

Perceptual Judgments of Accented Speech by Listeners from Different First Language Backgrounds

May 2016 – Volume 20, Number 1

Okim Kang

Northern Arizona University, USA.
<okim.kang@nau.edu>

Son Ca Thanh Vo

Iowa State University, USA.
<soncavo@iastate.edu>

Meghan Kerry Moran

Northern Arizona University, USA.
<mkm338@nau.edu>

Abstract

Research in second language speech has often focused on listeners' accent judgment and factors that affect their perception. However, the topic of listeners' application of specific sound categories in their own perceptual judgments has not been widely investigated. The current study explored how listeners from diverse language backgrounds weighed phonetic parameters (i.e., segmental features such as consonants and vowels and suprasegmental features such as word stress and sentence stress) differently when perceiving non-native speakers' accented speech. Two hundred forty listeners, including American, Vietnamese, and Arabic students, rated Vietnamese accented English for intelligibility, comprehensibility, and accentedness. Within this group of participants, 112 raters also provided interview responses to questions related to their perception of accented speech in general. The results suggest that listeners of English perceived degree of accent in fundamentally different ways, depending on factors such as their first language and their English instruction backgrounds. Features identified in this study can be useful both in the listeners' global judgments and in the communicative situations in which second language learners need to function.

Keywords: Accent judgment, listener background, segmental, suprasegmental

Introduction

Research on the effect of listeners' first language (L1) background on their perceptual judgments has been mixed thus far. Some studies have shown that L1 effects are small and not consistently observable (e.g., Munro, Derwing, & Morton, 2006), while others found significant differences between native speakers (NS) and non-native speakers (NNS; e.g., Riney, Takagi, & Inutsuka, 2005). An overall consensus seems to be that listeners' perceptions can be affected by speech properties of speakers or some

listeners' factors such as listeners' language experience (Munro, 2008). The current study sought to advance the understanding of factors that affect listeners' perceptions of non-native speech, particularly by investigating the impact of listeners' own language background on their perceptual judgments of accented English speech.

Previous research in speech perception often focused on global ratings of listeners, but not on listeners' application of specific segmentals or suprasegmentals. The relationship between NNSs' focus of pronunciation instruction (i.e., which segmental and/or suprasegmental features pronunciation teachers emphasized explicitly) and their accent perception have also been rarely investigated. In addition, with notable exceptions (e.g., Ortmeier & Boyle, 1985; Smith & Bisazza, 1982; Wilcox, 1978), listeners as research participants in the past have been either NSs or NNSs as English as a second language (ESL) students who have resided in the USA, but not necessarily speakers of English as a foreign language (EFL). Consequently, this study investigated how listeners from different first language and language learning backgrounds applied phonetic parameters differently when perceiving NNSs' accented speech in English.

The phonetic parameters in this study refer to segmentals and suprasegmentals. Segmentals are minimal units of sound (vowels and consonants) defined in phonetic terms (Pennington & Richards, 1986) while suprasegmentals refer to "a vocal effect which extends over more than one sound segment in an utterance, such as a pitch, stress or juncture pattern" (Crystal, 2003, p. 446). In this study, specific phonetic parameters (i.e., word stress, sentence stress, and particular consonants and vowels) in English were targeted and altered in such a way that is commonly heard in Vietnamese-accented speech. We were interested in understanding how untrained impressionistic judgments and the phonetic parameters that influence them differ by the listeners' background; that is, the actual accuracy of listeners' judgments was not the focus of this study.

Review of Literature

Different L1 Backgrounds in Listeners' Judgments of Accented Speech

Various research studies have examined listeners' accent judgments and factors that affect these evaluations. When the contribution of segmental and suprasegmental features on listeners' judgments of accented speech is discussed, one important factor to be taken into consideration is the listeners' backgrounds. Gass and Varonis (1984) demonstrated that listeners' judgments are affected by their language experience. They found that listeners' familiarity with the topic, accent, speaker, and L2 speech were strongly correlated with their judgments of intelligibility. One factor that may contribute to greater tolerance of listeners from particular language backgrounds for particular NNS accents is "the interlanguage speech intelligibility benefit" (Bent & Bradlow, 2003, p.1602), which predicts that a NNS listener may be better equipped to interpret specific acoustic-phonetic features of an L2 that are matched with his own L1 than a different L1. Although findings regarding the interlanguage speech intelligibility benefit were mixed in Major, Fitzmaurice, Bunta, and Balasubramanian's (2002) study, Spanish listeners seemed to benefit from their L1 accent, scoring better on a listening comprehension test featuring a Spanish speaker than did those of other L1 backgrounds. Also, Chinese and Japanese listeners were found to understand Spanish accented English rather well. Major et al. (2002) suggested that this phenomenon could be due to a

similar lack of vowel reduction found among Chinese, Japanese, and Spanish; however, other factors such as listener attitudes or the fact that the Spanish speakers had less of an accent are possible causes as well.

In contrasting studies, L1 effect on listeners' judgments has been shown to be minimal, if present. Listeners can show moderate to high correlation on global accent judgments regardless of L1 background (Munro et al., 2006). Few differences were found in the ratings of accented speech between NS and NNS listeners (MacKay, Flege, & Imai, 2006). In judgments of oral English performance, NS and NNS teachers exhibited similar severity patterns (Kim, 2009). Flege (1988) also confirmed that there was no consistent pattern found in ratings of perceived foreign accent among different groups of listeners (i.e., high-proficient experienced Chinese, low-proficient inexperienced Chinese, and experienced American) when listening to English sentences spoken by native speakers of English and Chinese. In light of this conflicting research, more empirical studies are clearly needed to provide further evidence on these issues (Munro, 2008).

In related studies, novice NNS raters appeared to be harsher than NSs in judgments of accented speech (Kang, 2008; 2012). These fundamental differences between NSs and NNSs' judgments might be based on the application of different phonetic parameters (i.e., segmentals vs. suprasegmentals) that raters utilized (Riney et al., 2005). For example, in Riney et al.'s (2005) study, two trained phoneticians conducted an auditory analysis on L2 sentences that untrained Japanese and American learners judged dissimilarly on the construct of accent. They found that the Japanese listeners used primarily nonsegmental parameters (specifically intonation, fluency, and speech rate) to make perceptual judgments, whereas segmental parameters had a relatively minor role. In contrast, the American listeners exhibited the opposite pattern; that is, they applied more segmental parameters (/l/ and /r/) but nonsegmentals played a minor role. These findings suggest that NS and NNS listeners perceive degree of accent in English in fundamentally different ways based on different phonetic parameters.

