Applied Psycholinguistics **28** (2007), 251–276
Printed in the United States of America
DOI: 10.1017.S0142716407070130

Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals

PAVEL TROFIMOVICH
Concordia University

WENDY BAKER
Brigham Young University

Received: March 8, 2006 Accepted for publication: July 3, 2006

ADDRESS FOR CORRESPONDENCE
Pavel Trofimovich, Department of Education/TESL Centre, Concordia University (EN 413), 1455 de Maisonneuve Boulevard West, Montréal, Québec H3G 1M8, Canada. E-mail: pavel@education.concordia.ca

ABSTRACT
This study examined second language (L2) experience effects on children's acquisition of fluency- (speech rate, frequency, and duration of pausing) and prosody-based (stress timing, peak alignment) suprasegmentals. Twenty Korean children (age of arrival in the United States = 7–11 years, length of US residence = 1 vs. 11 years) and 20 age-matched English monolinguals produced six English sentences in a sentence repetition task. Acoustic analyses and listener judgments were used to determine how accurately the suprasegmentals were produced and to what extent they contributed to foreign accent. Results indicated that the children with 11 years of US residence, unlike those with 1 year of US residence, produced all but one (speech rate) suprasegmentals natively. Overall, findings revealed similarities between L2 segmental and suprasegmental learning.

How do second language (L2) learners acquire the sound system of their L2? To date, most research addressing this question has focused on the acquisition of individual L2 segments (vowels and consonants) by adult learners in a variety of contexts, ranging from the speech laboratory (e.g., Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997), to the language classroom (e.g., Derwing, Munro, & Wiebe, 1998), to naturalistic L2-speaking environments (e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004). Much less research has, however, examined the learning of L2 suprasegmentals, which refer to tonal, temporal, and dynamic features associated with such aspects of phonology as stress, rhythm, and intonation (Trofimovich & Baker, 2006; see Chun, 2002, for review).

This paucity of research on the acquisition of L2 suprasegmentals is striking given the importance of suprasegmentals in language learning and use. For example, it has been shown that explicit instruction focusing on suprasegmentals, more so than segmental training, leads to improvements in spontaneous L2 speech (Derwing et al., 1998). Suprasegmentals have also been found to determine listener judgments of comprehensibility and accentedness in L2 speech (Munro & Derwing, 1998). The scarcity of research on the acquisition of L2 suprasegmentals is even more striking because there are virtually no studies on L2 suprasegmental learning by child L2 learners (see Guion, 2005; Guion, Harada, & Clark, 2004, for rare exceptions). The goal of the present study is to fill this gap by offering an investigation of child L2 learners' acquisition of five suprasegmentals.

Previous research on L2 phonological learning by children, particularly on the acquisition of individual segments, offers a point of departure for understanding L2 suprasegmental learning. There are several reasons, both theoretical and practical in nature, for using insights from L2 acquisition of individual segments to motivate the study of L2 suprasegmental learning. At the theoretical level, there is evidence that both segmental and suprasegmental learning may be driven by similar learning mechanisms (Pierrehumbert, 2003) and may be influenced by common factors (Trofimovich & Baker, 2006; Wang, Spence, Jongman, & Sereno, 1999). At the practical level, as discussed above, contextualizing the study of L2 suprasegmental learning within research on the acquisition of L2 segments may be requisite simply because relatively little is currently known about how L2 suprasegmentals are learned (see Chun, 2002).

To date, research on the acquisition of individual segments by child L2 learners has yielded two main findings. The first finding is that child learners, when given a sufficient amount of experience with an L2, often (but not always) attain native or near-native accuracy in L2 perception and production (Flege, MacKay, & Meador, 1999; Munro, Flege, & MacKay, 1996). For example, Flege et al. (1999) found that, after about 40 years of residence in an English-speaking country, native Italian speakers who were first exposed to English in childhood (before age 7) produced all of the 10 examined English vowels near natively. Other evidence, however, suggests that child learners do not attain nativelike perception and production accuracy for some L2 segments even after a substantial amount of L2 experience (Pallier, Bosch, & Sebastián-Gallés, 1997; Sebastián-Gallés & Soto-Faraco, 1999). Flege (1998), for instance, demonstrated that a similar group of native Italian speakers (first exposed to English between ages 2 and 8) produced some English vowels less accurately than native English speakers did, a finding that was attributed to differences in degree of cross-language similarity between English and Italian vowels (Baker & Trofimovich, 2005; see Flege, 1995, for review). Taken together, these findings suggest that, although child learners are overall successful in L2 segmental learning, their learning outcomes often depend on the particular L2 segment studied.

The second finding prompted by research on child learners' acquisition of L2 segments relates to the effect of experience on learning. There is evidence that child learners, at least in beginning stages of learning, improve in their perception and production of L2 segments after a relatively short amount of experience,

especially when compared to adult learners (Aoyama et al., 2004; but see Snow & Hoefnagel-Hoehle, 1978). Aoyama et al. (2004), for example, demonstrated that native Japanese children, but not adults, improved in their production of English /r/ after 1 year of residence in an English-speaking country. Other research, however, suggests that learning other L2 segments may require extensive amounts of L2 experience. Baker and Trofimovich (2005), for instance, showed that native Korean children (exposed to English between ages 7 and 13) did not produce some English vowels natively after an average of 8 years in an English-speaking environment (with the amount of experience ranging for individual children between 5 and 15 years). These findings suggest that the amount of L2 experience (and perhaps its nature) profoundly influences how accurately child learners perceive and produce L2 segments. In fact, at least for some segments, the amount of such experience required may approach and exceed 10 years of residence in an L2-speaking country (Baker & Trofimovich, 2005).

Does L2 suprasegmental learning resemble the acquisition of L2 segments? That is, do child L2 learners, with an increasing amount of L2 experience, produce L2 suprasegmentals like native speakers of that language? Can they produce all or only some L2 suprasegmentals like native speakers of that language? In the absence of research on child acquisition of L2 suprasegmentals, some tentative answers to these questions can be gleaned from research on native (first) language (L1) development (Halle, de Boysson-Bardies, & Vihman, 1991; Kehoe & Stoel-Gamon, 1997; Ramus, 2002; Snow & Balog, 2002). This research suggests that L1 suprasegmental learning occurs within the first several years of life, and that some suprasegmentals appear to be learned more rapidly than others (Kehoe & Stoel-Gammon, 1997; Snow & Balog, 2002). For example, infants as young as 3 to 4 months of age discriminate between syllable- and stressed-timed languages (Ramus, 2002) and produce language-appropriate contours and vowel durations in disyllabic vocalizations (Hallé et al., 1991).

In fact, it appears that children acquire such suprasegmentals as intonation and rhythm early in language development, long before grammar, lexicon, or even individual speech segments (Whalen, Levitt, & Wang, 1991). These early-acquired suprasegmentals are typically language specific and characterize the prosodic pattern of a particular language (Botinis, Granström, & Möbius, 2001). By contrast, other suprasegmentals, for example, speech rate and pausing, seem to take much longer to develop, their development continuing into early and late childhood (DeJoy & Gregory, 1985; Haselager, Slis, & Rietveld, 1991). These later-acquired suprasegmentals are usually not specific to a particular language, and characterize the fluidity with which the spoken language is used (Holmes, 1995). These findings suggest that there may be differences between L1 development of prosody-based and fluency-based suprasegmentals and these differences may also exist in L2 suprasegmental learning.

With an overall goal of examining whether child L2 phonological learning is similar at the segmental and suprasegmental levels, the present study investigated the acquisition of L2 suprasegmentals by child learners, both after relatively short (1 year) and relatively extensive (11 years) amounts of L2 experience. In particular, the present study sought to answer the following two questions. Does the amount of child learners' experience influence their production of L2 suprasegmentals?

If so, which suprasegmentals are likely to be influenced by their amount of L2 experience?

To address these questions, 20 native Korean child learners of English who differed in amount of L2 experience (1 or 11 years of residence in an English-speaking country) and 20 age-matched native English speakers were recorded saying six English declarative sentences. Three analyses were carried out to examine suprasegmentals in the recorded sentences. The first analysis investigated whether suprasegmentals contributed to a perceived accent in the child learners' speech at both levels of L2 experience, and whether nativelike production of suprasegmentals was attainable after a brief and/or more extended amount of L2 experience. The second analysis acoustically examined the child learners' production of five English suprasegmentals in the same sentences (stress timing, tonal peak alignment, speech rate, pause frequency, and pause duration) at both levels of L2 experience. Two of the examined suprasegmentals (stress timing, tonal peak alignment) characterize speech prosody, while the remaining three (speech rate, pause frequency, pause duration) characterize speech fluency. Examples of both prosody- and fluency-based suprasegmentals (all discussed in detail below) were included to determine if these types of suprasegmentals are acquired in an L2 at different rates and to different degrees of accuracy, as may be the case in L1 development (Haselager et al., 1991; Whalen et al., 1991). The third analysis investigated the relative contribution of each of the five suprasegmentals to an accent in the learners' speech. Taken together, these analyses sought to clarify how experience influences L2 suprasegmental learning by child learners.

