Intelligibility of Hindi-Accented English: The Role of Duration and Intonation

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Laura Powell Crocker

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

Submitted under the faculty supervision of Associate Professor Benjamin Munson

September 2010

Acknowledgements

I am deeply thankful to Dr. Peter Watson and Dr. Mark DeRuiter for their generous contributions of time and energy to this project. I am grateful for their perceptive comments, their patience, and their obvious dedication to excellence in research and teaching. I would like to thank my advisor, Dr. Ben Munson, for once again providing inspiration and support from the beginning to the end of a lengthy project like this one, and for encouraging me to do more than I thought I was capable of. It is also in large part because of him that I decided to undertake this graduate program, so without him, this thesis would not have been. To all of the professors and clinical supervisors I encountered during my graduate studies, thank you for helping me to learn fascinating, confusing, and difficult things that I otherwise never would have known. And finally, I would like to thank all of the family and friends who have supported me over the past two years, especially my husband Andy. Without all of them, I might well have lost track of my purpose in wanting to become a speech-language pathologist and in wanting to write this thesis.

Abstract

In order to examine the contribution of prosodic variables to the intelligibility of foreign-accented speech, this study asked listeners to provide intelligibility ratings of Hindi-accented English sentences that were acoustically manipulated along two parameters: duration (of segments and pauses) and F0. The manipulated sentences, which reflected typical American English patterns of duration and pitch, were presented to native American English listeners in the presence of background noise.

The findings were that for Hindi-accented English, manipulating F0 has a greater effect on intelligibility than does manipulating duration. When duration and F0 were each manipulated separately, both manipulations led to reduced intelligibility. Manipulating F0 led to a greater detriment in speech intelligibility than manipulating duration. Furthermore, the lowest intelligibility performance of all resulted from sentences where both duration and F0 had been modified. Even though the prosodic variables were manipulated to be native-like and to therefore enhance intelligibility, these manipulations resulted worse performance on the part of native listeners. Possible reasons for these findings and directions for future research are discussed.

Table of Contents

List of Tables

List of Figures

1. Introduction

When people speak to one another, many different layers of information are conveyed simultaneously. In addition to relaying the basic semantic content of a spoken message, a talker's speech also provides clues to the listener about characteristics such as the talker's age, sex, gender, emotional state, and regional dialect. Another piece of information that can be conveyed to a listener, and the one that is investigated in this study, is whether the talker is a native speaker of the language in question. That is, whether or not the talker has a foreign accent.

Speech that is foreign-accented has been shown to result in reduced speech intelligibility on the part of native listeners (i.e., Munro & Derwing, 1995a, 1995b; van Wijngaarden, Steeneken & Houtgast 2002, Burda et al., 2003; Rogers et al., 2004; Major et al., 2005). There are multiple factors that can contribute to a talker's speech sounding foreign-accented. It is not only mismatches in the pronunciation of specific speech sounds, but also prosodic differences that constitute a foreign accent. These prosodic differences include productions of stress, intonation, rhythm, and duration that are not in line with native speakers' productions (Flege, 1995).

Indeed, past research suggests that it is often these prosodic differences, as opposed to segmental mismatches, that result in a foreign-sounding accent (Anderson-Hsieh et al., 1992; Munro, 1995; Munro & Derwing, 1995a). For example, Bannert (1995) investigated Punjabi- and Persian-accented Swedish. Short, spoken phrases were synthetically manipulated in terms of duration and intonation, and were presented to 20 native Swedish listeners in quiet. Listeners rated the degree of foreign accent they

perceived. Bannert concluded that manipulating intonational properties of the speech to reflect native Swedish patterns resulted in a lower degree of perceived foreign accent than was achieved by modifying durational properties. A study by Magen (1998) presented the English utterances of two male Spanish speakers to ten native English listeners in quiet. The utterances either remained unaltered or had the F0 contour acoustically manipulated to reflect the F0 contour of an American English speaker. In this way, Magen hoped to investigate lexical and phrasal stress. She found that listeners rated the F0-modified utterances as significantly "closer to native English" than the unaltered ones.

Boula de Mareuil and Vieru-Dimulescu (2006) examined the role of prosody in the foreign-accented speech intelligibility of Spanish and Italian. They imposed the duration and F0 of 14 Spanish utterances onto Italian utterances nearly identical in terms of orthography and pronunciation, and vice versa. In addition to identifying Spanish utterances with Spanish prosody as "Spanish-accented," Spanish listeners also identified Italian utterances with Spanish prosody as "Spanish-accented." Italian listeners showed the opposite trend, identifying utterances as "Italian-accented" only when those utterances contained Italian prosody. Both groups identified utterances as foreign-accented when they contained non-native prosody, even when the utterances were spoken in the listeners' native language. That is, non-native prosody contributed more to the impression of a foreign accent than segmental information. This suggests that prosodic features are a robust property used by listeners when detecting a foreign accent, and that speech can be perceived as native-sounding in the presence of segmental distortions, as long as native prosody is present.

