Applied Psycholinguistics **37** (2016), 303–322
doi:10.1017/S0142716414000575

# The relative contribution of segments and intonation to the perception of foreign-accented speech

JOAN SERENO, LYNNE LAMMERS, and ALLARD JONGMAN
*University of Kansas*

Received: January 7, 2014 Accepted for publication: October 20, 2014

ADDRESS FOR CORRESPONDENCE
Joan Sereno, 425 Blake Hall, 1541 Lilac Lane, Linguistics Department, University of Kansas, Lawrence, KS 66045. E-mail: Sereno@ku.edu

ABSTRACT
The present study examines the relative impact of segments and intonation on accentedness, comprehensibility, and intelligibility, specifically investigating the separate contribution of segmental and intonational information to perceived foreign accent in Korean-accented English. Two English speakers and two Korean speakers recorded 40 English sentences. The sentences were manipulated by combining segments from one speaker with intonation (fundamental frequency contour and duration) from another speaker. Four versions of each sentence were created: one English control (English segments and English intonation), one Korean control (Korean segments and Korean intonation), and two Korean–English combinations (one with English segments and Korean intonation; the other with Korean segments and English intonation). Forty native English speakers transcribed the sentences for intelligibility and rated their comprehensibility and accentedness. The data show that segments had a significant effect on accentedness, comprehensibility, and intelligibility, but intonation only had an effect on intelligibility. Contrary to previous studies, the present study, separating segments from intonation, suggests that segmental information contributes substantially more to the perception of foreign accentedness than intonation. Native speakers seem to rely mainly on segments when determining foreign accentedness.

When adults learn to speak a second language, their nonnative speech is often identified as accented (Flege, Munro, & MacKay, 1995). Second language acquisition typically requires learning new motor patterns, creating new phoneme categories, and reorganizing existing phoneme boundaries. Adult learners' difficulty in mastering any of these aspects leads to pronunciations that deviate from the native norm, resulting in speech that sounds accented and may also be difficult to understand (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Edwards & Zampini, 2008; Munro & Derwing, 1995). These difficulties in second language pronunciation typically include both segmental and suprasegmental features (e.g., Edwards & Zampini, 2008; Major, 2001; van Els & de Bot, 1987).

The current study investigates the separate contribution of segmental and suprasegmental information, specifically intonation, to perceived foreign accent in Korean-accented English. Intonation refers to "the distinctive use of pitch over units larger than a single word" (Reetz & Jongman, 2009, p. 221). The goal is to determine the extent to which segmental and intonational information affect the perception of foreign-accented speech.

Evaluation of second language speech can involve measures of intelligibility, comprehensibility, or accentedness (Munro & Derwing, 1999). Accentedness and comprehensibility are both measured through listeners' judgments of speech samples, but intelligibility is an objective measure of how much of a sample was correctly understood. Intelligibility is broadly defined as the critical measure of the extent to which speech is understood by a listener (e.g., Munro & Derwing, 1995). While commonly used to assess the transmission of information in, for example, a noisy environment, intelligibility is gaining ground as one of the most useful measures to assess nonnative speech because making oneself understood in a foreign language is arguably the most important component of successful communication. Intelligibility measures range from orthographic transcription of sentences or key words to intelligibility ratings on a Likert scale (e.g., Fayer & Krasinski, 1987; Gass & Varonis, 1984). As pointed out by Munro and Derwing (1999), pronunciation experts have long emphasized improved intelligibility as the most important objective of pronunciation teaching. Comprehensibility is a measure of how difficult or easy an utterance is to understand, using a rating scale, without actually checking listeners' accuracy. It is therefore a measure of native speakers' *perception* of intelligibility (Derwing & Munro, 1997). Accentedness is a measure of how strong the foreign speaker's accent is perceived to be (Munro & Derwing, 1995). Accentedness is also typically measured along a rating scale and allows native listeners to indicate the extent to which they feel a nonnative speaker's utterance deviates from native-speaker norms. While intelligibility, comprehensibility, and accentedness are related, they are partially independent (e.g., Munro & Derwing, 1998). For example, studies have shown that moderately or highly accented utterances could nevertheless be highly intelligible and comprehensible (e.g., Derwing & Munro, 1997; Munro & Derwing, 1998). Such data emphasize the need for studies that assess all three components of nonnative speech production.

Pronunciation errors can include both segmental and suprasegmental deviations. Segmental errors are errors in the production of individual consonants and vowels (Anderson-Hsieh et al., 1992; Flege & Hillenbrand, 1984). Segmental errors include the substitution of one sound for another or the modification of a sound (e.g., Broselow, Chen, & Wang, 1998; Flege, Bohn, & Jang, 1997; Major & Faudree, 1996). Suprasegmental errors include errors in stress assignment, intonation, timing, phrasing, and rhythm (Anderson-Hsieh et al., 1992; Koster & Koet, 1993; Munro, 1995). Suprasegmental errors can further be divided into those that influence speech fluency, including speech rate, pause frequency, and pause duration, and those that influence speech melody, including tonal peak alignment and stress timing (Trofimovich & Baker, 2006). The impact of these errors on accentedness and comprehensibility has been examined in several studies (Anderson-Hsieh et al. 1992; Anderson-Hsieh & Koehler, 1988;

Derwing, Munro, & Wiebe, 1998; Munro & Derwing, 1999; Trofimovich & Baker, 2006).

In an early study, Anderson-Hsieh et al. (1992) studied the relationship between listeners' judgments of foreign-accented speech and the types of errors present in the speech. Native English listeners rated speech samples of nonnative speakers from a great variety of languages (Arabic, Armenian, Assamese, Chinese, Farsi, German, Greek, Hindi, Indonesian, Kannada, Korean, Malayalam, Punjabi, Serbo-Croatian, Spanish, and Tamil) and a variety of proficiency levels in terms of their English pronunciation. After auditory analysis of the speech samples for phoneme, syllable, and prosody errors, the correlation between these errors and the native listener ratings was determined. The researchers found that, although both segmental errors and prosody errors correlated with pronunciation rating, the prosody rating was more highly correlated with pronunciation ratings (Anderson-Hsieh et al., 1992). However, the scale used to judge overall pronunciation in the reported study did not distinguish between accentedness and comprehensibility, and instead combined the two. Furthermore, because the speech samples contained both prosodic and phonemic errors, it is not clear to what extent the pronunciation ratings were really able to separately assess the two types of information.

Munro and Derwing (1999) conducted a study with similar methods, but they specifically looked at the relationships among accentedness ratings, comprehensibility ratings, and intelligibility. Although the role of different types of errors in nonnative speech in these ratings was not the main focus of the study, it was statistically examined. In this study, native English listeners transcribed and rated accentedness and comprehensibility of sentences spoken in English by native Mandarin speakers. Later, linguists determined the number of phonemic errors and rated the intonation of each speech sample. In the analysis, the correlation between each of these error scores and accentedness, comprehensibility, and intelligibility was determined. The most relevant finding was that nonnative intonation correlated more with accentedness and comprehensibility than did the nonnative phonemic errors (Munro & Derwing, 1999). However, once again, it is difficult to separate segmental and prosodic contributions completely because both were present in the speech sample and may have affected each other in the ratings.

