Production benefits of childhood overhearing

Leah M. Knightly
Department of Psychology, University of California at Los Angeles, Los Angeles, California 90095

Sun-Ah Jun[a)]
Department of Linguistics, University of California at Los Angeles, Los Angeles, California 90095

Janet S. Oh
Department of Psychology, University of California at Los Angeles, Los Angeles, California 90095

Terry Kit-fong Au
Department of Psychology, University of Hong Kong, Hong Kong

(Received 15 August 2002; accepted for publication 7 April 2003)

The current study assessed whether overhearing Spanish during childhood helps later Spanish pronunciation in adulthood. Our preliminary report based on a subset of the data [Au *et al.*, Psychol. Sci. **13**, 238–243 (2002)] revealed that adults who overheard Spanish during childhood had better Spanish pronunciation, but not better morphosyntax, than adult learners of Spanish who had no childhood experience with Spanish. We now present data from the full sample with additional morphosyntax and pronunciation assessments, as well as measures to help rule out possible confounding prosodic factors such as speech rate, phrasing, and stress placement. Three groups of undergraduates were compared: 15 Spanish–English bilinguals (native Spanish speakers), 15 late learners of Spanish who overheard Spanish during childhood (childhood overhearers), 15 late learners of Spanish who had no regular experience with Spanish until middle or high school (typical late L2 learners). Results confirmed a pronunciation advantage for the childhood overhearers over the typical late L2 learners on all measures: phonetic analyses (VOT and degree of lenition), accent ratings (phoneme and story production), but no benefit in morphosyntax. Importantly, the pronunciation advantage did not seem attributable to prosodic factors. These findings illustrate the specificity of overhearers' advantage to phonological production. [DOI: 10.1121/1.1577560]

PACS numbers: 43.70.Ep, 43.70.Fq, 43.71.Hw [AL]

I. INTRODUCTION

Phonology is difficult for late second-language (L2) learners to master (Oyama, 1976; Tahta *et al.*, 1981; Flege and Fletcher, 1992; Flege *et al.*, 1995, 1999). Late learners' difficulty in producing nativelike accents in their L2 may in part be due to perceptual deficits (Flege, 1995). In fact, perceptual training in identification of L2 sounds seems to help late L2 learners pronounce the sounds of the target language better (Bradlow *et al.*, 1997, 1999).

Childhood exposure to the target language seems to benefit adult L2 perception and perhaps production as well. Case studies (Wode, 1981; Yamada, 1995) suggest that children learning English during visits to the U.S. maintained English-like production and perception after a two-year absence. Tees and Werker (1984) found that English-speaking adults who lived in a Hindi-speaking environment for the first year or two of their lives perceived Hindi phonemic contrasts reliably better than those who had no childhood experience with Hindi. However, it remains unclear what *type* of language experience (e.g., hearing, speaking) contributed to these perceptual and production benefits.

Infants learn much about the phonology of their ambient language simply by hearing it. They form language-specific vowel categories by 6 months of age and display ''language-specific phonetic'' perception by 10–12 months (Werker and Tees, 1984; Kuhl *et al.*, 1992; Werker, 1995). Impressively, 8-month-old infants learn word boundaries in a continuous speech stream after 2 min of exposure, using only statistical cues—where the only cue available for word boundaries is that sounds within a word are more likely to co-occur than sounds across word boundaries (Saffran *et al.*, 1996; Johnson and Jusczyk, 2001). Furthermore, such learning seems to be incidental rather than conscious and explicit. Saffran *et al.* (1997) asked adults and 6- and 7-year-old children to create pictures on a computer (cover task) while an artificial language played in the background. Even though participants were told they would not be tested on the acoustic material presented (i.e., the artificial language), both adults and children nonetheless demonstrated on a later test that they learned words of the artificial language after only 20 min of ''incidental'' exposure to it. Saffran *et al.* concluded that such ''incidental'' learning could be important in natural language acquisition.

Yet, to date few language acquisition studies have examined ''incidental'' learning *per se* outside the laboratory setting. Our study tries to fill this gap by focusing on ''incidental'' language learning in a *natural* language setting (i.e., overhearing everyday conversations). In an interim report of the current study, we explored possible effects of overhear-

[a)]Author to whom correspondence should be addressed. Electronic mail: jun@humnet.ucla.edu

ing Spanish during childhood (Au *et al.*, 2002). In that report, we compared three groups of English-speaking college students on their acquisition of Spanish phonology and morphosyntax. They included 10 Spanish-English bilinguals whose first language was Spanish (native speakers), 11 native English-speakers who regularly overheard Spanish during childhood but first learned Spanish in class around age 14 (childhood overhearers), and 12 native English-speakers who had minimal exposure to Spanish prior to learning it in class around age 14 (typical late L2 learners).

Compared to the typical late L2 learners, childhood overhearers' pronunciation of Spanish /p, t, k, b, d, g/ turned out to be more nativelike, as assessed by phonetic measures and accent ratings. Specifically, their VOT (voice-onset-time) for /p, t, k/ was reliably shorter for word initial position (e.g., *tacos*), thus closer to the norm of Spanish VOT, a short-lag VOT approximately 30–50 ms shorter than the long-lag VOT of English /p, t, k/ (Lisker and Abramson, 1964). Overhearers also produced Spanish intervocalic /b, d, g/, which are typically voiced and lenited (i.e., fricatives or approximants) but remain stops and are sometimes devoiced in English, more often as lenited consonants compared to the typical late learners. Importantly, overhearers' English pronunciation was not compromised; their pronunciation of English /p/ and /b/ was not reliably different from that of the typical late L2 learners.

Interestingly, the benefits of childhood overhearing did not extend to the area of morphosyntax. Au *et al.* found that childhood overhearers performed no better than typical late L2 learners in detecting morphosyntactic errors in a grammaticality judgment task or in their production of correct number and gender agreement in noun phrases. These interim findings suggest that simply hearing the language during childhood does not seem to benefit the mastery of morphosyntax, which like phonology is also challenging to late L2 learners (Snow and Hoefnagel-Hohle, 1978; Patkowski, 1980; Johnson and Newport, 1989; Johnson *et al.*, 1996; Flege *et al.*, 1999; Birdsong and Molis, 2001).

Although Au *et al.* (2002) found no benefits in morphosyntax, a more comprehensive assessment may yet show an overhearing advantage in this area. Additionally, the benefit in phonology reported by Au *et al.* could be due to confounding factors of prosody such as speech rate, placement of stress, and phrasing. Voiceless stop consonants tend to have longer VOT in slow speech, stressed syllables, and phrase initial position compared to those in fast speech, unstressed syllables, and phrase medial position, respectively (de Jong, 1995; Fougeron and Keating, 1997). Spanish intervocalic /b, d, g/ often become stops after a pause (Stockwell and Bowen, 1965), and slow speech typically contains more pauses than fast speech. In Au *et al.*'s study, participants were asked to produce word initial and medial /p, t, k, b, d, g/ in stressed syllables of Spanish words embedded in the carrier phrase *Diga* ___ *por favor* (meaning "Say ___ please"). Overhears could have produced more nativelike (i.e., shorter) VOT for word initial /p, t, k/ if the typical late learners happened to produce the carrier sentence more slowly than the overhearers, or if the overhearers misplaced stress on the Spanish words more often than the typical late learners (thereby producing more target phonemes without stress). Similarly, typical late learners' less frequent lenition of Spanish /b, d, g/ could have resulted from less fluent speech and hence more pauses before the target phonemes.