Overall, the question regarding how listeners from different L1 backgrounds perceive L2 speech still remains unclear. It is common for different groups of ESL/EFL speakers to use English for international communication, and their different perceptions of NNSs' speech continue to affect their interactions (Major, 2007). However, few studies have focused on the understanding of listeners' judgment process through their self-reports. The current study investigated how different groups of listeners differed in their judgments of accented speech. More specifically, the primary research question addressed is: When different groups of listeners (NSs, NNSs from the same L1 as the speaker, and NNSs from a different L1 than the speaker) perceive English accented speech, how do phonetic parameters influence their perceptual judgments?

Phonetic Parameters in Listeners' Judgments

Phonetic parameters used in the current study refer to specific pronunciation features, such as, vowels and consonants for segmentals and lexical and sentence stress for suprasegmentals. In particular, specific features from both segmental and suprasegmental components of English speech were chosen to be altered in accordance with typical Vietnamese-accented speech.

Researchers have argued over the roles that segmentals and suprasegmentals play in speech perception and intelligibility (Anderson-Hsieh, Johnson, & Koehler, 1992; Flege, 1981; Jenkins, 2002; Riney et al., 2005). First of all, segmental features can play an important role in speech perception. Segmental errors were found to contribute greatly to a foreign accent and to have detrimental effects on L2 comprehension (Fayer & Krasinski, 1987). Cutler and van Donselaar (2001) posited that although Dutch listeners used suprasegmental cues for word recognition in their native language, the contribution of segmental features was more important than that of suprasegmental features. According to Flege (1981), one of the most apparent features for a foreign accent is derived from segmental sound substitutions such as in French-accented I sink so or Arabic-accented I put my car in the barking lot. In short excerpts of speech produced by NNSs, the frequency of segmental substitutions was found to be highly correlated with NS judgments of accentedness (Brennan, Ryan, & Dawson, 1975). Although rare, especially among more proficient speakers, the extreme reduction or deletion of entire syllables can also interfere greatly with intelligibility (e.g., “decrating” instead of “decorating”; Kang & Moran, 2014). Johansson (1978) found that NSs judged mispronounced consonant errors more severely than vowel errors and that mispronounced sounds in isolated words contributed more to listeners’ comprehension than errors in sentence and text levels. He also compared phonological and grammatical errors in L2 speech and found that phonological errors played a significant role in listeners’ comprehension.

Jenkins (2002) also asserted that certain pronunciation features are more important to intelligibility than others and therefore deserve more pedagogical focus. In her *Lingua Franca Core* (LFC) model, segmentals have primacy over suprasegmentals and consonants over vowels in communication between NNSs and NNSs. Moreover, Gimson (1970) claimed that accurate production of consonants was more essential to L2 comprehension than native-like production of vowels, even though Schairer (1992) provided the opposite evidence for English-speaking learners of Spanish. In part because of these research findings regarding segmental importance, segmental accuracy has been stressed in pronunciation textbooks as well as ESL/EFL classrooms.

The relative impact of segmental errors on listeners’ judgments can be also determined by functional load (Brown, 1991; Catford, 1987). For example, interdental fricatives carry a low functional load and are thus not high-priority sounds in communication. Difficulty producing sounds with a high functional load such as /p/ and /b/ are more likely to cause a breakdown in communication than sounds with a low functional load (Brown, 1991; Catford, 1987). Likewise, it has been found that as ESL learners progress, their high functional load errors (both vowels and consonants) decrease significantly although their low functional load errors may not (Kang & Moran, 2014).

On the other hand, suprasegmental features of speech are associated with stretches that are larger than the segment (whether vowel or consonant), in particular pitch, stress, intonation, rhythm, or duration (Lehiste, 1970). Many studies have suggested that perceived foreign accent, intelligibility, and comprehensibility of NNSs’ English might be more greatly impacted by prosodic than segmental factors (Anderson-Hsieh, Johnson & Koehler, 1992; Derwing, Munro, & Wiebe, 1998; Hahn, 2004; Field, 2005; Isaacs, 2008;

Kang, 2010). Marslen-Wilson (1987) argued the low impact of segmental errors in L2 comprehension, stating that some phonemic errors might not be likely to disrupt communication due to more native-like suprasegmental features.

Anderson-Hsieh et al. (1992) investigated the relationship between different types of pronunciation errors (particularly in prosody and segmentals), syllable structure, and NS listeners' reactions in speech samples taken from the SPEAK Test. Although they found a strong correlation between the aforementioned pronunciation errors and global foreign accent, the prosodic variable proved to have the strongest effect. Other studies have further investigated different aspects of suprasegmental errors which could affect L2 perception, such as speech rate (Munro & Derwing, 1995; Issacs, 2008; Kang, 2010), voice quality (Munro, Derwing, & Burgess, 2003), several aspects of intonation (Wennerstrom, 2000), word/lexical stress (Field, 2005), and sentence (primary or nuclear) stress (Hahn, 2004; Kang, 2010). The contributions of these features to listeners' perception have varied widely.

NNSs from a variety of linguistic backgrounds seem to find the stress patterns of English particularly challenging. It is true that English learners often face problems such as misplacing word stress and sentence stress (Hahn, 2004), and stress patterns could easily cause communication breakdowns in the speech of NNSs (Gallego, 1990). In fact, according to Kang's (2010) study, stress measures best predicted untrained raters' accent ratings. In addition, the syllable structure associated with word stress is a critical component of intelligibility rating among ESL teachers (Zielinski, 2008). Therefore, in this study, suprasegmental errors mainly focused on word and sentence stress for an experimental purpose, in terms of their effects on NSs' judgments of intelligibility, comprehensibility, and accentedness. Segmental errors included vowel and consonant errors.

Methods

Listeners

Two hundred and forty university students (80 American, 80 Vietnamese, and 80 Arabic) participated as listeners and were assigned into three groups. The American university students (32 males and 48 females) were enrolled in undergraduate university courses at a southwestern university. Their age ranged from 18 to 45 ($M = 27.30$, $SD = 7.62$). The Vietnamese listeners (17 males and 63 females) were first-year university students from the English Department at a centrally located foreign language university in Vietnam. Their English proficiency proved to be upper intermediate with their age ranging from 18 to 20 ($M = 18.25$, $SD = .65$). Although these students had not taken the Test of English as a Foreign Language (TOEFL) or the International English Language Test System (IELTS), they had passed the English National Examination in order to be accepted into the university. This corresponds approximately to the B1 (CEFR) and/or a score of 4.5-6.0 on the IELTS. The Arabic students (27 males and 53 females) were upper intermediate and advanced ESL students from an intensive English program (IEP) at a southwestern university in the United States with an age range of 18 to 25 ($M = 18.85$, $SD = .65$). Proficiency levels were determined by the IEP's placement and achievement tests. The mean length of their U.S. residence was 5.6 months. Among the 240 students, 112 participated in short interviews (80 Vietnamese, 19 American,

and 13 Arabic students) after their speech ratings. Participants' responses to their background survey indicated that the American and Arabic listeners were not familiar with Vietnamese English L2 accent. All of the participants reported having normal hearing. All procedures were in accordance with the Institutional Review Board at the research university.

