The Effect of Age on the Acquisition of Second Language Prosody

Becky H. Huang and Sun-Ah Jun

Language and Speech 2011 54: 387 originally published online 5 May 2011
DOI: 10.1177/0023830911402599

Article

Language *and* Speech

The Effect of Age on the Acquisition of Second Language Prosody

Language *and* Speech
54(3) 387–414

DOI: 10.1177/0023830911402599
las.sagepub.com
SAGE

Becky H. Huang
Graduate School of Education, Harvard University

Sun-Ah Jun
Department of Linguistics, University of California, Los Angeles

Abstract

This study reports an exploratory analysis of the age of arrival (AoA) effect on the production of second language (L2) prosody. Three groups of Mandarin-speaking immigrants (N = 10 in each group) with varying AoA in the United States and ten native speakers of English as controls participated in the study. All participants read a paragraph of English, and their speech samples were subjected to three prosodic analyses: speech and articulation rates, native speakers' judgment of the prosody based on segment-filtered speech, and analyses of tones and prosodic groupings using the Mainstream American English Tones and Break Indices (MAE_ToBI) transcription conventions. The L2 groups also filled out a survey providing information about their demographic background, English input, and socio-psychological aspects of language learning. The results revealed that the AoA factor impacted different aspects of prosody to varying degrees. Group differences were statistically significant for speech rate, degree of foreign prosody, the frequency of pitch accents, and the frequency of high boundary tones (H-H%). However, group differences were not significant for articulation rate, prosodic groupings, and the rest of the ToBI-labeled phonological categories. Multiple regression analyses further confirmed the AoA effect on degree of foreign prosody, the frequency of pitch accents, and high boundary tones (H-H%); AoA remained a significant predictor controlling for the effects of other variables. However, speech rate was predicted by English media exposure and motivation variable but not by AoA.

Keywords

Age of learning, Mandarin speakers of L2 English, Second language prosody, Tones and Break Indices (ToBI)

Corresponding author:
Becky H. Huang, Laboratory for Developmental Studies, Shannon Hall 2nd Floor, 25 Francis Avenue, Cambridge, MA 02138, USA.
Email: bhuang@wjh.harvard.edu

Background information

Owing to its theoretical implications for the mechanism of second language (L2) acquisition and practical implications for L2 education, the age-related decline in ultimate second language (L2) attainment is one of the most controversial topics in the L2 acquisition field. Among the various L2 linguistic domains, phonological production is arguably the least controversial candidate for an age of learning effect. In fact, Scovel (1988) argued that the age effect exists only for phonology because the ability to master the sound patterns of an L2 is susceptible to neurological development. Although the number of empirical studies examining the age effect across multiple linguistic domains is small, their results agree that immigrants' age of arrival (AoA) in the L2-speaking country (generally used as the proxy variable for the age of learning effect) constrains long-term phonology outcomes more than it does morphosyntax outcomes (Flege, Yeni-Komshian, & Liu, 1999; Huang, 2009a; Oyama, 1973). However, previous studies on L2 phonological production have predominantly focused on segmental rather than suprasegmental properties. These studies reveal a clear negative relationship between AoA and native-like accuracy of segmental production (Flege, MacKay, & Meador, 1999; Flege, Munro, & MacKay, 1995; Tsukada, Birdsong, Mack, Sung, Bialystok, & Flege, 2004). For example, Flege and colleagues (Flege et al., 1995; Flege, MacKay, & Meador, 1999) found native speakers of Italian who arrived in the United States at a younger age produced English (L2) vowels and consonants more accurately than those who arrived later in life.

Phonological production, however, includes suprasegmental as well as segmental properties, and the acquisition of phonology is not complete without acquiring both of these properties. Suprasegmental or prosodic properties such as rhythm, stress, and intonation have been argued to differ fundamentally from segmental properties. According to Lehiste (1970), a suprasegmental feature is defined by reference to items in a sequence (syntagmatic comparison), whereas a segmental feature is defined by comparing itself with other items in the phonological inventory rather than with the sequence of segments in which it appears (paradigmatic comparison). In the current study, prosody refers to prominence relations and phrasing (i.e., the grouping of subunits in an utterance), marked phonetically by manipulating pitch, loudness, and duration (Beckman, 1996; Shattuck-Hufnagel & Turk, 1996). Specifically, prosody is operationalized in the current study to include the intonation contours of utterances, as marked by the existence and type of pitch accents and boundary tones, and by the tempo/timing of articulatory sequences, i.e., speech/articulation rate. Speech rate measured in terms of the number of syllables per second would reflect the speed of segmental articulation, but it also reflects a prosodic organization. As stated in Lehiste (1970), speech rate represents the linguistic function of duration on a sentence level. Sentences produced at a faster rate show a fewer number of pitch accents and phrase boundaries (Fougeron & Jun, 1998; Jun, 1993), and as shown in Ueyama and Jun (1998), more advanced L2 (English) speakers tend to have fewer pitch accents and prosodic boundaries. Speech rate data thus corroborate the measure of intonation as an indirect way to evaluate the acquisition of prosody.

The importance of prosody in language acquisition and processing has received empirical support in several research areas. Studies on first language (L1) acquisition have shown that, during the first postnatal year, language-specific perceptual organization occurs in prosody such as rhythm (Nazzi, Jusczyk, & Johnson, 2000), stress patterns (Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009; Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993; Skoruppa, Pons, Christophe, Bosch, Dupoux, Sebastián-Galles, Alves Limissuri, & Peperkamp, 2009), and lexical tones (Mattock & Burnham, 2006), as well as in consonants and vowels (Polka & Werker, 1994; Werker & Tees, 1983, 1984). Studies have also shown that some suprasegmental features, such as lexical tones and word stress, are acquired even earlier than

consonants and vowels (Golinkoff, 1983; Li & Thompson, 1977). Prosodic cues have also been found to play an important role in facilitating infants' word learning (Jusczyk, Houston, & Newsome, 1999; Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006), assisting adults in solving ambiguous syntactic structures (Kjelgaard & Speer, 1999; Millotte, Wales, & Christophe, 2007; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992), and helping speakers achieve discourse coherence via the use of increased pitch range to signal topic shift (Hirschberg & Pierrehumbert, 1986; Wennerstrom, 1998). Additionally, prosody has been shown to contribute to listeners' perception of global foreign accents (Magen, 1998); in a number of studies, these prosodic cues were found to contribute even more than segmental properties to the perception of foreign accents (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992).

However, despite the crucial role played by prosody in language development, comprehension, and processing, relatively little research has focused on L2 acquisition of suprasegmental properties except for speech rate. The body of literature covering L2 speech rate has focused on examining durational differences between native and non-native speakers or across L2 speakers of varying proficiency (Elsendoorn, 1984; Flege, 1979; Munro & Derwing, 1995), and on researching speech rate as a characteristic of L2 fluency (Chambers, 1997; Cucchiarini, Strik, & Boves, 2000, 2002; Pennington, 1992; Riggenbach, 1991; Towell, Hawkins, & Bazergui, 1996). The results of these investigations showed that speech rate is related to native speaker status (cf. Flege, 1979), more advanced L2 proficiency, and fluent speech.

In an investigation of the age effect on speech rate, Guion and colleagues (Guion, Flege, Liu, & Yeni-Komshian, 2000) examined the age effect on sentence duration (functionally equivalent to speech rate) with a sample of 240 immigrants in Canada who spoke either Korean or Italian as their first language. All immigrant participants repeated model sentences and the duration of each sentence was measured and averaged across sentences. The researchers employed partial correlation analysis to examine the relationships between speech rate outcomes, AoA, and three variables that were potentially confounded with AoA, i.e., chronological age, length of residence in Canada, and self-reported use of the first language. Results from the study demonstrated a strong AoA effect even after partialling out the confounding variables.

More recently, Trofimovich and Baker (2006) examined the effect of L2 experience on the acquisition of prosody by recording Korean speakers' production of English (L2). The recordings were subjected to acoustic analyses (i.e., speech rate, pause duration, pause frequency, stress timing, and tonal peak alignment) and the degree of foreign accent was rated by native English speakers based on low-pass filtered speech recordings. Participants' L2 experience, as measured by their length of stay in the United States, was found to correlate with differences in stress timing and with foreign accent rating, with more experience resulting in more native-like production. However, L2 experience was not correlated with the acquisition of tonal peak alignment accuracy or speech rate. Furthermore, greater L2 experience did not reduce pause frequency and pause duration. However, as this study included only post-pubescent L2 speakers, it did not address the experience effect in L2 speakers with different ages of immersion, an issue addressed by the current study.

Narrowing down the L2 prosody production literature specifically to studies of the age effect, the majority of long-term immersion studies focus on global foreign accent outcomes (e.g., Asher & Garcia, 1969; Flege, MacKay, & Meador, 1999; Flege, Yeni-Komshian, & Liu, 1999; Oyama, 1976; Tahta, Wood, & Loewenthal, 1981; Thompson, 1991). This line of research confirms the age effect found in segmental pronunciation research by showing that late arrivals are perceived to have a stronger global foreign accent than early arrivals based on their production of L2 sentences. To illustrate, Tahta and colleagues (1981) examined the factors associated with global foreign accent outcomes in a group of immigrants in the United Kingdom who had a variety of first language

backgrounds and whose AoA ranged from 6 to 15 years and beyond. The researchers also included variables such as gender, chronological age, numbers of languages spoken, length of residence, and language use at home. L2 participants were recorded reading an English paragraph aloud, and their speech samples were then judged by three English native speakers on a scale of 0 to 2 (0 = no foreign accent; 1 = detectable but slight accent; 2 = marked accent). Using multiple regression techniques, the researchers found AoA to be the strongest predictor of foreign accent ratings even after controlling for the effects of other variables. Additionally, immigrants with AoAs below 6 were found to be free of foreign accents, whereas immigrants with AoAs beyond 13 spoke with marked foreign accents. Although informative, the inferences were made from only the judgments of three native speakers and a rather crude rating scale. Neither acoustic measurement nor prosodic analysis was performed. Furthermore, a major problem with this study and the body of research on global foreign accents in general is that the outcome variable itself cannot be considered a "clean measure of prosody". A foreign accent is the product of deviations in sound, both segmental and suprasegmental (Magen, 1998), as well as perhaps non-linguistic variables such as confidence level (Piller, 2002). To obtain an accurate measure of prosodic production without the interference of segmental information, low-pass filtering has been used by some researchers in investigating L2 prosody (e.g., Munro, 1995; Trofimovich & Baker, 2006). In the current study, this technique was utilized to study specifically the age effect on L2 acquisition of prosody, apart from segmental information. The present study also included temporal measures (i.e., speech and articulation rates) and intonational analyses, thus advancing this topic both in methodology and in scope.