To reevaluate Au *et al.*'s (2002) interim findings, we report here the originally planned study in its entirety with the full sample, along with more in-depth phonetic analyses (degree of voicing in voiced stops, measurement of prosodic factors, as well as VOT and stop closure of voiceless stops, and lenition of voiced stops), additional accent rating data (narrative production), and additional assessment tasks in morphosyntax (verb morphology and story telling). In sum, the study presented here takes a more comprehensive look at whether there is a childhood overhearing advantage specific to phonology and not apparent in morphosyntax.

II. METHOD

A. Participants

Undergraduates enrolled in second year Spanish language courses at UCLA were recruited to complete a detailed language background questionnaire and follow-up interview about their experience with Spanish from birth to the time of testing. All participants completed a consent form and were paid for their participation. Out of the 238 participant responses received, 15 were identified as *childhood overhearers* and 100 were identified as *typical late L2 learners*. From the latter group, 15 were randomly selected with the constraint that they matched the *childhood overhearers* by gender and Spanish language instructor. Agreement between two trained coders on group assignment (childhood overhearer versus typical late L2 learner versus unclassifiable) was excellent (91% agreement; Cohen's Kappa$=0.90$). An additional 15 UCLA undergraduate native Spanish speakers were recruited for the *native speaker* group.[1] Three groups of 15 speakers (ten women, five men) made up the final sample.

Table I presents demographics and several language use measures for these three participant groups. Childhood overhearers and typical late L2 learners were all born in the U.S. and first learned Spanish in middle or high school. The overhearers reported overhearing Spanish during childhood from a parent/relative regularly and being spoken to in Spanish and speaking Spanish minimally (limited to short phrases and words in Spanish). Their amount of self-reported experience overhearing Spanish during ages 0–6 differed reliably from zero [$t(14)=2.33$, $p<0.05$] whereas their being addressed to in Spanish and speaking Spanish during ages 0–6 did not [t's$(14)<1.58$, n.s.]. Typical late L2 learners had minimal, if any, childhood exposure to Spanish. Native speakers were Spanish–English bilinguals who learned English as their L2 before age 10, and most of them were born in the U.S. Both native speakers and childhood overhearers were of Mexican or Central American descent. Overhearers were mostly mixed-Latino, with Spanish-English bilingual parents and relatives, whereas native speakers were mostly full-Latino, with monolingual Spanish-speaking parents and relatives. Typical late L2 learners were of non-Latino descent, with monolingual English-speaking parents and relatives.

TABLE I. Summary of speaker demographics and language use measures.

	Native speakers	Childhood overhearers	Typical late L2 learners
First language	Spanish	English	English
Age (years)	22.3[a] (0.7)	19.9[b] (0.6)	18.7[b] (0.3)
Hrs/wk heard Spanish			
age 0–6	33.0[a] (0.0)	5.3[b] (2.3)	0[c] (0)
age 6–12	30.1[a] (2.0)	1.8[b] (0.7)	0.13[b] (0.09)
Years taken Spanish classes			
middle/high school	2.4[a] (0.3)	3.9[b] (0.3)	4.4[b] (0.3)
college	0.9[a] (0.4)	1.0[a] (0.2)	0.6[a] (0.1)
No. visits to a Spanish-speaking country	6.7[a] (1.5)	1.5[b] (0.5)	0.1[b] (0.09)
Reported % use of Spanish (vs English)			
during high school years	41.7[a] (3.5)	9.3[b] (2.7)	9.9[b] (2.9)
during college years	30.4[a] (3.8)	7.5[b] (1.5)	8.0[b] (1.6)
Reported degree of (Phinney, 1992)			
identification with ethnic group	3.9[a] (0.04)	3.2[b] (0.1)	3.3[b] (0.2)
participation in Latino practices	3.5[a] (0.1)	2.9[b] (0.1)	2.4[c] (0.2)
Slang task performance (% correct)	80.9[a] (2.6)	19.8[b] (2.4)	2.7[c] (0.9)

Note. Table indicates group means with standard errors in parentheses. Means with different superscripts within a row were reliably different from each other according to Tukey's HSD test, $p<0.05$. Phinney Ethnic Identity Measures (1992) were on a four-point scale with higher numbers indicating more identification and participation.

Childhood overhearers and typical late L2 learners' self-reports were corroborated by reports from independent informants who knew the participants' prior experience with Spanish (e.g., parents). Reports on 40% of the research participants confirmed childhood overhearers' regular passive exposure to Spanish and limited spoken Spanish (single words, short phrases) and typical late L2 learners' lack of childhood exposure to Spanish. For further corroboration, participants' knowledge of Mexican/Central American household childhood expressions was tested to assess their childhood exposure to Spanish in the home. Participants read 20 English expressions (e.g., *cry baby, pacifier, dry crust in eyes*)—one at a time—on a computer screen and were asked to translate them into informal Spanish as they heard them at home, in the neighborhood, or in a schoolyard (slang production). They also heard 40 Spanish slang terms (e.g., *chiqueado* meaning ''spoiled child;'' *las escondidas* meaning ''hide-and-seek'') via a headset and were asked to translate them into English (slang comprehension). Participants' responses were audio-recorded and later independently transcribed and coded by two research assistants who were native Spanish speakers (average agreement between coders $=96\%$, disagreements were resolved by a third native Spanish speaker). As seen in Table I, the results corroborated participants' self-reports of childhood experience with Spanish quite well, suggesting that the childhood overhearers knew far less Spanish childhood slang than the native speakers, but nonetheless knew reliably more than the typical late L2 learners.

B. Phoneme production assessment

1. Stimuli and procedure

As in Au *et al.* (2002), pronunciation of Spanish phonemes /p, t, k, b, d, g/ was assessed. Twelve categories were created (3 places of articulation$\times$2 types of voicing$\times$2 positions in word) with three target words per category and each word containing the target phoneme in a stressed syllable (e.g., *base* ''base,'' *cabeza* ''head;'' see Table II for a complete list). Participants were asked to say each of the 36 target words in the sentence frame, ''*Diga* (target word) *por favor*,'' meaning ''Say (target word) please,'' thus producing 36 target sentences.

To see if overhearing Spanish during childhood might compromise speakers' English pronunciation, participants were asked to produce English voiceless consonant /p/ (in *pepper, pocket*) and voiced consonant /b/ (in *beggar, bonnet*) in the sentence frame, ''Take a (target word) once again.''