A study by Derwing et al. (1998) on the efficacy of different methods of foreign-language instruction revealed that segments and prosody may have varying impacts on accentedness and comprehensibility, depending on the type of speech sample under investigation. In this study, adult nonnative speakers in English as a second language classes were placed in one of three training groups: no specific training, training on segmental features, or training on global features such as speaking rate, intonation, rhythm, and stress (all suprasegmental features). Both before and after 12 weeks of training, native speakers listened to two types of speech samples from the speakers: simple sentences that were read aloud and an extemporaneous description of a picture story. For the single-sentence samples, comprehensibility improved for both groups who had training (either segmental or global training). However, accentedness improved much more for those who had segmental training than for those who had global training. A different pattern emerged for the extemporaneous speech samples, however. Accentedness ratings did not improve for any of the groups, and comprehensibility improved only for

those who had global training (Derwing et al., 1998). This study did separate segmental from suprasegmental features in training, finding that the nature of the training regime (segmental or suprasegmental) differentially affected ratings.

The issue of separately examining segmental and suprasegmental information was addressed by Trofimovich and Baker (2006) in a study that focused on effects of second language (L2) experience on suprasegmentals. In this study, the researchers recorded simple English sentences produced by native Korean speakers and removed segmental information through low-pass filtering. These sentences were then presented to English listeners, who rated them for accentedness. In addition, the researchers analyzed fluency factors (pause duration, pause frequency, and speech rate) and speech melody factors (tonal peak alignment and stress timing). This study found that even the nonnative speech of residents of over 10 years still sounded accented at the suprasegmental level, showing that suprasegmentals do carry some aspect of accent. Further, the factors that were most predictive of accentedness ratings were pause duration and speech rate (factors involved in speech fluency; Trofimovich & Baker, 2006). Although this study did separately examine suprasegmentals by eliminating segmental information from the presented utterances, only accentedness ratings were collected.

Most of the studies investigating accented speech productions have examined only accentedness or combined accentedness and comprehensibility. However, accentedness may not always overlap with comprehensibility. Several studies have shown that, although accentedness, comprehensibility, and intelligibility correlate, it is not always the case that a highly accented speaker has low intelligibility or comprehensibility (Derwing & Munro, 1997; Kashiwagi & Snyder, 2010; Munro & Derwing, 1995, 1999). In two studies by Munro and Derwing (1995, 1999), accentedness was often rated much harsher than comprehensibility, and for fully (100%) intelligible samples, accentedness ratings varied greatly (see also Kashiwagi & Snyder, 2010). Furthermore, in Derwing and Munro (1995), accentedness was not related to response times, showing that accentedness does not necessarily make the sentence less comprehensible. It is clear that the relation of segmentals and suprasegmentals to accentedness, comprehensibility, and intelligibility is something that needs to be further examined.

A few studies have employed signal manipulation to assess the role of prosody in the perception of a foreign accent. In an early study on Dutch-accented English, Willems (1982) collected acceptability ratings for three sets of synthesized sentences: native British English (BE), native BE with one fundamental frequency (F0) deviation, and Dutch-accented BE. Willems found that the direction and magnitude of F0 movement as well as the production of a rise at the end of *wh*-questions most strongly affected acceptability judgments. Jilka (2000) specifically focused on the role of intonation. He replicated and expanded Willems's approach with a focus on German and American English. His comparison of foreign accent ratings for original foreign-accented utterances and versions that were systematically manipulated on the basis of the category-oriented Tones and Break Indices framework of intonation description (e.g., Beckman, Hirschberg, & Shattuck-Hufnagel, 2005), for example, showed that the manipulated utterances were judged as having less of an accent. In a further extension, Jilka (2000) found that listeners were worse at deciding which language they heard in a monotonous

low-pass filtered condition, leading him to conclude that intonation per se is of crucial importance to the perception of a foreign accent.

To directly assess the contribution of intonation to the perception of foreign accentedness, Jilka (2000) compared accentedness ratings for pairs of sentences consisting of a foreign-accented utterance and the same utterance in which the intonation had been replaced by a correct, unaccented intonation pattern by means of rule-based generation and resynthesis. In addition, fully synthesized utterances were included that retained the original foreign-accented intonation contour but in which any segmental foreign accent was removed by using concatenative diphone synthesis. Thus, this evaluation involved a comparison between fully synthesized and original/resynthesized utterance pairs. Comparison of foreign accent ratings of the "natural" rule-generated version with those of the synthesized version of the same utterance showed that the synthesized stimuli were generally rated as less accented. Given these data, Jilka (2000) concluded that segmental information contributes more strongly to perceived foreign accent than intonation.

Winters and O'Brien (2013) recently investigated the relative contributions of F0 and duration to perceived accentedness and intelligibility, also examining typologically similar languages. In this study, English and German speakers produced both English and German sentences (both groups were highly proficient in their L2), and these utterances were evaluated by English monolingual listeners as well as English and German proficient L2 listeners. Native intonation contours and syllable durations or only syllable durations were combined with native or nonnative segmental productions. Most important for the present purpose are the conditions in which native English segments were combined with German-accented prosody (intonation and duration) and in which German-accented segments were combined with English prosody (intonation and duration). While the overall results showed that nonnative segments affected both accentedness and intelligibility, accentedness and intelligibility ratings diverged for nonnative segments with native prosody. It should be noted that sentences were presented in the clear for accentedness ratings but in pink noise for intelligibility ratings. Winters and O'Brien found that the combination of German-accented segments with English prosody reduced the perceived accentedness of the nonnative speech, suggesting that segments may contribute more to perceived foreign accent than prosody. However, for intelligibility, German-accented segments were less intelligible when combined with English prosody. Although Winters and O'Brien's overall conclusion is that segmental cues do seem to improve perceived accentedness in nonnative speech more than prosodic cues do, nonnative prosody does seem to play a role by reducing intelligibility.

Holm (2009) investigated the contribution of intonation and duration to the perception of foreign-accented Norwegian spoken by native speakers from seven different languages (English, French, German, Mandarin, Persian, Russian, and Tamil). In one condition, the duration of each segment in the L2 Norwegian productions was changed to that in native Norwegian productions. In a second condition, the native Norwegian intonation contour was simplified ("stylized") and copied onto the corresponding L2 Norwegian production. In a third condition, both duration and intonation were manipulated by copying the stylized native Norwegian intonation contour onto L2 Norwegian productions with native Norwegian

durations. Listeners had to indicate whether one sentence of the pair was far less accented, less accented, or equally accented relative to the other sentence. In general, results showed that the combined manipulation of duration and intonation significantly reduced the degree of perceived foreign accent for all native speakers, although certain native languages (e.g., Tamil and Mandarin) improved more than others (e.g., Russian and German). Holm also collected intelligibility ratings for different manipulations of the same sentence presented in pink noise. The results showed that the intonation manipulation improved intelligibility for English, German, Tamil, and Russian, while duration was most important for French. For Chinese and Persian, there was no difference among the intelligibility of the manipulations. However, it is somewhat difficult to interpret these intelligibility results and relate them to the accentedness results. Because the original accented nonmanipulated sentences were significantly more intelligible than the manipulated ones, they could not serve as a baseline condition. To remedy this, "close original" duration and intonation stimuli were created. Consequently, given these additions, the accentedness ratings and intelligibility ratings were not collected for the same set of manipulations.