As shown in the existing research on the age effect on L2 prosody, AoA appeared to be a robust variable even after controlling for the effects of other variables (e.g., Guion, Flege, Liu, & Yeni-Komshian, 2000; Piske, Mackay, & Flege, 2001; Tahta et al., 1981; Thompson, 1991). However, these studies also revealed the contributions of other variables, though to varying degrees, to L2 prosody outcomes. In particular, the L2 input factor, operationalized as L2 use (Tahta et al., 1981; Thompson, 1991), L1 use (Piske et al., 2001; Guion, Flege, & Loftin, 2000), or media exposure (Flege, Yeni-Komshian, & Liu, 1999), appeared to be a substantial contributor to the acquisition of native-like L2 prosody. However, due to the large variations in the selection of participants, study design, elicitation and rating techniques, and the variables surveyed in each study, it is hard to generalize across these studies and draw any conclusions (Piske et al., 2001). Therefore, the present study surveyed a wide range of variables in order to better investigate the contributing factors of various aspects of prosody.

Research examining the factors that contribute to listeners' perceptions of L2 production also provides valuable insights for understanding the basis of age effects on L2 prosody (Anderson-Hsieh et al., 1992; Mackay, Flege, & Imai, 2006; Magen, 1998; Munro, 1995; Munro & Derwing, 2001; Trofimovich & Baker, 2006). This line of research has established that the perception of foreign accents can be attributed to a combination of segmental and suprasegmental divergences from native speaker norms. However, none of the previous studies analyzed the phonological categories of intonation or examined the impact of their deviations on perceived foreign accents. It is conceivable that phonological categories of intonation may play crucial roles in the perception of foreign accents. Previous research also focused on the perception of global foreign accents rather than on foreign prosody specifically. The present study thus aims to fill this gap by exploring the contributions of both temporal measures and phonological categories in intonation in the perception of foreign prosody.

Finally, one remaining concern in the research on the age effect on L2 prosody (and in the research on L2 prosody in general) lies in the methodology of characterizing prosodic patterns. Research on segmental properties has well-established quantitative and qualitative analytic methods that are not only relatively straightforward but also readily available. In contrast, research on

prosodic properties has been limited to quantitative measurements of pitch and duration at a certain position in an utterance (e.g., utterance final syllable), and methods of analyzing prosody qualitatively, though well developed in theoretical linguistics as the Autosegmental-Metrical model of intonational phonology and the ToBI (Tones and Break Indices) transcription system (Beckman & Ayers-Elam, 1994; Beckman & Hirschberg, 1994; Ladd, 1996), have only been applied to a handful of empirical studies of L2 prosody (Chen & Mennen, 2008; Jun & Oh, 2000; Ueyama & Jun, 1998; Wennerstrom, 1998). The current study thus addresses this methodological issue by applying a qualitative analysis method, i.e., the Mainstream American English Tones and Break Indices (MAE_ToBI) prosodic transcription conventions, in the investigation of the age effect on L2 prosody.

2 Study goals and research questions

To summarize, the literature review above revealed the importance of researching the age effect on second language (L2) phonology, the dearth of studies focusing on L2 production of prosody, and the methodological shortcomings of L2 prosody research. The current study thus aims to improve our understanding of the age effect on L2 phonology by providing an empirical exploratory report that details the age effect on L2 prosody and investigates the prosodic categories that contribute to the perception of foreign prosody. The current study also utilizes two analytical approaches that have never been used in the investigation of the age effect on L2 prosody: the low-pass filtering technique and the MAE_ToBI (henceforth "ToBI") phonological model of prosodic transcription.

The research questions pursued by the current study are as follows:

1) Is there an age effect on the ultimate outcome of various categories of second language prosody? If so, which prosodic categories are impacted by the age effect?
2) What are the relative contributions of these prosodic categories to the perception of foreign-accented prosody in L2 speech?

To answer the first research question, the current study examined Mandarin-speaking immigrants ("L2 speakers" hereafter) with differing Ages of Arrival (AoA) in the United States as well as a group of native English speaker controls. Three approaches were employed to analyze different aspects of prosody, including: 1) analysis of speech rates and articulation rates; 2) native speaker listeners' rating of prosody foreignness based on low-pass filtered productions; and 3) analysis of the prosodic categories and structures based on ToBI transcription conventions. The current study also gathered information on a wide range of variables other than AoA that could potentially contribute to the ultimate L2 prosody outcomes, such as gender, second language exposure, motivation, and cultural identity, to test the validity of the age effect.

To understand the relative strength of the prosodic categories to the perception of a foreign prosody, the relationships between prosodic categories and native speakers' ratings of the filtered speech were examined via correlation tests and multiple regression models.

3 Methods

3.1 Participants

Participants included three groups of L2 speakers (ten participants in each L2 speaker group) who varied in their AoA (range = 5–27 years of age). All participants spoke Mandarin as their first language, had lived in the United States for at least five years, and had had their English language instruction

Table 1. Demographic information (AGE, AoA, LoR) by group

	Native speaker (*N* = 10)	L2 child (*N* = 10)	L2 adolescent (*N* = 10)	L2 adult (*N* = 10)
AGE[a]	25.00 (5.81)	20.10 (1.91)	28.40 (5.38)	34.40 (2.27)
AoA[b]	NA	7.60 (5–9)	13.60 (1–17)	23.10 (20–26)
LoR[c]	NA	12.40 (2.59)	13.83 (3.66)	11.36 (1.82)
GENDER	6f, 4m	7f, 3m	6f, 4m	6f, 4m

[a]Current age (standard deviation in parentheses).
[b]Age of arrival (range in parentheses).
[c]Length of residence (standard deviation in parentheses).

(if they had had any) conducted in a foreign language classroom prior to their arrival in the US. All participants also held a college degree or were current college students, and had never been diagnosed with hearing problems, language disorders, or learning disabilities. Based on their AoA, the three L2 speaker groups were labeled as "Child Arrivals" (AoA = 5–9 years old), "Adolescent Arrivals" (AoA = 12–17 years old), and "Adult Arrivals" (AoA = 20–26 years old). All three L2 speaker groups had lived in the United States for approximately ten years (range = 8–18), and the length of residence (LoR) did not differ among those groups, $F(2, 27) = 1.97$, $p = .159$, partial eta-squared = .127, NS. However, participants' current ages (AGE) were significantly different among AoA groups, $F(2, 27) = 40.98$, $p < .000$, partial eta-squared = .752. Tukey's post-hoc tests showed that all of the group comparisons were significant; *Child Arrivals* were on average significantly younger than either *Adolescent Arrivals* or *Adult Arrivals* ($p = .000$ for both results). *Adolescent Arrivals* were also significantly younger than *Adult Arrivals* ($p = .002$). Given the linear dependency of the AoA, LoR, and AGE variables, it was not possible to control for both length of residence and current age.

A group of ten native speakers (*NSs*) of American English also participated and served as the control group. All *NS* participants spoke English as their first language and had only been exposed to foreign languages in high school language courses. They were all affiliated with the same university in Southern California as students or staff. The inclusion of the *NS* group was to establish the native speaker norm. See Table 1 for a summary of demographic information for all participants in the current study.

3.2 Material

Previous research on L2 speech production revealed that task types played a role in production assessment. Specifically, scripted words and sentences tended to overestimate participants' proficiency level compared to scripted paragraphs or spontaneous speech (Moyer, 1999). To circumvent this limitation, an elicitation paragraph from the Speech Accent Archive website (http://accent.gmu.edu/) was used for the current study (see Table 2).[1] The English paragraph consists of sixty-nine words, seventy-seven syllables, and four sentences.[2] The paragraph encompasses practically the full inventory of American English vowels and consonants, and is composed of both declarative and imperative sentences. The paragraph was chosen to examine both segmental and prosodic properties, but the current paper presents only the prosodic data. Participants were asked to read the paragraph after having familiarized themselves with it.

As shown in previous studies (e.g., Munro & Derwing, 1995), participants' reading ability could be a confounding factor in a read-aloud task. In contrast to natural conversation or spontaneous speech, oral reading is a special kind of speech task, and thus speakers—whether native or

Table 2. Reading paragraph

Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

not—could vary in their performance. That is, the reading task alone may not reflect speakers' full prosodic production skills.[3] The inferences made from the current study are thus limited to speakers' production of prosody while reading aloud. However, although a reading task has the aforementioned limitations, its advantages in studying prosody on a smaller scale could outweigh these limitations. The limited pragmatics of the passage and the fixed text make it possible to control the variation of prosody to a great extent. Prosody changes depending on discourse structure and pragmatics as well as on the content of the passage. It would thus be extremely hard, if not impossible, to compare the prosodic features, phonological or otherwise, across speakers' spontaneous speech.

3.3 Procedure

All L2 speakers were recruited through a variety of venues and were individually tested in a quiet room at the university or their private residences. Native speaker controls, on the other hand, were all students and staff from the same university and were tested in a university laboratory. Each participant was given one minute to review the elicitation paragraph before the recording, and was instructed to read at their natural pace. They then read the paragraph twice into a high quality head-mounted microphone (Shure SM 10A), which was recorded in the *Audacity* program (Audacity Team, 2000, version 1.2.5). Since each participant read the paragraph twice, the researchers first listened to both readings and selected the one with higher fluency and better recording quality. In most cases, the second of the two recordings was selected. The selected recordings were then used for all prosodic analyses.

At the end of the language testing session, the L2 speakers filled out a survey about their demographic information, language learning history and other socio-psychological factors (see Appendix A for sample questions from the survey).

4 Data analysis and results

This section first presents the one-way ANOVA results for the prosodic outcomes in the order of speech rate, articulation rate, filtered speech rating, and phonological categories of tone and phrasing based on ToBI labeling. Results of the qualitative analysis of the phonological categories then follow. Based on the ANOVA results, bivariate correlations and regression analyses were conducted for prosodic outcomes with a significant age effect, with all variables from the survey data, to further investigate the validity of the age effect. Correlation results and regression models predicting the filtered speech rating are provided thereafter.