METHOD

Participants

There were 40 participants in this study ($n = 10$ per group). Twenty were native Korean speakers who started learning English as children. They were assigned to one of two groups based on their length of US residence. The participants in the first group had arrived in the United States as children ($M = 10.7$ years, $SD = 2.6$) and had resided in the United States for about 1 year ($M = 0.9$ years, $SD = 0.6$). At the time of testing, these child learners of English were children ("Korean children"). The participants in the second group had arrived in the Unired States also as children ($M = 9.0$ years, $SD = 1.2$) and had resided in the United States for about 11 years ($M = 11.1$ years, $SD = 1.9$). At the time of testing, these child learners of English were adults ("Korean adults"). Although the Korean children were slightly older than the Korean adults at the time of their arrival in the United States (10.7 vs. 9.0 years, respectively), the two groups did not differ statistically in their age upon arrival in the United States ($p = .07$).

A difference of over 10 years of residence in an L2-speaking country was used to differentiate between more and less experienced child L2 learners in this study because such an amount of time may be necessary to acquire at least some L2 suprasegmentals (e.g., Trofimovich & Baker, 2006). Length of residence in an L2-speaking country was used as a measure of child learners' L2 experience because it appears to be an adequate indicator of amount of experience in situations when

Table 1. *Stimulus sentences used in a delayed sentence repetition task*

Question (Prompt)	Answer (Response)
Did the boy get wet?	He **did**n't have a ***hood*** on his **coat**.
Where is my bed?	Your ***bed*** is by the ***win***dow.
Why is the boy sad?	He ***feels bad*** about the **news**.
Did the crowd boo the team?	***No***, they **booed** the **coach.**
Can I use your bat?	**No**, the ***bat*** is **mine**.
Is he feeling ok?	**No**, his ***head*** **hurts** a lot.

Note: Stressed and unstressed syllables used in the calculations of syllable–duration ratios are in bold and light text, respectively. Syllables used in the calculations of peak alignment are in italic.

learners use their L2 daily (Flege & Liu, 2001). All 20 child learners in this study arrived in the United States as children, attended public schools in English, and spoke English daily. When asked to estimate the amount of daily use of English, the Korean children and adults reported using English to a comparable degree in interactions at school (88% for children, 96% for adults) and with friends (74% for children, 81% for adults). The children and the adults differed, however, in the amount of daily English use at home; the children used predominantly Korean (97%) whereas the adults claimed to use mainly English (83%). This pattern of language use, especially in the home setting, was consistent with the amount of time the participants had resided in the United States. Unlike the Korean children, who were relatively new to the L2-speaking environment, the Korean adults (after about 11 years of US residence) almost exclusively switched to their L2, using it both in and outside the home (see Jia & Aaronson, 2003).

The remaining 20 participants were native English speakers matched in age to the groups of Korean children and adults. Ten participants were child native English speakers ($M = 10.5$ years, $SD = 1.9$), all raised in monolingual homes ("English children"). The other 10 participants were adult native English speakers ($M = 25.6$ years, $SD = 4.5$) with minimal or no knowledge of a language other than English ("English adults").

Materials and procedure

A delayed sentence repetition task was used in this study to elicit speech samples. In this task, the participants produced six English declarative sentences in response to question prompts (Table 1). The six target and 12 distractor sentences were presented in three randomized blocks, but only the sentences spoken in the last block were used in the analyses to ensure that the participants were familiar with the task and could repeat each sentence to the best of their ability. The analyses were based on 240 sentences (40 participants $\times$ six sentences). All audio recordings were made using a Shure head-mounted microphone (SM10A) and a Sony DAT tape recorder (TCD-D8) or a Marantz CDR-300 portable CD recorder. The digitally recorded sentences selected for analysis were normalized by

comparing and (where necessary) adjusting perceived loudness in each recorded sentence.

The participants were tested individually in a quiet location using a personal computer and stimulus presentation software (Smith, 1997). The participants first heard a question (prompt) spoken by a male native English speaker; the prompt was followed by a response spoken by a female native English speaker. Then, the participants listened to the prompt again and repeated the response they had heard. For example, to elicit the sentence, *He feels bad about the news*, the following prompt-response sequence was used:

Prompt (male voice):	Why is the boy sad? (pause)
Response (female voice):	He feels bad about the news. (pause)
Prompt (male voice):	Why is the boy sad? (longer pause)

Used in both L1 (e.g., Ratner, 2000) and L2 (e.g., Guion, Flege, Liu, & Yeni-Komshian, 2000) acquisition research, the delayed sentence repetition task allows for eliciting relatively fluent speech samples that are identical, and therefore maximally comparable, across all participants. In addition, the task likely imposes minimal demands on the participants to make lexical, syntactic, and pragmatic decisions while producing the sentences, all of which could affect the production of suprasegmentals (e.g., Derwing, Rossiter, Munro, & Thomson, 2004). Although repetition tasks often involve the limitation that participants may mimic both the segmental and the suprasegmental content of the utterance, the likelihood of the participants' direct mimicry in this task was decreased by the delay between the given response and the participants' repetition of it within 3 to 5 s.

Data analysis

The suprasegmentals in the recorded sentences were analyzed in three separate analyses. The first analysis examined the global accuracy with which the Korean children and adults as well as the English children and adults produced L2 suprasegmentals. For this analysis, 10 native English listeners performed a perceptual rating task, judging the degree of accentedness with which the 240 recorded sentences were spoken. The sentences had been low-pass filtered to preserve their suprasegmental aspects while removing most of their segmental content. The second analysis examined the accuracy with which the Korean children and adults as well as the English children and adults produced five *specific* L2 suprasegmentals, of which two are prosody based (stress timing, peak alignment) and three are fluency based (speech rate, pause length, pause frequency). In this analysis, acoustic measurements of the five suprasegmentals were carried out using the same 240 sentences. The final analysis examined the extent to which the Korean children and adults' L2 suprasegmental accuracy contributed to the native English listeners' judgments of accentedness in their speech. For all analyses below, the α level for significance was set at .05; the reported effect sizes are partial eta squared (η_p^2), calculated by dividing the effect sum of squares by the effect sum of squares plus the error sum of squares.

ACCENTEDNESS RATINGS

The first objective of the present study was to determine the degree of the Korean children's and adults' difficulty at acquiring L2 suprasegmentals, as a function of their amount of L2 experience: 1 year of US residence for the Korean children and 11 years of US residence for the Korean adults. This objective was addressed in a perceptual judgment task in which a separate group of 10 adult native English listeners ($M = 22$ years, $SD = 2.7$) judged the 240 recorded sentences (40 participants × six sentences) for degree of accentedness on a 9-point Likert scale (1 = *strong accent*, 9 = *no accent*). The sentences were low-pass filtered to remove all energy components of the speech signal above 450 Hz, a value that ensured a near-equivalent degree of low-pass filtering in both adults' and children's speech. Low-pass filtering preserved the suprasegmental aspects of the recorded sentence, including intonation (F0 contours or contours in lower frequency harmonics) and stress timing information, while removing most of their segmental content, thus increasing the likelihood that the listeners based their judgments on the suprasegmental, not segmental aspects of the sentences (Munro, 1995).

The sentences were presented over headphones in six randomized blocks, with each block containing all participants' renditions of the same sentence. Prior to and during listening, the listeners were shown all the sentences the participants were attempting to say and were able to refer back to the list throughout the experiment so that the listeners could compare what they heard to their expectations of what the sentence should sound like. The listeners, none of whom were trained in linguistics, phonetics, or teaching English as an L2, participated in a short training session before judging the sentences. In this session, they listened to several distractor sentences (i.e., sentences recorded in the delayed sentence repetition task but not used in data analyses) and judged them for degree of accentedness. The listeners were able to adjust the playback level during the training session and after each test block although none of the listeners actually did so. The dependent variable in this analysis was accentedness rating, computed by averaging the 10 English listeners' ratings for each participant's six sentences. Interrater reliability analyses comparing the 10 English listeners' ratings for each sentence yielded very high indices ($\alpha = 0.92$–0.97), suggesting that the listeners were consistent. Although the listeners' ratings for the native English speakers showed limited variance, they were normally distributed and met this assumption for parametric comparisons.

