false

Figure 1

Native Korean (top) and American English (bottom) children's and adults' accuracy in discrimination of English vowel contrasts. Brackets enclose ±1 SE. Dotted line represents chance performance



alarm rate: 1.3) and a low of .41 for the Korean children's discrimination of /u/–/ʊ/ (mean hit rate: 1.6; mean false alarm rate: 3.5). A preliminary analysis indicated, as expected, that the Korean adults and children did not differ significantly from age-matched native English speakers for /i/–/u/, $p > .10$. Nor did either the Korean or English children differ from either adult group, $p > .10$. This suggested that all the participants, including the Korean children, understood task directions and performed the task as intended.

The A' scores obtained for the remaining two contrasts were submitted to a mixed-design (2) Age × (2) Language × (2) Vowel contrast ANOVA, with repeated measures on the third factor. All three main effects reached significance: age $F(1, 60) = 8.33$, $p = .005$, $\eta_p^2 = .12$, language $F(1, 60) = 90.29$, $p = .0001$, $\eta_p^2 = .60$, and contrast $F(1, 60) = 28.51$, $p = .0001$, $\eta_p^2 = .32$. The analysis also yielded significant contrast × age, $F(1, 60) = 5.13$, $p = .027$, $\eta_p^2 = .08$, and contrast × language, $F(1, 60) = 4.17$, $p = .046$, $\eta_p^2 = .07$, interactions.

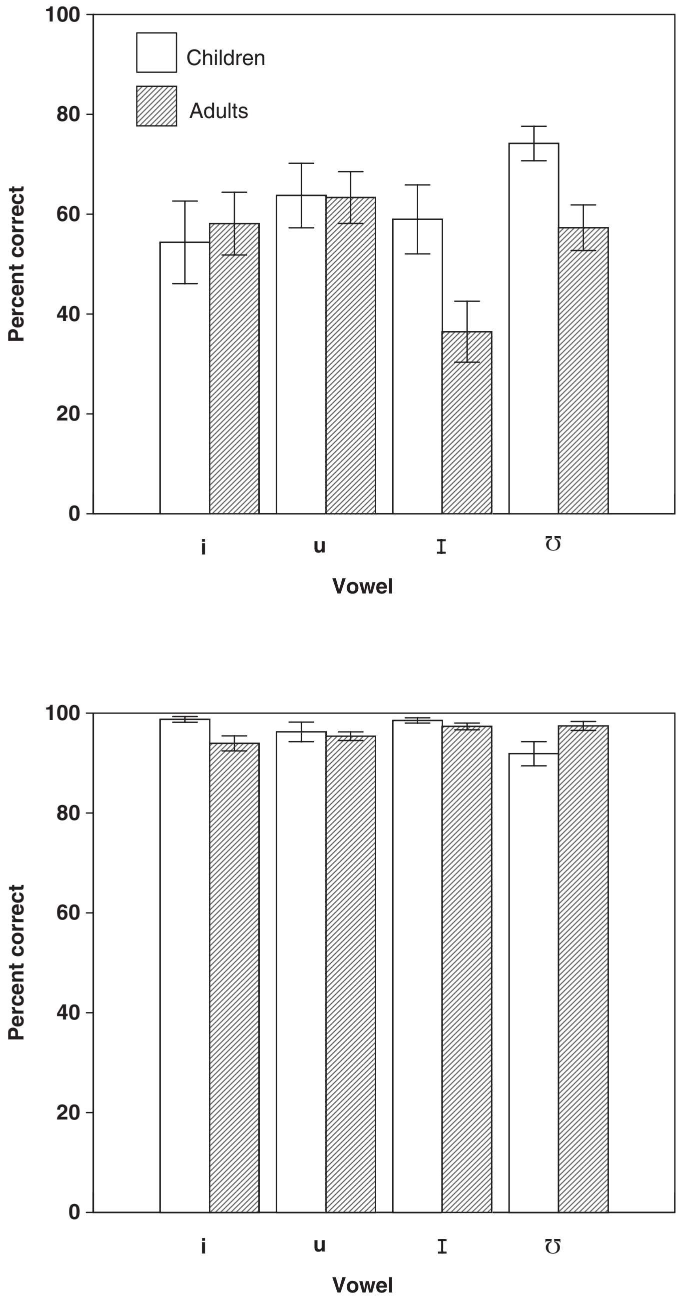
The interactions obtained in the above analysis were explored by testing the difference between the native English and Korean participants, and between the children and the adults. Between-language comparisons revealed that the native English speakers obtained significantly higher scores than did age-matched native Korean participants for both /i/–/ɪ/, $t(30) > 4.46$, $p < .001$, $r > .63$, and /u/–/ʊ/, $t(30) > 4.15$, $p < .001$, $r > .60$. Child–adult comparisons revealed higher scores for the native English adults than children for /i/–/ɪ/, $t(30) = 4.33$, $p < .001$, $r = .62$, and /u/–/ʊ/, $t(30) = 2.87$, $p < .007$, $r = .46$, but no significant differences between Korean children and adults for either contrast, $t(30) < 1.87$, $p > .07$, $r < .32$.