4.1 Speech rate analysis

Speech rate was calculated as the ratio of number of syllables produced per second, derived by dividing the number of syllables in the paragraph by the total duration of the speech file, including

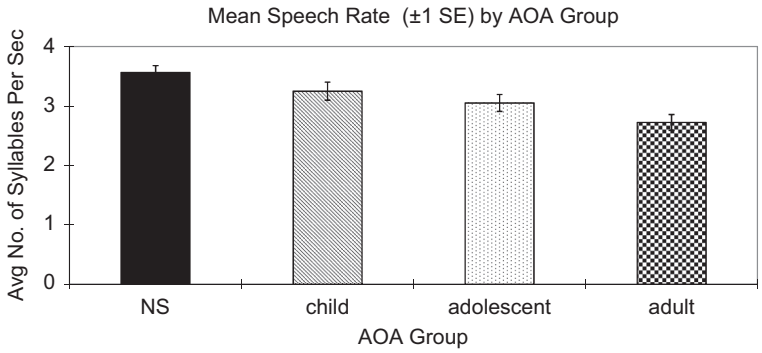

Figure 1. Group means for speech rate (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

pauses during the production of the paragraph.[4] This ratio was used as the dependent variable in the following ANOVA analysis.

A one-way ANOVA revealed a significant group effect with a small effect size, $F(3, 36) = 6.36$, $p = .001$, partial eta-squared = .347. Tukey's post-hoc tests showed that only the *Adult Arrivals*, not *Child* and *Adolescent Arrivals*, had a significantly slower speech rate than the *NS* ($p = .001$), although the difference between *Adolescent Arrivals* and the *NS* was marginally significant ($p = .062$). Among the three L2 speaker groups, *Adult Arrivals*' speech rates were also significantly slower than those of *Child Arrivals* ($p = .048$), but there was no significant difference between either *Child* and *Adolescent* or *Adolescent* and *Adult Arrivals*.

4.2 Articulation rate analysis

Articulation rate was calculated and derived manually by dividing the number of syllables in the paragraph by the total duration of the speech file, excluding pauses and disfluencies.[5] This ratio was used as the dependent variable in an ANOVA analysis that revealed a significant group effect with a small effect size, $F(3, 36) = 4.58$, $p = .008$, partial eta-squared = .277. Tukey's post-hoc tests showed that *Adult Arrivals* had a significantly slower articulation rate than the *NS* ($p = .005$), but *Child* and *Adolescent Arrivals* were not significantly different from the *NS*. None of the group comparisons among the three L2 speaker groups were significant, although a marginal effect was observed between *Adult Arrivals* and *Child Arrivals* ($p = .064$).

4.3 Analysis of foreign prosody rating via low-pass filtered speech

Recordings were trimmed to leave only one second of silence before and after the reading of the paragraph. The segmented paragraphs were then band-pass filtered (cutoff frequency = 50 Hz and 450 Hz) to remove all segmental information while preserving the prosodic information. To preserve the amplitude of the original signal, particularly in the vowels, the intensity curve of the original file was calculated and the filtered files were multiplied by the intensity curve. This procedure was conducted in *Praat* (Boersma & Weenink, 2009, version 5.0.45; http://www.praat.org/).

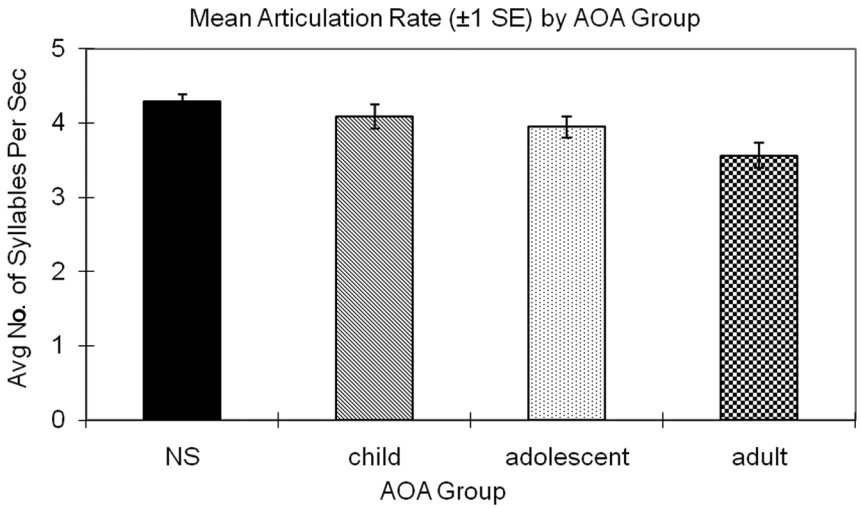

Figure 2. Group means for articulation rate (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

Twelve listeners between the ages of 18 and 33 ($M = 24$) were all recruited from the same university in Southern California to rate the prosody of the filtered files. All listeners reported normal hearing and spoke standard American English as their first language. There were equal numbers of females and males. Half of the listeners were linguistics majors or former ESL teachers, and half of them were naïve raters with no linguistics background.

To prevent fatigue from rating all forty filtered files in one session, the files were divided into two blocks with equal numbers of files from each of the four participant groups. The twenty files within each block were fully randomized via the MATLAB program (MathWorks Inc., 2003, version 6.5.1). The filtered files were played on a laptop computer (ThinkPad x60s), at a comfortable listening level via Sennheiser Model HD212 headsets. Since the low-pass filtered productions were barely intelligible out of context, the listeners were given a transcript of the paragraph against which they could compare the speech. The listeners were told that they would listen to filtered speech produced by both native speakers and non-native speakers, but the proportions of each population were unspecified. They were instructed to focus solely on the intonation, not on speech rate, of the filtered speech and disregard any segmental (i.e., vowels and consonants) information, if at all audible. The word "intonation" was not further defined.

Before the actual rating sessions, listeners were given five practice trials to familiarize themselves with the filtered speech and to ensure that they understood the task. The stimuli in the practice trials were different from those in the actual rating sessions. The listeners then heard each filtered file once and assigned a score of 1 to 9 (1 = strong foreign intonation; 9 = native-like intonation) to indicate the degree of foreignness for each recording. The listeners were encouraged to use the full scale for rating. The inter-stimulus interval was one second. Each rating session lasted about ten minutes, and a short break was provided between the two sessions.

Results showed that the twelve listeners' ratings were all highly and significantly correlated with one another. The intra-class correlation obtained for the raters was also high ($R = 0.93$, $p < .0001$), justifying the use of an average rating as the dependent variable for each participant. Figure

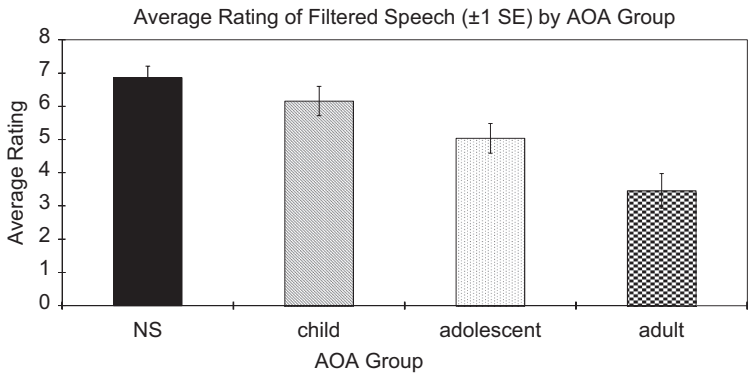

Figure 3. Group means for foreign prosody rating (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

3 presents group means for foreign prosody rating (±1 Standard Error) for the *Native Speaker* group (*NS*) and the three L2 speaker groups (*Child*, *Adolescent*, and *Adult Arrivals*).[6] A one-way ANOVA revealed a significant group effect with a moderate effect size, $F(3, 36) = 11.15$, $p < .0001$, partial eta-squared = .482. Tukey's tests comparing the L2 speaker groups with the *NSs* further revealed that both *Adult* and *Adolescent Arrivals* received significantly lower native fluency ratings than the *NSs* ($p < .0001$ and $p = .031$, respectively). *Child Arrivals*, however, were not significantly different from the *NSs*. Among the L2 speaker groups, *Adult Arrivals* were perceived to have a less native-like prosody than *Child Arrivals* ($p = .001$). The average ratings between *Child* and *Adolescent Arrivals* and between *Adolescent* and *Adult Arrivals* were not significantly different.

4.4 ToBI labeling and analysis

The current study adopts the Autosegmental-Metrical (AM) phonology model of intonation (Beckman & Pierrehumbert, 1986; Ladd, 1996; Liberman, 1975; Pierrehumbert, 1980) as a theoretical framework and utilizes the Mainstream American English Tones and Break Indices (MAE_ToBI) conventions (Beckman & Ayers-Elam, 1994; Beckman & Hirschberg, 1994; Beckman, Hirschberg, & Shattuck-Hufnagel, 2005) for prosodic transcription. The tone labels include an inventory of nine pitch accents (H*, !H*, H+!H*, !H+!H*, L*, L+H*, L*+H, L+!H*, L*+!H) marking prominence on stressed syllables, two phrase accents (H-, L-) marking small phrase (called Intermediate Phrase, ip) boundaries, and two boundary tones (H%, L%) marking large phrase (called Intonation Phrase, IP) boundaries. It has been claimed that the different types of tones convey different semantic/pragmatic meanings (Pierrehumbert & Hirschberg, 1990). For example, a word with H*, which represents prominent high pitch over the stressed syllable of the word, conveys that the word is newly introduced in the discourse, and a phrase ending with high pitch (H-, H%) conveys temporal or causal relationships between conjoined clauses, while a phrase ending with low pitch (L-, L%) does not.

All utterances were coded by two experienced ToBI labelers. Pitch tracks of all utterances were examined using the *Praat* speech analysis software, and tones were labeled on a textgrid tier created for each speech file. The tone labels were then subjected to frequency counts and ANOVA analyses to characterize and compare the intonation patterns of native speakers and L2 speakers with varying AoAs.

The following three analyses were conducted: 1) frequency counts and inferential statistical analysis of the pitch accents and prosodic groupings (i.e., ip and IP); 2) frequency counts and inferential statistical analysis of the pitch accents, phrase accents, and boundary tones by type;[7] and 3) dominant pattern analysis.

For the first and second analysis, the group means, rather than the group sums, were adopted as the measure because of the unequal numbers of tones produced in each group. To examine the age effect and the specific differences among the groups, one-way ANOVAs and Tukey's post-hoc tests were performed. These two analyses assessed participants' knowledge of English intonation and prosodic structures as well as their understanding of the complex relationship between semantic/pragmatic meaning and prosody.

The final analysis aimed to describe the differences in intonational patterns across all speakers and in particular the L2 speakers. It also aimed to characterize the non-native deviations from native prosody. A previous study conducted by Ross and colleagues (Ross, Ostendorf, & Shattuck-Hufnagel, 1992) found variations in pitch accent placement among several native speakers who all read the same news story, suggesting the potential difficulties of establishing norms even among native speakers. The current study thus sought to establish the common patterns of intonation produced by the majority of the speakers within each group, i.e., the dominant patterns. A dominant pattern is defined in this study as the pattern produced by at least half of the participants in each group. Based on this criterion, every word in the paragraph was assigned a dominant pattern within each group. After the dominant pattern was established for each word in the paragraph, the general patterns of the age-related deviations were described and the results were also compared with those of frequency counts and inferential statistical analysis of pitch accents and prosodic groupings.