Tajima, Port, and Dalby (1997) investigated the importance of temporal properties for intelligibility. They used linear predictive coding resynthesis and dynamic time warping to change the temporal properties of Chinese-accented English to match those of native English speech while leaving the original formant frequencies and F0 intact. Intelligibility was measured in a forced identification task with the target and three distracter sentences. Results showed a significant improvement in intelligibility for sentences with modified temporal properties as compared to the original Chinese-accented utterances. In addition, manipulation in the opposite direction, changing native English temporal attributes to match those of Chinese-accented English, significantly reduced intelligibility.

Quené and Van Delft (2010) recently replicated this latter finding in a sophisticated design. Using native Dutch and Polish-accented Dutch, segmental durations were swapped between the two versions while differences in speaking rate and intonation were removed. Pitch synchronous overlap and add (Moulines & Charpentier, 1990) was used to adjust the segment durations of one sentence version to match those of the other version, and vice versa. Using the speech reception threshold method (Plomp & Mimpen, 1979), sentences were presented in masking noise that matched the speech signal. Dutch listeners had to repeat each sentence out loud, and the signal to noise ratio was decreased or increased depending on whether their response was correct or incorrect, respectively, until accuracy had reached 50%. The results showed that Polish-accented sentences with Dutch segmental durations were slightly, but significantly, more intelligible than the original Polish-accented sentences. In a replication of Tajima et al. (1997), Quené and Van Delft (2010) also found that the original Dutch sentences were more intelligible than Dutch sentences with segment durations from the Polish-accented versions. The latter manipulation, native Dutch with Polish-accented durations, affected intelligibility more than the former, Polish-accented Dutch with native Dutch durations. This led Quené and Van Delft (2010) to conclude that hearing nativelike speech with inappropriate durational information may be relatively worse than hearing nonnativelike speech with appropriate

durational information. This is similar to Winter and O'Brien's (2013) finding for resynthesized German-accented English productions that intelligibility decreases when nonnative segments are combined with native prosody as compared to nonnative prosody. It is important to note that, while for Quené and Van Delft (2010) native durations improved intelligibility, their contribution was relatively small compared to the contribution of segmental information. That is, a much larger difference in intelligibility was observed between the original Dutch sentences and the Polish-accented sentences with Dutch durations. Because speaking rate and intonation were controlled for, this difference must be due to segmental deviations.

Previous research has established that nonnative suprasegmentals give rise to the perception of a foreign accent and that nonnative intonation and timing each contribute to this perception. While the extent to which different suprasegmental features contribute to a foreign accent depends on both the L2 learners' native language (Holm, 2009) and on the target language, there is little agreement on the relative contribution of segments and suprasegmentals to perceived foreign accent. While some studies suggest that suprasegmental aspects are more important (e.g., Anderson-Hsieh et al. 1992; Anderson-Hsieh & Koehler, 1988; Magen, 1998; Munro & Derwing, 1999), others suggest a primary role for segments (e.g., Jilka, 2000; Quené & Van Delft, 2010). The results of previous studies are not conclusive for a variety of reasons. In previous studies, listeners were asked to judge one aspect of nonnative speech while ignoring another, making it difficult to fully tease apart the individual contribution of segmental and suprasegmental information. For example, in Anderson-Hsieh et al. (1992), judges rated accented speech samples in terms of both segmental and suprasegmental errors; because both segments and suprasegmentals are simultaneously present in the samples, segmental errors may have affected suprasegmental ratings and vice versa. Moreover, Jilka's (2000) conclusion that segmental information contributes more strongly to perceived foreign accent than intonation was based on a comparison between fully synthetic and natural resynthesized sentences, which introduces naturalness of the materials as a potential confounding factor as well. Finally, because the focus of Quené and Van Delft (2010) was on nonnative durational patterns, their results can only indirectly suggest that segmental errors play a more significant role in foreign accentedness.

In the present study, we will attempt to separate segmental features from suprasegmental features and evaluate the contribution of each to perceived foreign accent. At the suprasegmental level, we will focus on intonation, which has been shown to play a role in the perception of foreign accent (e.g., Holm, 2009; Jilka, 2000; Willems, 1982). Specifically, using digital signal processing techniques, we will separate the intonation from the segmental content for both native and nonnative speech samples and superimpose the intonation of one onto the segmental content of the other. Furthermore, in the present study, we will examine accentedness, comprehensibility, and intelligibility separately to determine exactly how segmental and suprasegmental errors impact all three evaluation methods. We focus on Korean-accented English (Flege, 1999; Flege et al., 2006) because there are both substantial prosodic and segmental differences between Korean and English. While English is typically described as a stress-timed language, the rhythm

structure of Korean seems to fall in between that of syllable-timed and mora-timed languages (Arvaniti, 2012; Tark, 2012). The prosody of Korean has also been analyzed in detail (e.g., Jun, 1996) and the intonation of Korean-accented English has been documented in several studies (e.g., Jun, 1998; Trofimovich & Baker, 2006). For example, in their acoustic analysis of F0 peak alignment in Korean-accented English, Trofimovich and Baker (2006) report that the maximum F0 value occurs significantly later in the stressed syllable than in native English. This difference is presumably because while the pitch peak in English is usually aligned with the onset of the stressed syllable (Ladd, Mennen, & Schepman, 2000), the pitch peak in Korean is typically aligned with the offset of the stressed syllable (Jun, 1998; Lee, 2013). In their study of Korean-accented declarative English sentences, Kim and Kim (2001) also found that the pitch accent typically fell on the last syllable of the phonological word in focus. In addition, they report that Korean speakers exhibit distinct tonal patterns for phrases (e.g., the low-high-low-high tone sequence of Korean accentual phrases) that sometimes gives the impression that they are asking a question rather than making a statement in English. Kim and Kim conclude that these two instances of Korean-to-English transfer result in intonation patterns in Korean-accented English that are clearly distinct from those of natively produced English.

In terms of segments, there are a number of differences between Korean and English that may provide difficulty for Korean learners of English. Korean has alveolar and glottal but no interdental or postalveolar fricatives. Korean has no labial-velar or alveolar approximants. As for vowels, Korean lacks the lax counterparts of /i/ and /u/. It also lacks /æ/ (Lee, 1999). Finally, the syllable structure of English is more complex than that of Korean. For example, Korean does not allow syllable-initial consonant clusters and Korean learners of English may insert a vowel to break up a consonant cluster (e.g., Eckman & Iverson, 1993; Tarone, 1980).