Trofimovich and Baker (2006) presented short, declarative, low-pass filtered English phrases spoken by native Korean talkers to 10 native English listeners in quiet. One finding of the study was that the duration of pauses and speaking rate in the Korean talkers' speech, more so than melody-based characteristics (including stress and F0), were associated with the degree to which the speech was perceived as foreign-accented.

Gut (2007) conducted an investigation of foreign-accented German produced by 55 speakers from 24 different language backgrounds and foreign-accented English by 46 speakers from 17 different language backgrounds. A perception experiment used story retellings and read passages as stimuli. Seven native German listeners rated the degree of perceived foreign accent for the German L2 speakers, and 4 native British English listeners rated the degree of foreign accent for the English L2 speakers. Listener ratings were investigated for correlation with durational and intonational aspects of the foreign-accented speech. Gut found that speaking rate was the most important factor affecting the degree of perceived foreign accent.

A study by Munro et al. (2010) attempted to provide a clearer picture of the cues that may cause native listeners to identify speech as foreign-accented, and of the role specific prosodic variables play in such identifications. Munro et al. used content-masked speech, i.e., speech that was recorded and then played backwards, from native speakers of Mandarin, Cantonese, Czech, and English. Experiment 1 played content-masked prose passages from 10 Mandarin and 10 native English speakers to 10 native English listeners, and found that 90% of the native English speakers were correctly classified by over two thirds of listeners. All the Mandarin speakers were judged to have

Mandarin accents by more than half the listeners. Munro et al. concluded that this above-chance performance by listeners demonstrates that content-masked speech contains prosodic cues that allow listeners to distinguish native speech from foreign-accented speech. In Experiment II, the same speech materials were normalized for speaking rate and played to a new group of 27 native English listeners. The findings were that listeners again distinguished native from foreign-accented speech at above-chance levels, suggesting that speaking rate alone did not account for listeners' ability to distinguish the two groups of talkers. Experiment III investigated the effects of stimulus length and native language background on the ability of native listeners to detect a foreign accent in content-masked speech. Stimuli of one, three, and twelve words in length from speakers of Cantonese, Czech, and English were played to 36 native English listeners. Listeners' ability to distinguish native from foreign-accented speech was significant across different accents and different utterance lengths. In addition, sensitivity scores were comparable for two phonologically and prosodically dissimilar languages (Cantonese and Czech). Scores did not differ significantly between longer and shorter stimuli, although there was a trend toward lower sensitivity scores for shorter stimuli. Finally, Experiment IV evaluated the effects of random splicing (in which short sections of speech are re-ordered) and monotonization (in which speaker's productions are digitally altered to have a single, gender-appropriate pitch level) on content-masked Cantonese-accented and native English sentences. Despite these disruptions to the temporal properties and pitch contours of the speech, listeners in Experiment IV were once again able to distinguish native from foreign-accented speech at levels significantly

above chance. Scores for the temporally manipulated and pitch manipulated speech did not differ significantly. However, listeners' sensitivity was significantly lower in both these conditions than for the unmanipulated speech used in the previous three experiments. This suggests that both temporal and pitch characteristics do play a role in listeners' ability to differentiate native from foreign-accented speech.

One caveat regarding the studies described above, like others dealing with prosody and foreign-accent speech, is that they examine comprehensibility rather than intelligibility per se. In such studies, listeners are asked to rate how easy the accented speech is to understand or how heavy the foreign accent is, without being asked to transcribe what they heard or otherwise provide strict intelligibility responses. Although some research (Trofimovich & Baker, 2006) has found that listeners' comprehensibility ratings are correlated with more native-like productions of prosody on the part of accented speakers, other research has shown that the comprehensibility and intelligibility of accented speech differ in a statistically significant way. For example, Derwing and Munro (1997) investigated the effects of alterations in speaking rate on the foreign-accented English of 12 speakers each of Cantonese, Japanese, Polish, and Spanish. Twenty-six native English listeners rated perceived comprehensibility and also provided transcripts of the utterances they heard. Perceived comprehensibility was affected by speaking rate, whereas intelligibility remained unaffected. This is just one study demonstrating that measures of comprehensibility and intelligibility are not the same and should not be treated as such when examining the literature.

Studies that investigate the intelligibility proper of foreign-accented speech with regard to the role of prosodic variables are rare. Bent and Bradlow (2003) examined speakers of Chinese, Korean, and English, and found that the overall intelligibility of foreign-accented speech depends on the native language background of both the speaker and the listener. Thus, the studies presented below offer many insights but cannot necessarily be generalized. In Huckvale (2006), for example, one English speaker's English-accented Japanese was investigated. The segmental quality, pitch, and timing of single words were manipulated to match native Japanese pronunciation norms. Eight native Japanese listeners provided intelligibility ratings by writing down what they heard. Huckvale's main finding was that pitch was the only manipulated variable significantly affecting intelligibility, suggesting that this prosodic elements plays a role in the intelligibility of foreign-accented speech.

Kamiyama (2004) investigated the role of intonation and duration in the perceived degree of foreign accent in Japanese-accented French. Kamiyama used both synthesized speech and manipulated natural speech to examine these prosodic variables. Talkers were eleven native Japanese speakers and four native French speakers, each of whom read short phrases in French. Seventeen native French listeners judged the degree of foreign accent in the utterances. Kamiyama's primary finding was that manipulating intonation affected the degree of perceived foreign accent more than manipulating durational properties like speaking rate and pauses.