Speech Stimuli

Speech stimuli were prepared from several stages of the screening process after adopting methods from various sources (e.g., Gass & Varonis, 1994; Hahn, 2004; Munro & Derwing, 1995). Ten Vietnamese speakers (5 males and 5 females) who were highly proficient in English (TOEFL scores above 100 out of 120) were initially recruited; the high TOEFL score helped to ensure that speakers would make few, if any, unintended pronunciation errors. They were graduate students in the USA aged from 26 to 34. They were asked to read 40 English sentences which consisted of 20 with segmental pronunciation errors and 20 with suprasegmental errors common for Vietnamese speakers. In particular, they were asked to mispronounce highlighted sounds of words (vowels and consonants) in given sentences and to misplace stress in words and sentences according to guidelines provided by the authors. (See Speech Materials in the following section and the Appendix.)

Once the speech stimuli (400 sentences) made by 10 Vietnamese speakers were collected, the study recruited four linguistic experts, two native speakers and two non-native speakers of English (one Vietnamese L1 and one Korean L1) who had substantial linguistic/phonetic training as well as extensive experience in teaching ESL students. The linguistic experts were asked to test each of the sentences for its intended appropriateness and the accuracy (or inaccuracy) of the pronunciation. That is, while listening to speech files, they were asked to compare the scripts which had accurate sentences, to focus on words and sounds marked for the intended errors, and to determine whether or not the sentences included the intended errors properly made by the speakers. Lexical stress errors were verified by the location of stressed syllables and sentence stress by the placement of prominence on content words. Each sentence included two or three intended errors included. The experts were allowed to listen to the stimuli multiple times. They then selected the sentences which contained errors suitably made for the purpose of this study. Among 400 sentences, the process of this stimuli screening yielded 29 sentences (6 sentences with consonant errors, 4 with vowel errors, 8 with word stress errors, and 11 with sentence stress errors), all of which were agreed upon by all four experts for the precision of errors. In order to maintain the unity of the distribution for each phonetic category, however, the study chose 16 sentences only (i.e., four sentences for each parameter). These sentences did not contain any unintended pronunciation errors.

The speech stimuli were further tested by four additional listeners for their coherence to target objectives before they were played for primary ratings. The raters were two graduate and two undergraduate American students who did not have any linguistic training background.

Speech Materials

Pronunciation errors often found in Vietnamese speakers of English are final consonant substitutions, final consonant cluster deletions, or mispronunciation of lax/tense vowels (Avery & Ehrlich, 1992; Christian, Wolfram, & Hatfield, 1986; Osburne, 1996). Suprasegmental errors such as the misplacement of lexical stress or sentence stress are not uncommon. The stimuli materials, sentences with problematic sounds expected for Vietnamese speakers of English, were prepared after consulting Avery and Erlich (1992), Celce-Murcia, Brinton, and Goodwin (2010), Christian et al. (1986), and Morley (1992). The selected errors were further confirmed through personal contact of the second author with current Vietnamese teachers and students in Vietnam. Although characteristics of Vietnamese phonology may vary among regions of the country (Northern, Central, and Southern; Hwa-Froelich, Hodson, & Edwards, 2002), the difficult sounds (e.g., final consonants) chosen were mainly for Vietnamese speakers of English from Central Vietnam.

Word-final voiceless sounds included /p, t, k, f/, as they are often mispronounced as a mixture of /b, d, g, v/ by Vietnamese speakers. Vietnamese speakers do not often release those consonants in a final position or substitute those sounds with others (Hwa-Froelich, et al., 2002). Targeted word-final consonant clusters were /st, ts, ks, ft/. The vowel contrasts in focus were /i:/ vs. /i/, /e/ vs. /ɛ/, /u/ vs. /ʊ/, /ɔ/ vs. /ʌ/. Examples of suprasegmental errors were misplaced syllables in words (e.g., They are talking about last year's preSidential Election) and misplaced words in sentences, such as stressing function words instead of content words (e.g., THERE WAS A terrible car accident ON THE corner). Using a headset, recordings were made digitally on a computer. The samples varied from 5 to 13 words, with a mean of 8.6 words. Each sample was between 3.0 and 6.5 seconds long, with a mean length of 4.3 seconds. This speech rate (approximately 2 words/second) is at the low end of what previous research has found to be indicative of natural speech of native English speakers (i.e., 125-225 words/minute or 2.08-3.75 words/second; Jones, Berry, & Stevens, 2007).

Rating Instruments

The study yielded four outcome measurements for listeners' perceptual judgments: intelligibility, comprehensibility, accentedness, and global judgments. First, listeners were asked to listen to the entire set of 16 sentences initially and to rate the global comprehensibility and accentedness. These global measures were intended to assess listeners' overall impression on the entire 16 sentences, but not for a specific category. The global comprehensibility was measured on a 9-point Likert scale (1 = *hard to understand*; 9 = *easy to understand*) and the global accentedness was assessed with another 9-point scale of 1 = *has a strong accent* and 9 = *has no accent*.

Next, comprehensibility and accentedness were individually assessed for each of the sentences. The comprehensibility measure also employed a 9-point bipolar scale adopting Munro and Derwing's (1995) and Kang's (2010) instruments. The listeners were asked to listen to the 16 sentences and to assign perceived comprehensibility (1 = *hard to understand*; 9 = *easy to understand*) for each sentence. The accentedness rating scale (1=*has a strong accent*; 9 = *has no accent/native-like accent*) was also adopted from Kang's (2010) and Munro and Derwing's (1995) accent standardness rating scale.

After the global measurements, the individual accentedness and comprehensibility measurements, intelligibility was measured employing Derwing and Munro's (1997) approach. All the 16 sentences were orthographically transcribed with multiple checks for accuracy. The listeners listened to each utterance and then wrote out in standard orthography exactly what they heard. For this task, the recording was only played once; however, listeners had heard the utterance twice previously during the global measurement tasks. Intelligibility was calculated by the percentage of words exactly matching the original transcription. Overall intelligibility scores for the four categories were calculated by counting the mean of each group of listeners for their correct words in sentences. The mean scores for each sentence ranged from 20% to 74%.

Procedures

Listener participants completed a language background questionnaire in which questions asked listeners' language learning experience and their familiarity with the Vietnamese accent. Listeners were then asked to listen to the entire set of 16 speech sentences as a whole and to complete global ratings of comprehensibility and accentedness. Two to five meetings for each L1 group were arranged in quiet classrooms for these rating tasks. Each meeting consisted of 15-40 listeners.

After a break, listeners were asked to listen to each of the speech samples individually for the ratings of comprehensibility and accentedness. They assigned rating scores to each of these two rating constructs for each sentence. All speech samples were randomly presented. Subsequently, for transcriptions of sentences that served to measure intelligibility, listeners were given booklets with numbered spaces. The participants were instructed to listen to each utterance and to write out in standard orthography exactly what they heard. There were approximately 1.5 minutes of pause between sentences. The stimuli were played only once.