Results

The obtained accentedness ratings were submitted to a two-way analysis of variance (ANOVA) with language (Korean, English) and age (children, adults) as between-subjects factors. This analysis revealed a significant effect for language, $F\ (1, 36) = 62.08$, $p < .0001$, $\eta_p^2 = .63$, and age, $F\ (1, 36) = 21.27$, $p < .0001$, $\eta_p^2 = .37$, as well as a significant Language × Age interaction, $F\ (1, 36) = 19.84$, $p < .0001$, $\eta_p^2 = .36$. Follow-up Bonferroni tests (with $\alpha = .0125$ to adjust for four pairwise comparisons in this and all subsequent analyses) exploring the simple main effect of age revealed that the English children's and adults' accentedness ratings did not differ ($p = .91$), whereas the Korean children's ratings were

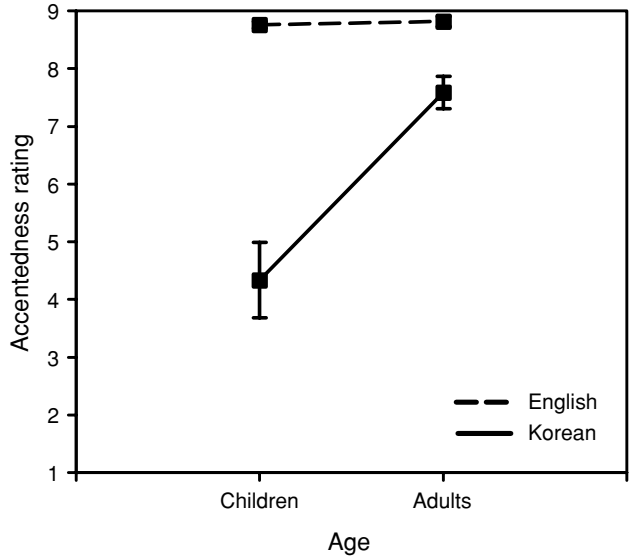

Figure 1. Means for accentedness ratings (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

significantly lower (less nativelike) than the Korean adults' ratings ($p < .0001$). Bonferroni tests exploring the simple main effect of language further revealed that the Korean children's ratings were significantly lower than the English children's ratings ($p < .0001$), although the English and Korean adults' ratings did not differ significantly ($p = .02$).

These findings indicated that the Korean adults (i.e., the child learners with 11 years of US residence), but not the Korean children (i.e., the child learners with 1 year US residence), could produce L2 suprasegmentals in a nativelike manner, as measured by the English listeners' global accentedness ratings (Figure 1). Overall, these findings suggested that at least some degree of accentedness in the child learners' speech, particularly those with little L2 experience, resides at the level of suprasegmentals. The next analyses sought to identify which suprasegmentals contribute to accentedness and to determine the extent to which the child learners produce these suprasegmentals accurately.

ACOUSTIC ANALYSES OF SUPRASEGMENTALS

Examined in the acoustic analyses were five specific suprasegmentals that may have contributed to accentedness in the Korean children's speech: stress timing, peak alignment, speech rate, pause frequency, and pause duration. These suprasegmentals characterize the general flow of speech and represent potential sources of difficulty for L2 learners of English, likely contributing to foreign accent (Mennen, 1998; Towell, 2002). These suprasegmentals were acoustically

analyzed in the 240 recorded sentences and compared across the four groups. The objective was to examine how accurately, in comparison to age-matched native English speakers, the Korean children and adults produced these five suprasegmentals. The purpose here was to determine whether and to what extent these suprasegmentals contributed to the difference in accentedness ratings obtained in the preceding analysis. The acoustic analyses are presented in turn, preceded by a brief review of relevant research on each suprasegmental and a description of the methodology used.

Stress timing

As a stress-timed language, English has a rhythm characterized by alternations in degree of stress, with stressed syllables significantly longer (by as much as 50%) than unstressed ones (Nakatani, O'Connor, & Aston, 1981) and most vowels in unstressed syllables reducing to a schwa (Bolinger, 1965). By contrast, as a syllable-timed language, Korean does not exhibit alternations in tone patterns in the way English exhibits alternations in stress; Korean syllables therefore all have approximately the same duration (Jun, 1996). Although the distinction between syllable- and stress-timed languages has been debated (Dauer, 1983), it is still generally considered that most languages of the world fall somewhere along the syllable- and stress-timed continuum (Ramus, Nespor, & Mehler, 1999).

Adult L2 learners' difficulty in acquiring L2 stress timing (variation in stress and in degree of vowel reduction from syllable to syllable) has been documented for adult speakers of syllable-timed Spanish (Gutiérrez-Díez, 2001) and Korean (Trofimovich & Baker, 2006) acquiring stress-timed English. Similar difficulty has been noted for adult speakers of stress-timed English acquiring syllable-timed French (Freland-Ricard, 1996) and Portugese (Ciancio, 2001). For example, relatively advanced Spanish learners of English have been shown to produce English with a ratio of unstressed to stressed syllables that was intermediate between those obtained for native speakers of Spanish and English in their L1 (Gutiérrez-Díez, 2001), which suggests that learning L2 stress timing might pose a problem for adult learners. Do child L2 learners also have difficulty in acquiring L2 stress timing, both after a relatively short and a relatively extensive L2 experience? To our knowledge, there is no research that directly addresses this question (for studies on child L2 acquisition of word stress, see, however, Guion et al., 2004; Guion, 2005).

Method. This analysis examined the duration of stressed and unstressed syllables in the 240 recorded sentences to determine how accurately English stress timing was produced. The duration of stressed and unstressed syllables was measured from the display of digital speech-analysis software (*Praat*) between two cursors placed at the onset and offset of voicing in each syllable, which included the length of all consonants and vowels (see Deterding, 2001, for a description of this procedure). The syllables were designated as either stressed or unstressed, first, by examining English stress placement as described in published sources (e.g., Couper-Kuhlen, 1986, p. 39), and second, by conducting an acoustic analysis of

the sentences produced by the English child and adult groups. This procedure resulted in a unanimous decision about the English-appropriate stress status of each syllable in the six sentences. Included in all subsequent analyses were only those syllables that were designated as stressed using the two-step procedure described above. (See Table 1 for stressed and unstressed syllables used in this analysis.)

The obtained stressed versus unstressed syllable durations were averaged for each participant across the six declarative sentences and a single ratio of unstressed to stressed syllables, used as the dependent variable in this analysis, was computed for each participant. If the Korean children and adults used English-like stress timing, the ratio of unstressed to stressed syllables should be lower (e.g., closer to .5 or .4, the actual average ratios obtained for the 10 English adults and children, respectively) because unstressed syllables are shorter than stressed ones in nativelike English syllable-to-syllable stress alternations (Nakatani et al., 1981). If the Korean children and adults used Korean-like syllable timing, the ratio should be higher (closer to 1) because unstressed and stressed syllables are roughly equal in duration.

Results. The obtained ratios were submitted to a two-way (Language × Age) ANOVA, which yielded a significant effect for language, $F\ (1, 36) = 49.15$, $p = .0001$, $\eta_p^2 = .58$, and a significant Language × Age interaction, $F\ (1, 36) = 33.27$, $p = .0001$, $\eta_p^2 = .48$, but no significant effect for age. Follow-up Bonferroni tests exploring the simple main effect of age revealed that the English children's syllable-duration ratios were significantly smaller than the English adults' ratios ($p < .001$), and that the Korean children's ratios were significantly larger (i.e., more Korean-like) than Korean adults' ratios ($p < .0001$). Bonferroni tests exploring the simple main effect of language further revealed that the Korean children's syllable-duration ratios were significantly larger than the English children's ratios ($p < .0001$) while the English and Korean adults' ratios did not differ significantly ($p = .39$). Taken together, these findings suggested that amount of L2 experience influenced the production of stress timing. That is, the Korean adults (i.e., child learners with more L2 experience) were better able than the Korean children (i.e., child learners with less L2 experience) to produce English stress timing like the age-matched English speakers (Figure 2). These findings also suggested that English-speaking children may differ from English-speaking adults in their production of stress timing (see Cunningham, Nicol, Zecker, & Kraus, 2000, for evidence of developmental differences in L1 speech processing and learning).