3.3.2
Vowel production

Figure 2 shows the mean intelligibility scores obtained for /i/, /ɪ/, /u/, /ʊ/ (i.e., the percentage of times the vowels were heard as intended by the 10 native English-speaking listeners). These scores ranged between a high of 99% correct for the English children's production of /i/ and a low of 37% correct for the Korean adults' production of /ɪ/. When the listeners misidentified the intended vowel, they did so a great majority of the time by selecting the other vowel in each vowel contrast (i.e., /i/ for /ɪ/, and vice versa, or /u/ for /ʊ/, and vice versa). Other vowels chosen in place of the intended targets accounted for less than 8% of the responses. The mixed-design (2) Age × (2) Language × (4) Vowel ANOVA examining these scores yielded significant main effects of age, $F(1, 60) = 4.42$, $p = .04$, $\eta_p^2 = .07$, and language, $F(1, 60) = 291.17$, $p = .0001$, $\eta_p^2 = .83$, as well as significant vowel × language, $F(3, 180) = 4.66$, $p = .004$, $\eta_p^2 = .07$, and vowel × age × language, $F(3, 180) = 3.16$, $p = .026$, $\eta_p^2 = .05$, interactions. The interactions appear to have been due to between-vowel differences as a function of L1 background. Between-language comparisons revealed that all four vowels produced by the English children and adults received higher scores than those produced by the age-matched native Korean groups, $t(30) > 2.66$, $p < .012$, $r > .44$. Child–adult comparisons revealed higher scores for the English adults than children for just one of the four target vowels, (/i/), $t(30) = 3.04$, $p = .005$, $r = .49$, but not for the others, $t(30) < 1.04$, $p > .31$, $r < .19$. These comparisons also revealed that the intelligibility scores were higher for the Korean children than the adults for /ɪ/ and /ʊ/,

Figure 2

Native Korean (top) and American English (bottom) children's and adults' accuracy in production of English vowels. Brackets enclose ±1 SE



*t*s(30) > 2.61, *p*s < .014, *r*s > .43, but not for /i/ and /u/, *t*s(30) < .65, *p*s > .52, *r*s < .12. This final result was unlike the findings reported in the vowel discrimination task in which the Korean children and adults did not differ in their perception accuracy.

3.4 Discussion

These results suggest that, following a relatively short (< 1 year) period of residence in the U.S.A., the Korean children did not discriminate English vowels more accurately than the Korean adults, but did produce two of the four English vowels examined more accurately than the Korean adults. One might conclude from this that, contrary to the claim of the Speech Learning Model (Flege, 1995), segmental production in an L2 develops more rapidly than segmental perception. However, as discussed by Tsukada et al. (2005, p. 286), there is reason for caution in reaching such a conclusion. The perception task (categorial discrimination) may have been too difficult for the children and the adults (e.g., Walley & Flege, 1999). To obtain a perfect score in this task, it was necessary for participants to disregard irrelevant within-category variation (i.e., differences in speaker voices) while focusing on relevant between-category variation (i.e., differences in vowel quality) across vowel tokens. The high scores (> .95, maximum = 1.0) obtained by the native English adults suggested that listeners having mature language-specific categories can do both. After only several months in contact with English in the U.S.A., one would not expect either Korean adults or children to have developed robust perceptuomotor categories for English vowels, and thus to have obtained native-like discrimination scores for pairs of English vowels that do not readily map onto vowels in the L1 vowel inventory. By contrast, vowel production, assessed in an ostensibly easier picture description task, was more sensitive to detecting the children's initial advantage at accurately producing at least some L2 vowels.