The target sentences were presented to each listener over earphones. We controlled the CD by pressing a pause button at the end of each utterance. A new stimulus was not presented until all the listeners had finished their rating of the previous one. Each meeting lasted approximately 1.5 hours. After the listeners completed their ratings, they took part in 5-10 minute interviews answering questions such as "When you listen to accented speech, to what pronunciation errors do you react most sensitively (e.g., vowels, consonants, word stress, sentence stress, intonation, and rhythm)? Why?" While 19 out of the 80 American participants and 13 of the 80 Arabic listeners volunteered to participate in the interviews, all 80 Vietnamese participants contributed to this interview process. The American and Arabic students received course credit for their interview participation. For the Vietnamese students, participating in the interview was considered part of their English practice activity as well as an extra credit opportunity. All responses were recorded and notes were taken when necessary.

Data Analysis

The study yielded four dependent variables: global ratings, comprehensibility, accentedness, and intelligibility. A total of 16 sentences were divided into four sections: consonants, vowels, word stress, and sentence stress. Each section was composed of four sentences and each sentence included two to three category-specific errors. For

convenience of subsequent analysis, the means of the four sentence judgment scores in each section (i.e., the sum of the four sentence scores divided by four) were utilized as composite measures (except for the scores of the global ratings). Reliability coefficients (Cronbach's alpha) were .87, .91, and .90 for intelligibility, comprehensibility, and accentedness, respectively. Quantitative analysis included one-way ANOVAs, correlations, and multiple regressions along with post hoc pair-wise Tukey tests. Interview data were used as supportive evidence for the quantitative data results (Creswell & Plano Clark, 2007).

Results

The study aimed to examine differences in perceptual ratings by three groups of listeners (NSs, NNSs from the same L1 as the speaker, and NNSs from a different L1 than the speaker) on Vietnamese accented speech in terms of the degree of perceived comprehensibility, accentedness, and intelligibility. High values in ratings indicate listeners' positive judgments of the speakers (i.e., high intelligibility, high comprehensibility, and native-like accent). Table 1 displays the mean scores of three groups of comprehensibility ratings in four different categories of pronunciation errors. It also demonstrates how different pronunciation parameters affect listeners' perceptual judgments.

Table 1. Scores of Three Groups of Comprehensibility Ratings in Different Categories of Pronunciation Errors

Listeners	Consonants Mean (SD)	Vowels Mean (SD)	Word Stress Mean (SD)	Sentence Stress Mean (SD)
American	5.00 (1.38)	2.41 (1.02)	2.92 (1.30)	2.14 (1.03)
Vietnamese	2.50 (1.58)	3.57 (1.61)	3.46 (1.30)	3.67 (1.59)
Arabic	3.04 (1.85)	3.25 (1.38)	3.74 (1.40)	4.52 (1.70)

Note. Comprehensibility measure: 1 = hard to understand; 9 = easy to understand

Listeners found Vietnamese speech generally hard to understand, as shown in mean scores lower than Likert score 5 in all categories, because all sentences had certain pronunciation errors. Vietnamese listeners viewed the speech as less comprehensible when there were consonant errors in pronunciation. Conversely, American listeners reacted more sensitively to vowels and suprasegmental errors, but the consonant errors were the least influential factor for their comprehensibility judgments. Arabic listeners had trouble with both segmental and suprasegmental errors when listening to Vietnamese-accented speech.

One-way ANOVA results revealed that all comparisons of the three groups of rating scores for each of the pronunciation error categories were statistically significant: $F_{(2, 237)} = 60.30, p < .0005$, partial eta squared = .34 for consonant errors; $F_{(2, 237)} = 16.67, p < .0005$, partial eta squared = .12 for vowel errors; $F_{(2, 237)} = 7.46, p < .001$, partial eta squared = .06 for word stress errors; and $F_{(2, 237)} = 49.52, p < .0005$, partial eta squared = .29 for sentence stress errors. According to post hoc Tukey test results, all comparisons of mean scores of ratings between the American and Vietnamese listeners were

statistically significant ($p < .0005$). The mean difference of rating scores between American and Arabic listeners were also significant for all the categories of pronunciation errors ($p < .001$), while a significant difference in ratings between Vietnamese and Arabic listeners were only found in the sentence error section ($p < .001$). A similar pattern was found in the results of accentedness ratings. As shown in Table 2, all three groups of listeners found the speech samples relatively accented with mean scores of 5 or lower in the 9-point Likert scale. The American listeners reacted less sensitively to consonant errors than to other pronunciation errors in their accent judgments, whereas Vietnamese listeners treated the speech as more accented when there were consonant errors. As for Arabic listeners as non-native speakers listening to unfamiliar Vietnamese-accented speech, they perceived sentences with lexical stress errors as more accented than those with other errors.

Table 2. Mean Scores of Three Groups of Accentedness Ratings in Different Categories of Pronunciation Errors

Listeners	Consonants Mean (SD)	Vowels Mean (SD)	Word Stress Mean (SD)	Sentence Stress Mean (SD)
American	4.77 (1.54)	2.56 (1.11)	2.54 (1.10)	2.30(1.03)
Vietnamese	2.69 (1.35)	3.88 (1.80)	3.71 (1.69)	4.38 (1.59)
Arabic	4.72 (1.87)	4.22 (1.92)	2.88 (2.10)	4.74 (1.70)

Note. Accentedness measure: 1=has a strong accent.... 9= has no accent/native-like accent

Among the three groups of listeners, statistical differences were found in mean scores of accentedness ratings. The results of one-way ANOVAs were $F_{(2, 237)} = 43.34, p < .0005$, partial eta squared = .27 for consonant errors; $F_{(2, 237)} = 22.36, p < .0005$, partial eta squared = .16 for vowel errors; $F_{(2, 237)} = 14.64, p < .001$, partial eta squared = .11 for word stress errors; and $F_{(2, 237)} = 47.46, p < .0005$, partial eta squared = .29 for sentence stress errors. For each of the parameters, ratings of American listeners were statistically different from those of the Vietnamese ($p < .0005$). When it comes to accent ratings between American and Arabic listeners, significant differences were found in the categories of vowel and word stress errors. That is, the U.S. listeners found the Vietnamese speech slightly more accented than the Vietnamese listeners, when speech had pronunciation problems with vowels and word stress. Due to Vietnamese raters' sensitivity to consonant errors, accent ratings of Vietnamese listeners were significantly lower than those of Arabic listeners when sentences had consonant problems ($p < .001$). Intelligibility ratings also revealed a similar pattern in terms of how listeners apply their pronunciation parameters for their perceptual judgments. Table 3 shows mean scores of three groups of intelligibility ratings in different categories of pronunciation errors.