In order to quantify and compare the amount of variation in tonal categories across groups, the proportion of dominant patterns (i.e., consensus) was used as an index and was derived for each group from dividing the number of words with a consensus (i.e., numbers for which at least half of the participants in the group produced the same intonation categories) by the total number of words in the paragraph (n = 69). On the other hand, to quantify the amount of deviations from the *NS* group, the proportion of consensus that overlapped with the *NS* group's consensus was also calculated for the three L2 speaker groups by dividing the number of consensus words in each group that matched the *NS* consensus words by the total number of *NS* consensus words.

The following sections present the results from the three analyses.

4.4.1. Total number of pitch accents and prosodic groupings. One-way ANOVAs revealed a significant group effect with a small effect size only for the frequency of total number of pitch accents, $F(3, 36) = 6.05$, $p = .002$, partial eta-squared = .335, and not for the frequency of intermediate phrases $F(3, 36) = 1.92$, $p = .144$ or intonational phrases $F(3, 36) = 2.43$, $p = .080$ (see Figure 4). Tukey's post-hoc tests comparing the four groups on the frequency of pitch accents further suggested that *Adult Arrivals* produced a significantly larger number of pitch accents than did the other three groups. However, no significant differences were observed among the *NSs*, *Child Arrivals* and *Adolescent Arrivals*.

4.4.2. Types of pitch accents, phrase accents, and boundary tones. The mean frequency of the six common pitch accent types (H*, !H*, H+!H*, L*, L+H*, L*+H) were submitted to one-way ANOVAs, which revealed a significant group effect only for the frequency of high (H*) tones[8] with a small effect size, $F(3, 36) = 3.01$, $p = .043$, partial eta-squared = .201 (see Figure 5). Post-hoc comparisons for the high tone (H*) category did not yield conventional significance results (i.e., $p < .05$). The small sample size might have limited the power to detect significance in post-hoc tests. However, marginal effects were observed between the *NSs* and *Adult Arrivals* and the *Child* and *Adult*

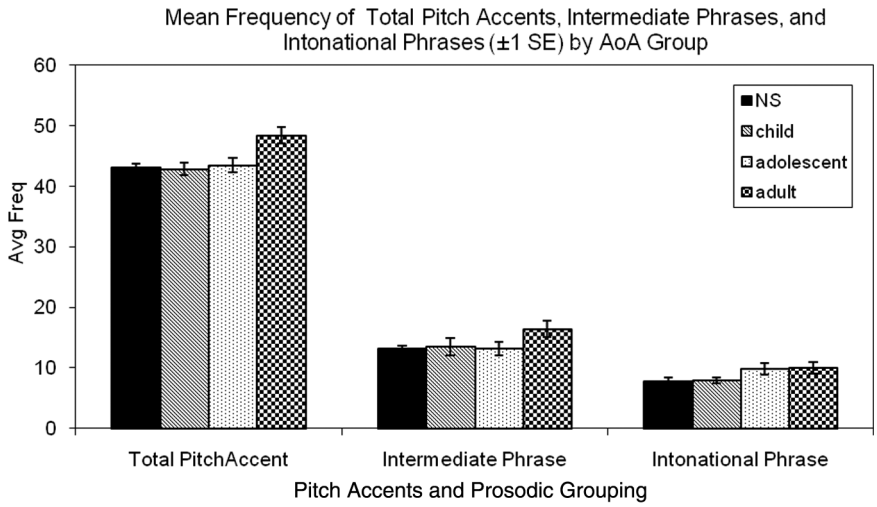

Figure 4. Group mean frequencies for total pitch accents, intermediate phrases and intonational phrases (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

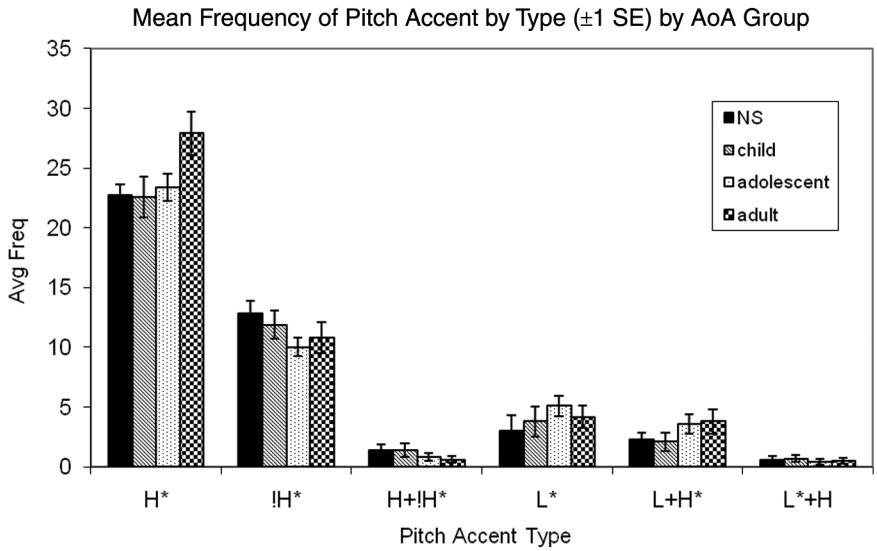

Figure 5. Group mean frequencies for pitch accents by type (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

Arrivals (p = .072 and .065, respectively). In both cases, *Adult Arrivals* produced more high tones than the comparison group.

The mean frequencies of phrase accents (H-, L-) and boundary tones (H-L%, H-H%, L-H%, L-L%) are presented in Figures 6 and 7. One-way ANOVAs revealed a significant group effect

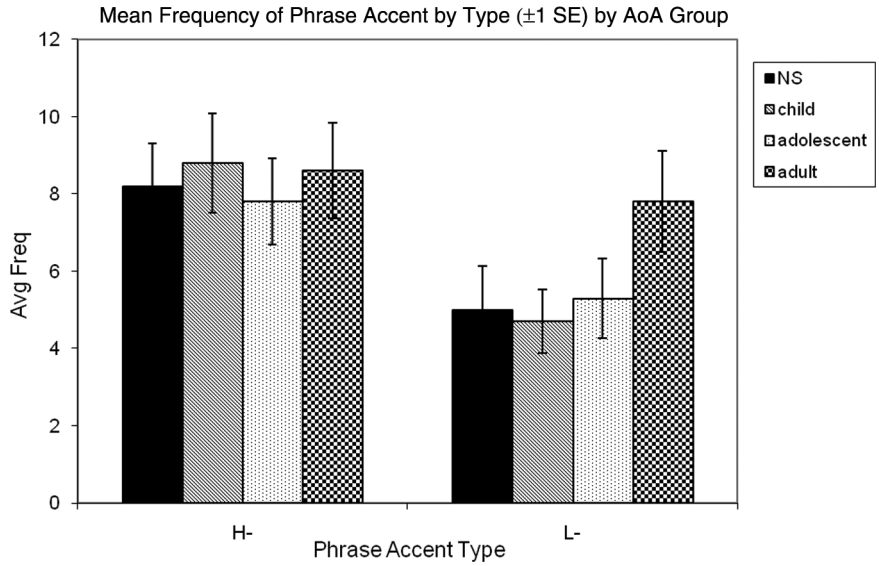

Figure 6. Group mean frequencies for phrase accents by type (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

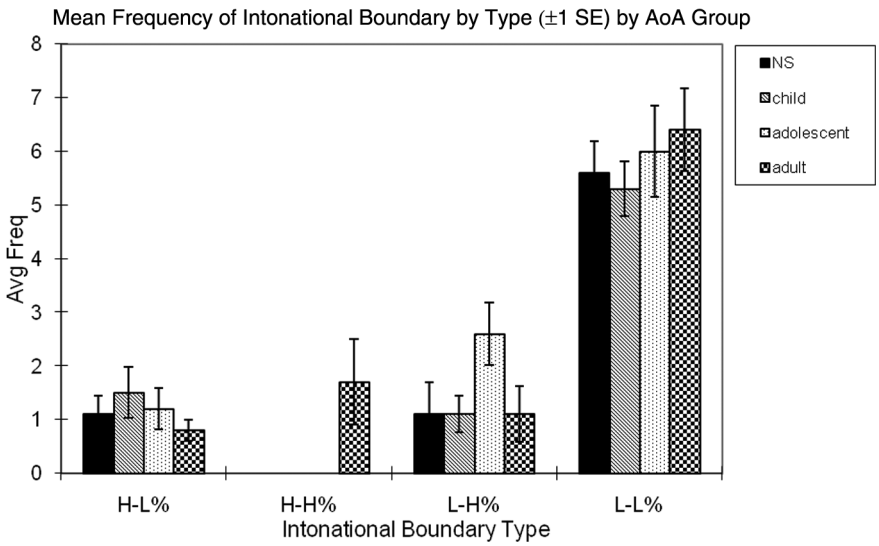

Figure 7. Group mean frequencies for boundary contours by type (±1 Standard Error) for the Native Speaker group (*NS*) and the three L2 speaker groups (*Child, Adolescent, and Adult Arrivals*)

with a small effect size only for the H-H% boundary tone $F(3, 36) = 4.47$, $p = .009$, partial eta-squared = .272, and post-hoc tests indicated that *Adult Arrivals* was the group that drove the effect; *Adult Arrivals* produced many H-H% while the other three groups produced none. Additionally,

although no significant group effect was found for any other tone types, there were some observed trends of incremental deviations from the *NSs* as the AoA increased.

4.4.3. Dominant pattern analysis. The results of the dominant patterns in the *NS* group showed variations even among native speakers, corroborating previous studies (Ross et al., 1992) (see Table 3). Among the three L2 speaker groups (*Child*, *Adolescent*, and *Adult Arrivals*), the amount of within-group variations and deviations from the *NS* patterns increased with AoA. Generally speaking, *Child Arrivals* patterned more closely with *NSs* and *Adult Arrivals* deviated from the other three groups. To illustrate this finding, for the paragraph-ending compound noun *train station*, the *NSs* produced a high pitch accent (H*) on *train* and a low phrase accent plus a low boundary tone (L-L%) during *station*, patterns which are considered conventional in Mainstream American English and whic h were replicated by the *Child* and *Adolescent Arrivals*. Although correctly assigning a H* on the first word, *train*, *Adult Arrivals* also placed a variety of pitch accents on the second word, *station*, revealing their lack of knowledge to produce a compound noun with native-like prosody.