One interesting finding of this experiment, likely reflective of differences between the two tasks, pertains to the relative difficulty of certain L2 contrasts for children and adults. In the perception task (see Figure 1), the Korean adults attained higher accuracy than the Korean children in their perception of the /u/–/ʊ/ contrast (although this difference failed to reach statistical significance). However, these same adults did not differ from the Korean children in their perception of the /i/–/ɪ/ contrast. (Note a similar pattern in the English children's and adults' responses.) One factor likely contributing to these differences between contrasts relates to the frequency with which both contrasts occur in English. Indeed, English /i/ and /ɪ/ occur in more contrastive environments than English /u/ and /ʊ/, which may cause the /u/–/ʊ/ contrast to be more difficult for children to learn (Edwards et al., 2004). For adults, who have learned to distinguish these vowels through drills and other explicit methods of learning, such a difference in lexical frequency between the two contrasts may be less apparent.

The main finding of this experiment was, however, that the Korean children outperformed the Korean adults in production of two English vowels. This finding is striking because children frequently perform more poorly than adults on a variety of linguistic (both L1 and L2) and non-linguistic perceptuomotor tasks (Hazan & Barrett, 2000; Snow & Hoefnagel-Höhle, 1978). In fact, the English children in this experiment were less accurate than the English adults in their perception of two vowel contrasts and in their production of one vowel. Flege et al. (2006), who compared English sentences spoken by Korean adults and children to those spoken by native

English-speaking adults and children, also reported significantly lower ratings for native English children than for adults (see also Johnson, 2000). By contrast, the Korean children tested here did not differ from the Korean adults in their perception of any vowel contrast examined and, in fact, outperformed the Korean adults in production of two English vowels (/ɪ/, /ʊ/), those that do not exist in the vowel inventory of Korean. This finding may represent one of the very few demonstrations of children outperforming adults on L2 speech tasks in relatively early stages of L2 phonological learning, a finding that may not have been revealed by earlier studies (e.g., Snow & Hoefnagel-Höhle, 1978) because these studies did not compare children's and adults' L2 perception and production as a function of L1–L2 perceptual similarity.

4 Comparison of Experiments 1 and 2

The IH (Flege, 1992, 1999; Walley & Flege, 1999) offers at least one explanation for the findings of this study. As discussed above, two predictions based on the IH were evaluated here: (1) children are less likely than adults to perceptually associate L2 sounds with L1 sounds and, (2) because of this, children are more likely to perceive and produce L2 sounds accurately. The experiments reported above yielded partial support for both predictions. The children, compared with the adults, appeared to perceive fewer tokens of three of the four English vowels examined in this study as unique and unambiguous instances of a single L1 sound category. The children also equaled the adults in perception of both vowel contrasts examined, and even surpassed them in production of two of the four target vowels.

At least one way to interpret these findings would be to posit a relationship between children's being less likely than adults to perceive L2 vowels as instances of a single L1 vowel category and their ability to perceive and produce L2 vowels accurately (Aoyama et al., 2004; Guion et al., 2000). This relationship was examined in a series of Spearman rank-order correlation analyses. The correlations were computed (a) between the Korean participants', $n = 32$, cross-language identification rates from Experiment 1 and their production-accuracy scores for each of the four English vowels from Experiment 2, and (b) between cross-language identification rates and production-accuracy scores, on the one hand, and the participants' age of exposure to English (AOE) and their length of residence (LOR) in the U.S.A., on the other. The use of both AOE and LOR in these analyses was not arbitrary. AOE, as an index of the participants' age at the time of exposure to English, was a variable of primary interest here. In turn, LOR was an estimate of the participants' amount of exposure to English. Because the Korean children had resided in the U.S.A. slightly longer than the Korean adults (9 vs. 6 months, on average), it was possible that the findings of this study might have been influenced, at least in part, by differences in the participants' LOR. Cross-language similarity-rating scores and discrimination (perception) accuracy rates were not used in these analyses because the groups of Korean children and adults did not differ on these variables for any of the examined vowels.