In order to lessen the potential prosodic confounds discussed earlier, participants were asked to stress the target word in each sentence and avoid pausing between words. Participants were given two practice sentences before reading the target sentences. Each sentence was presented three times in random order, and displayed on the computer screen for three seconds.

Instructions and stimuli for all tasks were presented on a Macintosh Powerbook G3 or 3400c/200 using PsyScope (Cohen *et al.*, 1993), with auditory stimuli presented via a headset, and participants responding via a button box. Participants were tested individually in a soundproof room, and their utterances were recorded using a Sennheiser HMD 25-1 microphone headset and a Marantz PMD-222 or PMD-430 professional recorder.

2. Measurement

Speech recordings were digitized at 12.5 kHz, and phonetic measurements were made using KAY Elemetrics speech analysis programs *CSL* and *MultiSpeech*. All measurers/coders (one primary, two secondary) were blind to the speakers' language backgrounds and analyzed roughly the same percentage of speakers from each of the three groups (native speakers, childhood overhearers, typical late

TABLE II. Spanish word list.

Target phoneme	Word initial	English gloss	Word medial	English gloss
/b/	base	base	cabeza	head
	beca	scholarship	jabón	soap
	beso	kiss	sabor	taste
/d/	datos	data	nadar	swim
	deja	to leave	pedido	an order
	día	day	rodar	roll
/g/	gallo	rooster	hogar	home
	gato	cat	pagó	paid
	goma	glue	pegó	hit
/p/	pase	to pass/pass	zapeta	diaper
	pena	embarrassed	vapor	vapor
	peso	weight	tapón	stopper
/t/	tacos	tacos	matar	to kill
	teja	shingling/weave	metido	it's in
	tía	aunt	notar	to notice
/k/	callo	a blister	tocar	to touch
	caso	case/pay attention	sacó	took out
	coma	command to eat/eat	pecó	sinned

L2 learners). To assess agreement, two speakers analyzed by each secondary measurer/coder were randomly selected to be analyzed by the primary measurer/coder.[2]

a. VOT and stop closure. For /p, t, k/, VOT was measured from the stop release to the onset of the second formant (F2) of the following vowel. Stop closure duration was measured from the offset of F2 of the vowel before the target phoneme to the release of the target phoneme. Mean differences between the primary and secondary measurer(s) for individual speakers were minimal for VOT (ranging from 2.6 to 3.6 ms for Spanish; 2.3 to 2.7 ms for English) and stop closure duration (ranging from 3.9 to 13.9 ms for Spanish; 2.1 to 2.3 ms for English).

b. Lenition. Spanish voiced consonants /b, d, g/ were categorized as either stops or lenited consonants, distinguished by an abrupt or a gradual change, respectively, in amplitude between the consonant and the following vowel. Percent agreement on these categorical assessments was high, ranging from 91% to 100% (Cohen's Kappas between 0.70 and 1.0).

c. Voicing. To assess the production of voiced consonants more comprehensively, we examined the degree of voicing during the consonant by categorizing all target consonants as having one of three voicing types: voiceless, partial voicing, or full voicing. For stops, voicing type was determined by voicing during the closure; for lenited consonants, by voicing during the duration of the consonant. Tokens were classified as ''voiceless'' if there was no voicing (i.e., no voice bar in spectrogram), as ''full voicing'' if voicing was present throughout the entire consonant duration, and as ''partial voicing'' if they displayed some voicing but not full voicing. Percent agreement on voicing among the coders was high for Spanish, ranging from 82% to 94% (Cohen's Kappas between 0.78 and 0.92) and for English, 83% to 92% (Cohen's Kappas 0.71 to 0.80).

d. Prosodic factors (Speech rate, phrasing, and stress placement). Speech rate was assessed by the duration from the F2 onset of the vowel after the release burst of the first word in the carrier sentence (e.g., /i/ in ''diga,'' /e/ in ''take a'') to the F2 offset of the vowel before the target word (e.g., /a/ in ''diga,'' /ə/ in ''take a''). Mean differences between measurers on individual speakers for speech rate were on average 9.4 ms for Spanish and 2.6 ms for English.

To measure phrasing differences, a coder listened to each Spanish token and judged whether a phrase boundary had been inserted before the target phoneme (subjective measure). As a quantitative measure a ratio of stop closure duration to speech rate duration was calculated for each token categorized as a stop (see Sec. II B 2 b). Since stop consonants in phrase initial position tend to have longer closure duration, most of the tokens judged to have a prosodic boundary had a ratio around 0.75, and those without around 0.50. Using this criterion, all tokens were reevaluated so that final classification was based on ratio (objective measure); only tokens with a ratio of 0.75 or greater were classified as having a phrase boundary before the target sound.

For stress misplacement, a coder listened to each Spanish token and noted whether speakers placed stress on the incorrect syllable (e.g., correct—*PAse*, incorrect—*paSE*).

C. Accent ratings

As in Au *et al.* (2002), we asked another group of native Spanish speakers to rate the participants' pronunciation of the target phonemes /p, t, k, b, d, g/ in the target sentences (phoneme accent ratings). To assess their accents in more natural speech, we asked yet another group of native Spanish speakers to rate participants' accents in narrative production (narrative accent ratings).[3]

a. Phonemic accent ratings. Forty-eight native Spanish speakers[4] were recruited to rate participants' pronunciation of the target phonemes in the 36 target sentences (e.g., ''*Diga*

pase por favor.''). The second of the three tokens produced by each speaker for each of the 36 target sentences was selected for rating. To avoid rater fatigue, 36 blocks were created whereby each block consisted of the same target sentence spoken by all speakers. Blocks were arranged into three sets so that each set included 12 blocks of target sentences containing the six target phonemes (/p, t, k, b, d, g/) in word-initial and medial position. In this way, three tokens of a category (e.g., tokens *pase*, *pena*, *peso* of word-initial /p/ category) were dispersed among the three sets (see ''stimuli and procedure,'' Sec. II B 1). Each rater was asked to rate each speakers' pronunciation of the target sound (e.g., the /p/ in *pase*)[5] for one of the three sets. Prior to rating, raters listened to each speaker say, ''*Diga teja por favor*,'' to familiarize themselves with the range of speakers' pronunciation abilities. A rating scale was presented on the computer screen during the familiarization and accent rating trials [1 = very strong foreign accent, definitely non-native; 2 = strong foreign accent; 3 = noticeable foreign accent; 4 = slight foreign accent; 5 = no foreign accent, definitely native; adopted from Bongaerts *et al*. (1997)]. Sentences within blocks were randomized and the test was self-paced. Interrater reliability was assessed using all of the ratings to compute, for each rater, an average rating for each speaker. An intraclass correlation on these averages revealed excellent agreement (average accent scores for individual participants: intraclass $R = 0.98$, $p < 0.0001$).

b. Narrative accent ratings. To assess participants' accents in more natural speech, we elicited narratives using a 10-page abridged version of a wordless children's picture book in Mercer Mayer's ''Frog, Where are you?'' series—a widely-used task for eliciting narratives from children and adults in different languages (e.g., Berman *et al*., 1994). Participants were given 2 min to scroll through the storybook pictures on a computer before coming up with a story. They then saw each page for 12 s, allowing time to say one or two sentences in Spanish per page. Audio recordings were independently rated by two native Spanish speakers[6] using the same scale adopted for phonemic accent ratings. The intraclass correlation between the two raters was excellent ($R = 0.91$).