Peak alignment

Another prosody-based suprasegmental examined in this study is peak alignment, that is, the location of the highest value (peak) of pitch (or of its acoustic correlate, fundamental frequency) relative to the accented syllable in an intonation phrase. Along with the shape of pitch peak and the size of pitch movement, peak alignment defines a prosodic unit of a particular language and, by extension, characterizes its

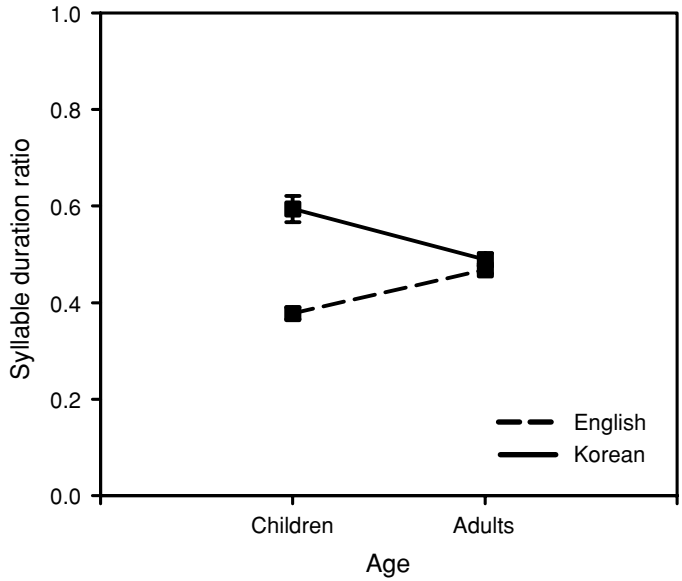

Figure 2. Means for syllable-duration ratios (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

rhythm (Botinis et al., 2001). Several studies have documented that languages differ considerably in their use of pitch rises and falls to indicate emphasis, emotion, or syntactic information (e.g., Chen, Gussenhoven, & Rietveld, 2004). In English, for example, pitch peak often corresponds to a high-value tonal accent associated with a prominent syllable, usually in the most prominent word in an intonation phrase (see Pierrehumbert, 1980, for review). Thus, in English, pitch peak is aligned with the onset (beginning) of the stressed syllable of the most prominent word in an intonation phrase (Ladd, Mennen, & Schepman, 2000), this word often being the last content word in an intonation phrase. In Korean, however, variations in pitch prominence are used not for the purpose of accent or prominent syllable marking, but for the purposes of marking the boundary of an accentual phrase, a tonally demarcated unit containing more than one lexical item (Pierrehumbert, 1980). Thus, in Korean, pitch peak is phrasal and is therefore not specific to a particular word (Jun, 1998). That is, pitch peak is usually aligned with the offset (end) of the syllable occurring on the last word in a phrase as a marker of its boundary (de Jong, 1994).

To date, only two studies have specifically investigated peak alignment in adult L2 learners' speech (Mennen, 1998; Trofimovich & Baker, 2006), and none has examined peak alignment in child L2 learners. Mennen (1998), for example, reported that even after an extensive amount of experience with Greek (12–35 years), adult Dutch speakers did not "reach native-like values for this aspect of intonation" (p. 337), suggesting that peak alignment might be difficult for adults to master (see also Trofimovich & Baker, 2006). Will this same difficulty hold for child learners,

after a relatively short and a relatively extensive L2 experience? Earlier studies of L2 intonation have revealed that children have an advantage over adults in their ability to imitate L2 intonation contours (Tahta, Wood, & Loewenthal, 1981) and to produce L2 tonal distinctions (Ioup & Tansomboon, 1987). The following analysis was conducted to determine if this advantage extends to child learners' ability to produce L2 peak alignment.

Method. This analysis examined peak alignment in the 240 recorded sentences to determine how accurately peak alignment was produced. First, pitch contours in the six English sentences were examined for each participant using digital speech-analysis software (*Praat*) to identify pitch peaks. Second, sentence-final pitch peaks were excluded from those identified because Korean learners of English are less likely to displace pitch peaks in sentence-final and prepausal contexts, that is, in situations where no segments follow the stressed syllable and where English pitch peak is possibly identical to Korean pitch peak. Third, peak alignment was calculated for the remaining pitch peaks (see Table 1 for the eight syllables used in these calculations). The location of peak alignment was defined as the distance (milliseconds) between the onset of the vowel in the stressed syllable and the point in the pitch contour with the highest value of fundamental frequency (Ladd et al., 2000). With rare exception (in all but 3 of the 280 measurements), pitch peak occurred within a close proximity to the stressed syllable. The obtained peak alignment values were averaged for each participant across the six declarative sentences, yielding the dependent variable in this analysis. If the child learners used English-like peak alignment, then pitch peak should be aligned with the onset of the vowel of the stressed syllable (i.e., the peak alignment value should be close to 0). If the child learners used Korean-like peak alignment, then the peak should be aligned with the offset of the stressed syllable (i.e., the peak alignment value should be larger than 0).

Results. The obtained peak alignment values were submitted to a two-way (Language × Age) ANOVA, which yielded a significant effect for language, $F\ (1, 36) = 31.38$, $p < .0001$, $\eta_p^2 = .47$, and age, $F\ (1, 36) = 9.23$, $p < .01$, $\eta_p^2 = .20$, as well as a significant Language × Age interaction, $F\ (1, 36) = 21.64$, $p < .0001$, $\eta_p^2 = .38$. Follow-up Bonferroni tests examining the simple main effect of age revealed that the English children's and adults' peak alignment values did not differ ($p = .26$), whereas the Korean children's peak alignment values were significantly longer (less nativelike) than the Korean adults' values ($p < .0001$). Bonferroni tests exploring the simple main effect of language further revealed that the Korean children's peak alignment values were significantly longer than the English children's values ($p < .0001$), although the English and Korean adults' values did not differ significantly ($p = .51$). These findings indicated that the amount of L2 experience influenced the production of peak alignment. That is, the Korean adults (i.e., child learners with more L2 experience), but not the Korean children (i.e., child learners with less L2 experience), produced English peak alignment in a nativelike manner, that is, with a pitch peak that was aligned with the onset of the stressed syllable (Figure 3).

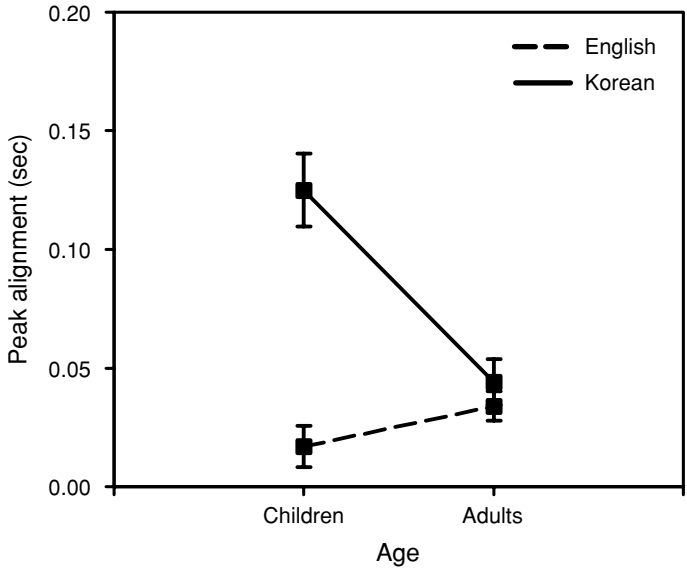

Figure 3. Means for peak alignment values (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

Speech rate

Speech rate is one of the most studied suprasegmentals in research on adult L2 acquisition (e.g., Munro & Derwing, 1998; Trofimovich & Baker, 2006). Previous investigations have yielded a common finding that adult L2 learners often produce L2 speech at a slower rate than do native speakers (Munro & Derwing, 1998), perhaps because of constraints related to the processing, encoding, and retrieval of phonological information, or because of difficulties in articulation of L2 speech, or both (Munro & Derwing, 2001). Speech rate that is either too slow (Munro & Derwing, 2001) or too fast (Janse, 2004) is often seen as an impediment to comprehensibility and perhaps intelligibility, as a sign of nonnativeness (Anderson-Hsieh & Koehler, 1988), and even as a source of stereotypes and evaluative judgments (Zuengler, 1988) and frustration with nonnative speech (Brennan & Brennan, 1981). Although with experience adults improve in their L2 fluency, as measured by speech rate (Towell, 2002; Towell, Hawkins, & Bazergui, 1996), their ability to produce L2 speech at a nativelike rate appears to be strongly predicted by age at which L2 learning begins (Guion et al., 2000; Trofimovich & Baker, 2006).