The correlation analyses yielded four findings. First, there appeared to be no obvious relationship between the participants' cross-language identification judgments and their production of English vowels. In fact, only one significant, albeit relatively weak, correlation emerged in these analyses (for English /u/, $r = .30$, $p < .05$). Second, there appeared to be a relationship between the participants' cross-language identification judgments and their AOE. This finding was based on weak-to-medium-strength positive correlations between AOE and cross-language identification rates obtained for all four English vowels, *r*s = .30 – .53, *p*s < .05, indicating that those participants who were younger at the time of their exposure to English in the U.S.A. were also those who were less likely to perceive English vowels as instances of a single Korean vowel category. Third, AOE appeared to be related to the participants' production of only two English vowels (/ɪ/, /ʊ/), that is, those vowels that are not part of the Korean phonemic vowel inventory (although /ɪ/ is an allophone of /i/ in Korean). This finding was based on medium-strength negative correlations between the participants' AOE and their production of English /ɪ/ and /ʊ/, *r*s = –.38 – –.46, *p*s < .05, suggesting that those participants who were younger at the time of their exposure to English were also those who were more accurate in their production of these vowels. Finally, there was no significant association between the participants' LOR and their cross-language identification judgments or their production of English vowels. This finding helped rule out differences in participants' LOR as a factor explaining the obtained findings. In sum, the correlation analyses indicated that the degree to which children and adults perceive L2 vowels as instances of L1 vowel categories, as indexed by cross-language identification rates, was not directly related to L2 production accuracy; however, both these variables were related to AOE, at least for English /ɪ/ and /ʊ/.

One way to conceptualize these findings is to suggest that age at the time of exposure to the L2 (indexed here by AOE) serves as a variable mediating the relationship between the two abilities: being able to perceive L1–L2 phonetic differences and being able to accurately produce L2 sounds. In other words, an individual's age at the time of exposure to the L2 appears to determine the degree to which these abilities are related. In younger learners, they are related more closely, allowing these learners to perceive L1–L2 phonetic differences and to accurately produce L2 sounds, particularly those that are not present in the L1 sound inventory (i.e., English /ɪ/, /ʊ/). In older learners, producing L2 sounds and perceiving cross-language phonetic differences represent abilities that are associated loosely (if at all). Whatever the exact relationship between L2 production and perception of L1–L2 differences, a point which should be investigated further, it appears that both abilities enable children to be more successful than adults in L2 phonological learning.

5 General discussion

The purpose of the present study was to examine one explanation, based on the IH of Flege's Speech Learning Model (although the Speech Learning Model was not constructed to account for beginning learners such as those tested in this study), for why children are often more successful than adults at L2 learning. The obtained results in part provided evidence in support of the IH. In Experiment 1, the Korean children were found to be less likely than the Korean adults to perceive English vowels as instances of a single Korean vowel category. In Experiment 2, the same children were more accurate than the adults in production of two of the four English vowels, notably, those absent from the L1 vowel inventory, and equaled them in discrimination of both examined English vowel contrasts. Both skills—being able to perceptually disassociate

L2 sounds from L1 sound categories and being able to produce L2 sounds—appear to be associated with the age at which the children and the adults were first exposed to the L2, indicating that age may mediate the relationship between these two skills. Taken together, these findings suggest that children, perhaps due to the developmental state of their L1 at the time of their exposure to the L2, are able to perceptually distinguish L2 sounds from similar L1 sounds and to produce and (perhaps) perceive such L2 sounds more accurately than adults (Aoyama et al., 2004; Flege, 1987, 1999, 2002, 2003; Guion et al., 2000).