D. Morphosyntax assessment

In addition to the grammaticality judgment task and noun-phrase production task reported in Au *et al*., the narrative production task just described and a verb-phrase production task were included to yield a more comprehensive picture of participants' mastery of Spanish grammar. For all tasks reported in this section, stimuli are available upon request.

a. Grammaticality judgment task. Participants listened to 33 grammatical–ungrammatical sentence pairs spoken by a native Spanish speaker. They heard the sentences in random order; they heard each sentence twice and pressed a button to indicate whether it was grammatical or ungrammatical. Both decision and response time were recorded. To minimize fatigue, the 66 sentences were presented in two blocks intermixed with other tasks. Ungrammatical sentences contained an error in one of the following categories: number/gender agreement in noun phrases (e.g.,**la flores*,**el carro blanca*), number agreement in verbs (e.g.,**Marta corren*,**mi mamá toman*), tense-aspect marking in verbs (e.g.,**Dentro de cuatro años*,**soy un abogado*), negation (e.g.,**El conoce a nadie*), indirect object (e.g.,**El enseña a nosotros*), or person agreement (e.g.,**nosotros comienzan*,**A qué hora llegué usted*).

b. Noun-phrase production task. Participants were asked to verbally complete five simple four-piece jigsaw puzzles designed to elicit four combinations of number and gender markers [adapted from Plann (1979)]. Each puzzle appeared on the computer screen for 18 s with four puzzle pieces and a puzzle frame (showing numbered spaces for the pieces). For example, pieces in one puzzle depicted two white pianos (*los pianos blancos*: plural masculine), two white cows (*las vacas blancas*: plural feminine), a black piano (*el piano negro*: singular masculine), and a black cow (*la vaca negra*: singular feminine). To complete the puzzle properly, participants had to specify the number and gender of the nouns used for naming the puzzle pieces (e.g., ''Pon los pianos blancos en cuatro, pon la vaca negra en tres,...'' meaning ''Put the white pianos in four, put the black cow in three,...''). Two native speakers of Spanish independently transcribed the audiotaped responses and coded them for number and gender agreement. A third native speaker resolved any discrepancies between the two transcribers/coders. Percent agreement between transcribers/coders was greater than 95%.

c. Verb-phrase production task. The task was adapted from Curtiss and Yamada's (1987) CYCLE test to elicit verb morphology (tense, aspect, person, and number) production. Participants heard 20 incomplete sentences, one at a time, illustrated with pictures presented on a computer. They were asked to offer sensible completions. For instance, they might hear ''Ayer fuí a la tienda, y yo...'' (meaning ''Yesterday I went to the store, and I...'') and see a picture of someone in a store. Next they would see a picture of the person buying milk. They were then given 6 s to complete the sentence. To be counted as an acceptable completion, appropriate morphosyntactic markers had to be used for the verb, as constrained by the lead-in clause (e.g., first person and singular form in the preterite tense in Spanish in the example just given). The 20 items were designed to elicit a variety of tense/aspect, number, and person markings. Agreement between the two transcribers/coders on participants' audiotaped responses was excellent (>90% agreement), and discrepancies were resolved by a third native speaker of Spanish.

d. Narrative production. Each frog story produced for the narrative accent rating task (see Sec. II C b) was rated by two native speakers of Spanish on a grammatical well-formedness scale, with 1 = definitely nonnative and 5 = definitely native. Interrater reliability was excellent (intraclass $R = 0.90$, $p < 0.0001$).

III. RESULTS

A. Phonology

1. Spanish phonemes

a. Prosodic factors. One-way ANOVAs revealed no reliable differences between the three speaker groups in any of

TABLE III. Prosodic factors results.

Measure	Native speakers	Childhood overhearers	Typical late L2 learners
Speech rate (ms)			
word initial	209 (5)	203 (7)	196 (7)
word medial	195 (6)	203 (7)	198 (8)
Misplaced stress (%)			
word initial	0.5 (0.4)	1.5 (0.5)	2.2 (1.1)
word medial	5.6 (2.0)	7.2 (1.9)	10.4 (2.5)
Phrase boundary before target word (%)			
word initial	16.2 (5.8)	25.2 (7.0)	15.7 (5.0)
word medial	7.8 (2.0)	8.6 (2.7)	7.2 (2.2)

Note. Standard errors are given in parentheses.

the prosodic factors examined [speech rate, percentage of misplaced stress, and percentage of prosodic boundary before the target word; F's$(2,42)<1.30$, n.s.; see Table III]. These findings suggest that any differences found among the three speaker groups in VOT, percent lenition, and voicing cannot be attributed to these prosodic factors. Rather, the phonetic values seem to reflect the degree of mastery of Spanish phoneme production.

b. Voiceless consonants. Childhood overhearers produced word initial /p, t, k/ with shorter VOT (i.e., more nativelike) than did the typical late learners. An ANOVA with place of articulation (i.e., bilabial/alveolar/velar) and consonant position (i.e., word initial/medial) as within-subject factors and speaker group as a between-subject factor on VOT revealed a main effect of speaker group [$F(2,42)=6.28$, $p<0.01$], place [$F(2,84)=149.91$, $p<0.001$], and a reliable position by group interaction [$F(2,42)=8.44$, $p<0.01$]. No other reliable main effect or interactions were found.

Figure 1 shows the mean VOT for word initial and medial /p, t, k/ for each speaker group. The typical late L2 learners produced reliably longer VOT in word initial position than did both the native speakers and overhearers (by HSDs, p's<0.01), but the three groups did not differ reliably in VOT for word medial /p, t, k/, thereby yielding the reliable group by position interaction. The main effect of place of articulation was due to longer VOT as the target sounds moved from a bilabial to velar position for all three groups (by Bonferroni, p's<0.001).

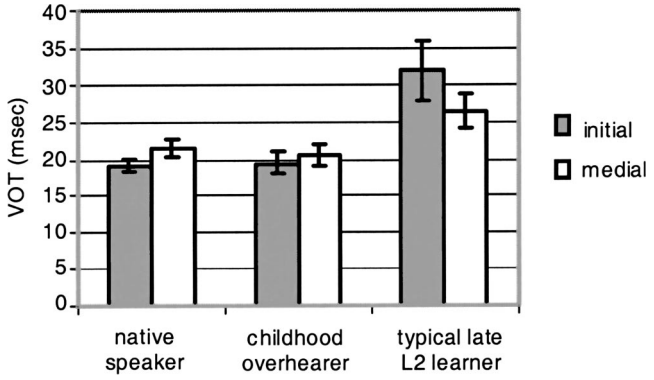

FIG. 1. Mean VOT duration of Spanish word initial and medial /p, t, k/ for each speaker group.