A Norwegian doctoral dissertation by Holm (2008) represents one of the more detailed investigations of the effect of prosodic variables on the intelligibility of foreign-

accented speech available at this time. Holm used two speakers from each of seven different language backgrounds (English, French, German, Mandarin, Russian, Persian and Tamil) who were highly educated and communicated regularly in Norwegian. Speakers read a list of 60 phonetically balanced sentences in Norwegian, which were then acoustically modified in terms of segmental duration and intonation contour to match the productions of one native Norwegian speaker. Four different conditions resulted: original sentences, duration modified, intonation modified, and duration-intonation modified. The sentences were presented to 103 native Norwegian listeners in the presence of pink noise. Holm's findings were that duration manipulation significantly enhanced intelligibility compared to the original sentences, but that this was due to effects only in the French, Tamil and Persian groups. Intonation manipulation also enhanced the intelligibility of the foreign-accented Norwegian speech significantly, but when the effect was investigated within language backgrounds, significant effects were found only for the English and German groups. Duration-intonation manipulations significantly enhanced intelligibility for two groups, German and Russian, while they reduced intelligibility for another two groups, Mandarin and Persian. For the German, French, and Russian groups, the duration-intonation manipulated stimuli had higher intelligibility scores than the intonation manipulated stimuli, while the English, Tamil, Mandarin, and Persian groups showed the opposite tendencies. Holm concluded that intonation manipulated stimuli were more intelligible than duration-intonation manipulated stimuli, despite the hypothesis that the combined manipulation of both prosodic features would make the utterances more intelligible than the intonation manipulation alone. Holm also found that

intonation was the most important aspect enhancing intelligibility for the English, German, Tamil and Russian groups, while duration was the most important aspect for the French group. For the Mandarin and Persian groups, there was no difference among the intelligibility of the manipulations. This investigation demonstrates that prosodic variables like duration and intonation affect the intelligibility of foreign-accented speech in different ways and to different degrees depending upon the native language background of the talker.

Rogers et al. (2004) investigated the intelligibility of Chinese-accented English in different levels of background noise. They recorded four native Chinese speakers and two native American English speakers reading a list of 50 sentences. These stimuli were presented to 48 native American English speaking listeners in quiet and in the presence of multi-talker babble at three different S/N ratios: + 10 dB, 0 dB, and -5 dB. Listeners provided intelligibility ratings by writing down what they heard. Rogers et al. found, as in many previous studies, that the native English speakers were the easiest to understand. Furthermore, it was found that adding background noise to the speech signal reduced the intelligibility of accented speech significantly more than it reduced the intelligibility of native speech. This suggests that non-native speech, even that of high-proficiency speakers like those used in the Rogers et al. study, is less robust than native speech when it is heard in compromised listening conditions.

Several additional studies have found that the intelligibility and/or comprehensibility of foreign-accented speech presented in noise is lower than that of identical speech materials presented in quiet (Nábelek & Donahue, 1984; Takata &

Nábelek, 1990; Munro, 1998; Cutler et al., 2004). Yet, overall, relatively few studies have investigated the intelligibility of foreign-accented speech in noise. This is significant because very little interaction between native and non-native speakers takes places in a quiet, controlled environment. Rather, communication must normally occur in settings where a certain amount of background noise is present, such as in classrooms, places of work, and any number of other public spaces. Therefore, research looking at foreign-accented speech in noise is more ecologically valid and more studies along these lines are needed.

Having described the state of the research dealing with the intelligibility of foreign-accented English, it should be noted that there are difficulties that come with having a foreign accent. For example, several studies have shown that native listeners may hold negative attitudes and perceptions about talkers whose speech is foreign-accented (Bresnahan et al., 2002; Lindemann, 2005; Boula de Mareuil & Vieru-Dimulescu, 2006, Gluszek & Dovidio, 2010). Job discrimination, communication difficulties, and social alienation are some of the consequences that can result from such attitudes toward non-native speakers (Lindemann, 2005: Gluszek & Dovidio, 2010). Accented speech is most certainly not disordered speech; however, as noted above, the consequences of it in society can sometimes be similar to those of disordered speech. In part because of this, it is often the goal of speech-language pathologists working in the area of accent modification and of ESL/EFL teachers to help non-native speakers increase their speech intelligibility. The aim of such professionals is not to erase the unique identity of the many individuals who speak a language with a foreign accent.

Rather, it is to give people, when needed, the tools to communicate more flexibly and successfully within the many different contexts they encounter on a daily basis.