This age-bound relationship between the ability to perceive L1–L2 differences and L2 perception and production should be examined in future research in greater detail, with a goal of identifying and describing cognitive and neurobiological factors underlying this relationship. One such factor may pertain to neurobiologically based age-related changes in the plasticity of brain structures underlying language learning and use. That is, younger L2 learners, as compared to older L2 learners, may rely on different language processing and learning mechanisms and/or may draw on different brain structures in the course of language learning and use. For example, in a functional magnetic resonance study, Kim et al. (1997) documented that the brain areas involved in the processing of L1 and L2 overlapped in younger learners but did not do so in older learners. In another study, Mechelli et al. (2004) demonstrated that younger learners appeared to have a larger composition of grey matter in areas of the brain devoted to language processing than did older learners.

Although revealing, the results of such studies are often difficult to interpret (see, e.g., Perani et al., 1998) primarily because they do not indicate *exactly* how differences in the scope and localization of brain activity in younger and older learners relate to child–adult differences in the ability to perceive cross-language differences or to perceive and produce L2 sounds. To begin investigating questions such as this, future research should include both measures of children's and adults' ability to perceive cross-language differences and neuropsychological measures of age-related changes in children's and adults' L2 speech processing (see Cunningham et al., 2000).

In the meantime, before valid neuropsychological evidence is available to document age-related changes in children's and adults' L2 speech processing, the IH provides one possible account of child–adult differences in L2 learning. This account is compatible with the "interference" explanations of adults' difficulties in acquiring an L2 (Bever, 1981; Flege, 1992, 1999; Iverson et al., 2003; McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002; Pallier et al., 2003). Such explanations posit that the act of prior learning itself, as opposed to age-bound neurobiological limitations *alone*, imposes constraints on individuals' ability to learn new information. Thus, for example, adults' difficulties in L2 learning may be traceable to age-based developmental processes that render speech perception and production mechanisms more specialized for the processing of L1 input and therefore functionally autonomous (Bever, 1981). Likewise, adults' difficulties in L2 learning may be attributable to perceptual distortions or to a loss of perceptual sensitivities due to adults' extensive prior experience in processing L1 input (Iverson et al., 2003; McCandliss et al., 2002; Pallier et al., 2003). Such perceptual interferences are believed to be "self-reinforcing" (Iverson et al., 2003: B54), imposed on adults through their long-lasting exposure to and experience with their L1.

Whether or not the results obtained in the present study represent such interference explanations, the results reported here lend support to the claim that child–adult differences in L2 phonological learning are at least in part the result of the still-developing phonetic system of children's L1. Children thus appear to have an initial advantage over adults, an advantage that may become even greater with time given subsequent differences in the quality and quantity of input children and adults receive (Flege, Yeni-Komshian et al., 1999; Jia & Aaronson, 2003). Whatever the ultimate cause of such child–adult differences, it is hoped that this study invites further investigation into child and adult L2 speech learning in particular, and into the processes underlying language learning in general.

References

ABU-RABIA, S., & KEHAT, S. (2004). The critical period for second language pronunciation: is there such a thing? Ten case studies of late starters who attained a native-like Hebrew accent. *Educational Psychology*, **24**, 77–98.

AOYAMA, K., FLEGE, J. E., GUION, S. G., 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.

ATTNEAVE, F. (1959). *Application of information theory to psychology: A summary of basic concepts, methods, and results*. New York: Holt, Rinehart and Wilson.

BAKER, W., & TROFIMOVICH, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, **48**, 1–27.

BAKER, W., TROFIMOVICH, P., MACK, M., & FLEGE, J. E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings of the Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla Press.

BEST, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: the transition from speech sounds to spoken words*. Cambridge, MA: MIT Press.

BEST, C. T. (1995). A direct-realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp.171–204). Timonium, MD: York Press.

BEVER, T. G. (1981). Normal acquisition processes explain the critical period for language learning. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp.48–95). London: Newbury.

BIRDSONG, D., & MOLIS, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, **44**, 235–249.

BLAUSDEL, C. G., MITCHELL, D. E., MUIR, D. W., & PETTIGREW, J. D. (1977). A physiological and behavioural study in cats of the effect of early visual experience with contours of a single orientation. *Journal of Physiology*, **265**, 615–636.

BOND, Z., & ADAMESCU, L. (1979). Identification of novel phonetic segments by children, adolescents and adults. *Phonetica*, **36**, 182–186.