Table 3. Mean Scores of Three Groups of Intelligibility Ratings in Different Categories of Pronunciation Errors

Listeners	Consonants Mean (%) (SD %)	Vowels Mean (%) (SD%)	Word Stress Mean (%) (SD%)	Sentence Stress Mean (%) (SD%)
American	82 (12)	63 (11)	45 (12)	29 (11)
Vietnamese	9 (10)	25 (11)	23 (11)	28 (10)
Arabic	25 (12)	40 (14)	27 (10)	13 (10)

Note. Intelligibility scores: the percentage of words exactly matching the original transcription

Intelligibility scores appeared generally lower in NNS listeners (Arabic and Vietnamese listeners) compared to those in American listeners, perhaps due to NNSs' command of the English language itself. Listeners were required to transcribe the entire sentence after listening to the stimuli only once during that task. It is possible that transcribing might not have been an easy task for NNS participants in this study. Notwithstanding this proficiency issue, there was a noticeable contrast found between ratings of American listeners and Vietnamese listeners in terms of their reaction to pronunciation errors. When the speech had consonant errors, sentences were transcribed up to 82% correctly by American listeners, but only 9% by Vietnamese listeners. American listeners' intelligibility scores decreased with speech which had suprasegmental errors. These suprasegmental errors affected Arabic listeners in a similar manner. In contrast, Vietnamese intelligibility scores increased with such errors.

One-way ANOVA results showed statistically significant differences in the intelligibility ratings by the three groups of listeners: $F_{(2, 237)} = 63.12, p < .0005$, partial eta squared = .34 for consonant errors; $F_{(2, 237)} = 23.67, p < .0005$, partial eta squared = .16 for vowel errors; $F_{(2, 237)} = 45.41, p < .001$, partial eta squared = .27 for word stress errors; and $F_{(2, 237)} = 59.52, p < .0005$, partial eta squared = .32 for sentence stress errors. The post hoc Tukey results indicated that, except for the sentence stress category, the other three comparisons of mean scores of ratings between the American and Vietnamese listeners were statistically significant ($p < .0005$). Mean difference of rating scores between American and Arabic listeners were also significant for all the categories of pronunciation errors ($p < .001$). Ratings between Vietnamese and Arabic listeners were statistically different in all the error sections ($p < .001$), except the word stress section. As described in the Methods section, participants were initially asked to listen to the entire speech samples for the global judgments of comprehensibility and accentedness before any specific ratings. Global comprehensibility and accent ratings were collected to determine whether students' overall perceptions of the speech differed with their individual assessments. Overall, listeners of all the three groups found the speech very accented and hard to comprehend (see Table 4). Although statistical significance (family-wise F-values) existed in mean comparisons of all those three groups ($F_{(2, 237)} = 12.30, p < .0005$ for global comprehensibility ratings; and $F_{(2, 237)} = 12.61, p < .0005$ for global accentedness ratings), the actual differences in scores were relatively minimal.

Nevertheless, Vietnamese listeners, listening to their own L1-accented English speech, tended to be slightly more lenient than other groups of listeners both in ratings of global comprehensibility (post hoc Turkey tests, $p < .001$) and accentedness (post hoc Tukey tests, $p < .001$).

Table 4. Mean Scores of Three Groups of Global Accentedness and Comprehensibility Ratings Errors

Listeners	Global Comprehensibility Ratings	Global Accentedness Ratings
American	3.36 (1.73)	2.61 (1.35)
Vietnamese	3.61 (1.89)	3.51 (2.15)
Arabic	2.17 (1.04)	2.23 (1.29)

Statistical analysis showed that Arabic listeners (NNSs from the different L1 from the speaker), who were not familiar with the Vietnamese accented speech, were harsher than NSs (American listeners) or NNSs from the same L1 as the speaker. Tukey tests confirmed that their rating scores for global comprehensibility were significantly lower than the other two groups of listeners ($p < .0005$). Likewise, in terms of global accentedness ratings, both American listeners and Arabic listeners found the Vietnamese speech more accented than their own Vietnamese listeners. Although no significant correlation was found among the three groups of listeners for their comprehensibility ratings, American NSs were moderately correlated with Arabic listeners ($r = .41$).

Furthermore, Vietnamese listeners appeared to be somewhat more distinct from the other two groups of listeners with regard to their speech perception and application of pronunciation parameters. The results of multiple regression analyses confirmed the phenomenon that for global comprehensibility and accentedness ratings, the sentence stress error variable was a significant and potent predictor for both American listeners ($\beta = .56$ and higher, $p < .005$) and Arabic listeners ($\beta = .47$ and higher, $p < .005$), but for Vietnamese listeners the consonant error variable was the strongest predictor of their global judgments ($\beta = .31$ and higher, $p < .01$).

Interview Responses

One hundred twelve participants (80 Vietnamese, 19 American, and 13 Arabic students) took part in short interviews directly after their speech ratings. The interviews were in group format and informal; they lasted between five and ten minutes. The interview questions were asked in English for both the American and Arabic L1 listeners. For the Vietnamese group, the questions were first asked in English and then translated to Vietnamese to ensure full comprehension. Each interview session was videotaped by the researcher. Listener participants answered questions generally related to their perceptual judgments and processes of evaluating accented speech.

The interviews were the primary means by which the focus of pronunciation instruction was ascertained and revealed the connection between listeners' judgments and explicit pronunciation teaching. For example, the Vietnamese strong sensitivity to consonant errors was found in these interview reports. In response to a question about

pronunciation features that may affect their accent judgments, approximately 90% of respondents among 80 participants expressed consonant-related errors and their importance in their pronunciation learning and evaluation (i.e., 51% for only consonants, 10% for consonants and vowels, 27% for consonants and other features, 3% for vowels only, and 9% for others). A following comment from one of the Vietnamese respondents further supports this pattern: “consonants because my teachers teach me a lot of consonant errors compared to other types.” In fact, there was a clear tendency found in Vietnamese listeners that their judgmental decisions were closely intertwined with their current speaking/conversation class curricula. Almost all of the respondents identified the link between these two, adding that their English (EFL) teachers often emphasized the significance of consonants, followed by vowels, which were still limited to segmental features only.

The influence of teachers’ instruction on their perceptual judgments was also found from Arabic listeners as ESL learners at an Intensive English Program in the USA. For example, the following comment was made by one of the Arabic listeners: “currently we are learning about stress a lot. I know that I have to pay more attention to stress.” Others explicitly remarked on the effect that consonant instruction had made on their evaluations (note that although the Arabic students as a whole attended more to suprasegmental features, this was not necessarily true for each Arabic listener). For instance, one Arabic student responded, “When listening to accented speech, I react most sensitively to consonant errors because I practice consonants a lot with my teacher so I know them and I know very quickly who is speaking with consonant errors.” Although Arabic respondents’ comments varied, most of them commented on their current pronunciation curriculum and its influence on their perception.