The results also complemented the quantitative analysis of phonological categories presented earlier. Specifically, the inferential statistical results revealed that *Adult Arrivals* produced a significantly larger number of pitch accents than did all the other groups. The dominant pattern results showed that the positions of these pitch accents were quite varied and sometimes produced at inappropriate locations. For example, for the sentence "*ask her to bring these things with her*", both *NSs* and *Child Arrivals* produced a delayed high pitch accent (H* <) on *ask* and no pitch accent on *her* in the beginning phrase "*ask her*". In contrast, *Adult Arrivals* produced a H* on both *ask* and *her*. *Adult Arrivals* were also the only group that assigned a H* to the preposition *with*, a function word that is not expected to receive any prominence in this context.

The quantitative analysis also revealed an age-related trend, though not statistically significant, in the frequency of phrasal breaks, which increased with AoA. The dominant pattern analysis complemented this finding and revealed later AoA groups assigned phrasal breaks in locations that would be considered inappropriate or less-ideal in native English prosody. For example, a phrasal break after the preposition *of* in "*six spoons of fresh snow peas*" would not be expected, especially for a reading task in the current study. However, *Adolescent Arrivals* produced a dominant pattern of phrase break (L-), and some *Adult Arrivals* also produced L- and even put a pitch accent on the function word *of*. Furthermore, for the phrase "*we will go meet her*", only *Adult Arrivals* produced a dominant pattern of phrase break (L-) after *will*, separating the auxiliary verb from the main verb.

Turning now to the analyses of intonation variation within each group, the results suggested increasing variation in the production of tones and phrasing as AoA increased. As shown in Table 4, the *NS* group agreed on a specific tone pattern for approximately 80% of the 69 words in the paragraph. The *Child* and *Adolescent Arrival* groups also reached consensus for about 70%–75% of the words while *Adult Arrivals* averaged 60%. An additional analysis to quantify and compare the deviations from the *NS* consensus patterns revealed wider divergence from the *NSs*' patterns as AoA increased across the three L2 groups. Specifically, the consensus patterns of the *Child Arrival* group were mostly NS-like with 76% overlap. The deviations from the *NS* patterns slightly increased with a later AoA, with the *Adolescent* and *Adult Arrival* groups producing respectively 72% and 69% overlapping consensus with the *NS* patterns.

4.5 Survey data analysis: putting the age effect to test

The survey data were obtained from L2 speakers' self-reports and self-ratings in the survey. The original variables went through a series of data reductions (see Huang, 2009b) and 27 predictors,

Table 3. Dominant tone patterns and numbers of agreements by word by group[a]

	Please	**call**	**Stella**	**Ask**	**her**	**to**	**bring**	**these**	**things**	**with**	**her**
1	H* (10)	!H* (9)	!H*L-L%(8)	H* < (8)	0 (9)	0 (10)	None	H* (7)	!H* H- (5)	0 (7)[a]	None[b]
2	H* (9)	!H* (6)	!H*L-L%(8)	H* < (7)	0 (8)	0 (10)	None	0 (6)	None	0 (7)	None
3	H* (9)	!H* (5)	None	None	0 (6)	0 (6)	None	0 (6)	None	0 (8)	None
4	H* (10)	!H* (9)	None	H*(10)	H*(6)	0 (7)	None	H* (6)	None	H*(5)/0(5)	None
	from	**the**	**store**	**Six**	**spoons**	**of**	**fresh**	**snow**	**peas**	**Five**	**thick**
1	0 (10)	0 (10)	L+H* L-L% (5)	H* (10)	!H* H-(6)	0 (9)	H* (9)	!H* (8)	None	H*H- (5)	H* (10)
2	0 (10)	0 (10)	H*L-L%(6)	H* (10)	!H* H-(7)	0 (8)	H* (5)	H* (6)	None	H* (6)	H* (7)
3	0 (9)	0 (10)	None	H* (9)	None	L-(5)/0 (5)	H* (7)	H* (7)	None	H* (9)	H* (7)
4	0 (7)	0 (9)	None	H* (9)	None	None	H* (5)	None	None	H* (8)	H* (6)
	slabs	**of**	**blue**	**cheese**	**and**	**maybe**	**a**	**snack**			
1	None	0 (6)	H* (10)	None	0 (10)	H* (8)	0 (10)	None			
2	!H* (5)	0 (6)	H* (7)	None	0 (10)	H* (6)	0 (10)	None			
3	None	0 (5)	H* (9)	L*L-H% (5)	0 (10)	H* (7)	0 (10)	None			
4	None	None	H* (7)	None	0 (9)	H* (5)	0 (8)	None			
	for	**her**	**brother**	**Bob**	**We**	**also**	**need**	**a**	**small**	**plastic**	**snake**
1	0 (9)	0 (8)	H* (6)	None	0 (10)	H* (6)	None	0 (7)	H* (8)	None	None
2	0 (10)	0 (7)	None	L*L-L%(5)	0 (8)	H* (5)	None	0 (10)	H* (5)	H* (5)	None
3	0 (10)	0 (8)	H* (5)	L*L-L%(6)	0 (9)	H* (9)	None	0 (8)	None	H* (5)	None
4	0 (8)	0 (6)	None	None	0 (9)	H* (8)	None	0 (8)	H* (6)	H* (6)	None
	and	**a**	**big**	**toy**	**frog**	**for**	**the**	**kids**			
1	0 (10)	0 (10)	H* (8)	H* (5)	!H* H-(5)	0 (9)	0 (10)	None			
2	0 (10)	0 (10)	H* (8)	H* (6)	None	0 (9)	0 (9)	H*L-L% (5)			
3	0 (8)	0 (9)	H* (9)	H* (6)	!H* H-(5)	0 (8)	0 (10)	L+H*L-L% (6)			
4	0 (7)	0 (9)	H* (5)	H* (6)	None	0 (9)	0 (10)	None			
	she	**can**	**scoop**	**these**	**things**	**into**	**three**	**red**	**bags**		
1	H* (6)	L- (6)	H* (8)	H* (6)	None	0 (10)	H* (9)	H* (8)	None		
2	0 (6)	None	None	None	None	0 (5)	H* (8)	!H* (5)	None		
3	H* (7)	L- (5)	None	None[c]	!H* H-(5)	None	H* (8)	!H* (6)	None		
4	H* (9)	L-(5)	H* (6)	None	0 (6)	0 (7)	H* (8)	!H* (7)	None		
	and	**we**	**will**	**go**	**meet**	**her**	**Wednesday**	**at**	**the**	**train**	**station**
1	0 (10)	0 (7)	0 (5)	H* (9)	H* (5)	None	None	0 (9)	0 (10)	H* (6)	L-L%(8)
2	0 (8)	H* (8)	None	H* (5)	H* (5)	H* H-(5)	None	0 (10)	0 (10)	H* (6)	L-L%(7)
3	0 (9)	None	None	H* (6)	H* (7)	None	None	0 (10)	0 (9)	H* (6)	L-L%(5)
4	0 (10)[d]	H* (7)	L- (5)	None	H* (6)	None	None	0 (7)	0 (8)	H* (6)	None

[a]The number of agreements, i.e. the number of participants in each group who produced the dominant tone patterns, was included in the parentheses (0 = no tones assigned; None = no dominant patterns established). Group 1 = *Native Speaker*, Group 2 = *Child Arrivals*, Group 3 = *Adolescent Arrivals*, Group 4 = *Adult Arrivals*.

[b]One *NS* participant skipped reading the word.

[c, d]Disfluency was observed for one participant.

Table 4. Proportions of consensus and consensus with *NS* patterns by group

	Proportion of consensus (within-group variation)	Proportion of overlapping consensus with *NS* (comparison with the *NS*s)
NS	78%	NA
Child	75%	76%
Adolescent	71%	72%
Adult	62%	69%

including age of arrival (AoA) and length of residence (LoR) as continuous variables, were derived from the survey data. See Table 5 for the list of predictors.

The earlier ANOVA results revealed an age effect among the L2 speaker groups for four prosodic features: speech rate, filtered speech rating, total number of pitch accents, and frequency of H-H%. However, since H-H% was only produced in *Adult Arrivals*, it was apparently a deviance among *Adult Arrivals*. The H-H% feature was thus excluded from further analyses.

To explore the validity of the age effect, bivariate correlation tests were first conducted between the 27 predictors and the three prosodic outcomes. Correlation results suggested that AoA was not the only variable significantly associated with the outcomes. Other predictors, such as the number of years of education received in the United States (Years of US Education) and their English language input, among others, were also correlated with the outcomes.

To understand the relative strength of these predictors in the outcomes, a stepwise multiple regression was selected to accommodate the small sample size and multiple predictors. Given the relatively small sample size and the numerous predictors, the results should be best conceived as exploratory in nature.

Since AoA and Years of US Education were fully confounded in the current study ($r = .89$, $p < .0001$),[9] keeping both variables in the regression models was conceptually and statistically problematic. Years of US Education was excluded from the regression analyses because the aim of the study was to investigate the AoA effect. As shown in Table 6, the final 2-predictor regression model (Model 2) accounted for approximately half of the variance in the speech rate (adjusted $R^2 = .51$, $p = .003$). AoA was no longer a significant predictor. The two significant predictors were Initial Motivation ($\beta = .58$, $p < .0001$) and Current English Media Input ($\beta = -.42$, $p = .003$).

The final model (Model 2) predicting filtered speech rating (see Table 7) accounted for approximately half the variance in the ratings (adjusted $R^2 = .46$, $p = .02$). AoA remained a significant predictor while controlling for the influence of other predictors ($\beta = -.49$, $p = .003$), and Initial Motivation was also a significant predictor of the filtered speech rating ($\beta = -.36$, $p = .02$).

Finally, Table 8 presents the results of the multiple regression models predicting total number of pitch accents. Both AoA ($\beta = .37$, $p = .034$) and Current L2 (English) Media Input ($\beta = -.37$, $p = .034$) were significant predictors. The two predictors accounted for one-third of the variances in the sample (adjusted $R^2 = .32$, $p = .034$).