Although there is no uniform agreement about the relative contributions of segments and intonation, the bulk of the evidence leads to the expectation that intonation errors will be the factor that most influences ratings of both comprehensibility and accentedness, with sentences that have native segments and nonnative intonation rated as less comprehensible and more accented than the sentences that have nonnative segments and native intonation. We also hypothesize based on some previous literature that nonnative intonation may also reduce intelligibility. This overall pattern of data would suggest that intonation errors contribute more to comprehensibility, accentedness, and intelligibility than do segmental errors.

## METHODS

### *Production*

*Speakers.* The participants included two female Korean speakers of the Seoul dialect and two female native English speakers. According to a linguistic background questionnaire, Korean Speaker 1, age 29, started learning English at the age of 14 and had a total of 10 years of English education. Korean Speaker 2, age 28, started learning English at the age of 13 and had a total of 7 years of English

education. Because our primary focus was not on potential individual differences but on the relative contributions of segments and intonation across native and nonnative speakers, we limited our investigation to these two nonnative speakers. We selected these two speakers because they were of the same gender, had started learning English at approximately the same age, and were judged to have a similar degree of foreign accent. Both speakers had a moderately strong foreign accent as judged (on a 4-point scale ranging from *no accent* to a *very strong accent*) by all three authors. The two English speakers (age 22 and 20) were native speakers of Midwestern American English, with no knowledge of Korean and with minimal knowledge of a L2. One speaker had limited knowledge of Spanish and French; the other speaker had limited knowledge of German. All of the participants were students at the University of Kansas who volunteered to participate in the study.

*Materials.* The materials consisted of 40 sentences (lists B, F, I, and J) from the Central Institute for the Deaf everyday sentences (Davis & Silverman, 1970; see Appendix A for a complete list of stimuli). The everyday sentences were specifically chosen for two primary reasons: first, they are designed to be representative of the vocabulary and syntactic structures of everyday spoken English; and second, they represent multiple sentence types that differ in their intonational properties, including a combination of declaratives, imperatives, and both yes/no and *wh*-questions. In order to determine if there were acoustic differences between the original speech samples, five measures were collected. Duration (ms), mean F0 (Hz), minimum F0 (Hz), maximum F0 (Hz), and F0 range (Hz) for each of the sentences were all measured using Praat (Boersma & Weenink, 2010).

*Procedure.* The speakers read the sentences one at a time from a sheet of paper. The sentences were written in four lists of 10 with filler sentences at the beginning and end of each list to avoid a list intonation. All recordings were made in an anechoic chamber with an ElectroVoice 767 microphone and a Marantz PMD 671 solid-state recorder at a sampling rate of 22,050 Hz.

Next, the sentences were manipulated using Praat (Boersma & Weenink, 2010). The average amplitude of each sentence was equated to 70 dB. Sentences were manipulated in pairs of the same sentence produced by two different speakers: one who provided the segments and the other who provided the intonation. Our decision to take segments from one speaker and intonation from the other speaker was based on random choice. The duration of the segmental sample was shortened or lengthened to match the duration of the intonation sample. Finally, the F0 contour of the segmental sample was replaced with the F0 contour from the intonation sample. All three authors, as well as an additional three lab members, informally rated the naturalness of all original and manipulated stimuli. The results suggested no differences in naturalness between original and manipulated sentences.

Four versions of each sentence were presented: native segments from English Speaker 1 with native intonation from English Speaker 2 (Es+Ei); native segments from English Speaker 1 with nonnative intonation from Korean Speaker 1 (Es+Ki); nonnative segments from Korean Speaker 1 with native intonation from English

Speaker 1 (Ks+Ei); and nonnative segments from Korean Speaker 1 with nonnative intonation from Korean Speaker 2 (Ks+Ki).

### *Perception*

*Listeners.* The participants were 40 native English speakers (27 females, 13 males). These participants did not speak Korean and were not familiar with Korean-accented English. The ages ranged from 18 to 46, with an average age of 21. The participants consisted of University of Kansas students without any known hearing disorders; they received extra credit in an introductory linguistics course.

*Materials.* The 160 sentences (40 sentences × 4 manipulated conditions) were presented to listeners in a Latin square design. Each listener heard all 40 sentences with 10 sentences produced by each of the four speaker combinations (Es+Ei; Es+Ki; Ks+Ei; and Ks+Ki) so that no listener heard the same sentence more than once.

*Procedure.* After filling out a linguistic background questionnaire, each participant began the experiment. The experiment was set up in Paradigm (Tagliaferri, 2008). Each participant sat in front of a monitor and listened to the sentences with headphones. All listeners were students who were naive to the purpose of the experiment and were not explicitly trained. The experiment was presented in three blocks. In the first block, examining intelligibility, participants listened to and transcribed the sentences, which were presented in randomized order, on a sheet of paper. After the sheet of paper was taken by the researcher, the participants began the second block. In this block, examining comprehensibility, they listened to the sentences in a new randomized order and rated them for comprehensibility by clicking on a scale from 1 to 5 (1 = *very easy to understand*, 5 = *very difficult to understand*). The participants then began the third block, examining accentedness, in which they again listened to the sentences in a new randomized order and rated them for accentedness on a scale from 1 to 5 (1 = *no accent*, 5 = *strong accent*). We selected this order so that we would start with intelligibility, because it does not require a subjective judgment, and finish with an accentedness judgment, which would require an explicit judgment about the degree of accentedness.

## RESULTS

### *Intelligibility*

For intelligibility, the total number of errors was calculated for each speaker combination (Derwing & Munro, 1997; Winters & O'Brien, 2013). Errors included omission of words, replacement of one word for another, addition of words, splitting of contractions, and relocation of a word from one location to another within a sentence (Derwing & Munro, 1997). Examples of errors include "Why don't they paint their words some other color?" for "Why don't they paint their walls some other color?"; "Core me a little later" for "Call me a little later"; and "You will get fed eating candy" for "You will get fat eating candy." The total