An additional 19 U.S. undergraduate listeners also provided various responses ranging from ratings grounded in segmentals to those grounded in suprasegmentals. However, their comments included mostly features related to vowels and other suprasegmental parameters, but not necessarily to consonants. For example, as one participant noted, “I think what I notice first in accents are things that make them markedly different from my own American accent. For example, generally with Italians, what I notice first are the rhythm of speaking and the different syllable stresses...” The U.S. undergraduate responses were especially interesting because this group was not receiving pronunciation instruction and therefore could not link their judgments to pronunciation pedagogy. It is clear from these interview responses that listeners in different groups attend to different aspects of pronunciation when they listen to NNS’s speech.

Discussion

The purpose of this study was to examine how different groups of untrained listeners differ in using phonetic parameters (segmentals vs. suprasegmentals) to make their perceptual judgments of accented speech. The study also aimed to offer more empirical evidence to support claims about how segmentals and suprasegmentals affect the native’s and nonnative’s comprehension of nonnative speech. In this particular study, NSs (American) and NNSs from a different L1 than the speaker (Arabic) listeners’ judgments were somewhat more sensitive to suprasegmental errors such as sentence stress errors, whereas NNSs from the same L1 as the speaker (Vietnamese) reacted

more perceptively to segmental errors (e.g., consonant clusters) when listening to their Vietnamese-accented English. These findings suggest that listeners of English perceive accented speech in fundamentally different ways, depending on their L1 backgrounds and the focus of their pronunciation instruction. However, this conclusion should be considered with caution; without additional combinations of NS-NNS and NNS-NNS, it is not possible to determine that this is a result solely of the relationship between speakers' L1s or if it is catalyzed by the specific features of the languages targeted in this study.

Our overall findings are somewhat opposite to those of Riney et al.'s (2005) study, in which untrained Japanese listeners used primarily non-segmental parameters to make perceptual judgments and untrained American listeners applied segmental parameters more. There are a couple of possible explanations for this. First, as suggested by Riney et al., it is possible that suprasegmental features "sounded louder" (2005, p. 460) to Japanese listeners because the listeners did not make the same segmental distinctions that American listeners did. Specifically, many Japanese learners of English (even those with advanced proficiency) may not hear the English /r/ versus /l/ distinction in the same way that American listeners do (Takagi, 1993, 2002). Thus, this segmental feature may have served as a signal of accentedness to the American listeners, but Japanese listeners may not have attended to it. Second, the focus of pronunciation instruction was not ascertained in Riney et al. (2005). It is possible that the Japanese learners weighted suprasegmental parameters greater when evaluating accent because their pronunciation instruction had a suprasegmental emphasis.

The results of multiple regression analyses indicated that sentence stress was the most salient predictor of global perceptual judgments for American and Arabic listeners, whereas the consonant related variable most significantly predicted their global judgment scores when Vietnamese listeners rated the Vietnamese accented-speech. The high correlation between overall comprehensibility scores and prosody features has been well documented (e.g., Anderson-Hsieh et al., 1992; Field, 2005; Kang, 2010; Munro & Derwing, 1995). The results were also in line with Hahn's (2004) conclusion that the sentence stress errors of the NNS utterances made it difficult for native listeners to comprehend NNSs' speech.

What was not expected, however, is a distinctive pattern found among Vietnamese learners of English performing as listeners who evaluated their Vietnamese accented speech. Segmental deviance, particularly with consonant errors, affected their speech evaluations more adversely than did suprasegmental deviance. Interview responses gathered from each of the 80 respondents supported this tendency in that more than 90% of Vietnamese listeners addressed segment (consonants especially) related issues (i.e., consonant features were of their main concern, but not other pronunciation characteristics). Interestingly, according to Vietnamese respondents, this judgment pattern originated from the current pronunciation curriculum (i.e., mainly segment (consonant) focused pronunciation instruction) that they had received in Vietnam. In fact, Vietnamese listeners found terms such as lexical stress or sentence stress somewhat foreign, as they often seemed to have conceived pronunciation only as vowels

and consonants. Therefore, the focus of pronunciation instruction seems to contribute to how speech is understood and evaluated.

Another finding is dissimilarity in global rating judgments among the three groups of listeners (NSs, NNSs from the same L1 as the speaker, NNSs from the different L1 from the speaker). As the Vietnamese listeners might have benefited from listening to Vietnamese-accented English, their scores for global comprehensibility and accentedness were slightly higher than the NSs (American listeners) or other NNSs (Arabic listeners). This result concurred with findings of previous studies where the Japanese listeners rated the Japanese speakers as easier to understand than the Cantonese speakers (Munro et al., 2006; Smith & Bisazza, 1982). Thus, when looking at comprehensibility scores, the current finding seems to add additional evidence to support an intelligibility benefit (Bent & Bradlow, 2003) for speech produced in their L1 accent. However, the intelligibility scores do not reflect the Vietnamese listeners' confidence in comprehension. In fact, American L1 and even Arabic L1 listeners surpassed Vietnamese listeners on the intelligibility measure on all but the sentences with misplaced stress, in which the Vietnamese listeners outscored the Arabic listeners.

Factors that affect listeners' judgments of accented speech have been broadly studied, particularly with regard to L1 effect or accent familiarity (e.g., Bent & Bradlow, 2003; Gass & Varonis, 1984; Kang, in press; Munro, 2008). Findings have been mixed so far, however. While some studies found that prior exposure to varieties of accent does facilitate speech comprehension (Field, 2003; Gass & Varonis, 1984), others found no such effect (e.g., Munro, Derwing, & Morton, 2006). In Munro et al.'s (2006) study, for example, the listener groups of different L1s showed moderate to high correlations on intelligibility scores and comprehensibility and accentedness ratings, regardless of native language background. The current study exactly exhibited such a complexity of listeners' perception. Global accent ratings yielded a relatively moderate correlation ($r = .41$) between NS listeners and NNS Arabic listeners, but the three groups of listeners were not significantly correlated in their global comprehensibility ratings. This means that listeners' background (native) language factor did play a considerable role in their comprehensibility judgments of accented speech.

In line with the listener's L1 background, listeners' native English language status is another factor to consider. Findings of previous research on this topic are also inconclusive. In some studies (Fayer & Krasinski, 1987; Kang, in press), NNS listeners tend to be more severe in their assessments than NS listeners. In others, NS raters are harsher than NNS raters (Brown, 1995) or NS and NNS raters exhibit similar severity patterns (Kim, 2009). In this study, findings also differed depending on the constructs of the ratings. The NNSs that possessed a different L1 than the speaker (Arabic listeners) who were not familiar with the Vietnamese accented speech were harsher than NSs (American listeners) or NNSs from the same L1 as the speaker, especially in global comprehensibility ratings, but no significant rating difference between U.S. listeners and Arabic listeners emerged in global accent ratings.