Practical approaches in the literature for how to achieve such goals are somewhat lacking. For example, at an accent modification workshop for international students at a large university, attended by the author, activities focused on "tone," which encompassed some aspects of prosody. Good tone was defined as the combination of a wide pitch range, a loud speaking volume, and the use of facial expressions and gestures while communicating. The instructions and practice assignments for implementing these elements in one's own speech, however, were somewhat vague and inconsistent, and did not appear to increase the naturalness or intelligibility of the accented speakers' utterances. Rather, these suggestions seemed designed to allow speakers to convey a greater sense of confidence to their listeners. Methods such as these point to the fact that, despite the finding that prosodic variables typically contribute more to a foreign-sounding accent than segmental features, and despite many claims in the foreign language pedagogy literature that teaching prosody is essential (Hahn, 2004), relatively little empirical research regarding the teaching of prosody to non-native speakers exists. The current study does not directly investigate practical techniques that could be used by speech-language pathologists or ESL/EFL teachers. It does, however, attempt to contribute to the body of evidence that has examined the role of prosodic features in the intelligibility of foreign-accented speech. Its findings could then be extended to pedagogically relevant contexts, with the aim of allowing native and non-native speakers to achieve better communication based on practical approaches.

1.1 The present study

The purpose of the present study is twofold. First, given the divergent findings of previous research, this study aims to shed light on which aspects of prosody in accented speech affect intelligibility, and in what ways. As in the studies described above, the focus of the current study is duration and intonation. Second, with regard to these prosodic aspects of duration and intonation, this study investigates a native language background that is virtually unrepresented in the foreign-accent intelligibility literature by looking at Hindi-accented English.

The present study focuses on Hindi-accented English for several reasons. One is that the intelligibility of Hindi-accented English has not been well investigated to date. A study by Black et al. (1965) examined the speech of 72 university students whose native language was Hindi, Japanese, or Spanish. Speakers read short passages of prose, which were then recorded and presented to listeners in the presence of white noise. Black et al. found that, regardless of the level of English proficiency of the foreign speakers, the Hindi accent was the most difficult for native English listeners to understand. Apart from the Black et al. (1965) paper, there are no other published studies that investigate the intelligibility of Hindi-accented English. Another reason this study looks at Hindi-accented English is that the prosodic properties of Hindi are very distinct from those of English. For example, there are some Hindi words in which primary stress is not rigidly fixed on a particular syllable, but may fall on either the first or second syllable of the word depending upon the style and speech rate of the speaker (Shukla, 1990). Anecdotal evidence suggests that native Hindi speakers using English often exhibit stress-related

mismatches that compromise intelligibility on the part of native English speaking listeners. Finally, the interaction between Hindi speakers and English speakers is long-standing and rich, playing a vital role in the global marketplace as well as in countless scientific, academic and medical communities around the world. For this reason, an examination of Hindi-accented English has high ecological validity. There are approximately 182 million people speaking Hindi as a first language today (Lewis, 2009). Given the fact that English is indisputably a global language, there is a high likelihood that many native Hindi speakers will continue to learn English in the coming years. This language background is thus well suited for a study like the present one, whose results may potentially benefit both speakers of Hindi who acquire English for academic, social, or professional reasons, and native English speakers whose daily activities find them interacting with a diverse group of people speaking English with a foreign accent.

In the interest of examining the prosodic features of duration and intonation in relation to Hindi-accented English, the current study uses as inspiration a paper by Tajima, Port, and Dalby (1997). Tajima et al. manipulated the speech of a native Chinese speaker producing eleven short English phrases, such that the duration of individual sound segments was identical to that of a native English speaker. Similarly, native English phrases were manipulated to reflect the durational patterns of Chinese-accented English. Each phrase was grouped with three additional distracter phrases that were similar in terms of phonetic content and number of syllables, creating a four-alternative forced-choice identification task. The stimuli were presented to the listeners in the presence of recorded cafeteria noise at three different S/N ratios: +5 dB, 0 dB, and -5 dB.

Thirty-six native English listeners provided intelligibility responses. Tajima et al. found that, regardless of the S/N ratio, the temporally modified Chinese-accented English was significantly more intelligible to native English-speaking listeners than unmodified Chinese-accented phrases. Similarly, English phrases modified to match Chinese-accented timing were significantly less intelligible than native English productions. The current study is a follow up to Tajima et al., examining whether the findings hold when a typologically different accent is examined. It builds upon the methods of Tajima et al. to see whether modifying the duration of accented speech affects the intelligibility of Hindi-accented English in the same way as Chinese-accented English. In addition, this study expands on the work of Tajima et al. by introducing intonation as a variable for study. The prosodic elements of duration and intonation, along with their impact on speech intelligibility, are thus investigated in terms of the relative contribution of each. Finally, the current study uses an open-set response rather than the four-choice forced-alternatives used by Tajima et al. An open-set methodology more closely resembles real world communicative exchanges than does a forced-choice test, and is therefore more ecologically valid.

In response to findings and gaps in past research, the present study investigates intelligibility proper of an underrepresented language background, namely Hindi-accented English. It focuses on the prosodic variables of duration and intonation, because past research has shown these elements to be paramount in regards to what constitutes a foreign accent and what affects intelligibility. It investigates intelligibility in the presence of background noise, as this is more ecologically valid than investigating

it in quiet. This study imposes American English prosody in Hindi-accented English utterances because of the Tajima et al. finding that doing so with Chinese-accented utterances aids intelligibility on the part of native listeners. It makes intuitive sense that prosody that is more native-like would be easier for native speakers to understand. Imposing the American English prosody is also influenced by the findings of Boula de Mareuil and Vieru-Dimulescu (2006), in which patterns of duration and F0 alone allowed listeners to identify a speech sample as being from their own native language. While perhaps not generalizable, their investigation does suggest that the prosody of one's native language is markedly different from non-native prosody and is identifiable as such. The working hypothesis is therefore that replacing the prosodic variables of duration and intonation in Hindi-accented English with patterns of a native American English speaker will result in increased speech intelligibility by native American English listeners.