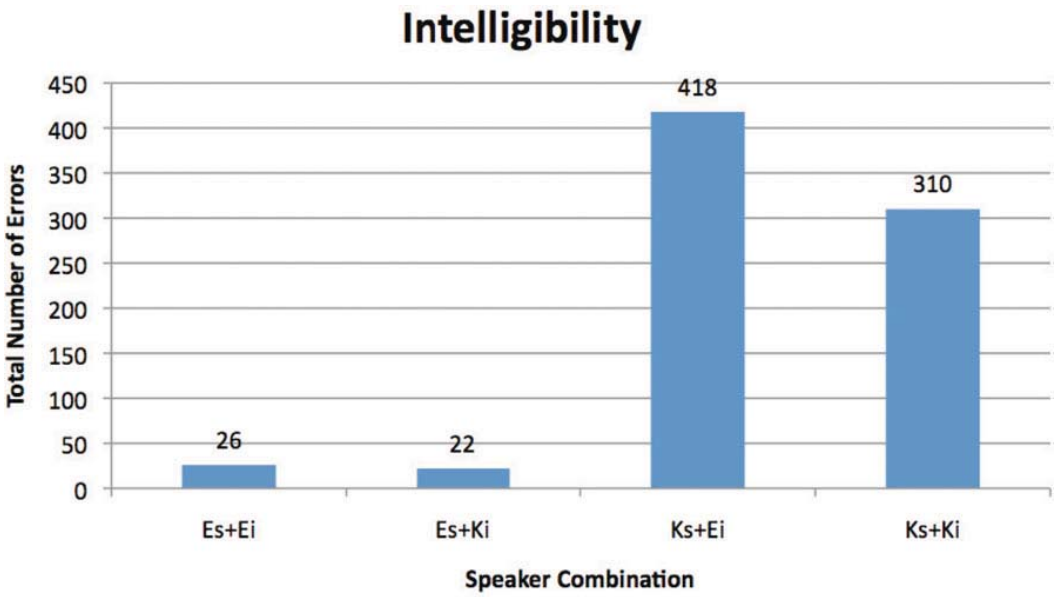


Figure 1. (Color online) Total number of errors for English segments with English intonation (Es+Ei), English segments with Korean intonation (Es+Ki), Korean segments with English intonation (Ks+Ei), and Korean segments with Korean intonation (Ks+Ki).

number of errors for each speaker combination is shown in Figure 1. A two-way repeated measures analysis of variance (ANOVA; Segments × Intonation) showed a significant effect of segments, $F(1, 9) = 210.532$, $p < .001$, $\eta_p^2 = 0.96$. Sentences with English segments (48 errors; mean number of errors per sentence = 0.6, $SD = 1.2$) had fewer errors than sentences with Korean segments (728 errors; mean number of errors per sentence = 9.1; $SD = 12.1$). Overall, sentences with English segments were more intelligible. This held true for segments with both English and Korean intonation.

There was also a significant effect of intonation, $F$ (1, 9) = 6.279, $p = .034$, $\eta_p^2 = 0.41$. Sentences with English intonation (444 errors; mean number of errors per sentence = 5.6, $SD = 10.8$) had more errors than sentences with Korean intonation (331; mean number of errors per sentence = 4.2, $SD = 8.2$). Overall, sentences with English intonation were less intelligible. This main effect varied depending on the speaker of the segments, as shown by the trend for the interaction between segments and intonation, $F$ (1, 9) = 4.007, $p = .076$, $\eta_p^2 = 0.31$. This interaction effect was carried by the sentences with Korean segments. Sentences with Korean segments and English intonation (418 errors; mean number of errors per sentence = 10.5, $SD = 13.6$; 14.0% error rate) produced significantly more errors compared to sentences with Korean segments and Korean intonation (310 errors; mean number of errors per sentence = 7.8, $SD = 10.4$; 10.5% error rate), $t$ (9) = 2.241, $p = .052$. There was no significant difference, $t$ (9) = 0.768, $p = .462$, in the total number of errors between sentences with English segments and English intonation (26 errors; mean number of errors per sentence = 0.65, $SD = 1.5$; 0.9% error rate) compared to sentences with English segments and Korean

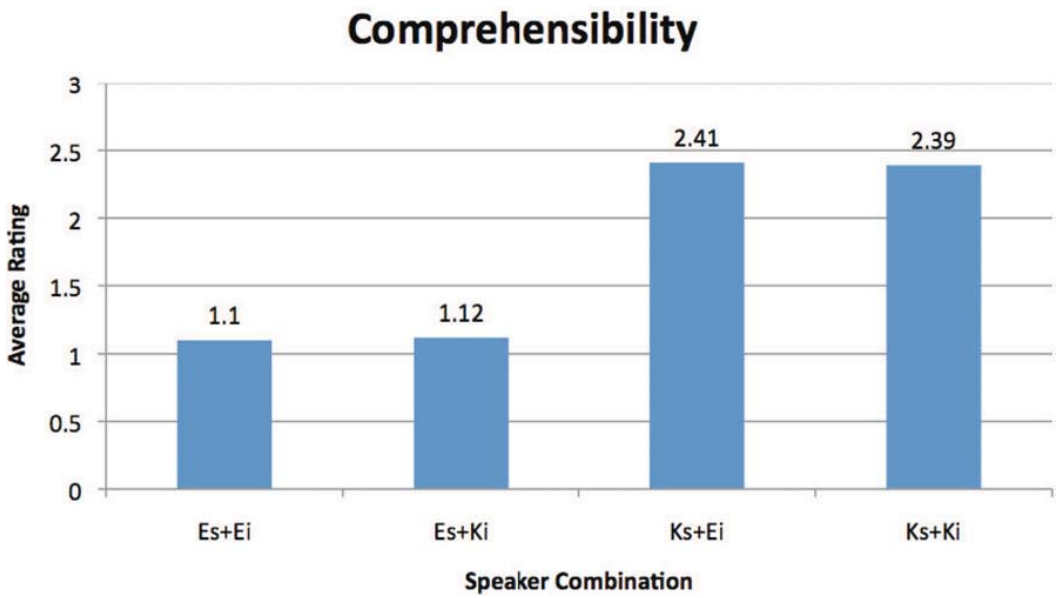


Figure 2. (Color online) Average comprehensibility ratings for English segments with English intonation (Es+Ei), English segments with Korean intonation (Es+Ki), Korean segments with English intonation (Ks+Ei), and Korean segments with Korean intonation (Ks+Ki) on a scale from 1 to 5, with 1 = *very easy to understand* and 5 = *very difficult to understand*.

intonation (22 errors; mean number of errors per sentence = 0.55, *SD* = 1.0; 0.9% error rate).

A paired-samples *t* test showed a significant difference between the number of errors for sentences with English segments and Korean intonation compared to the number of errors for sentences with Korean segments and English intonation, $t$ (9) = –9.492, $p < .001$. The sentences with English segments and Korean intonation (22 errors) had significantly fewer errors than the sentences with Korean segments and English intonation (418 errors).

### *Comprehensibility*

For comprehensibility, the average comprehensibility rating for each speaker combination was calculated, on a scale from 1 to 5 (1 = *very easy to understand*, 5 = *very difficult to understand*). The results are shown in Figure 2. A two-way repeated measures ANOVA showed a significant effect of segments, $F$ (1, 9) = 774.501, $p < .001$, $\eta_p^2 = 0.99$. Sentences with English segments (mean comprehensibility rating = 1.11, *SD* = 0.15) were rated as significantly more comprehensible than sentences with Korean segments (mean comprehensibility rating = 2.40, *SD* = 0.88). There was no significant effect of intonation, $F$ (1, 9) = 0.012, $p = .914$, $\eta_p^2 = 0.001$, and no significant interaction between segments and intonation, $F$ (1, 9) = 0.874, $p = .374$, $\eta_p^2 = 0.09$.