The ESL/EFL distinction may play a role here as well, as the Arabic listeners had received instruction in an ESL environment while the Vietnamese listeners were EFL students. The need for EFL teachers' pronunciation training has been particularly

emphasized (Breitkreutz, Derwing, & Rossiter, 2001; Burgess & Spencer, 2000; MacDonald, 2002; Wang & Munro, 2004). Good pronunciation programs taught by professionally trained instructors may not be often available, and teachers themselves may be confused about what is possible or desirable in pronunciation instruction (Derwing & Munro, 2005). However, another urgent issue to address in matters of pronunciation instruction is appropriate training in pronunciation pedagogy in EFL contexts. In listening to listeners' voices through this study, teachers' instructional approach in pronunciation could play a critical role in shaping learners' perception of accented speech. Learners' pronunciation issues might not only be caused by students', but also by teachers' lack of awareness in functional features of L2 speech and their relationship with listeners' perception. Note that American listeners as NS listeners or even Arabic listeners as NNS listeners in the ESL environment who did not share the same L1 as the speaker in this study reported that they mainly attended to suprasegmental features (sentence or word stress) when they listened to accented speech. On the contrary, Vietnamese learners of English tended to prioritize segmentals (i.e., consonant features only). It is possible that Vietnamese learners of English may encounter disadvantages in international communicative situations in which L2 learners need to function. A suggestion that emerges from these findings is that in the NS-NNS listener research, a distinction needs to be made among NNS listeners: (1) NNSs in an ESL setting and (2) NNSs in an EFL setting.

Conclusion

Several implications can be drawn from our findings. We saw that listeners' factors (L1 background and their language learning experience) could affect their perceptual judgments of accented speech. Nevertheless, a question might still remain regarding whether or not this background factor plays a more important role in an EFL context rather than in an ESL context due to different instructional methods. This question can be further investigated in future research. In addition, findings emphasize the importance of teachers' roles in pronunciation instruction, particularly shaping learners' perceptual judgments of L2 speech. As for individual speech properties, three groups of listeners (NSs, NNSs from the same L1 as the speaker, and NNSs from a different L1 than the speaker) applied different phonetic parameters to their perceptual judgments. However, some correlations were found in global judgments (i.e., accentedness) among different L1 groups, which imply that listeners perhaps attend to different speech properties depending on types of speech rating constructs. Further research is called for concerning the effect of listening assessment constructs on listeners' use of phonetic parameters.

The implications of these findings can extend to the argument of Jenkins' (2002) *Lingua Franca Core (LFC)*. According to our findings, the LFC assertion that segmentals trump suprasegmentals and consonants trump vowels in NS/NNS communication can be true with Vietnamese speakers of English in an EFL environment, but not with Arabic speakers of English in an ESL environment. Jenkins' LFC is known for a model that explains communication between NNSs and NNSs. However, the results of this study suggest that within NNSs' communication, a more specific categorization of speech features may be needed to better understand successful oral communication. In

addition, NNS listeners' language learning background should be taken into consideration before involving NNSs' in any speech ratings.

Finally, the findings of this study provide support to both suprasegmental and segmental focus in pronunciation teaching (Anderson-Hsieh et al., 1992; Derwing, Munro, & Wiebe, 1998). EFL/ESL teachers should develop their pronunciation curriculum considering functional features of L2 speech and their relationship with listeners' judgments of intelligibility, comprehensibility, and accentedness. Reviewing comments by Vietnamese learners of English, students in that setting very much desire feature-balanced, curriculum-efficient pronunciation instruction.

Despite the implications listed above, the study can be further improved by overcoming a few limitations and expanding the scope of the study in the future. First, it would be beneficial to include more L1 backgrounds for both speakers and listeners in order to lessen the possibility that the results are based on language-specific characteristics. Also, the study treated post-secondary Vietnamese listeners as upper-intermediate English speakers, but no official English proficiency scores of these Vietnamese listeners were collected. Their English proficiency could have affected the results of this study. One particularly interesting facet would be to see if there is any difference among Vietnamese speakers of English from two different English-spoken contexts: (1) ESL and (2) EFL. That is, do listeners from the same L1 perceive their L1-related accented speech differently in different contexts? Additionally, grouping participants so that there were different focal points of pronunciation instruction within each of the listener groups would ensure that language background and emphasis of pronunciation pedagogy were not confounding variables. As a final point, phonetic parameters examined in this study were somewhat limited to four features (consonants, vowels, lexical stress, and sentence stress). A more comprehensive approach including other features of pronunciation (e.g., rhythm and intonation) or lexico-grammar is recommended.

It is important to bear in mind that this study lends just one piece to the puzzle of intelligibility of accented speech. Listeners' comprehension, which is integral to communication as well as assessment, was not measured. Because of this limitation, it would be difficult to justify any conclusion that a listener's specific performance was a direct result of his or her perception of the phonological features of accented speech. In fact, we do not know from the current study how the speaker's perceived accent would affect the listener's performance, an issue that is crucial for standardized English tests such as the TOEFL or the IELTS. Clearly, more work will need to be done in order to better understand the connections between pronunciation instruction, phonology, perception, and performance.

About the Authors

Okim Kang (PhD) is an associate professor of applied linguistics at Northern Arizona University. Her research focuses on second language pronunciation, oral language proficiency assessment, speech production and perception, language attitudes, and World Englishes. She is the recipient of the 2013 TOEFL Outstanding Young Scholar

Award and the Christopher Brumfit PhD/EdD Thesis 2009 Award by Cambridge University Press and Journal of Language Teaching.

Son Ca Thanh Vo is currently a PhD student in the Applied Linguistics and Technology program at Iowa State University, USA. She was a Fulbright recipient from Vietnam and received her MA in TESL from Northern Arizona University, Flagstaff, AZ, USA in 2011. She has been a lecturer of English at University of Foreign Languages – the University of Danang, Vietnam. Her research interests include pronunciation, corpus linguistics, assessment, curriculum and materials development, and second language teaching methods.

Meghan Moran is a third year doctoral student in the program of applied linguistics at Northern Arizona University. She received a master's degree from The Pennsylvania State University in Teaching English as a Second Language in 2008, after which she taught ESOL in a public school in western New York. Her interests include language planning and policy, pronunciation assessment, accept perception and discrimination, language education policy, and sociolinguistics.

References

- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42, 529-555.
- Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114(3), 1600-1610.
- Brennan, E., Ryan, E., & Dawson, W. (1975). Scaling of apparent accentedness by magnitude estimation and sensory modality matching. *Journal of Psycholinguistic Research*, 4, 27-36.
- Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19, 51-61.
- Brown, A. (1991). Functional load and the teaching of pronunciation. In A. Brown (Ed.), *Teaching English pronunciation: A book of readings*(pp. 211-224). London: Routledge.
- Brown, A. (1995). The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing*, 12(1), 1-15.
- Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System*, 28, 191-215.
- Catford, J.C. (1987). Phonetics and the teaching of pronunciation. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 83-100). Washington, D.C.: TESOL.