2. Listeners

Thirty-six listeners participated in this study, all of whom were native speakers of American English. Almost all of them spoke Midwestern English exclusively, while a few had either grown up with or been exposed to other American English dialects before living for several years in the Midwest. The participants included 24 women and 12 men. They ranged in age from 20 to 49 years, with a mean age of 27. Before completing the experiment, participants filled out a questionnaire verifying that they had no history of speech, language, or hearing disorders. The questionnaire also invited participants to share whether or not they spoke or had studied languages other than English. This

question was included in case it proved significant in post-hoc analyses. The hypothesis was that participants with foreign language experience might be more familiar with speakers of languages other than English, and may have had more contact, either socially, academically, or through travel, with such speakers. This increased familiarity with non-native English speakers on the part of participants with foreign language experience, it was hypothesized, might prove to be a significant variable affecting intelligibility performance. Exactly half of the participants reported speaking or having studied another language, while the other half had no experience with languages other than English. None of the participants who had experience with foreign languages reported being familiar with Hindi.

Participants were recruited through fliers advertising the study placed around a university campus, as well as in other community locations. In addition, participants in past studies conducted in the speech research laboratory were contacted through an email list-serve. These techniques may have resulted in a wider cross-section of participant ages and backgrounds than would have been attainable with on-campus recruitment alone. None of the recruitment materials specified that the experiment would involve Hindi-accented English. Instead, the study was advertised as an experiment where native English speaking listeners were needed to judge which aspects of speech facilitate better intelligibility of a foreign accent. Participants were compensated $10.00 for their completion of the experiment.

3. Methods

3.1 Speakers and speech material

The Hindi-accented speech material used in this study was obtained by consent from the BuckeyeGTA Corpus of Accented English created by Jocelyn Hardman at The Ohio State University. Eighteen native Hindi speakers were recorded. All speakers were international graduate students who had achieved a level of oral English proficiency consistent with "graduate TA certification," which is defined as a SPEAK test score of 230 or higher out of 300 or as an "unconditional pass" on the Mock Teaching Test (MTT), a teaching performance test developed in Columbus, OH. Recording was done in a sound-attenuated booth, using a Shure SM10-A head-mounted microphone. The microphone was connected to a Symetrix SX302 Dual Microphone Pre-amplifier (gain ~ 50 dB). The talkers' utterances were digitally recorded using an Audigy 2 sound blaster card. Cool Edit (Syntrillium Co.) software was used to save the sound files (Microsoft PCM.wav) on the computer hard drive, using a 22,050 kHz sampling rate with 16-bit resolution. These were then saved to a secure server, and finally to an external hard drive.

Speech material from two female and two male speakers was selected for use in the present experiment. The precise age and background of the native Hindi speakers is not known; however, they were each informally judged to be approximately 20 to 45 years old. They were chosen because their speech was judged to represent the two least intelligible females and the two least intelligible males out of all the speakers present in the Hindi-accented corpus.

The abovementioned corpus also included recordings of several native American English speakers. From these speakers, one female and one male were chosen to provide the reference prosody for the acoustic manipulation of the Hindi-accented speech. Similarly, the exact age and background of these native English speakers is unknown. They were selected because, like the native Hindi speakers, they were judged to be approximately 20 to 45 years old. In addition, they were selected because their speech most closely resembled a combination of Standard American English and Midwestern English. This was desirable because this dialect background was anticipated to be most similar to that of the listeners who would participate in the study.

It should be noted that no intelligibility data exist for these two American English speakers. Having precise intelligibility scores at the outset of this experiment would have been preferable, given the role of these speakers as suppliers of the reference prosody for Hindi-accented utterances later in the intelligibility task. However, since virtually all past research has found that the speech of native speakers is more intelligible and/or comprehensible than accented speech for native listeners (Munro & Derwing, 1995a, 1995b; van Wijngaarden, Steeneken & Houtgast 2002, Burda et al., 2003; Rogers et al., 2004; Major et al., 2005), it can at least be surmised that the speech of these American English talkers would be more intelligible than the Hindi-accented speech for the listeners in this experiment. Furthermore, the American speakers had no apparent deviations from the norm in terms of pronunciation, speaking rate, speaking volume, stress patterns, or any other speech parameter that might affect intelligibility for native listeners.