A paired-samples *t* test showed there was no significant difference in the comprehensibility ratings between the sentences with English segments and English intonation compared to sentences with English segments and Korean intonation,

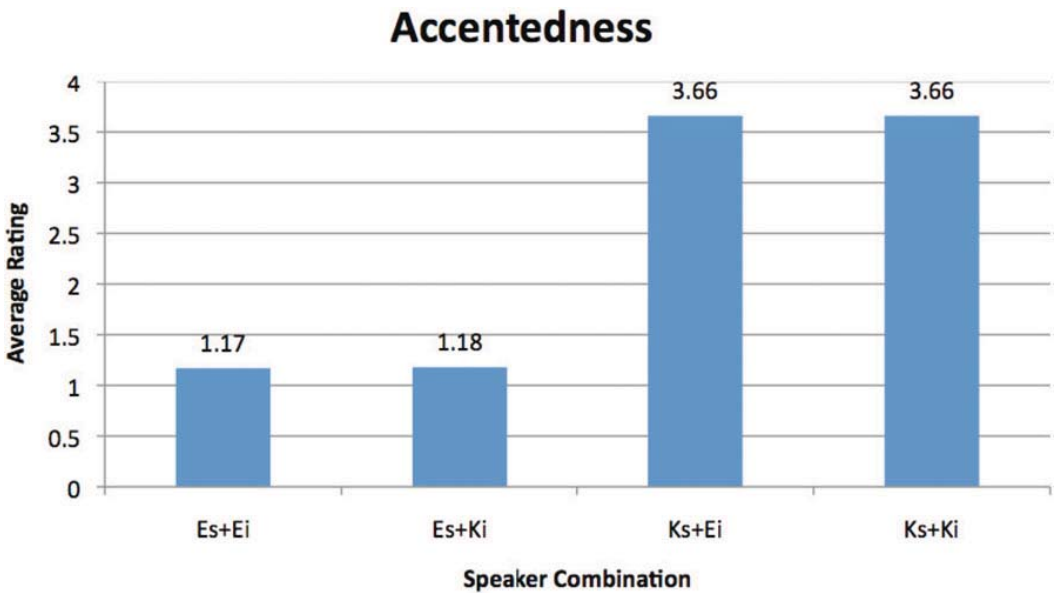


Figure 3. (Color online) Average accentedness ratings for English segments with English intonation (Es+Ei), English segments with Korean intonation (Es+Ki), Korean segments with English intonation (Ks+Ei), and Korean segments with Korean intonation (Ks+Ki) on a scale from 1 to 5, with 1 = no *accent* and 5 = *strong accent*.

$t$ (9) = 1.272, $p$ = .235, or between sentences with Korean segments and English intonation compared to sentences with Korean segments and Korean intonation, $t$ (9) = 0.569, $p$ = .583.

A paired-samples $t$ test showed a significant difference between the comprehensibility ratings for sentences with English segments and Korean intonation compared to the comprehensibility ratings for sentences with Korean segments and English intonation, $t$ (9) = –26.511, $p < .001$. The sentences with English segments and Korean intonation (mean comprehensibility rating = 1.12) were more comprehensible than the sentences with Korean segments and English intonation (mean comprehensibility rating = 2.41).

### *Accentedness*

For accentedness, the average accentedness rating for each speaker combination was calculated, on a scale from 1 to 5 (1 = *no accent*, 5 = *strong accent*). The results are shown in Figure 3. A two-way repeated measures ANOVA showed a significant effect of segments, $F$ (1, 9) = 492.266, $p < .001$, $\eta_p^2 = 0.98$. Sentences with English segments (mean accentedness rating = 1.17, $SD$ = 0.20) were rated as less accented than sentences with Korean segments (mean accentedness rating = 3.66, $SD$ = 0.74). There was no significant effect of intonation, $F$ (1, 9) = 0.020, $p$ = .891, $\eta_p^2 = 0.002$, and no significant Segment × Intonation interaction, $F$ (1, 9) = 0.009, $p$ = .927, $\eta_p^2 = 0.001$.

A paired-samples $t$ test showed there was no significant difference in the accentedness ratings between sentences with English segments and English

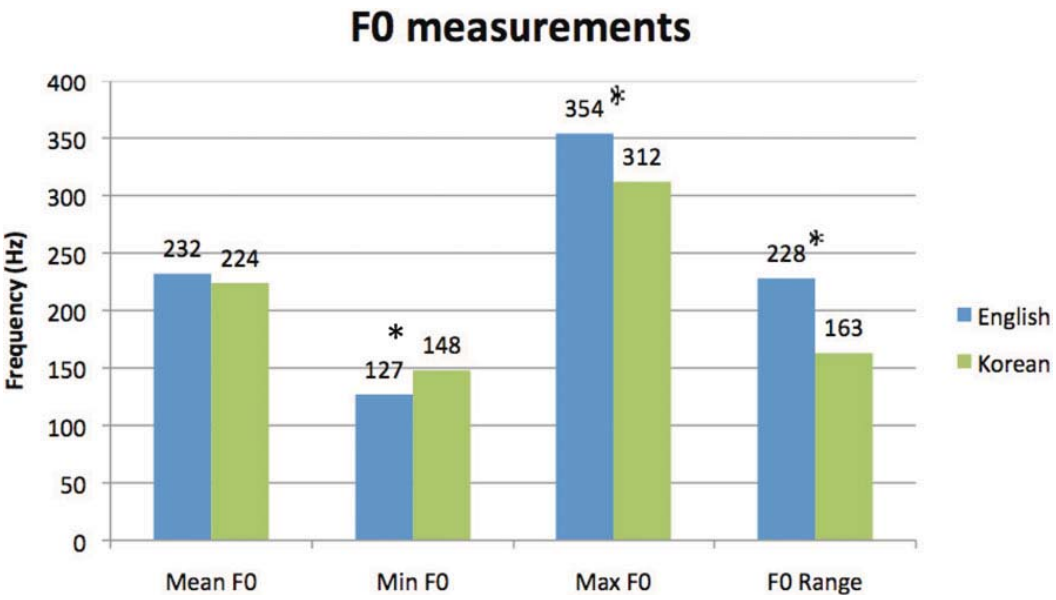


Figure 4. (Color online) Mean fundamental frequency (Mean F0), minimum fundamental frequency (Min F0), maximum fundamental frequency (Max F0), and fundamental frequency range (F0 range) for native English and native Korean speakers. Asterisks indicate significant differences.

intonation compared to sentences with English segments and Korean intonation, $t$ (9) = –0.307, $p$ = .766, or between sentences with Korean segments and English intonation compared to sentences with Korean segments and Korean intonation, $t$ (9) = –0.032, $p$ = .975.