4.6 Predicting the perception of foreign prosody

Bivariate correlation tests and stepwise regression analyses were conducted to explore the relationships between the prosodic features and listeners' perception of foreign prosody. Correlation tests revealed that seven variables correlated significantly with the filtered speech rating, including speech rate ($r = .66$, $p < .0001$), articulation rate ($r = .69$, $p < .0001$), total number of pitch accents

Table 5. Bivariate correlations between three prosodic outcomes (speech rate, filtered speech rating, and total number of pitch accents) and predictors by construct[a]

	Speech rate	Filtered speech rating	Number of pitch accents
Age of Arrival (AoA)	–0.51**	–0.62**	0.50**
Age	–0.31	–0.50**	0.32
Length of Residence (LoR)	0.37*	0.27	–0.30
Years of US Education	0.51**	0.63**	–0.55**
Total Years of Education	–0.13	–0.26	0.18
Years of ESL Instruction in the US	–0.20	0.03	0.02
Gender	–0.09	0.10	–0.03
Age of First Exposure to English	–0.39*	–0.53**	0.40*
Second Language Input			
Parents' English Proficiency Composite	0.28	0.57**	–0.28
Initial English Media Input Variable	0.25	0.22	–0.31
Current English Media Input Variable	0.61**	0.40*	–0.50**
Initial English Literacy Input Variable	0.27	0.21	–0.15
Current English Literacy Input Variable	0.47**	0.42*	–0.36
Initial Oral English Input Variable	–0.01	–0.27	0.07
Current Oral English Input Variable	0.47**	0.43*	–0.49**
Initial English Proficiency			
Initial English Proficiency Composite	–0.29	–0.39*	0.33
Language Aptitude			
Language Aptitude Composite	0.08	0.26	–0.10
Motivation			
Initial Motivation Composite	–0.46*	–0.54**	0.24
Current Motivation Composite	–0.31	–0.23	0.26
Use of Language Learning Strategy			
Initial Use of Language Learning Strategy Composite	–0.43*	–0.50**	0.27
Current Use of Language Learning Strategy Composite	–0.35	–0.25	0.19
Cultural Affiliation			
Initial American Culture Identification Variable	–0.33	–0.23	0.07
Current American Culture Identification Variable	0.09	0.14	0.01
Initial Heritage Culture Appreciation Variable	–0.37*	–0.10	–0.12
Current Heritage Culture Appreciation Variable	0.05	0.13	–0.04
Psychological-Affective			
Initial Embarrassment Variable	–0.34	–0.10	–0.05
Current Embarrassment Variable	–0.56**	–0.35	0.39*

[a] Predictors include both single variables and composites. Composites are averages of multiple variables measuring the same construct. *Parents' English Proficiency Composite* is an average of ratings of L2 speakers' father's and mother's English proficiency. The *Second Language Input* variables are estimated English input in disparate modalities during the initial few years and most recent few years of L2 speakers' residence in the US. *Initial English Proficiency Composite* is an average of English proficiency upon L2 speakers' arrival in the US in five domains (i.e., listening, speaking, reading, writing, and pronunciation). *Language Aptitude Composite* is an average of ratings of sound processing ability and language learning ability. *Initial/Current Motivation Composite* is averages of motivation to learn English and concerns about English proficiency in the initial years and the most recent few years. *Initial/Current Use of Language Learning Strategy Composite* is also averages of attention to linguistic forms in the aural and written language input. The *Cultural Affiliation* variables are ratings of the extent to which L2 speakers identified with American and Chinese cultures. The *Psychological-Affective* variables are ratings of affective responses toward communication breakdowns.
*$p < .05$; **$p < .01$.

($r = –.57$, $p < .0001$), total number of Intermediate phrases ($r = –.37$, $p = .020$), total number of Intonational phrases ($r = –.42$, $p = .008$), frequency of H* ($r = –.32$, $p = .046$), and frequency of H-H% ($r = –.54$, $p < .0001$).

Table 6. Summary of stepwise multiple regression analyses predicting speech rate with AoA and covariates

Model	R^2	Adjusted R^2	R^2 change	F change	p	Predictors	B	SE β	β
1	.37	.35	.37	16.32	.000	Current L2 Media Input	.01	.00	.61**
2	.54	.51	.18	10.28	.003	Current L2 Media Input	.01	.00	.58**
						Initial Motivation	–.09	.03	–.42**

Table 7. Summary of stepwise multiple regression analyses predicting filtered speech rating with AoA and covariates

Model	R^2	Adjusted R^2	R^2 change	F change	p	Predictors	B	SE B	β
1	.39	.36	.39	17.52	.000	AoA	–.17	.04	–.62**
2	.50	.46	.11	6.11	.020	AoA	–.13	.04	–.49**
						Initial Motivation	–.32	.13	–.36*

Table 8. Summary of stepwise multiple regression analyses predicting total number of pitch accents with AoA and covariates

Model	R^2	Adjusted R^2	R^2 change	F change	p	Predictors	B	SE B	β
1	.25	.22	.24	9.10	.005	AoA	.33	.11	.50**
2	.36	.32	.12	5.01	.034	AoA	.24	.11	.37*
						Current L2 Media Input	–.06.	.03	–.37*

Table 9. Summary of stepwise multiple regression analyses predicting filtered speech rating (excluding articulation rate) with prosodic variables

Model	R^2	Adjusted R^2	R^2 change	F change	p	Predictors	B	SE B	β
1	.66	.43	.43	29.17	.000	Speech Rate	2.40	.45	.66**
2	.76	.57	.14	12.16	.001	Speech Rate	2.02	.41	.55**
						H-H%	–.51	.15	–.39**

Given that speech rate and articulation rate were highly correlated (r = .95), the two variables could not be examined in the same model. Two separate models were thus built with each variable. Table 9 presents the results with speech rate and other prosodic variables predicting filtered rating. The final model (Model 2) included two significant predictors, speech rate (β = .55, p < .001) and H-H% (β = –.39, p = .001). The two predictors combined accounted for more than half of the variances in the filtered ratings (adjusted R^2 = .57, p = .001).

Replacing speech rate with articulation rate in the regression analysis yielded similar results (see Table 10). The final model included articulation rate (β = .57, p < .0001) and H-H% (β = –.33, p = .007) as significant predictors, and accounted for similar amount of variances in the ratings (adjusted R^2 = .55, p = .007).

Table 10. Summary of stepwise multiple regression analyses predicting filtered speech rating (excluding speech rate) with prosodic variables

Model	R^2	Adjusted R^2	R^2 change	*F* change	*p*	Predictors	B	SE B	β
1	.48	.47	.48	35.21	.000	Articulation Rate	2.53	.43	.69**
2	.58	.55	.10	8.30	.007	Articulation Rate	2.09	.42	.57**
						H-H%	−.43	.15	−.33**

5 Discussion

The current study aimed to investigate the age of learning effect on the ultimate outcome of second language (L2) prosody. Past research on L2 phonological acquisition has primarily focused on the segmental properties, such as consonant production (e.g., Birdsong, 2007; Flege, Frieda, Walley, & Randazza, 1998; Flege et al., 1995), vowel quality (Flege, MacKay, & Meador, 1999; Piske, Flege, MacKay, & Meador, 2002) or global foreign accents (e.g., Flege, Yeni-Komshian, & Liu, 1999). Little research has been devoted to examining the age effect on the production of L2 prosody. The present study fills this gap in the literature and explores the age effect on the production of L2 prosody with speech rate and articulation rate analysis, the low-pass filtering technique, and ToBI—the phonological transcription system of prosody. To explore the validity of the age effect, the current study included a wide range of covariates that may also explain variations in L2 prosody outcomes (e.g., amount of L2 exposure, motivation, and cultural identity) and employed multiple regression models to evaluate the relative strength of AoA and other variables. Additionally, the study examined the relationships between the prosodic properties and native-speaker listeners' ratings of the filtered speech to understand the basis of the perception of foreign prosody.

Results from the study suggest that the strength of the AoA constraints vary depending on the specific aspects of prosody under investigation, and *Adult Arrivals* overall appear to deviate most from the native speakers (*NSs*) and the other two L2 speaker groups. Specifically, the analysis of speech rate revealed that *Adult Arrivals* read slower than *NSs* and *Child Arrivals*, but were similar to *Adolescent Arrivals*. *Child* and *Adolescent Arrivals* were not significantly different from *NSs*. Despite these observed trends, the regression analysis revealed that the differences in speech rates were predicted by differences in English media exposure and motivation to learn English rather than by AoA.

The current results diverged from previous work by Guion and colleagues (Guion, Flege, Liu, & Yeni-Komshian, 2000), where the speech rate differences were predicted by AoA. The divergence may be explained by the differences in the stimuli, study design, or sample sizes. Guion and colleagues included a larger sample ($n = 240$) than did the current study, and used sentences rather than a passage to elicit speech samples. Although the selected sentences were composed of short words, they appeared to contain more low frequency words, such as "rook", "soot", and "teak" (word frequency = 1.39, 1.08, and 0.39 per million words, respectively[10]), than the passage used in the present study. In addition, they only controlled for three other covariates (i.e., chronological age, length of residence, and use of L1) rather than conducting a comprehensive evaluation of the age effect. Media exposure, for example, was not examined in their study, and may have been a strong predictor had it been included. In fact, in another study by Flege, Yeni-Komshian, and Liu (1999), media exposure and language use were argued to be two independent variables in the prediction of immigrants' foreign accents. Results from the current study also indicate that media exposure significantly predicted foreign accent outcomes whereas language use did not.

The articulation rate analysis showed that *Adult Arrivals* were slower than *NSs* in their speech rates, but not significantly different from *Child* and *Adolescent Arrivals*. Results combined from the speech and articulation rate analyses suggested that *Adult Arrivals* produced more pauses and disfluencies than the other two L2 speaker groups, but were not necessarily articulating segments slower than the other two L2 groups.

The analysis of filtered speech ratings indicates that the prosody productions of both *Adolescent Arrivals* and *Adult Arrivals* were perceived to deviate significantly from *NS* prosody. *Adult Arrivals* were also perceived to have a stronger foreign prosody than the *Child Arrivals*. However, *Adult Arrivals* were not significantly different from *Adolescent Arrivals*, and *Child Arrivals* were similar to *NSs*. The AoA effect remained significant even after controlling for the effects of other variables. This particular finding paralleled the robust age effect in the global foreign accent literature (e.g., Flege, Yeni-Komshian, & Liu, 1999; Oyama, 1976; Tahta et al., 1981; Thompson, 1991). In addition, L2 speakers' motivation to learn English in the initial years of their residence was also a significant predictor of the filtered speech ratings, though in an interestingly negative direction. L2 speakers who were more motivated to learn English may have made more efforts to articulate each segment at the cost of prosodic fluency. Alternatively, the unexpected negative predictive power may suggest that the motivation variable was in fact a proxy for other variables that contribute to the filtered speech ratings, such as L2 speakers' initial English proficiency. That is, L2 speakers with lower English proficiency may have a stronger motivation to learn English.