The speech material provided by both the native Hindi and the native English speakers consisted of four lists (lists 7, 8, 9, and 10) from the Bamford-Kowal-Bench sentence protocol, revised for American English (BKB-R). Each list contained 16 sentences, for a total of 64 unique sentences spoken by each speaker. The sentences were originally developed in Britain as an audiometric tool for the evaluation of hard-of-hearing children (Bench, Kowal, & Bamford, 1979) and were later revised by the Cochlear Corporation for use with American children. They were constructed, based on spontaneous speech samples collected from children, to contain simple, familiar vocabulary and uncomplicated sentence structures. Thus, the sentences range in length from four to seven words (and four to nine syllables) and contain declarative statements exclusively. Lists 7, 8, 9, and 10 in particular were selected because of their equivalent intelligibility scores for normal-hearing children as reported in Bamford and Wilson (1979). These lists can be found in the Appendix. The BKB-R sentences, though not administered here in a standardized way for audiometric purposes, represent a commonly available speech-in-noise evaluation tool and have been used in previous research examining the intelligibility of foreign-accented English (Bent & Bradlow, 2003).

3.2 Acoustic manipulations

Hindi-accented sentences in the form of .wav files were acoustically manipulated along two parameters: duration (of segments and pauses) and F0 contour. The manipulations reflected typical American English patterns of duration and F0, as derived directly from the speech of the two native English speakers described above. This resulted in 4 different conditions under which listeners heard sentences and provided

intelligibility responses: unmanipulated Hindi-accented sentences, Hindi-accented sentences with English duration, Hindi-accented sentences with English F0, and Hindi-accented sentences with both English duration and F0. The female native English speaker provided the reference prosody used in all manipulations for both female Hindi-accented speakers; the male native English speaker likewise provided the reference prosody for the male Hindi-accented speakers. Prosody swapping was performed only on identical sentences spoken by same-sex native/non-native speaker pairs, and never across different BKB-R sentences or between native/native or non-native/non-native speaker pairs.

The prosody-swapping process was completed using the Praat Digital Signal Processing System, version 5.1.25 (Boersma and Weenick, 2009). The procedure was based on a technique for replacing prosodic features of non-native speakers' utterances with those of native speakers, developed by Yoon (2006). Yoon's process implements speech-waveform operations in the time domain, formalized by Moulines and Charpentier (1990) as the pitch synchronous overlap-and-add (PSOLA) technique, via a Praat script that targets duration first and F0 second. PSOLA was developed as a tool for the modification of duration and pitch of natural speech, and past research has found it to be a practical method for speech modification (Moulines & Laroche, 1995; Kortekaas & Kohlrausch, 1999). Yoon's Praat script requires that there be the same number of segments in the non-native utterance is in the native one. The script then operates by aligning the segmental duration of the two utterances. The result is that some segments are elongated and others are shortened. Once the duration alignment has been performed,

the script substitutes the F0 contour of the non-native utterance with that of the native one. The non-native utterance now has the same segmental duration and F0 contour as the native utterance. The process is illustrated in Figures 1 and 2 below, which are taken directly from Yoon (2006):

Figure 1

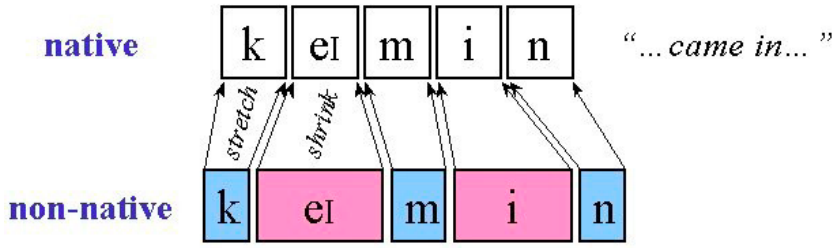

Figure 1. Illustration of the process for imposing native segmental duration on a non-native utterance

Figure 2

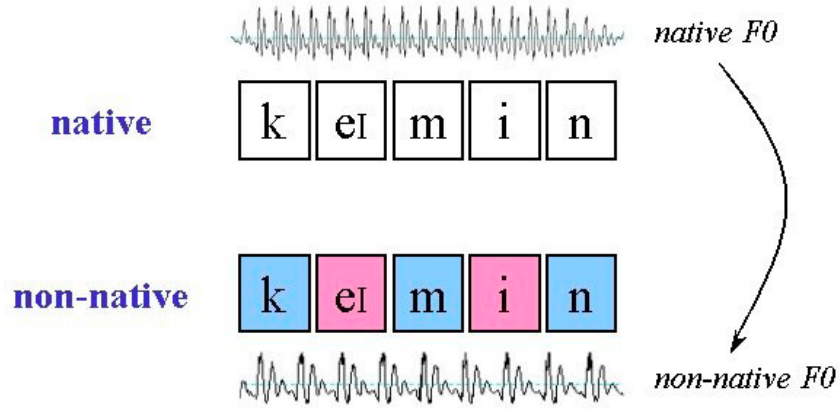

Figure 2. Illustration of the process for imposing native F0 contour on a non-native utterance

For this experiment, the prosody-swapping technique described above was implemented in separate steps, due to the need to manipulate the Hindi-accented stimuli for duration and F0 separately in addition to creating sentences with both duration and F0 manipulated. When swapping only the F0 contours, the matching up of individual segments was still performed as a first step to ensure better results. Yoon's script contains a third possible prosody-swapping variable: the intensity contour of an utterance. However, as intensity is not under investigation in this study, it was not manipulated. It should also be noted that, on a test manipulation where the intensity contour was swapped, the result was an extremely distorted and nearly unintelligible utterance.