- Celce-Murcia, M., Brinton, D., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide*. New York: Cambridge University Press.
- Christian, D., Wolfram, W., & Hatfield, D. (1986). The English of adolescent and young adult Vietnamese refugees in the United States. *World Englishes*, 5(1), 47-60.
- Creswell, J., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Crystal, D. (2003). *A dictionary of linguistics & phonetics*. Malden, MA: Blackwell.
- Cutler, A., & van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech*, 44 (2), 171-195.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 1, 1-16.
- Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379-397.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, 393-410.
- Fayer, J.M. and Krasinski, E.K. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning* 37, 313-26.
- Field, J. (2003). The fuzzy notion of 'intelligibility': A headache for pronunciation teachers and oral testers. *IATEFL Special Interest Groups Newsletter*, 34-38.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399-423.
- Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly*, 15(4), 443-455.
- Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, 84, 70-79.
- Gallego, J. C. (1990). The intelligibility of three nonnative English-speaking teaching assistants: An analysis of student-reported communication breakdowns. *Issues in Applied Linguistics*, 1, 219-237.
- Gass, S. M., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34, 65-89.
- Gass, S. M., & Varonis, E. (1994). Input, interaction, and second language production. *Studies In Second Language Acquisition*, 16(3), 283-302.
- Gimson, A. C. (1970). *An introduction to the pronunciation of English*. London: E. Arnold.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Hwa-Froelich, D., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology*, 11(3), 264.

- Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *The Canadian Modern Language Review*, 64, 555-580.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23(1), 83-103. doi: 10.1093/applin/23.1.83
- Johansson, S. (1978). *Studies of error gravity: Native reactions to errors produced by Swedish learners of English*. Göteborg: Acta Universitatis Gothoburgensis.
- Jones, C., Berry, L., & Stevens, C. (2007). Synthesized speech intelligibility and persuasion: Speech rate and non-native listeners. *Computer Speech & Language*, 21(4), 641-651. doi:10.1016/j.csl.2007.03.001
- Kang, O. (2008). The effect of rater background characteristics on the rating of international teaching assistants' speaking proficiency. *Spain Fellow Working Papers*, 6, 181-205.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301-315.
- Kang, O. (2012). Impact of rater characteristics on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly*, 9, 249-269.
- Kang, O., & Moran, M. (2014). Pronunciation features in non-native speakers' oral performances. *TESOL Quarterly*, 48, 173-184.
- Kim, Y-H. (2009). An investigation into native and non-native teachers' judgments of oral English performance: A mixed methods approach. *Language Testing*, 26, 187-217.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass: M.I.T. Press.
- Major, R. C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, 29, 539-556.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173-190.
- Macdonald, S. (2002). Pronunciation – Views and practices of reluctant teachers. *Prospect: An Australian Journal of TESOL*, 17(3), 3-18.
- MacKay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, 27, 157-183.
- Marslen-Wilson, W. (1987). *Functional parallelism in spoken word recognition*. *Cognition*, 25, 71-102.
- Morley, J. (1992). *Rapid review of vowel and prosodic contexts: Improving spoken English: Consonants in context*. Ann Arbor, MI: University of Michigan Press.

- Munro, M. J. (2008). Foreign accent and speech intelligibility. In J.G. Hansen Edwards & M.L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193-218). Amsterdam: John Benjamins.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 45, 73-97.
- Munro, M., & Derwing, T.M. (1995b). Processing time, accent, and comprehensibility in the perception of foreign-accented speech. *Language and Speech*, 38, 289-306.
- Munro, M. J., Derwing, T. M., & Burgess, C. S. (2003). The detection of foreign accent in backwards speech. In M-J. Sole, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 535-538). Barcelona, Spain: Universitat Autònoma de Barcelona.
- Munro, M. J., Derwing, T. M., & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28, 111-131.
- Ortmeyer, C., & Boyle, J. P. (1985). The effect of accent differences on comprehension. *RELC Journal*, 16(2), 48-53
- Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics*, 17 (2), 164-81.
- Pennington, M., & Richards, J. (1986). Pronunciation revisited. *TESOL Quarterly*, 20(2), 207-25.
- Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics*, 26, 219-233.
- Riney, T., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly*, 39(3), 441-466.
- Schairer, K. (1992). Native speaker reaction to non-native speech. *Modern Language Journal*, 76, 309-319.
- Takagi, N. (1993). *Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view*. Unpublished doctoral dissertation, University of California, Irvine.
- Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America*, 111, 2887-2896.
- Smith, L., & Bisazza, J. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*, 32, 259-269.
- Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, 32, 539-552.
- Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 102-127). Ann Arbor, MI: University of Michigan Press.

Wilcox, G. K. (1978). The effect of accent on listening comprehension: A Singapore study. *English Language Teaching Journal*, 22(2), 118-127.

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69-84.

© Copyright rests with authors. Please cite TESL-EJ appropriately.

Appendix

Materials Used for Speech Stimuli

Consonant

1. 1. What do “tripe” and “bet” mean?
2. 2. The roof was broken after the worst storm one week ago.
3. 3. John told his parents the truth which gave them shocks.
4. 4. Before he left, he washed all the plates.

Vowel

5. 1. He put a red sheep on a red ship.
6. 2. The lady set the pepper on the paper.
7. 3. There is black soot on a black suit.
8. 4. Dirk’s duck was on the dock.

Word Stress

9. 1. They are talking about last year’s preSidential Election.
10. 2. ReCENTly, there has been an increase in car imPORTS in Vietnam.
11. 3. She’s a wonDERful MUsician.
12. 4. We will proBABly go TOgether.

Sentence Stress

13. 1. THERE WAS A terrible car accident ON THE corner.
14. 2. MY landlord collects THE rent payment ON THE FIRST OF THE month.
15. 3. PATIENCE is THE KEY TO joy; but HASTE is THE KEY TO sorrow.
16. 4. He ate A lettuce AND tomato salad FOR lunch.

Measure of speaker comprehensibility (*adapted from Kang, 2010*)

The utterance I just listened ...

was easy to understand ___/___/___/___/___/___/___/___ was hard to understand

Measure of speaker accentedness (*adapted from Kang, 2010*)

The utterance I just listened ...

has no accent ___/___/___/___/___/___/___/___ has a strong accent

Global Judgments (*adapted from Kang, 2010*)

Measure of speaker global comprehensibility

The speaker to whom I just listened ...

was easy to understand ___/___/___/___/___/___/___/___ was hard to understand

Measure of speaker global accentedness

The speaker to whom I just listened ...

has no accent ___/___/___/___/___/___/___/___/___ has a strong accent

© Copyright rests with authors. Please cite TESL-EJ appropriately.