The analyses of prosodic grouping and intonation patterns using ToBI labeling provided further specificities of the age of learning effect. Specifically, *Child Arrivals* produced similar numbers of intermediate phrases and intonational phrases as the *NSs* whereas *Adult Arrivals* diverged from the *NSs* by producing superfluous prosodic groupings. *Adult Arrivals*' excessive uses of phrasal breaks indicated that they divided the paragraph into smaller chunks instead of linking the words in a native-like way to create a meaningful flow. Some of the phrasal breaks were also placed in locations not typical of American English, suggesting that *Adult Arrivals* had trouble processing the sentence and/or cuing the information structure of the sentence prosodically when reading aloud.

Similarly, the analysis of the different types of tones revealed a robust AoA effect on the total frequency of pitch accents even after controlling for the effects of other variables. Later arrivals produced superfluous pitch accents compared to earlier arrivals and assigned pitch accents in normatively inappropriate locations. Exposure to English media was again found to be a significant predictor, reaffirming its facilitating role in L2 prosody learning.

Among the various tone types, an age effect was found only in the production of the high pitch accent (H*) and the high boundary tones (H-H%), although descriptive trends were also observed in a few other tone types with later arrivals deviating more from *NSs*. In particular, *Adult Arrivals* produced more high pitch accents (H*) than both *NSs* and *Adolescent Arrivals*. The excessive use of H* by *Adult Arrivals* suggests that they tend to put emphasis on words regardless of their information status. The findings show that the *Adult Arrivals* had not mastered the relationship between meaning and prosody, and thus defaulted to H* even for words that could have other legitimate tone assignments or for words that should not bear a pitch accent at all given the context (e.g., function words). Additionally, *Adult Arrivals* was the only group that produced H-H% boundary tones, normally seen in question intonation or continuation rises and thus inappropriate for the current paragraph consisting of only declarative and imperative sentences in read speech. The production of H-H% was not limited to a certain individual but was observed among half of the *Adult Arrivals*, again suggesting a common lack of knowledge of the complex prosody–meaning relationship.

The results of dominant pattern analyses corroborated the quantitative analyses of prosodic groupings and intonation categories; the amount of both the within-group variation and deviations from the *NS* patterns increased as the AoA advanced. The age-related trends paralleled the L2 proficiency-related trends found in Ueyama and Jun (1998) and Jun and Oh (2000), both of which utilized ToBI to analyze L2 prosody. Specifically, beginning Japanese and Korean learners of English in Ueyama and Jun (1998) produced more pitch accents and phrasal breaks than more advanced learners. In a similar vein, beginning English learners of L2 Korean in Jun and Oh's (2000) study also produced more accentual phrases (APs) in Korean—which are analogous to the distribution of pitch accents in English—than advanced learners. Furthermore, some of the AP boundaries produced by the beginning learners of L2 Korean were inappropriate in native Korean prosody. This particular finding is consistent with the findings of dominant patterns in the current study, where later AoA groups assigned pitch accents and phrasal breaks in locations that do not reflect the patterns of *NSs*.

Finally, the exploratory investigation of the contributing factors to the filtered speech ratings revealed that speech rate and articulation rate were the major contributors to the listeners' perception of a speaker's prosody as foreign. The other significant predictor was the frequency of H-H% boundary tones, which were, however, observed only among *Adult Arrivals*. The foreign prosody perception thus appeared to be mainly based on speech and articulation rate. The results were consistent with previous research revealing speech rate as a significant contributor of fluency outcomes (e.g., Cucchiarini et al., 2002; Pennington, 1992) and global foreign accent ratings (e.g., Munro, 1995; but cf. Munro & Derwing, 2001[11]). A few explanations may be offered for the current results: 1) speech/articulation rate could indeed play a crucial role in the perception of foreign prosody accentedness, and exert more influence on the perception of foreign prosody than the accuracy of producing phonological categories of prosody; 2) alternatively, despite being explicitly instructed to rate the *intonation* of the filtered speech, the listeners may have been unable to focus on the tonal aspect and instead judged the speech based on the fluency of the speech; or 3) the filtered material may not be appropriate for judging the goodness of prosody. Native-like intonation would come not only from the right choice of pitch accent or boundary tones but also from the right timing of tone realizations relative to the segmental property. Evaluating tone–text timing or alignment would not be easy in filtered speech.

Taken together, the current study confirms the validity of the age of learning effect on the acquisition of three L2 prosodic properties, i.e., global foreignness in prosody, the frequency of pitch accents, and the production of H-H% boundary tones. Although group differences were not statistically significant for articulation rate, prosodic groupings, and the rest of the ToBI-labeled phonological categories, age-related trends were observed for all of these properties of L2 prosody. Overall, *Child* and *Adolescent Arrivals* patterned more similarly to the *NSs* than *Adult Arrivals* did. However, *Adult Arrivals*' performance was better than expected, results which may be attributed to the constraints of the stimuli and the method of analysis employed in the present study. As mentioned in the methods section, the paragraph stimulus consists of declarative and imperative sentences only and is thus quite limited in eliciting and assessing a wide variety of intonation patterns. Furthermore, ToBI transcribes phonological tonal categories but does not distinguish different ways of realization of the same category if the categories themselves do not interfere with semantic/pragmatic meaning. For example, an f0 contour consisting of a pitch peak followed by falling pitch at the end of a phrase is labeled "H* L- (L%)" regardless of the exact shape of the falling slope after the peak. Thus, the prosodic foreignness perceptible through the realization of certain categories may not have been captured by the ToBI analyses.

6 Conclusion and directions for future research

To conclude, the current exploratory study provides evidence for an overall advantage in the early learning of L2 prosody and demonstrates the heterogeneous effects of the age variable on different aspects of L2 prosody. The results also appear to suggest the prominent roles of media exposure and motivation in the ultimate outcomes of certain prosodic features. However, the generalizability of the current results is limited by the relatively small sample size and the sample's homogeneous L1 background. Future empirical replications with larger sample sizes and with different linguistic origins would help verify and clarify the contributions of AoA and other variables to L2 prosody learning. Additionally, exploring or creating novel analytical techniques to characterize prosody, such as using the ToBI model to analyze phonological tonal categories as in the present study, would contribute to advancing methodology in L2 prosody research. In our study, however, the ToBI analysis of phonological categories revealed the general developmental trends, but did not capture the group differences in prosody as reflected in the speech rate analysis and the native speakers' ratings of segment-filtered speech. Future research is clearly needed to examine the phonetic realizations of the phonological tonal categories in order to reach a comprehensive understanding of the age effect on the acquisition of L2 prosody (see Ueyama & Jun, 1998 for the comparison among groups based on phonetic realizations of the H- phrase accent in English).[12]

Future research would also greatly benefit from methodological improvements in establishing native speaker baseline patterns as well as by quantifying L2 speakers' prosodic deviances from the native speaker patterns. For example, in order to yield more consistent baseline intonation patterns for comparison, stimuli could be constructed with constrained contexts to reduce the variations in native speakers' intonation patterns.

Finally, it will be equally beneficial to research the sources and causes of the perception of accentedness. The present results from correlation and multiple regression tests, though based on a small sample size, suggest that the perception of non-native prosody in segment-filtered speech appears to be largely based on speech or articulation rate. However, correlation results do not indicate causation. Future research with experimental methodology is needed to verify the causal link between speech or articulation rate and the perception of foreignness in prosody. The accuracy of the inferences is also constrained by the small sample size. Further research with larger samples is also needed to understand the relative importance of segmental and suprasegmental features in the perception of global foreign accents (e.g., Boula de Mareüil & Vieru-Dimulescu, 2006; Vicenik & Sundara, 2008).

The current study serves as a preliminary effort in the investigation of the age constraints in ultimate L2 prosody attainment. Further research in the aforementioned directions is needed before a comprehensive picture can be revealed.

Acknowledgements

We are grateful to Peggy Li, James Flege, Sameer Khan, and the members of the UCLA Phonetics Laboratory for their assistance and feedback at various stages of this study. We also thank Marissa Tse, Alex Wang, Tammy Ko, Tia Fang, Tiffany Tseng, and Eric Williams for their help with data collection, speech segmentation, and ToBI labeling. Special thanks to all the participants in the study for their time and patience. Thierry Nazzi, Mitsuhiko Ota, and one anonymous reviewer provided very insightful comments, which helped improve this paper. All the remaining errors are ours.

Data in the current study were collected as part of the first author's dissertation research at UCLA. Portions of the data were presented at the 33rd Boston University Conference on Language Development in October 2008, and were later published in the online supplements of the conference proceedings. The complete results were also presented at the Biennial Meeting for the Society for Research in Child Development in April 2009.

The study was funded by dissertation grants from the Chiang Ching-Kuo Foundation for International Scholarly Exchange and from UCLA to the first author, and a UCLA Faculty Research Grant to the second author.

Notes

1 We thank Steven Weinberger at George Mason University for giving us permission to use the paragraph.
2 All words in the paragraph are common and high frequency words except for "scoop" and "slab" (5.67 and 1.84 per million words according to the SUBTLEXus corpus (Brysbaert & New, 2009)). However, they are phonologically simple for decoding, and no disfluency for these two words was observed among the participants.
3 We thank one of the reviewers for bringing this issue to our attention.
4 The number of syllables produced was adjusted for participants who skipped ($N = 2$), repeated ($N = 2$), or added words ($N = 1$) in reading the paragraph. Note that only complete words were considered for the adjustment. Fillers such as "hm", "er", "ah", and single syllables or words truncated due to disfluency were not counted as additional syllables.
5 Gaps in the waveform that contain a stop or affricate closure were considered pauses if longer than 200ms.
6 The overall mean rating is lower (range: 3.45–6.86) than that reported in the literature based on unfiltered speech because—as indicated by the low rating for the *NS* group ($M = 6.86$)—rating low-pass filtered speech is much harder than rating unfiltered speech possibly due to the fact that the recordings not only deviate from natural speech but lack segmental information to judge the timing of tone realization and alignment between tones and text.
7 Since the edge of an IP always corresponds to the edge of an ip, an IP boundary tone is always preceded by an ip phrase accent, creating four tonal contours (i.e., L-L%, L-H%, H-L%, H-H%) at the end of an IP. Thus, the analysis was performed on the nine pitch accent types, two phrase accent types, and four boundary contour types.
8 We only present the mean frequencies for six out of the nine pitch accent types because the other three types (i.e., !H+!H*, L+!H*, L*+!H) had very low frequencies ($M < 0.5$ across all four groups).
9 Partial correlational tests were conducted to tease apart the contributions of AoA and Years of US Education to the outcomes. However, after removing the influence of the two variables on each other, neither variable remained significantly associated with the outcomes, suggesting substantial overlaps between the two variables in the current sample.
10 The word frequency indexes are also based on the SUBTLEXus corpus (Brysbaert & New, 2009).
11 Munro and Derwing (2001) manipulated speaking rates in L2 speech using speech compression-expansion software, and found a curvilinear relationship between speech rate and foreign accent ratings. L2 speech with a slightly faster rate than the average L2 speaking rate was perceived as least accented whereas very fast and very slow speech were both perceived as more foreign-accented. We thank the anonymous reviewer for bringing this issue to our attention.
12 In the current MAE_ToBI, the only tonal labels reflecting the differences in phonetic realizations are diacritics for a delayed peak ("<") and an early peak (">"). In our data, *NS* and *Child Arrivals* produced a delayed peak of H* on *ask* (in *ask her to bring*) more often than *Adolescent Arrivals*, while *Adult Arrivals* never produced this pattern (see the dominant patterns in Table 3).