Below is a series of figures detailing the acoustic manipulation process. These examples represent a case where the prosody-swapping algorithm resulted in particularly successful (i.e., natural and distortion-free) utterances. The figures show the American English male talker's production of, "The floor looked clean," as well as a Hindi-accented male talker's production of the same sentence in all four conditions used in the experiment.

Figure 3

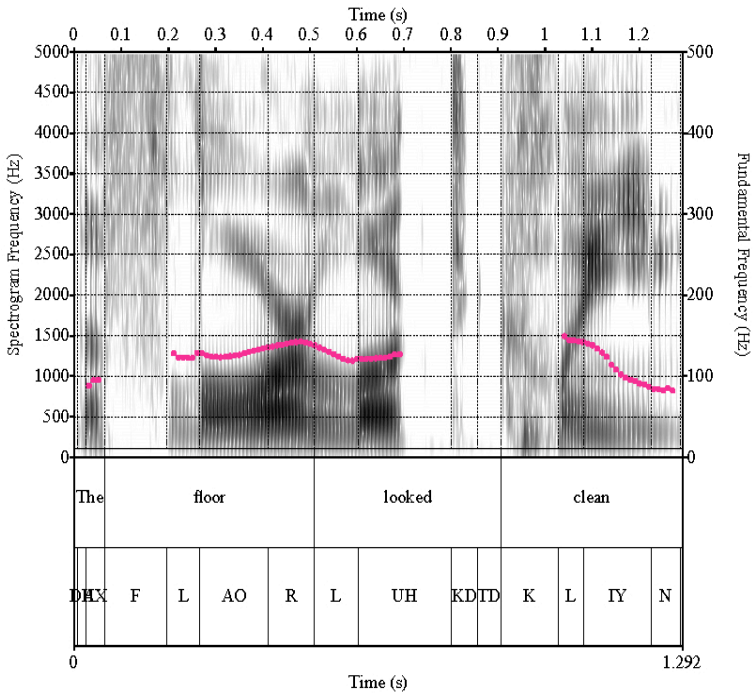

Figure 3. Spectrogram showing the segmental duration and F0 contour of the original American English reference utterance spoken by a male talker.

Figure 4

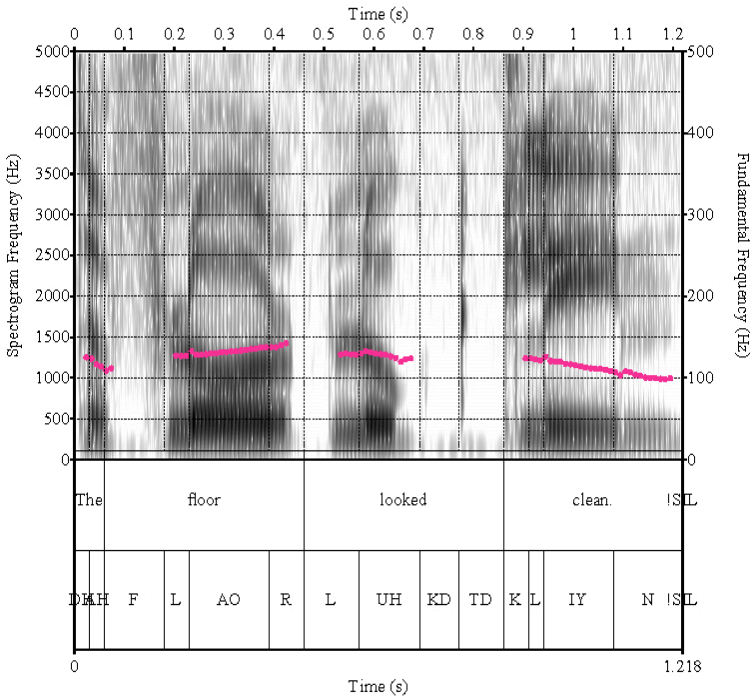

Figure 4. Spectrogram showing the segmental duration and F0 contour of an unmanipulated Hindi-accented utterance spoken by a male talker.