BRADLOW, A. R., & PISONI, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *Journal of Acoustical Society of America*, **106**, 2074–2085.

BUTCHER, A. (1976). *The influence of native language on the perception of vowel quality*. Arbeitsberichte des Instituts für Phonetik. Institut für Phonetik: Kiel University.

CUNNINGHAM, J., NICOL, T., ZECKER, S., & KRAUS, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: development and behavioral correlates of perception. *Ear & Hearing*, **21**, 554–568.

EDWARDS,. J., BECKMAN, M. E., & MUNSON, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, **47**, 421–436.

FIFKOVA, E. (1968). Changes in the visual cortex of rats after unilateral deprivation. *Nature*, **220**, 379–381.

FLEGE, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, **15**, 47–65.

FLEGE, J. E. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications* (pp. 565–604). Timonium, MD: York Press.

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

FLEGE, J. E. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp.117–137). Mahwah, NJ: Lawrence Erlbaum.

FLEGE, J. E. (2002). Interactions between the native and second-language phonetic systems. In. P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp.234–257). Trier: Wissenschaftilcher Verlag.

FLEGE, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: differences and similarities* (pp. 319–345). Berlin: Mouton de Gruyter.

FLEGE, J. E., BIRDSONG, D., BIALYSTOK, E., MACK, M., SUNG, H., & TSUKADA, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, **34**, 153–175.

FLEGE, J. E., BOHN, O.-S., & JANG, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, **25**, 437–470.

FLEGE, J. E., & EFTING, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica*, **43**, 155–171.

FLEGE, J. E., MacKAY, I. R. A., & MEADOR, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, **106**, 2973–2987.

FLEGE, J. E., YENI-KOMSHIAN, G. H., & LIU, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, **41**, 78–104.

GATBONTON, E., & TROFIMOVICH, P. (2008). The ethnic group affiliation and L2 proficiency link: Empirical evidence. *Language Awareness*, **17**, 229–245.

GOTTFRIED, T. L., JENKINS, J. J., & STRANGE, W. (1985). Categorial discrimination of vowels produced in syllable context and in isolation. *Bulletin of the Psychonomic Society*, **23**, 101–104.

GUION, S. G., FLEGE, J. E., AKAHANE-YAMADA, R., & PRUITT, J. C. (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**, 2711–2724.

GUION, S. G., FLEGE, J. E., LIN, S. H. YENI-KOMSHIAN, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, **21**(2), 205–228.

HAKUTA, K., BIALYSTOK, E., & WILEY, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, **14**, 31–37.

HAZAN, V., & BARRETT, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, **28**, 377–396.

HILLENBRAND, J., GETTY, L. A., CLARK, M. J., & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *Journal of Acoustical Society of America*, **97**, 3099–3111.

HIRSH, K. W., MORRISON, C. M., GASET, S., & CARNICER, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, **6**, 117–128.

IVERSON, P., KUHL, P. K., AKAHANE-YAMADA, R., DIESCH, E., TOHKURA, Y., KETTERMANN, A., ET AL. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, **87**, B47–B57.

JIA, G., & AARONSON, D. A. (2003). Longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, **24**, 131–61.

JOHNSON, C. E. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, **43**, 144–157.

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

KIM, K., RELKIN, N., LEE, K.-M., & HIRSCH, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, **388**, 171–174.

KUHL, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, **50**, 93–107.

KUHL, P. K. (1998). The development of speech and language. In T. C. Carew, R. Mezel, & C. J. Shatz (Eds.), *Mechanistic relationships between development and learning* (pp.53–74). New York: Wiley.

KUHL, P. K., WILLIAMS, K. A., LACERDA, F., STEVENS, K. N., & LINDBLOM, B. (1992). Language experience alters phonetic perception in infants by 6 months of age. *Science*, **255**, 606–608.

LENNEBERG, E. (1967). *Biological foundations of language*. New York: Wiley.

MACK, M. (1990). Phonetic transfer in a French–English bilingual child. In P. H. Nelde (Ed.), *Language attitudes and language conflict* (pp.107–124). Bonn: Dümmler.

MACK, M. (2003). The phonetic systems of bilinguals. In M. Banich & M. Mack (Eds.), *Mind, brain, and language: multidisciplinary perspectives* (pp.309–349). Mahwah, NJ: Lawrence Erlbaum.