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529–555.

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53*, 334–341.

Audacity Team (2000). Audacity (Version 1.2.5) [Computer program]. Retrieved October 31, 2009 from http://audacity.sourceforge.net/
Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes, 11*, 17–67.
Beckman, M. E., & Ayers-Elam, G. (1994). *Guidelines for ToBI labeling*. Retrieved October 31, 2009, from Ohio University Linguistics Department Website: http://www.ling.ohiostate.edu/research/phonetics/E_ToBI/singer_tobi.html
Beckman, M. E., & Hirschberg, J. (1994). *The ToBI annotation conventions*. Unpublished manuscript, Ohio State University.
Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 9–54). Oxford: Oxford University Press.
Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in English and Japanese. *Phonology Yearbook, 3*, 255–310.
Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning* (pp. 99–116). Amsterdam: John Benjamins.
Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.0.45) [Computer program]. Retrieved October 31, 2009 from http://www.praat.org/
Boula de Mareüil, P., & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica, 63*, 247–267.
Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 997–990.
Chambers, F. (1997). What do we mean by fluency? *System, 4*, 535–544.
Chen, A., & Mennen, I. (2008). Encoding interrogativity intonationally in a second language. In *Proceedings of the 4th International Conference on Speech Prosody*. 6–9 May, Campinas, Brazil.
Cucchiarini, C., Strik, H., & Boves, L. (2000). Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology. *Journal of the Acoustical Society of America, 107*, 989–999.
Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America, 111*, 2862–2873.
Elsendoorn, B. (1984). *Tolerances of durational properties in British English vowels*. Unpublished doctoral dissertation. Institute of Phonetics, University of Utrecht.
Flege, J. (1979). *Phonetic interference in second language acquisition*. Unpublished doctoral dissertation. Department of Linguistics, Indiana University, Bloomington.
Flege, J., Frieda, E., Walley, A., & Randazza, L. (1998). Lexical factors and segmental accuracy in second-language speech production. *Studies in Second Language Acquisition, 20*, 155–188.
Flege, J. E., MacKay, I., & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America, 106*, 2973–2987.
Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication, 16*, 1–26.
Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language, 41*, 78–104.
Fougeron, C., & Jun, S.-A. (1998). Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics, 26*, 45–69.
Golinkoff, R. M. (1983). *The transition from prelinguistic to linguistic communication*. Hillsdale, NJ: Lawrence Erlbaum.
Guion, S. G., Flege, J. E., Liu, S., & Yeni-Komshian, G. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21*, 205–228.
Guion, S., Flege, J., & Loftin, J. (2000). The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *Journal of Phonetics, 28*, 27–42.

Hirschberg, J., & Pierrehumbert, J. (1986). Intonational structuring of discourse. *Proceedings of the 24th Meeting of the Association for Computational Linguistics*, New York (pp. 126–144).
Höhle, B., Bijeljac-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). The development of language specific prosodic preferences during the first half year of life: Evidence from German and French. *Infant Behavior and Development, 32*, 262–274.
Huang, B. H. (2009a). The earlier the better? Age effect on second language phonology and grammar. In *Proceedings of the 2008 International Conference on Applied Linguistics*, Chiayi, Taiwan (pp. 137–158).
Huang, B. H. (2009b). *Age-related effects on the acquisition of second language phonology and grammar*. Unpublished doctoral dissertation. The University of California, Los Angeles, CA.
Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Unpublished doctoral dissertation. The Ohio State University, Columbus, OH. [Published in 1996, Garland Publishing Inc., New York.]
Jun, S.-A., & Oh, M. (2000). Acquisition of second language intonation. *Proceedings of International Conference on Spoken Language Processing, 4*, 76–79.
Jusczyk, P. W., Cutler, A., & Redanz, N. (1993). Preference for the predominant stress patterns of English words. *Child Development, 64*, 675–687.
Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*(3), 402–420.
Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159–207.
Kjelgaard, D., & Speer, S. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153–194.
Ladd, D. R. (1996). *Intonational phonology*. New York: Cambridge University Press.
Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4*, 185–199.
Liberman, M. (1975). *The intonational system of English*. Unpublished doctoral dissertation. MIT, Cambridge, MA.
MacKay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics, 27*, 157–183.
Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*, 381–400.
MathWorks Inc. (2003). MATLAB version 6.5.1. Natick, MA.
Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*(3), 241–265.
Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes, 22*, 898–909.
Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*, 81–108.
Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition, 17*, 17–34.
Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*, 289–306.
Munro, M., & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition, 23*, 451–468.
Nazzi, T., Iakimova, I., Bertoncini, J., Frédonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for cross-linguistic differences. *Journal of Memory and Language, 54*, 283–299.
Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language, 43*, 1–19.
Oyama, S. (1973). *A sensitive period for the acquisition of a second language*. Unpublished doctoral dissertation. Harvard University.
Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, *5*, 261–283.

Pennington, M. C. (1992). Discourse factors related to L2 phonological proficiency: An exploratory study. In J. Leather & A. James (Eds.), *Proceedings of the 1992 Amsterdam Symposium on the acquisition of second-language speech* (pp. 137–155).
Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation.* Unpublished doctoral dissertation. MIT, Cambridge, MA.
Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonation in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.
Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics, 6*(2), 179–208.
Piske, T., Flege, J., MacKay, I., & Meador, D. (2002). The production of English vowels by fluent early and late Italian–English bilinguals. *Phonetica, 59*, 49–71.
Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191–215.
Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421–435.
Price, P., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America, 90*(6), 2956–2970.
Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of non-native speaker conversations. *Discourse Process, 14*, 423–441.
Ross, K., Ostendorf, M., & Shattuck-Hufnagel, S. (1992). Factors affecting pitch accent placement. In *Proceedings of the 2nd International Conference on Spoken Language Processing* (pp. 365–368).
Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House.
Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigations of auditory sentence processing. *Journal of Psycholinguistic Research, 25*(2), 193–247.
Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., Alves Limissuri, R., & Peperkamp, S. (2009). Language-specific stress perception by 9-month- old French and Spanish infants. *Developmental Science, 12*, 914–919.
Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech, 24*, 265–272.
Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41*, 177–204.
Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 1*, 84–119.
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.
Tsukada, K., Birdsong, D., Mack, M., Sung, H., Bialystok, E., & Flege, J. (2004). Release bursts in English word-final voiceless stops: A cross-sectional and longitudinal study of Korean adults' and children's speech production. *Phonetica, 61*, 67–83.
Ueyama, M., & Jun, S.-A. (1998). Focus realization in Japanese English and Korean English intonation. *Japanese and Korean Linguistics, 7*, 629–645.
Vicenik, C., & Sundara, M. (2008). *Role of segmental and intonational cues in dialect discrimination*. Poster presented at Acoustics '08, the 155th meeting of the Acoustical Society of America, Paris, France.
Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition, 20*(1), 1– 25.
Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of nonnative speech sounds. *Canadian Journal of Psychology, 37*, 278–286.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization in the first year of life. *Infant Behavior and Development, 7*, 49–63.
Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*, 1707–1717.

Appendix A: Sample survey questions

Full Name:_____ Interviewer Initial:____ Date:_____ (mm/dd/yy) Main lang:____Chi ____Eng ID:_____

Language Learning History

1. What language(s) do you and your parents speak? That is, languages that you can carry a conversation. (Check all that apply)

You	__Mandarin	__Taiwanese	__Cantonese	__English	__Other (specify languages/dialects & age of learning)
Mother	__Mandarin	__Taiwanese	__Cantonese	__English	__Other (specify languages/dialects & age of learning)
Father	__Mandarin	__Taiwanese	__Cantonese	__English	__Other (specify languages/dialects & age of learning)

2. Please rate your parents' English proficiency. (Circle)

MOTHER											
Listening & Speaking	No proficiency	1	2	3	4	5	6	7	8	9	Native-like
Pronunciation/Accent	No proficiency	1	2	3	4	5	6	7	8	9	Native-like
Reading & Writing	No proficiency	1	2	3	4	5	6	7	8	9	Native-like
FATHER											
Listening & Speaking	No proficiency	1	2	3	4	5	6	7	8	9	Native-like
Pronunciation/Accent	No proficiency	1	2	3	4	5	6	7	8	9	Native-like
Reading & Writing	No proficiency	1	2	3	4	5	6	7	8	9	Native-like

3. What language or languages do your parents usually speak to each other at home? (If not applicable, write NA in the "Other")

__Mandarin __Taiwanese __Cantonese __English __Other languages/dialects (specify)

4. At what age did you start learning English? (i.e. first study of one semester or more) Age _____

5. Where did you start to learn English? _____ In my native country _____ In the US _____ Other (specify)

6. How did you start to learn English? (please check all the appropriate answers)

_____ In school class taught by a native speaker
_____ In school class taught by a non native speaker
_____ Taught by a native speaker (_____)
_____ Taught by a non native speaker (_____)
_____ Self learning (please specify how (_____)
_____ Picked it up naturally

7. How frequent is the English instruction, and for how long? (Write NA if not applicable)

8. Have you received any intensive training in English pronunciation/accent correction? _____Yes _____No

If yes, please describe the training you received (when/where/details of the training):

(Continued)

Appendix A: (Continued)

9. Please list all places (city, country) in which you have lived for more than 3 months EXCEPT for the U.S. and your native country. If you have not lived in other places for more than 3 months, please just leave the question blank.

(a) _____ from _____ (month, year) to _____ (month, year)
(b) _____ from _____ (month, year) to _____ (month, year)

10. Please provide information about your formal education (schooling) below, including the number of years, the country where you received the education, and the primary language the courses were taught in.

Level	*Country (circle)*	*Grade (or Number of Years)*	*Primary Language Used in Class (check)*			*Field of Study (if applicable)*
Preschool & Kindergarten	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	
Elementary School	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	
Secondary School	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	
College	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	
Post-College I	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	
Post-College II	Native Country		___Native language	___English	___Other (specify)	
	U.S.A.		___Native language	___English	___Other (specify)	
	Other		___Native language	___English	___Other (specify)	