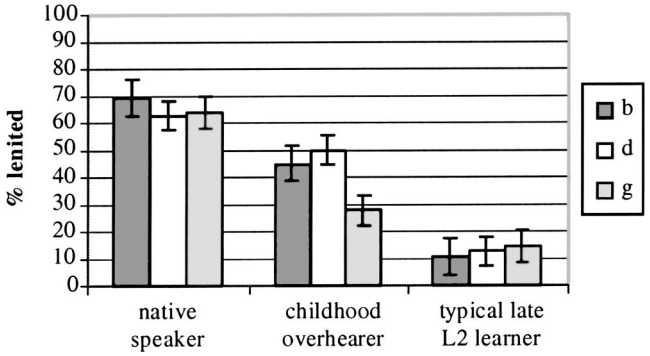

FIG. 2. Mean percentage use of lenited consonants in each place of articulation (i.e., /b/, /d/, /g/) for each speaker group.

One-way ANOVAs on speakers' stop closure duration for initial and medial /p, t, k/ revealed no reliable effects [F's$(2,42)<2.18$, n.s.].

c. Voiced consonants: degree of lenition. Speakers' tendency to produce /b, d, g/ as lenited consonants was computed based on the total number of phonetically voiced tokens for a particular phoneme (e.g., /b/, /d/, or /g/) and the number of tokens produced as lenited consonants. An ANOVA was then performed on these lenition percentages with place of articulation and position in word as within-subject factors, and speaker group as a between-subject factor. There were reliable main effects of speaker group [$F(2,42)=24.03$, $p<0.001$], position [$F(1,42)=74.78$, $p<0.001$], and place [$F(2,84)=3.76$, $p<0.05$], and a reliable interaction between place and group [$F(4,84)=5.65$, $p<0.01$].

Bonferroni *posthoc* tests showed that all speaker groups produced reliably more lenited consonants in word medial position ($M=52.9\%$, s.e.$=3.4\%$) than in word initial position ($M=26.4\%$, s.e.$=3.5\%$; $p<0.001$). As can be seen in Fig. 2, the place by group interaction was primarily due to overhearers being less likely to produce /g/ as a lenited consonant than /b/ and /d/ [t's$(14)=2.65$, p's<0.05]. This trend was not seen among the native speakers or the typical late L2 learners. Additional follow-up ANOVAs revealed that native speakers outperformed overhearers in producing intervocalic lenited /b/s, who in turn outperformed the typical late L2 learners [HSDs, p's<0.05; $F(2,42)=20.84$, $p<0.001$]. For intervocalic /d/, the native speakers and overhearers were comparable, and both outperformed the typical late L2 learners [HSDs, p's<0.001; $F(2,42)=21.91$, $p<0.001$]. In contrast, overhearers were no different from the typical late L2 learners in producing intervocalic lenited /g/s, and both did worse than the native speakers [by HSDs, p's<0.001; $F(2,42)=18.98$, $p<0.001$].

d. Voiced consonants: degree of voicing. Degree of voicing was assessed by the percentage of /b, d, g/ tokens produced by each speaker in each of the three voicing categories (i.e., voiceless, partial voicing, and full voicing), with the total number of /b, d, g/ tokens as the denominator. One-way ANOVAs revealed reliable group differences for partial [$F(2,42)=7.85$, $p<0.01$] and full voicing [$F(2,42)=6.40$, $p<0.01$] but not for voiceless. Tukey *posthoc* tests

TABLE IV. Accent rating assessment results.

Rated speech type	Native speakers	Childhood overhearers	Typical late L2 learners
Phonemic			
/p, t, k/	4.4[a] (0.08)	3.6[b] (0.08)	3.0[c] (0.09)
/b, d, g/	4.4[a] (0.10)	3.4[b] (0.10)	2.8[c] (0.10)
Narrative	5.0[a] (0.0)	3.0[b] (0.23)	2.4[c] (0.17)

Note. Accent ratings were made using a five-point scale, with higher ratings indicating more nativelike pronunciation. Standard errors are given in parentheses. Within a row, means with different superscripts were reliably different from each other according to Tukey's HSD test with $p<0.001$ for phonemic ratings, and with $p<0.05$ for narrative ratings.

revealed that native speakers produced /b, d, g/ with full voicing ($M=81.7\%$, s.d.$=15.2\%$) more often than typical late L2 learners ($M=55.5\%$, s.d.$=21.0\%$; by HSD, $p<0.01$). Overhearers' percent use of full voicing ($M=70.5\%$, s.d.$=23.4\%$) was in between these two groups, although not reliably different from either. Native speakers used partial voicing less often than typical late L2 learners ($M=12.0\%$, s.d.$=3.0\%$; $M=29.5\%$, s.d.$=3.7\%$, respectively; by HSD, $p<0.01$). Again, overhearers were in between these two groups but differed from neither reliably ($M=13.7\%$, s.d.$=3.6\%$).

2. English phonemes

English production data for word-initial /p/ and /b/ were collected from 14 native speakers, 12 childhood overhearers, and 14 typical late L2 learners.[7] One-way ANOVAs revealed no reliable differences among the three groups in speech rate, closure duration, or VOT. All groups produced English /p/ as an aspirated stop with mean VOT values ranging from 50 to 60 ms and English /b/ at 10 ms. The three groups of speakers were also similar in their voicing of English /b/, producing it as a voiceless stop approximately 25% of the time.

B. Accent ratings

1. Phonemic accent ratings

Given the excellent interrater agreement (see Sec. II C a), a rating averaged across raters for each speaker was calculated for both voiced and voiceless consonants in word initial and medial position. ANOVAs on these average ratings with group as a between-subject factor and position as a within-subject factor revealed a reliable effect of group for both voiceless [$F(2,41)=71.76$, $p<0.001$] and voiced consonants [$F(2,41)=72.32$, $p<0.001$]. No other reliable main effect or interaction was found. Posthoc tests revealed that native speakers received reliably better accent ratings than overhearers, who in turn received reliably better ratings than typical late L2 learners [by HSDs, p's<0.001; see Table IV]. Pearson correlations between phonemic accent ratings and phonetic measurements, namely VOT and lenition, were substantial when averaged across tokens (VOT, $r=-0.53$, $p<0.001$; lenition, $r=0.78$, $p<0.001$) as well as at the level of individual tokens (i.e., second token; VOT, $r=-0.50$, $p<0.01$; lenition, $r=0.77$, $p<0.001$).

TABLE V. Participants' performance on Spanish morphosyntax assessment tasks.