A paired-samples $t$ test did show a significant difference between the accentedness ratings for sentences with English segments and Korean intonation compared to the accentedness ratings for sentences with Korean segments and English intonation, $t$ (9) = –19.948, $p < .001$. The sentences with English segments and Korean intonation (mean accentedness rating = 1.18) were less accented than the sentences with Korean segments and English intonation (mean accentedness rating = 3.66).

### *Acoustic analyses*

The acoustic measures for the stimuli were also examined. For duration, an independent-samples $t$ test revealed that the mean duration of sentences produced by the English speakers (1681 ms) was significantly shorter than that produced by the native Korean speakers (2009 ms), $t$ (158) = –2.558, $p$ = .011. This confirms previous studies that also reported slower speaking rates for nonnative as compared to native speech (e.g., Munro & Derwing, 1995, 1999).

There was no significant difference between the English speakers' mean F0 (232 Hz) and the Korean speakers' mean F0 (224 Hz), with overall F0 of the two English speakers similar to the overall F0 of the two Korean speakers. The F0 measurements are shown in Figure 4. The independent-samples $t$ test showed significant differences between English and Korean speakers for minimum F0, $t$ (158) = –3.034, $p$ = .003, maximum F0, $t$ (158) = 5.212, $p < .001$, and F0

range, $t\,(158) = 6.095$, $p < .001$. English speakers had a lower minimum F0 (127 Hz) than Korean speakers (148 Hz) and a higher maximum F0 (354 Hz) than Korean speakers (312 Hz), resulting in a greater F0 range for English speakers (228 Hz) than for Korean speakers (163 Hz).

## DISCUSSION

The current study investigated the separate contribution of segments and suprasegmentals to the perception of accented speech, specifically examining the contribution of segments and intonation to intelligibility, comprehensibility, and accentedness ratings. Sentences were created with the segments of a native speaker combined with the intonation of a nonnative speaker or the segments of a nonnative speaker combined with the intonation of a native speaker. The present study focused on Korean-accented English because there are substantial prosodic and segmental differences between English and Korean. While the majority of previous studies on the effect of segments and suprasegmentals on foreign-accented speech found that the suprasegmentals, or prosody, have more of an impact on comprehensibility and accentedness than segments, the results of the present study, focusing on one aspect of prosody, intonation, point in a different direction. The present study found that segments, rather than intonation, had a significant effect on the perception of accented speech. Moreover, this pattern was observed across the intelligibility, comprehensibility, and accentedness ratings. Overall, sentences with native segments were rated as more intelligible, more comprehensible, and less accented than were sentences with nonnative segments, holding true for sentences with either native or nonnative intonation. Winters and O'Brien (2013) found comparable results for two prosodically similar languages, German and English, in that the combination of native segments with nonnative prosody is perceived as less accented than the combination of nonnative segments with native prosody, further supporting the conclusion that segments contribute more to perceived foreign accent than prosody. For the present study, segments had a significant effect on intelligibility, comprehensibility, and accentedness. For all three measures, sentences containing English segments and Korean nonnative intonation were judged more intelligible, more comprehensible, and less accented than sentences with Korean nonnative segments and English intonation. It appears that native speakers rely on segments when judging foreign-accented speech.

Previous studies have observed differences among intelligibility, comprehensibility, and accentedness in terms of how well they correlate with different types of errors. In their study of Mandarin-accented English, Munro and Derwing (1995) found that for over 70% of the native English judges, accentedness ratings correlated significantly with the accent features of phonemic errors, phonetic errors, grammatical errors, and intonation. Comprehensibility showed significant correlations with intonation and grammatical errors for a majority of judges, while intelligibility showed significant correlations for only a small percentage of the judges. A subsequent follow-up study (Derwing & Munro, 1997), with English learners from a variety of native languages, found substantially different results. In Derwing and Munro (1997), there is only one type of error that showed a significant correlation for the majority of judges: grammatical errors were significantly correlated with comprehensibility for 54% of the judges. Derwing and Munro (1997)

offer several potential reasons for these discrepancies between the two studies, including a difference in proficiency level of the nonnative speakers (speakers in Derwing and Munro, 1997, were more proficient than in Munro and Derwing, 1995) and that intonation ratings were based on natural samples in Munro and Derwing (1995) but on low-pass filtered samples in Derwing and Munro (1997). Low-pass filtering is often used to remove segmental information from the speech signal while preserving prosodic information. That this manipulation may have yielded very different results suggests that listeners do not selectively focus on segmental or prosodic information but that the two are closely intertwined and that it is therefore difficult, if not impossible, to focus on one or the other level in rating tasks when both are simultaneously present. Moreover, low-pass filtering may not be a suitable technique when studying accentedness because it has been shown to affect ratings of native speech as well, resulting in an overall impression of "foreign-ness" (van Els & de Bot, 1987).

As was evident in the present acoustic analyses, the Korean samples were significantly longer than the English samples. Therefore, when the Korean segments with English intonation samples were created, the duration of the Korean segments was often shortened. This increase in speaking rate may have made the speech more difficult to understand, leading to the significant effect of intonation that we observed in our intelligibility ratings. Anderson-Hsieh and Koehler (1988) found that nonnative speech spoken at an increased rate (approximately 30% faster than the normal rate) was less comprehensible than nonnative speech presented at a regular rate. However, differences in speaking rate did not affect other measures of foreign accentedness. In the present study, we found that while an increase in speaking rate (by approximately 20%) of the Korean-accented segments did have a significant effect on intelligibility, this increase did not affect comprehensibility or accentedness. That is, there were no differences in either comprehensibility or accentedness between sentences with Korean segments and Korean intonation and sentences with Korean segments and English intonation in which the Korean segments were typically shortened to match the duration of the English intonation contour.

Although segment duration and speaking rate were adjusted at a sentential level in the present study by shortening or lengthening utterances so that segmental productions matched the durations of the intonation contours, previous research suggests that individual segment duration and overall speaking rate may not be the primary contributors to intelligibility or the perception of a foreign accent. Quené and Van Delft (2010) found that foreign-accented sentences with native durations were much less intelligible than natively produced sentences with foreign durations. Perception of native speech with nonnative durations was more accurate than perception of nonnative speech with native durations, a finding similar to the present study. Because speaking rate and intonation were controlled in these experiments (Quené & Van Delft, 2010), this result points to a minor role for segment duration. Trofimovitch and Baker (2006) do report that speaking rate contributed significantly to the perception of a foreign accent, but the $R^2$ change of 9% suggests a modest contribution relative to pause duration, which accounted for 37% of the variance. Moreover, while Trofimovitch and Baker (2006) found no differences in speaking rate among the three proficiency levels of Korean learners of English, these three learner groups did differ in terms of ratings of foreign

accentedness, with the more inexperienced learners' speech receiving significantly stronger accentedness ratings. Finally, Derwing and Munro (1997) also suggest that speaking rate is not a good predictor of foreign accentedness because speaking rate was significantly correlated with only a few listeners' judgments (ranging from 38% of the judges for comprehensibility to only 8% for intelligibility).