Much less is known, however, about the development of the temporal organization of speech, including speech rate, in child L2 learners. In fact, compared to a sizeable body of research on the development of speech rate in monolingual children (e.g., Smith, Kenney, & Hussain, 1996), literature on speech rate in child L2 learners (or bilingual children) is small. To our knowledge, the only study examining the development of speech rate in child L2 learners is Menyhárt's (2002) investigation of temporal characteristics of speech in Hungarian and Bulgarian

monolingual and bilingual 9- to 10- and 12- to 13-year-olds exposed to their L2 from early infancy. Menyhárt (2002) found that bilingual children produced L2 speech at a slower rate than aged-matched monolingual children, and that older children produced L2 speech at a faster rate than younger children. Although revealing, these findings likely characterize bilingual children's nearly simultaneous acquisition of two languages. Do children exposed to their L2 in childhood (as opposed to infancy) demonstrate the same pattern of L2 fluency development? The following analysis was therefore conducted to determine how speech rate develops in child L2 learners, both after brief and substantial L2 experience.

Method. This analysis examined speech rate in the 240 recorded sentences to determine if speech rate was produced in a nativelike manner. Following previous investigations of speech rate (e.g., Munro, 1995), speech rate in this study was measured by dividing the number of uttered syllables by their total duration (i.e., total length of utterance, including pauses). The obtained speech-rate ratio (syllable/s), averaged for each participant across the six sentences, was the dependent variable in this analysis. If the Korean children and adults produced the sentences at a nativelike rate, then their speech-rate ratio should be similar to that obtained for the aged-matched English speakers.

Results. The obtained speech-rate ratios were submitted to a two-way (Language × Age) ANOVA, which yielded a significant effect for language, $F\ (1, 36) = 51.93$, $p < .0001$, $\eta_p^2 = .59$, and age, $F\ (1, 36) = 33.03$, $p < .0001$, $\eta_p^2 = .48$, as well as a significant Language × Age interaction, $F\ (1, 36) = 8.87$, $p < .01$, $\eta_p^2 = .20$. Follow-up Bonferroni tests exploring the simple main effect of age revealed that the English children's and adults' speech-rate rates did not differ ($p = .058$), whereas the Korean children's speech rate was significantly slower than the Korean adults' rate ($p < .0001$). Bonferroni tests exploring the simple main effect of language further revealed that the Korean children's speech rate was significantly slower than the English children's rate ($p < .0001$), and that the Korean adults' speech rate was significantly slower than the English adults' rate ($p = .005$). These findings indicated that both the Korean children (i.e., child learners with less L2 experience) and the Korean adults (i.e., child learners with more L2 experience) produced English sentences at a slower speech rate than the aged-matched native English speakers did (Figure 4). In other words, the amount of L2 experience did not seem to influence the production of speech rate in the same way that it influenced the production of stress timing and peak alignment.

Pause frequency and duration

Another aspect of speech production usually associated with fluency-based suprasegmentals is related to hesitation phenomena, often characterized by the frequency and duration of pauses. Pause duration and pause frequency affect listeners' ratings of foreign accent in L2 speech and are viewed as determinants of both fluency and intelligibility (Albrechsten, Henriksen, & Faerch, 1980). Pauses, like speech rate, may indicate the speaker's difficulty with the task, perhaps reflecting processing or memory constraints unique to L2 speech (Schachter, Christenfeld,

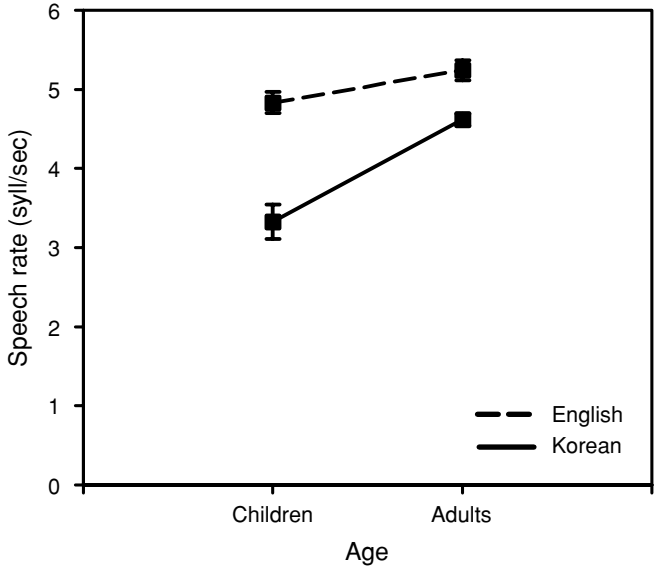

Figure 4. Means for speech rate (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

Ravina, & Bilous, 1991) as well as those associated with language development in general (Rispoli & Hadley, 2001). Learners also commonly pause more in their L2 than in their L1 (Riazantseva, 2001), regardless of crosslanguage and crosscultural differences in pause frequency and pause duration (Holmes, 1995).

The development of speech fluency, as measured by frequency and duration of pauses, is relatively well understood in adults acquiring an L2 (Trofimovich & Baker, 2006; Towell et al., 1996). For example, the frequency and duration of pausing in adult L2 learners' speech typically decreases as learners gain more experience with their L2 (Trofimovich & Baker, 2006). However, such changes may require an amount of L2 experience that is more extensive than 3 years of a college foreign language program (Towell, 2002) or more intensive than 6 months of study-abroad experience (Towell et al., 1996). In fact, it appears that adult L2 learners' ability to produce L2 speech in a nativelike manner (without frequent and long pauses) may depend more on learners' age at the time of first exposure to their L2 than on the amount of their experience with the L2, such that more fluent speech is associated with an early exposure to an L2 (Olynyk, D'Anglejan, & Sankoff, 1987).

To date, only one study known to us has examined the development of speech fluency, as measured by frequency and duration of pauses, in child L2 learners (Menyhárt, 2002). Menyhárt observed that the bilingual children who were exposed to two languages from infancy did not differ in pause duration in their two languages from age-matched monolingual children, and that older children (bilingual and monolingual alike) produced speech with shorter pauses than younger

children. Although indicative of a general pattern of fluency development in these (nearly simultaneous) bilinguals, these findings leave unanswered questions about the development of fluency in children who are exposed to their L2 in childhood.

Method. The following analysis therefore examined pause frequency and pause duration in the 240 recorded sentences to determine if the sentences were produced fluently (i.e., without frequent and long pauses). Pauses were defined as any break in the speech stream longer than 100 ms in length, a measure commonly used in pausing research (see Riazantseva, 2001, for review). Pause frequency was calculated by averaging the number of pauses for each participant across the six sentences; pause duration was computed by averaging pause durations for each participant across the six sentences. If the Korean children and adults produced the sentences in a nativelike manner, then their pausing (as measured by pause frequency and pause duration) should not differ from the pausing of age-matched native English speakers.

Results. The frequency and duration of pauses were submitted to two separate two-way (Language × Age) ANOVAs. The ANOVA comparing the frequency of pausing yielded a significant effect for language, $F\ (1, 36) = 12.74$, $p < .001$, $\eta_p^2 = .26$, and age, $F\ (1, 36) = 8.90$, $p < .005$, $\eta_p^2 = .20$, as well as a significant Language × Age interaction, $F\ (1, 36) = 4.82$, $p < .05$, $\eta_p^2 = .12$. Follow-up Bonferroni tests exploring the simple main effect of age revealed that the English children and adults did not differ in duration of pauses produced ($p = .58$), whereas the Korean children produced significantly more pauses than the Korean adults did ($p < .0001$). Bonferroni tests exploring the simple main effect of language revealed that the Korean children produced significantly more pauses than the English children did ($p < .0001$), whereas the Korean adults and the English adults did not differ in number of pauses produced ($p = .34$).

The ANOVA comparing the duration of pausing yielded a significant effect for language, $F\ (1, 36) = 10.79$, $p < .01$, $\eta_p^2 = .23$, and age, $F\ (1, 36) = 10.90$, $p < .01$, $\eta_p^2 = .23$, as well as a significant Language × Age interaction, $F\ (1, 36) = 9.38$, $p < .01$, $\eta_p^2 = .21$. Follow-up Bonferroni tests exploring the simple main effect of age revealed that the English children and adults did not differ in number of pauses produced ($p = .86$), whereas the Korean children produced significantly longer pauses than the Korean adults did ($p < .0001$). Bonferroni tests exploring the simple main effect of language further revealed that the Korean children produced significantly longer pauses than the English children did ($p < .0001$), whereas the Korean adults and the English adults did not differ in duration of pausing ($p = .87$). Taken together, these findings indicated that the Korean adults (i.e., child learners with more L2 experience), but not the Korean children (i.e., child learners with less L2 experience), produced English sentences similarly to the age-matched English speakers, that is, with relatively infrequent and short pauses (Figures 5 and 6). In other words, amount of L2 experience influenced the pause frequency and pause duration with which child learners produced L2 speech.