Measure	Native speakers	Childhood overhearers	Typical late L2 learners
Grammaticality judgment			
percentage correct	91.8[a] (0.97)	63.6[b] (1.6)	62.5[b] (2.6)
reaction time (ms)	1201[a] (124)	2661[b] (283)	1936[a] (307)
Noun-phrase production			
gender agreement	94.8[a] (2.6)	66.3[b] (6.2)	72.7[b] (2.8)
number agreement	93.8[a] (2.6)	82.5[a] (5.3)	92.3[a] (1.7)
Verb-phrase production			
tense/aspect	94.1[a] (1.6)	50.4[b] (5.1)	50.0[b] (5.1)
person	98.2[a] (1.0)	68.6[b] (4.4)	72.3[b] (5.1)
number	98.6[a] (1.0)	70.7[b] (4.8)	80.3[b] (4.9)
Narrative production	4.93[a] (6.7)	2.5[b] (0.17)	2.6[b] (0.14)

Note. Narrative production ratings were on a 5-point scale, with high ratings indicating better formulated sentences. For all other measures, the table indicates the mean percentage correct unless otherwise specified. Standard errors are given in parentheses. Within a row, means with different superscripts were reliably different from each other according to Tukey's HSD test, $p<0.01$. Numbers with the same superscript were not reliably different from each other.

2. Narrative accent ratings

As shown in Table IV, similar results were found for the frog story accent ratings.[8] An ANOVA on the ratings averaged across raters for each speaker revealed reliable group differences [$F(2,41)=73.57$, $p<0.001$]. Native speakers were rated as having a better whole-sentence accent than the overhearers, who in turn were rated more favorably than the typical late L2 learners (by Tukey's HSDs, p's<0.05).

C. Morphosyntax

1. Grammaticality judgment

There were reliable group differences for grammaticality judgment both in terms of accuracy and reaction time [$F(2,41)=77.41$, $p<0.001$ and $F(2,41)=8.01$, $p<0.01$, respectively; see Table V]. Native speakers outperformed both childhood overhearers and typical late L2 learners in making correct grammaticality judgments (Tukey's HSDs, p's <0.01). However, the latter two did not differ reliably from each other. The overhearers, but not the typical late L2 learners, took longer than the native speakers in deciding whether a sentence was grammatical (Tukey's HSD, $p<0.01$) whereas the latter two did not differ reliably from each other.

2. Noun-phrase production

The native speakers outperformed both the overhearers and the typical late L2 learners in marking gender agreement, but not in marking number agreement [$F(2,35)=14.38$, $p<0.001$ and $F(2,35)=3.07$, $p=0.06$, respectively; see Table V for Tukey's HSD test results]. The childhood overhearers and the typical late L2 learners did not differ reliably from each other.

3. Verb-phrase production

For verb morphology production, the native speakers outperformed the childhood overhearers and the typical late

L2 learners, who again did not differ reliably from each other [$F(2,37)=9.6$, $p<0.001$; see Table V for Tukey's HSD test results].

4. Narrative production

Native speakers' frog stories were rated as containing fewer morphosyntactic errors than the overhearers' and the typical late L2 learners' stories. The latter two did not differ reliably from each other [$F(2,41)=118.27$, $p<0.001$; see Table V for Tukey's HSD test results].

Taken together, the results of this study reveal considerable benefits of childhood overhearing in phonology and no measurable benefits in morphosyntax.

IV. DISCUSSION AND CONCLUSIONS

Childhood overhearers of Spanish were found to have more nativelike Spanish pronunciation than typical late L2 learners according to phonetic analyses (i.e., VOT, lenition), phonemic accent ratings, and narrative accent ratings. Importantly, the overhearers' advantage in Spanish pronunciation was not found to be attributable to prosodic factors such as speech rate, phrasing of the carrier sentence, and/or the location of stress on the target word. Native speakers, childhood overhearers, and typical late L2 learners were all comparable on these prosodic characteristics. That said, these findings do not necessarily mean that the three groups acquired prosodic features of Spanish to the same degree. Studies on prosodic transfer (e.g., Ueyama and Jun, 1998; Jun and Oh, 2000) show that the degree of proficiency in an L2 is reflected in the realization of L2 intonation patterns and prosodic grouping of words. No reliable group differences were uncovered in the present study perhaps because the sentences evaluated for prosodic factors (i.e., the target sentences) were short and simple. The groups might have shown prosodic differences if we had evaluated production in the story-telling task, which elicited longer and more natural sentences with more complex syntactic structures.

Our study speaks to the relation between perception and production. Flege's speech learning model (Flege, 1995) postulates that a speaker's pronunciation ability should be only as good as his or her perceptual abilities. When compared to the typical late L2 learners, the childhood overhearers' better accents may reflect more nativelike perceptual discrimination of Spanish phonemes. Although our study did not test this hypothesis directly, a similar study on Korean L2 acquisition (Oh *et al.*, 2003) suggests that this may indeed be the case. Oh *et al.* found that adults who regularly heard Korean during childhood performed better in their perception of Korean stops than those who had no exposure to Korean until college. Childhood overhearing, then, may have helped improve speech perception. Such perceptual abilities can be put to use when childhood overhearers try to learn the overheard language later in life (e.g., in high school and college), resulting in a better accent. If this account is on the right track, our findings of childhood overhearing benefits in speech production can be added to the growing body of evidence for a perception-production link. That is, early overhearing triggers better perception, which can be translated into better production even for late language learners (e.g., Bradlow *et al.*, 1997, 1999).

However, the logical inference of "if better perception, then better production" needs elaboration. Note that the overhearers in this study, like the typical late L2 learners, rarely produced intervocalic /g/ as a lenited consonant. Nonetheless, they were nativelike for intervocalic /d/, and their intervocalic /b/ fell between these two levels of performance. This pattern of results may be due to the frequency with which the overhearers were exposed to lenited /b, d, g/. For example, lenited /g/ (i.e., [ɣ]) is not in the English sound system either phonemically or allophonically, so lack of practice may account for the overhearers' difficulty in producing lenited /g/. By contrast, the overhearers were nativelike for lenited /d/ (i.e., [ð]), which happens to be a high-frequency phoneme in English (e.g., "th" in *this, other*). For lenited /b/ (i.e., [β]), which is not an English phoneme but is close in articulation and similar in voicing to the English /v/, the overhearers produced it better than the typical late learners but not yet nativelike.

The puzzle then is why typical late L2 learners had great difficulty producing lenited sounds even when one of them occurs in English (namely, [ð]).[9] It may have to do with Spanish and English orthography. Since the letters "b, d, g" are produced as stop consonants in English, the typical late L2 learners could be more vulnerable to being misled by the orthographic presentation of Spanish /b, d, g/ than the childhood overhearers. Another possibility is that the typical late L2 learners had not yet acquired the allophones of Spanish /d/ in intervocalic position. That is, even though they knew how to produce the dental fricative, i.e., a lenited /d/, they were not aware that the same sound was also an allophone of /d/ in Spanish. Note that typical late L2 learners also produced Spanish voiced stops with less voicing than did native speakers. Childhood overhearers' voicing was in between these two groups although not reliably different from either. Combined, these findings hint at the possibility that childhood overhearing of Spanish may lead to better awareness of Spanish allophones.