It is also possible that the overall length of the sentential stimuli may have impacted the results. Munro and Derwing (1995) used 36 samples that ranged from 4 to 17 words, with a mean of 10.7 words per sample. They found no effect of utterance length on any of their measures and concluded that their samples were of suitable length for the raters' tasks. In the present study, we used 40 samples, ranging in length from 2 to 12 words (2 to 17 syllables), with a mean of 7.2 words (8.4 syllables) per sample. The present samples were thus shorter than the materials used by Munro and Derwing (1995) but substantially longer and more varied than those used by Trofimovich and Baker (2006). While the latter study included 6 short sentences, ranging in length from 5 to 8 words (5 to 8 syllables) with a mean of 6 words (6.3 syllables) per sentence, Trofimovich and Baker (2006) showed that these sentences were of sufficient length to obtain significant correlations between accentedness ratings and suprasegmental accuracy scores. It is therefore unlikely that the length of our stimuli affected the results.

Finally, we considered our acoustic analyses. The Korean speech samples did not differ from the English samples in mean overall F0, indicating that the Korean and English speakers had a similar voice pitch. That the samples did differ significantly in minimum F0, maximum F0, and F0 range in addition to duration shows that the Korean and English intonation do differ, and the lack of an effect of intonation on the accentedness and comprehensibility ratings is not due to a lack of variability between the two speaker groups.

The present study is one of the first that attempts to truly separate segments from intonation, showing that segmental information contributes substantially more to the perception of foreign accentedness than intonation. Perhaps having the segmental and intonation errors both present in previous studies created confounding effects, which were avoided in the present study by teasing apart segments and suprasegmental information, focusing on a single native language, and examining intonation as one aspect of prosody. Future studies should include both spectral and temporal prosodic manipulations to fully determine which aspects of prosody affect the perception of a foreign accent. In addition, future studies should also examine a variety of L2 learners with a range of language proficiencies.

## APPENDIX A

### List B

B1. The water's too cold for swimming.
B2. Why should I get up so early in the morning?
B3. Here are your shoes.
B4. It's raining.
B5. Where are you going?
B6. Come here when I call you!

B7. Don't try to get out of it this time!
B8. Should we let little children go to the movies by themselves?
B9. There isn't enough paint to finish the room.
B10. Do you want an egg for breakfast?

List F

F1. Music always cheers me up.
F2. My brother's in town for a short while on business.
F3. We live a few miles from the main road.
F4. This suit needs to go to the cleaners.
F5. They ate enough green apples to make them sick for a week.
F6. Where have you been all this time?
F7. Have you been working hard lately?
F8. There's not enough room in the kitchen for a new table.
F9. Where is he?
F10. Look out!

List I

I1. Where can I find a place to park?
I2. I like those big red apples we always get in the fall.
I3. You'll get fat eating candy.
I4. The show's over.
I5. Why don't they paint their walls some other color?
I6. What's new?
I7. What are you hiding under your coat?
I8. How come I should always be the one to go first?
I9. I'll take sugar and cream in my coffee.
I10. Wait just a minute!

List J

J1. Breakfast is ready.
J2. I don't know what's wrong with the car, but it won't start.
J3. It sure takes a sharp knife to cut this meat.
J4. I haven't read a newspaper since we bought a television set.
J5. Weeds are spoiling the yard.
J6. Call me a little later!
J7. Do you have change for a five-dollar bill?
J8. How are you?
J9. I'd like some ice cream with my pie.
J10. I don't think I'll have any dessert.

## REFERENCES

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

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

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*, 351–373.

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.

Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer, Version 5.1.31 [Computer program]. Retrieved April 4, 2010, from http://www.praat.org/

Broselow, E., Chen, S.-I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*, 261–280.

Davis, H., & Silverman, S. (1970). *Hearing and deafness*. New York: Holt, Rinehart & Winston.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 20*, 1–16.

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

Eckman, F. R., & Iverson, G. K. (1993). Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research, 9*, 234–252.

Edwards, J., & Zampini, M. (2008). *Phonology and second language acquisition*. Philadelphia, PA: John Benjamins.

Fayer, J., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*, 313–326.

Flege, J. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–132). Hillsdale, NJ: Erlbaum.

Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34*, 153–175.

Flege, J., Bohn, O., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics, 25*, 437–470.

Flege, J., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America, 76*, 708–721.

Flege, J., Munro, M., & MacKay, D. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125–3134.

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34*, 65–89.

Holm, S. (2009). *Intonational and durational contributions to the perception of foreign-accented Norwegian*. Unpublished doctoral dissertation, Norwegian University of Science and Technology.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent.* Unpublished doctoral dissertation, University of Stuttgart.

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

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

Kashiwagi, A., & Snyder, M. (2010). Speech characteristics of Japanese speakers affecting American and Japanese listener evaluations. *Working Papers in TESOL & Applied Linguistics, 10*, 1–14.

Kim, J. S., & Kim, S.-H. (2001). Remarks on Korean speakers' realization of English intonation: Focusing on declarative sentences. *Pan-Pacific Association of Applied Linguistics, 5*, 187–207.

Koster, C., & Koet, T. (1993). The evaluation of accent in the English of Dutchmen. *Language Learning, 43*, 69–92.

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.

Lee, H. B. (1999). *Korean.* In *Handbook of the International Phonetic Association* (pp. 120–123). New York: Cambridge University Press.

Lee, H. (2013). *Evidence of diachronic sound change: A comparative acoustic study of Seoul and Kyungsang Korean*. Unpublished doctoral dissertation, University of Kansas.

Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*, 381–400.

Major, R. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Erlbaum.

Major, R., & Faudree, M. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition, 18*, 69–90.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication, 9*, 453–468.

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 foreign-accented speech. *Language and Speech, 38*, 289–306.

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. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*, 285–310.

Plomp, R., & Mimpen, A. M. (1979). Improving the reliability of testing the speech reception threshold for sentences. *International Journal of Audiology, 18*, 43–52.

Quené, H., & Van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication, 52*, 911–918.

Reetz, H., & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics, and perception*. Hoboken, NJ: Wiley–Blackwell.

Tagliaferri, B. (2008). Paradigm: Perception research systems [Computer software]. Retrieved from http://www.paradigmexperiments.com/

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics, 25*, 1–24.

Tark, E. (2012). *An experimental study on Korean rhythm structure on the basis of rhythm metrics*. Unpublished master's thesis, University of Kansas.

Tarone, E. (1980). Some influences on the syllable structure of interlanguage phonology. *International Review of Applied Linguistics, 18*, 139–152.

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.

van Els, T., & de Bot, K. (1987). The role of intonation in foreign accent. *Modern Language Journal, 71*, 147–155.

Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris.

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*, 486–507.