In summary, the results of the preceding five analyses revealed that the child learners' amount of L2 experience influenced the production of all five

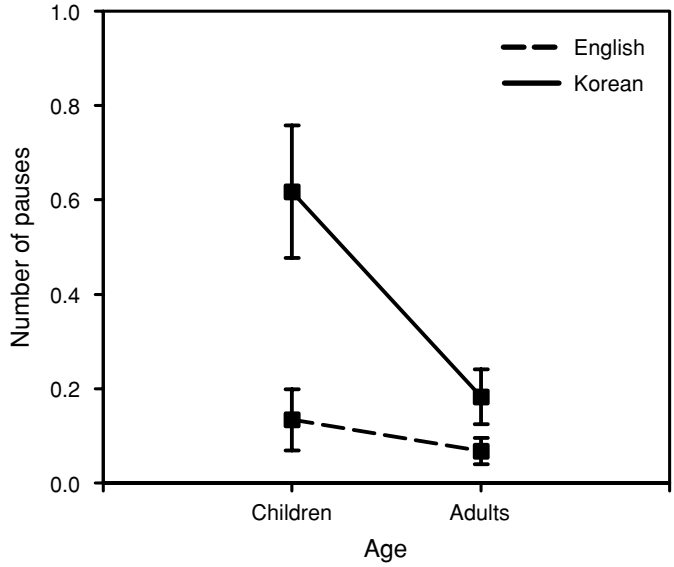

Figure 5. Means for frequency of pausing (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

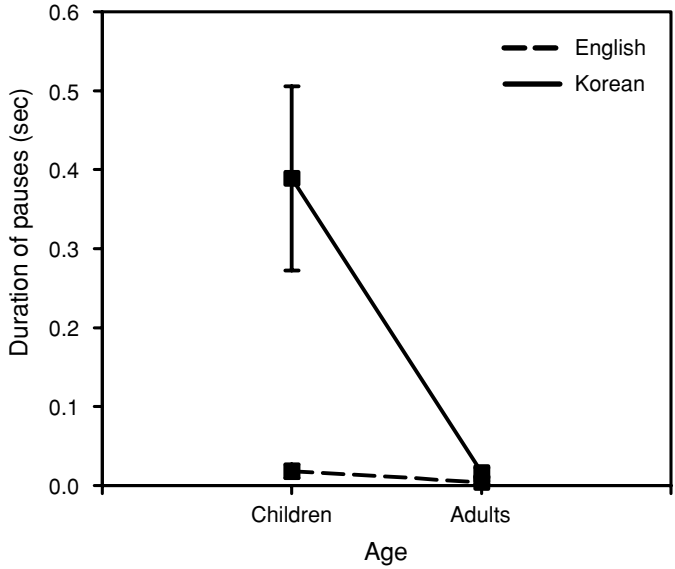

Figure 6. Means for duration of pausing (±1 *SE*) for groups of Korean and English children and adults. The invisible error bars are indicative of very small variances.

Table 2. *Summary of correlation analyses among acoustic measurements*

Measures	1	2	3	4	5
1. Stress timing					
2. Peak alignment	.43				
3. Speech rate	−.65*	−.69**			
4. Pause frequency	.48	.70**	−.84***		
5. Pause duration	.53	.64*	−.88***	.88***	

Note: The asterisks identify correlation coefficients that remained significant after a Bonferroni correction ($\alpha = .005$).
*$p < .005$. **$p < .001$. ***$p < .0001$ (two tailed).

examined suprasegmentals. That is, the Korean adults (i.e., child learners with about 11 years of US residence) produced all five suprasegmentals in a more nativelike manner than the Korean children (i.e., child learners with about 1 year of US residence) did. In all but one case, the Korean adults reached nativelike levels of L2 suprasegmental accuracy after about 11 years of US residence, producing English stress timing and peak alignment as well as speaking with infrequent and brief pauses akin to the age-matched English speakers. The only exception to this pattern of results pertained to speech rate. Despite extensive L2 experience, the Korean adults who were first exposed to their L2 as children (between ages 7 and 11) produced English sentences at a significantly slower rate than the age-matched English speakers did.

RELATIONSHIP BETWEEN SUPRASEGMENTAL ACCURACY AND ACCENTEDNESS

The previous analyses established that amount of L2 experience influences child learners' acquisition of L2 suprasegmentals, those that characterize the prosody (stress timing, peak alignment) and fluency (frequency, duration of pausing) of L2 speech. What these analyses did not establish, however, is the importance of these suprasegmentals to the native English listeners' ratings of accentedness in L2 speech. To answer this question, the Korean children's and adults' accentedness ratings and their accuracy scores for the five suprasegmentals examined here were submitted to correlation and regression analyses. First, zero-order correlations were computed among the five suprasegmental accuracy scores (syllable duration ratios, peak alignment values, speech rate ratios, frequency, and duration of pauses) for the groups of Korean children and adults ($n = 20$). This analysis indicated that most of the scores were significantly correlated with one another, suggesting that there was a complex relationship among the suprasegmental measures (Table 2). What possibly underlies this relationship is a temporal dimension, common to all five suprasegmental measures examined in this study. This dimension may characterize a speaker's ability to execute appropriate temporal control at different levels of speech production (Smith, 1978). This ability, which is discussed in greater detail below, may be specific to a particular suprasegmental (stress timing, peak alignment), or may reflect efficient functioning of several psycholinguistic processes (speech rate, frequency and duration of pausing).

Table 3. *Summary of regression analyses for acoustic measurements as predictors of accentedness ratings*

	B	*SE B*	β	R^2	*t*
Stress timing					
Constant	14.74	2.72			5.42***
Stress timing	−16.20	4.96	−.61	.37	−3.27*
Speech rate					
Constant	−3.69	1.20			−3.07*
Speech rate	2.43	0.30	.89	.79	8.20***
Pause frequency					
Constant	7.52	0.55			13.75***
Pause frequency	−3.90	0.98	−.68	.47	−3.98**
Pause duration					
Constant	7.11	0.38			18.52***
Pause duration	−5.65	1.04	−.79	.62	−5.45***

*$p < .01$. **$p < .001$. ***$p < .0001$.

Because there existed strong associations among the five suprasegmental accuracy scores (see Table 2), which resulted in multicollinearity, a multiple regression analysis with accentedness ratings as the criterion measure and the five suprasegmental accuracy scores as predictors was not possible. Instead, the accuracy scores for each of the five suprasegmentals were individually regressed on the accentedness ratings. The goal of these analyses was to estimate the degree to which each of the five suprasegmentals examined here predicted the Korean children's and adults' accentedness ratings. These five separate linear regression analyses (Bonferroni adjusted $\alpha = .01$) allowed for determining the amount of variance that each of the suprasegmental accuracy scores shared with the acccentedness ratings. Results of these analyses indicated that four of the five suprasegmental accuracy scores (stress timing, speech rate, pause frequency, pause duration) significantly predicted accentedness ratings: these variables explained between 37% (stress timing) and 79% (speech rate) of the variance in the accentedness ratings (Table 3). It remains to be determined why peak alignment scores did not appear to have any shared variance with the Korean children's and adults' accentedness ratings. It may be that fine-grained temporal differences in pitch peak alignment are so subtle that they are virtually imperceptible to the listener, particularly in low-pass filtered sentences, and therefore, have little impact on accentedness judgments. Whatever the reasons for this finding (some of which we revisit below), the pattern of the obtained results suggested that a combination of prosody-based (stress timing) and fluency-based (speech rate, pause frequency, pause duration) characteristics of L2 speech seemed to influence native English listeners' perceptions of accentedness in the Korean children's and adults' speech.