Our findings are also relevant to the nature of input in early childhood bilinguals' L2 phonology. As in Au *et al.* (2002), we found that overhearers' pronunciation advantage did not come at the cost of producing English /p/ and /b/ with a Spanish accent. Nor did the English of our native Spanish speakers (i.e., Spanish–English bilinguals) seem compromised since they also produced English /p/ and /b/ with values comparable to that of typical late L2 learners. These findings are consistent with prior findings where Spanish–English bilinguals who learned English by age 5–6 produced English /t/ (Flege, 1991), /p/ (Williams, 1980), and /b/ (Flege and Eefting, 1988) like English monolinguals. However, our results contrast with findings where Spanish–English bilinguals produced English /p, t, k/ with "compromised" VOT values (Flege and Eefting, 1987) and English /b/ with Spanish-like voicing lead (Flege and Eefting, 1988).

One way to make sense of these conflicting results is to focus on the nature of English input. Flege and Eefting (1987) speculated that the early childhood bilinguals in their

study had ''compromised'' VOT values because much of their English input was Spanish-accented; they were born in the U.S. but were attending a university in a Spanish-speaking country at the time of testing. In contrast, the bilinguals in Williams (1980) and Flege (1991) were more like the Spanish–English bilinguals in our study in that they were living in the U.S. at the time of testing and used *both* languages in their daily lives. Thus, the nature (hearing good models of American English) as well as the timing of L2 input may play an important role in childhood bilinguals' eventual L2 pronunciation.

An important conclusion of the current study is that the childhood overhearing advantage seems to be domain specific: it is very robust in phonology but so far not detectable in morphosyntax—a domain, like phonology, that is easy for children to acquire but difficult for adults to master. Even with our rather comprehensive assessment of morphosyntax with additional tasks not included in Au *et al.*'s (2002) study (i.e., elicited verb-phrase production, story production), no childhood overhearing advantage in morphosyntax was detected. Perhaps it takes more than merely overhearing a language during childhood to gain an edge in morphosyntax acquisition. Future research can help determine whether other kinds of childhood language experience, such as speaking a language for a few years during childhood, results in measurable benefits in mastery of morphosyntax as well as phonology for adults (re-)learning a childhood language.

Our results can have important applied implications for immigrant children in a predominantly monolingual nation such as the U.S. Most immigrant children in such linguistic environments tend to lose, or severely limit their use of, their heritage language (e.g., Fillmore, 1991; Au, in press). Our study can help immigrant parents and early childhood educators make more informed decisions about what kind of childhood language environment and experience they want to foster for their children. Specifically, this study suggests that overhearing a language during childhood can probably translate into a measurable accent advantage when children try to acquire the overheard language later in life.

ACKNOWLEDGMENTS

We thank Celia Adame, Marilyn Alvarado, Brenda Alvarez, Argelia Andrade, Cecilia Arriaza, J. Patrick Barjam, Jessica Brauner, Mauricio Carvallo, Ivonne Cisneros, Susan Curtiss, Kristina Cutura, Olga Escamilla, Laura España, Catherine Fountain, Ingrid Gooding, John Grinstead, Vilma Hernandez, Gerardo Ibarra, Patricia Keating, Sahyang Kim, Hyuck-Joon Lee, Reuben Lim, Ann Marroquin, Olivia Martínez, Karla Mayen, Nancy Miranda, Claudia Parodi, Jaime Paz, Katie Polsky, Irasema Ramos, Salome Rincón, Krista Rodriguez-Bruno, Laura Romo, Norma Salazar, Susan Schaffer, Sylvia Sherno, Minjung Son, Henry Tehrani, Motoko Ueyama, and John Whalen for their help with various aspects of this research. We thank the research participants for their time. We also thank Anders Löfqvist and two anonymous reviewers for their insightful and valuable comments on an earlier version of this paper. This work was supported by NIMH Grant No. MH56118 (Au and Jun), an NIMH supplement (Au and Knightly), an NSF predoctoral fellowship (Oh), NIMH traineeships in Developmental Cognitive Science 5T32MH10926 (Knightly, Oh), and seed grants from University of California Linguistic Minority Research Institute, UCLA Institute of American Cultures, and UCLA Academic Senate (Au).

[1]Among the 45 speakers participating in the current study, 18 speakers (5 of the 15 native speakers, 6 of the 15 childhood overhearers, and 7 of the 15 typical late L2 learners) contributed data to Au *et al.*'s (2002) interim report, which also included 15 other speakers from a pilot study. (i.e., 5 natives, 5 childhood overhearers, and 5 typical late L2 learners).

[2]To assess inter-measurer/coder agreement, the primary measurer/coder analyzed the Spanish data on two speakers measured by each of the two secondary measurers/coders, and the English data on two speakers measured by the only secondary measurer/coder. Thus, the reported ranges of agreement are based on the four values for Spanish and two values for English.

[3]Except where noted, all native Spanish-speaking raters were of Mexican or Central American descent. In addition, speech samples assessed by two native Spanish speakers (Mexican) verified self-reports of raters' native speaker status.

[4]Of the 48 raters, 36 were of Mexican or Central American descent, 5 were from Spain, and 7 were of South American descent; 3 from Argentina and 1 each from Columbia, Peru, Ecuador and Bolivia.

[5]Even though raters were instructed to focus on the target segment, their ratings could be influenced by the carrier sentence. However, since the carrier sentence was short and repeated for every target sentence, its effect may be minimal. Compared to narrative accent ratings based on more varied and longer sentences, the phonemic accent ratings are likely to reflect the quality of the target phonemes.

[6]Since the narrative accent ratings involved a much smaller data set (45 story sessions) compared to the phonemic accent ratings, only two raters of Mexican American descent, who readily achieved excellent interrater reliability, were recruited for this task.

[7]One native speaker, three overhearers, and one typical late L2 learner were not available for testing when English data were collected about two months after the Spanish data collection.

[8]The relatively low phonemic accent ratings for native speakers may reflect the difficulty in rating just the target sounds embedded in the carrier sentence. The raters may therefore have been more conservative, using a narrower range of the rating scale and staying closer to the mid-point. On the other hand, raters may have been more confident about their narrative accent ratings based on the longer speech samples (i.e., stories) and hence showed more range in their ratings.

[9]Examination of individual lenition percentages revealed that two typical late L2 learners failed to produce any lenited consonants.

Au, T. K. (**in press**). ''Salving heritage languages,'' in *Heritage Language Acquisition: A New Field Emerging*, edited by D. Brinton and O. Kagan (Erlbaum, Mahwah, NJ).

Au, T. K., Knightly, L. M., Jun, S., and Oh, J. S. (**2002**). ''Overhearing a language during childhood,'' Psychol. Sci. **13**, 238–243.

Berman, R., Slobin, D. I., Aksu-Koc, A. A., Bamberg, M., Dasinger, L., Marchman, V., Neeman, Y., Rodkin, P. C., and Sebastian, E. (**1994**). *Relating Events in Narrative: A Cross Linguistic Developmental Study* (Erlbaum, Hillsdale, NJ).