MAYE, J., WEISS, D. J., & ASLIN, R. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, **11**, 122–134.

MAYE, J., WERKER, J. F., & GERKEN, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**(3), B101–B111.

McCANDLISS, B. D., FIEZ, J. A., PROTOPAPAS, A., CONWAY, M., & McCLELLAND, J. L. (2002). Success and failure in teaching the r-l contrast to Japanese adults: predictions of a hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, and Behavioral Neuroscience*, **2**, 89–108.

MECHELLI, A., CRINION, J. T., NOPPENEY, U., O'DOHERTY, J., ASHBURNER, J., FRACKOWIAK, R. S., ET AL. (2004). Structural plasticity in the bilingual brain. *Nature*, **431**, 757.

MOORMAN, S. J., CORDOVA, R., & DAVIES, S. A. (2002). A critical period for functional vestibular development in zebrafish. *Developmental Dynamics*, **223**, 285–291.

MUNRO, M. J., FLEGE, J. E., & MacKAY, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, **17**, 313–334.

NEWPORT, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, **14**, 11–28.

OHDE, T. N., & SHARF, D. J. (1988). Perceptual categorization and consistency of synthesized /r-w/ continua by adults, normal children and /r/-misarticulating children. *Journal of Speech and Hearing Research*, **31**, 556–568.

OLIVE, J. P., GREENWOOD, A., & COLEMAN, J. (1993). *Acoustics of American English speech: A dynamic approach*. New York: Springer-Verlag.

OYAMA, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, **5**, 261–283.

PALLIER, C., BOSCH, L., & SEBASTIÁN-GALLÉS, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, **64**, B9–B17.

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.

PERANI, D., PAULESU, E., SEBASTIÁN-GALLÉS, N., DUPOUX, E., DEHAENE, S., BETTINARDI, V., ET AL. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, **121**, 1841–1852.

PIERREHUMBERT, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, **46**, 115–154.

PISONI, D. B., LIVELY, S. E., & LOGAN, J. S. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp.121–166). Cambridge, MA: MIT Press.

ROZIN, P. (1976). The evolution of intelligence and access to the cognitive unconscious. *Progress in Psychobiology and Physiological Psychology*, **6**, 245–280.

SHARMA, A., KRAUS, N., McGEE, T. J., & NICOL, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant–vowel syllables. *Electroencephalography and Clinical Neurophysiology*, **104**, 540–545.

SMITH, S. C. (1997). *UAB software*. Department of Rehabilitation Sciences: University of Alabama at Birmingham.

SNODGRASS, J. G., LEVY-BERGER, G., & HAYDON, M. (1985). *Human experimental psychology*. London, OUP.

SNOW, C. E., & HOEFNAGEL-HÖHLE, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, **49**, 1114–1128.

STEVENS, G. (1999). Age at immigration and second language proficiency among foreign-born adults. *Language in Society*, **28**, 555–578.

STRANGE, W., AKAHANE-YAMADA, R., KUBO, R., TRENT, S., & NISHI, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, **109**, 1691–1704.

SUSSMAN, J. E., & CARNEY, A. E. (1989). Effect of transition length on the perception of stop consonants by children and adults. *Journal of Speech and Hearing Research*, **32**, 151–160.

TAKAGI, N. (1995). Signal detection modeling of Japanese listeners' /r/-/l/ labeling behavior in a one-interval identification task. *Journal of the Acoustical Society of America*, **97**, 563–574.

TROFIMOVICH, P., & BAKER, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, **28**, 1–30.

TROFIMOVICH, P., BAKER, W., & MACK, M. (2001). Context- and experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences*, 31, 167–186.

TSUKADA, K., BIRDSONG, D., BIALYSTOK, E., MACK, M., SUNG, H., & FLEGE, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, **33**, 263–290.

WALLEY, A. C., & FLEGE, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, **26**, 307–331.

WERKER, J. F. (1994). Cross-language speech perception: Development change does not involve loss. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp.93–120). Cambridge, MA: MIT Press.

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.

YANG, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, **24**, 245–261.