GENERAL DISCUSSION

This study was conceptualized as an investigation of child learners' acquisition of L2 suprasegmentals. Its goal was to determine how relatively brief and more

substantial L2 experience (defined here as 1 and 11 years of residence in an L2-speaking country) influences child learners' ability to produce five suprasegmentals (those that characterize speech prosody and fluency), and to establish the degree to which these suprasegmentals contribute to perceived accent in child learners' speech. Results indicated that, after about 1 year of L2 experience, the Korean children did not produce any of the five examined suprasegmentals with nativelike accuracy. This brief experience was clearly insufficient for these learners to acquire fully the examined L2 suprasegmentals. However, after about 11 years of L2 experience, the Korean adults (who had begun their L2 learning in childhood) produced four of the five examined suprasegmentals with nativelike accuracy. These suprasegmentals included those that define speech prosody (stress timing, peak alignment) and that characterize speech fluency (frequency and duration of pausing). The only exception to this learning pattern pertained to speech rate, a temporal, fluency-based measure of L2 speech. Even after 11 years of L2 experience, these learners produced English sentences at a slower rate than aged-matched native English speakers did. Results also indicated that both fluency-based and prosody-based suprasegmentals appeared to determine the degree to which Korean children's and adults' speech was perceived as being accented. When asked to judge the degree of accent specific to L2 suprasegmentals (to the extent that the use of low-pass filtered sentences made it possible), native English listeners appeared to consider a combination of suprasegmentals: those that characterize speech prosody (stress timing) and speech fluency (speech rate, frequency and duration of pausing).

Taken together, the results of this study prompt a general conclusion that child L2 phonological learning appears to be similar at the segmental and suprasegmental levels. Learning L2 segments and suprasegmentals alike represents a gradual process, often requiring extended amounts of L2 experience (e.g., Baker & Trofimovich, 2005; Munro et al., 1996), a process whose outcomes vary according to the suprasegmental studied (Guion, 2005; Guion et al., 2000). For example, the experienced child learners in this study were generally successful (when compared to age-matched native English speakers) at producing most of the L2 suprasegmentals examined (stress timing, peak alignment, pause frequency and duration). This finding parallels results of several recent investigations of child L2 learning at the segmental level (i.e., Aoyama et al., 2004; Baker & Trofimovich, 2005; Flege et al., 1999). Yet these child learners' success did not extend to at least one suprasegmental (speech rate). Apparently, even a substantial amount of L2 experience (11 years of residence in an L2-speaking country in this study) was not enough for these learners, exposed to their L2 between 7 and 14 years of age, to produce speech rate in a nativelike manner (Guion et al., 2000). This finding, discussed in more detail below, agrees with the results of recent investigations of child L2 learning at the segmental level, particularly those documenting some "limits" to child L2 learners' ability to perceive and produce L2 segments nearly natively, even despite massive amounts of L2 exposure (Flege, 1998; Sebastián-Gallés & Soto-Faraco, 1999).

Although this study does not yield compelling evidence of limits to child learners' acquisition of L2 suprasegmentals (in part, because obtaining such evidence requires testing child learners after amounts of L2 experience substantially greater

than 11 years of residence in an L2-speaking country), its findings are suggestive. In particular, they indicate that learners' age, typically defined as age at the time of L2 onset, is a factor determining learners' success in L2 phonological learning. In fact, an early exposure to an L2 is often associated with nativelike L2 segmental (Baker & Trofimovich, 2005; Flege at al., 1996) and suprasegmental (Tahta et al., 1981) accuracy. If this is indeed the case, then it is possible that child learners exposed to an L2 prior to age 7 (i.e., prior to the "earliest" age of exposure in this study) may attain nativelike accuracy in their production of all suprasegmentals examined, including speech rate. In support of this claim, at least one study examining the learning of intonational patterns in simultaneous bilinguals (children exposed to two languages from birth) has revealed that these bilinguals develop two ostensibly nativelike suprasegmental systems that nonetheless heavily influence each other (Queen, 2001). By contrast, the few other studies examining the acquisition of L2 word stress patterns by young children (aged 4 at the time of L2 onset and tested an average of 18 years later) have shown that even an early exposure (at or prior to age 4) does not necessarily result in nativelike perception and production of L2 word stress (Guion 2005; Guion et al., 2004; see also Sebastián-Gallés & Soto-Faraco, 1999). Clearly, what needs to be addressed in future studies of L2 suprasegmental learning by young children are individual and combined effects of the factors known to contribute to age-related differences in L2 phonological learning, factors characterizing neurobiological maturation (Kim, Relkin, Lee, & Hirsch, 1997), nature and type of L2 input (Jia & Aaronson, 2003), amount and extent of L2 practice (Flege & Liu, 2001), sociological/motivational variables (Gatbonton, Trofimovich, & Magid, 2005), and cognitive processing (Cochran, McDonald, & Paralaut, 1999).

Perhaps more interestingly, the findings of the present study suggest a possible difference in the development of prosody- versus fluency-based L2 suprasegmentals. The child learners in this study attained nativelike accuracy in their production of the suprasegmentals that characterize speech prosody (stress timing, peak alignment). However, they were seemingly unable to attain nativelike accuracy in their production of at least one suprasegmental that characterizes speech fluency (speech rate). Because pausing and speech rate are closely related as temporal determinants of fluency (Trofimovich & Baker, 2006), this obtained L1–L2 difference may not be restricted to speech rate alone, and may extend to the other two fluency-based suprasegmentals (pause frequency, pause duration). Although the experienced child learners in this study did not appear to differ from the age-matched native speakers in the number and duration of pauses produced, the utterances they spoke may have been effectively too short and easy to articulate to reveal the differences that would otherwise be detectable if longer and therefore more cognitively demanding speech materials had been used. Regardless of the extent to which the child learners were able to produce all of the fluency-based suprasegmentals examined here, this obtained dissociation is reminiscent of different developmental trajectories in L1 acquisition of fluency- and prosody-based suprasegmentals (DeJoy & Gregory, 1985; Whalen et al., 1991), and is akin to differences found in adult L2 suprasegmental learning (Trofimovich & Baker, 2006).

One theoretical implication of this finding is that different processing mechanisms underlie the processing and learning of fluency- and prosody-based

suprasegmentals. The processing and learning of the suprasegmentals characterizing speech prosody, including stress timing and peak alignment, likely reflect linguistic knowledge that differs from language to language and must be processed and stored in a language-specific manner (Botinis et al., 2001). By contrast, the suprasegmentals characterizing speech fluency (e.g., speech rate) reflect rapid and efficient functioning of several psycholinguistic mechanisms at multiple levels of processing, including those of lexical access and conversion of a speech plan into articulatory output (Costa & Santesteban, 2004; Zsiga, 2003).

One question that needs to be addressed in future studies of L2 suprasegmental learning by children pertains to clarifying the sources of developmental differences in learning fluency- and prosody-based L2 suprasegmentals. Such differences may reflect an age-related reduction in learners' processing efficiency, causing a slower speech rate, due to the competition between learners' two languages, requiring older learners (whose L1 is already well established), as opposed to younger learners (whose L1 is still developing), to allocate greater processing resources to suppress their L1 (Guion et al., 2000; Meuter & Allport, 1999). Such differences may also reflect older learners' difficulty in proceduralizing (automatizing) their L2 (Towell, 2002), possibly linked to their overreliance on declarative memory (i.e., memory for factual information) but not procedural memory (i.e., memory for cognitive operations or procedures) in their processing of L2 input (Ullman, 2001).

Although appealing, the finding of this study regarding a possible difference in the development of prosody-based versus fluency-based L2 suprasegmentals must remain speculative until investigated further. For example, this difference may be restricted to the development of speech rate alone (and thus may not apply to the development of other fluency-based suprasegmentals, such as pausing), or may arise as an artifact of the particular suprasegmentals examined, task employed, and/or speech materials used. In addition, it is important to investigate the development of prosody- and fluency-based L2 suprasegmentals in learners of different L1 backgrounds. It may well be that the results of this study are specific to Korean learners' acquisition of English as an L2. Certainly, more research needs to be done with child speakers of other languages, for example, those languages that are tonal, or pitch accented and those whose intonational patterns are either more or less similar to English than is Korean.

In addition, the development of L2 suprasegmentals needs to be examined in other, more naturalistic situations and tasks: those that allow researchers to estimate the effects of lexical access, syntactic encoding, and pragmatic decisions (among many other factors) on the production of L2 suprasegmentals. Last, but not least, future research needs to clarify the precise contribution of prosody- and fluency-based suprasegmentals to determining ratings of foreign accent in L2 speech. The results of this study suggested that both fluency-based (speech rate, frequency, and duration of pausing) and prosody-based (stress timing) suprasegmentals determine the perception of foreign accent in child learners' speech. Based on low-pass filtered speech, that is, speech that likely sounds unnatural to a casual listener, these ratings may not reflect perceptions of foreign accent in face-to-face interaction or in situations when clear speech is rated.