Birdsong, D., and Molis, M. (**2001**). ''On the evidence for maturational constraints in second-language acquisition,'' J. Memory Lang. **44**, 235–249.

Bongaerts, T., Van Summeren, C., Planken, B., and Schils, E. (**1997**). ''Age and ultimate attainment in the pronunciation of a foreign language,'' Studies in Second Lang. Acquis. **19**, 447–465.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., and Tohkura, Y. (**1999**). ''Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production,'' Percept. Psychophys. **61**, 977–985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (**1997**). ''Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production,'' J. Acoust. Soc. Am. **101**, 2299–2310.

Cohen, J. D., MacWhinney, B., Flatt, M., and Provost, J. (**1993**). ''PsyScope: A new graphic interactive environment for designing psychology

experiments,'' Behav. Res. Methods Instrum. Comput. **25**, 257–271.
Curtiss, S., and Yamada, J. (**1987**). ''The Curtiss-Yamada Comprehensive Language Evaluation (CYCLE),'' University of California, Los Angeles.
de Jong, K. (**1995**). ''The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation,'' J. Acoust. Soc. Am. **97**, 491–504.
Fillmore, L. W. (**1991**). ''When learning a second language means losing the first,'' Early Childhood Res. Qu. **6**, 323–346.
Flege, J. E. (**1991**). ''Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language,'' J. Acoust. Soc. Am. **89**, 395–411.
Flege, J. E. (**1995**). ''Second language speech learning: Theory, findings, and problems,'' in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Perception*, edited by W. Strange (York, Baltimore, MD), pp. 233–272.
Flege, J. E., and Eefting, W. (**1987**). ''Production and perception of English stops by native Spanish speakers,'' J. Phonetics **15**, 67–83.
Flege, J. E., and Eefting, W. (**1988**). ''Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation,'' J. Acoust. Soc. Am. **83**, 729–740.
Flege, J. E., and Fletcher, K. L. (**1992**). ''Talker and listener effects on degree of perceived foreign accent,'' J. Acoust. Soc. Am. **91**, 370–389.
Flege, J. E., Munro, M. J., and MacKay, I. R. A. (**1995**). ''Factors affecting strength of perceived foreign accent in a second language,'' J. Acoust. Soc. Am. **97**, 3125–3134.
Flege, J. E., Yeni-Komshian, G. H., and Liu, S. (**1999**). ''Age constraints on second-language acquisition,'' J. Memory Lang. **41**, 78–104.
Fougeron, C., and Keating, P. A. (**1997**). ''Articulatory strengthening at edges of prosodic domains,'' J. Acoust. Soc. Am. **101**, 3728–3740.
Johnson, E. K., and Jusczyk, P. W. (**2001**). ''Word segmentation by 8-month-olds: When speech cues count more than statistics,'' J. Memory Lang. **44**, 548–567.
Johnson, J. S., and 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,'' Cognit Psychol. **21**, 60–99.
Johnson, J. S., Shenkman, K. D., Newport, E. L., and Medin, D. L. (**1996**). ''Indeterminacy in the grammar of adult language learners,'' J. Memory Lang. **35**, 335–352.
Jun, S.-A., and Oh, M. (**2000**). ''Acquisition of 2nd Language Intonation,'' in *Proceedings of International Conference on Spoken Language Processing* (Beijing, China), Vol. 4, pp. 76–79.
Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., and Lindblom, B. (**1992**). ''Linguistic experience alters phonetic perception in infants by 6 months of age,'' Science (Washington, DC, U.S.) **255**, 606–608.
Lisker, L., and Abramson, A. (**1964**). ''A cross-language study of voicing in initial stops: Acoustical measurements,'' Word **20**, 384–422.
Oh, J. S., Jun, S., Knightly, L. M., and Au, T. K. (**2003**). ''Holding on to childhood language memory,'' Cognition **86**, B53–B64.
Oyama, S. (**1976**). ''A sensitive period for the acquisition of a nonnative phonological system,'' J. Psycholinguist. Res. **5**, 261–283.
Patkowski, M. (**1980**). ''The sensitive period for the acquisition of syntax in second language,'' Lang. Learning **30**, 449–472.
Phinney, J. S. (**1992**). ''The multigroup ethnic identity measure: a new scale for use with diverse groups,'' J. Adolescent Res. **7**, 156–176.
Plann, S. J. (**1979**). ''Morphological problems in the acquisition of Spanish in an immersion classroom,'' in *The Acquisition and Use of Spanish and English as First and Second Languages*, edited by R. W. Anderson (TESOL, Washington, DC), pp. 119–132.
Saffran, J. R., Aslin, R. N., and Newport, E. L. (**1996**). ''Statistical learning by 8-month-old infants,'' Science (Washington, DC, U.S.) **274**, 1926–1928.
Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., and Barrueco, S. (**1997**). ''Incidental language learning: Listening (and Learning) out of the corner of your ear,'' Psychol. Sci. **8**, 101–105.
Snow, C. E., and Hoefnagel-Hohle, M. (**1978**). ''The critical period for language acquisition: evidence from second language learning,'' Child Dev. **49**, 1114–1128.
Stockwell, R. P., and Bowen, J. D. (**1965**). *The Sounds of English and Spanish* (Univ. of Chicago, Chicago).
Tahta, S., Wood, M., and Loewenthal, K. (**1981**). ''Foreign accents: Factors relating to transfer of accent from the first language to a second language,'' Lang. Speech **24**, 265–272.
Tees, R. C., and Werker, J. F. (**1984**). ''Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds,'' Can. J. Psychol. **38**, 579–590.
Ueyama, M., and Jun, S.-A. (**1998**). ''Focus realization in Japanese English and Korean English intonation,'' in *Japanese and Korean Linguistics* (CSLI, Cambridge U.P., Cambridge), Vol. 7, pp. 629–645.
Werker, J. F. (**1995**). ''Age-related changes in cross-language speech perception: Standing at the crossroads,'' in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Perception*, edited by W. Strange (York, Baltimore, MD), pp. 155–169.
Werker, J. F., and Tees, R. C. (**1984**). ''Cross-language speech perception: Evidence for perceptual reorganization during the first year of life,'' Infant Behav. Dev. **7**, 49–63.
Williams, L. (**1980**). ''Phonetic variation as a function of second-language learning,'' in *Child Phonology Volume 2: Perception*, edited by G. H. Yeni-Komshian, J. F. Kavanagh, and C. A. Ferguson (Academic, New York), pp. 185–215.
Wode, H. (**1981**). *Learning a Second Language: An Integrated View of Language Acquisition* (Gunter Narr Verlag, Tubingen).
Yamada, R. A. (**1995**). ''Age and acquisition of second language speech sounds: perception of American English /r/ and /l/ by native speakers of Japanese,'' in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, edited by W. Strange (York, Baltimore), pp. 305–320.