Despite these possible constraints on the interpretation of the results obtained here, these results point toward some similarities between child phonological learning at the level of sentences and at the level of individual segments or sounds. For example, by comparing the findings of this study, which did not focus on segmental learning, with the results of those investigations that did (Baker & Trofimovich, 2005; Flege et al., 1999), it is possible to conclude that, in both segmental and suprasegmental learning, reaching nativelike accuracy depends on extended amounts of L2 experience. The needed experience is likely more extensive than a few years of classroom instruction (Lepetit, 1987) and more intensive than 6 months of a study-abroad experience (Towell et al., 1996). Moreover, in both cases, the outcomes of such learning entail some aspects of L2 phonology, both segmental and suprasegmental (Baker & Trofimovich, 2005; Flege, 1998; Guion, 2005; Munro et al., 1996), that are not entirely nativelike. At a general level, these findings are not only relevant to theoretical conceptualizations of L2 phonological learning (e.g., Flege, 1995) but also informative to language testing specialists, who design assessment materials for children, and to clinicians, who work with diverse (and increasingly bilingual) populations of child learners.

ACKNOWLEDGMENTS

This research was partially supported by research grants from the University of Illinois at Urbana–Champaign and Brigham Young University. Many thanks are extended to Youngju Hong for her help in testing the Korean participants and to Molly Mack and James E. Flege for their advice throughout this research project. An earlier version of this paper was presented at the 5th International Symposium on Bilingualism in Barcelona, Spain. The authors gratefully acknowledge Randall Halter and three anonymous reviewers for their helpful suggestions on earlier drafts of this paper. Thanks also to Randall Halter for his invaluable statistical assistance.

REFERENCES

Albrechtsen, D., Henriksen, B., & Faerch, C. (1980). Native speaker reactions to learners' spoken interlanguage. *Language Learning, 30*, 365–396.

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38*, 561–613.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*, 233–250.

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

Bolinger, D. L. (1965). Pitch accent and sentence rhythm. In I. Abe & T. Kanekiyo (Eds.), *Forms of English: Accent, morpheme, order* (pp. 139–180). Cambridge, MA: Harvard University Press.

Botinis, A., Granström, B., & Möbius, B. (2001). Developments and paradigms in intonation research. *Speech Communication, 33*, 263–296.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299–2310.

Brennan, E., & Brennan, J. (1981). Measurement of accent and attitude toward Mexican–American speech. *Journal of Psycholinguistic Research, 10*, 487–501.

Chen, A. J., Gussenhoven, C., & Rietveld, T. (2004). Language specificity in the perception of paralinguistic intonational meaning. *Language and Speech, 47*, 311–349.

Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice.* Amsterdam, The Netherlands: John Benjamins.

Ciancio, J. (2001). Phonological interference and transference of a native English speaker when speaking Portuguese. *CTJ Journal, 40*, 25–38.

Cochran, B. P., McDonald, J. L., & Parault, S. J. (1999). Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language, 41*, 30–58.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficiency bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491–511.

Couper-Kuhlen, E. (1986). *An introduction to English prosody*. London: Edward Arnold.

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: Development and behavioral correlates of perception. *Ear and Hearing, 21*, 554–568.

Dauer, R. M. (1983). Stress timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51–62.

de Jong, K. (1994). Initial tones and prominence in Seoul Korean. *OSU Working Papers in Linguistics, 43*, 1–14.

DeJoy, D. A., & Gregory, H. H. (1985). The relationship between age and frequency of disfluency in preschool children. *Journal of Fluency Disorders, 10*, 107–122.

Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*, 393–410.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*, 655–679.

Deterding, D. (2001). The measurement of rhythm: A comparison of Singapore and British English. *Journal of Phonetics, 29*, 217–230.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.

Flege, J. E. (1998). The role of subject and phonetic variables in L2 speech acquisition. In M. Gruber, D. Higgins, K. Olsen, & T. Wysocki (Eds.), *Papers from the 34th Annual Meeting of the Chicago Linguistic Society: Vol. 2. The Panels* (pp. 213–232). Chicago: Chicago Linguistic Society.

Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23*, 527–552.

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

Freland-Ricard, M. (1996). Second-language learners' temporal and rhythmic organizations: A multilingual study. *Revue de Phonétique Appliquée, 118–119*, 61–91.

Gatbonton, E., Trofimovich, P., & Magid, M. (2005). Learners' ethnic group loyalty and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly, 39*, 489–511.

Guion, S. G. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition, 27*, 503–533.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21*, 205–228.

Guion, S. G., Harada, T., & Clark, J. J. (2004). Early and late Spanish–English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7*, 207–226.

Gutiérrez-Díez, F. (2001). The acquisition of English syllable timing by native Spanish learners of English. An empirical study. *International Journal of English Studies, 1*, 93–113.

Hallé, P. A., de Boysson-Bardies, B., & Vihman, M. M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language and Speech, 34*, 299–318.

Haselager, G. J. T., Slis, I. H., & Rietveld, A. C. M. (1991). An alternative method of studying the development of speech rate. *Clinical Linguistics and Phonetics, 5*, 53–63.

Holmes, M. (1995). A crosslinguistic comparison of the production of utterances in discourse. *Cognition, 54*, 169–207.

Ioup, G., & Tansomboon, A. (1987). The acquisition of tone: A maturational perspective. In G. Ioup & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 333–349). Cambridge, MA: Newbury House.

Janse, E. (2004). Word perception in fast speech. *Speech Communication, 42*, 155–173.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics, 24*, 131–161.

Jun, S.-A. (1996). *The phonetics and phonology of Korean prosody: Intonational phonology and prosodic structure*. New York: Garland Publications.

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology, 15*, 189–226.

Kehoe, M., & Stoel-Gammon, C. (1997). The acquisition of prosodic structure: An investigation of current accounts of children's prosodic development. *Language, 73*, 113–144.

Kim, K., Relkin, N., Lee, K.-M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*, 171–174.

Ladd, D. R., Mennen, I., & Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America, 107*, 2685–2696.

Lepetit, D. (1987). Cross-linguistic influence in intonation: French/Japanese and French/English. *Language Learning, 39*, 397–413.

Mennen, I. (1998). Second language acquisition of intonation: The case of peak alignment. *Chicago Linguistic Symposium, 34*, 327–341.

Menyhárt, K. (2002). The temporal organization of speech in monolingual and bilingual children. *Acta Linguistica Hungarica, 49*, 347–362.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25–40.

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. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning, 48*, 159–182.

Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition, 23*, 451–468.

Munro, M. J., Flege, J. E., & MacKay, I. R. A. (1996). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics, 17*, 313–334.

Nakatani, L. H., O'Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American English speech rhythm. *Phonetica, 38*, 84–106.

Olynyk, M., D'Anglejan, A., & Sankoff, D. (1987). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. *Applied Psycholinguistics, 8*, 121–136.

Pallier, C., Bosch, L., & Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64*, B9–B17.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Unpublished doctoral dissertation, MIT, Cambridge, MA.

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

Queen, R. M. (2001). Bilingual intonation patterns: Evidence of language change from Turkish–German bilingual children. *Language in Society, 30*, 55–80.

Ramus, F. (2002). Language discrimination in newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition, 2*, 85–115.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265–292.

Ratner, N. B. (2000). Elicited imitation and other methods for the analysis of trade-offs between speech and language skills in children. In L. Menn & N. B. Ratner (Eds.), *Methods for studying language production* (pp. 291–312). Mahwah, NJ: Erlbaum.

Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*, 497–526.

Rispoli, M., & Hadley, P. (2001). The leading-edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research, 44*, 1131–1143.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60*, 362–367.

Sebastián-Gallés, N., & Soto-Faraco. S. (1999). On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition, 72*, 111–123.

Smith, B. L. (1978). Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics, 6*, 37–67.

Smith, B. L., Kenney, M. K., & Hussain, S. (1996). A longitudinal investigation of duration and temporal variability in children's speech production. *Journal of the Acoustical Society of America, 99*, 2344–2349.

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

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

Snow, D., & Balog, H. L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua, 112*, 1025–1058.

Tahta, S., Wood, M., & Loewenthal, K. (1981). Age changes in the ability to replicate foreign pronunciation and intonation. *Language and Speech, 24*, 363–373.

Towell, R. (2002). Relative degrees of fluency: A comparative case study of advanced learners of French. *International Review of Applied Linguistics, 40*, 117–150.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*, 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.

Ullman, M. T. (2001). The neural basis of lexical and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105–122.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106*, 3649–3658.

Whalen, D. H., Levitt, A. G., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language, 18*, 501–516.

Zsiga, E. C. (2003). Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition, 25*, 399–432.

Zuengler, J. (1988). Identity markers and L2 pronunciation. *Studies in Second Language Acquisition, 10*, 33–49.