Figure 5

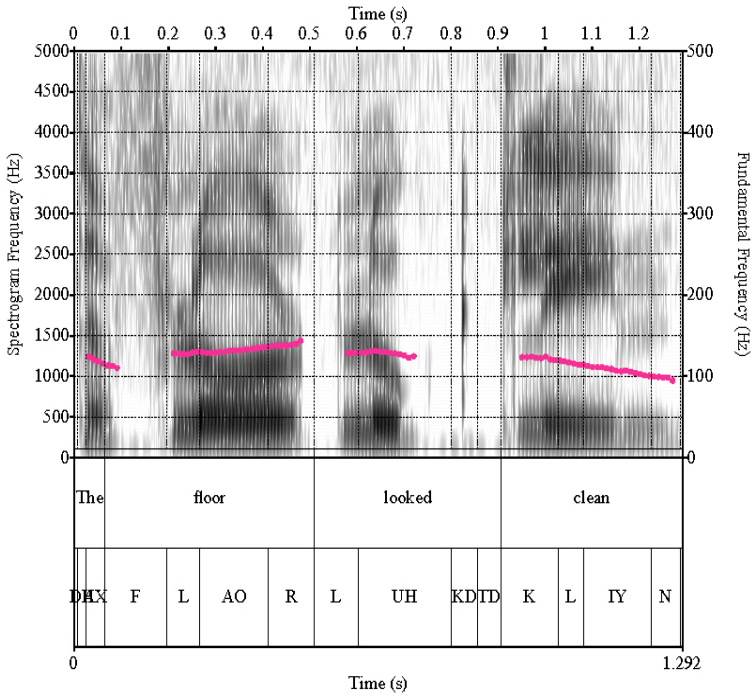

Figure 5. Spectrogram showing the successful duration manipulation of a Hindi-accented utterance spoken by a male talker.

Figure 6

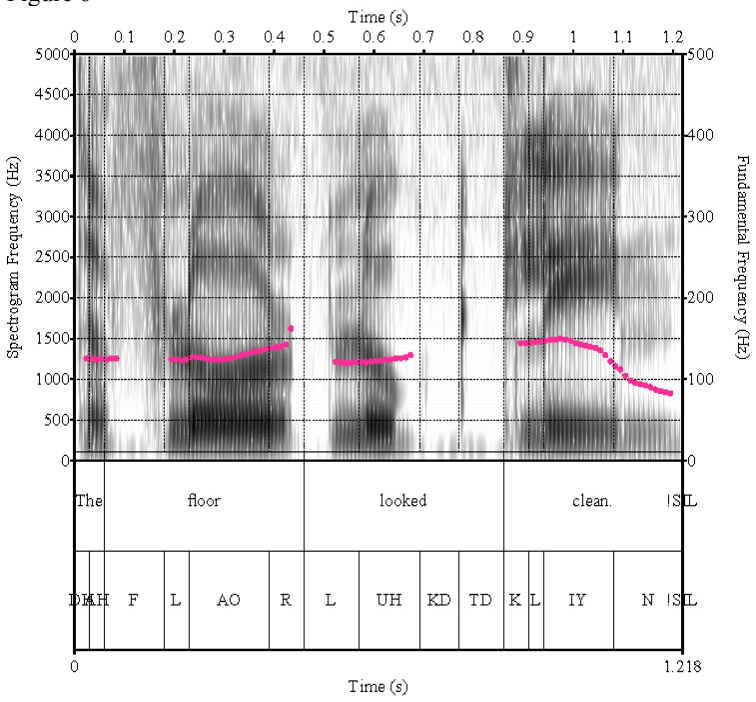

Figure 6. Spectrogram showing the successful F0 manipulation of a Hindi-accented utterance spoken by a male talker.

Figure 7

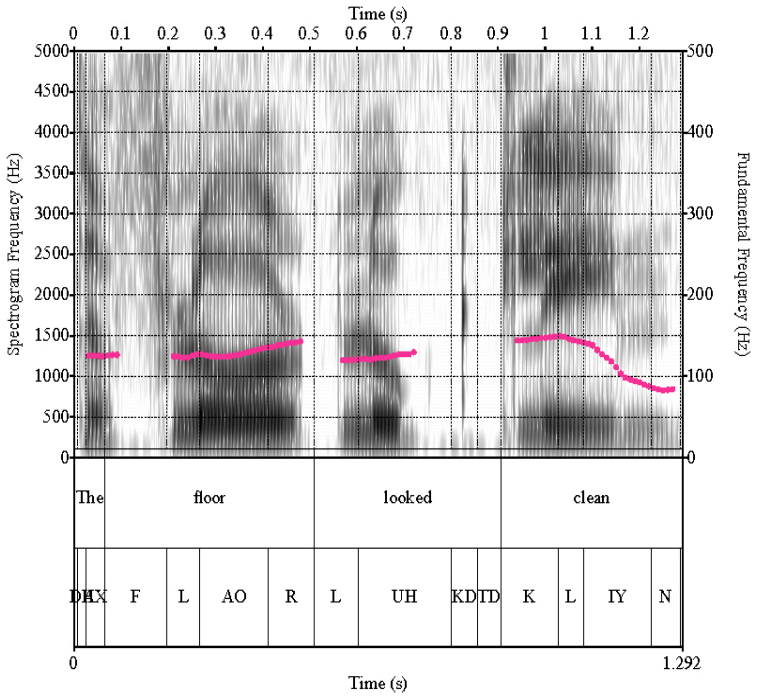

Figure 7. Spectrogram showing the successful duration-and-F0 manipulation of a Hindi-accented utterance spoken by a male talker.

After the acoustic manipulations were completed, any resulting sentences that sounded obviously distorted or unnatural were discarded. There were 83 such sentences that were unusable and were not included in the experiment. In most cases, the source of the distortion was the duration-swapping algorithm inappropriately parsing pauses following word boundaries as part of the duration of the preceding word. This resulted in certain words either being unnaturally elongated and therefore significantly misaligned with the intonation contour of the sentence, or being truncated or omitted from the resulting sentence completely.

3.3 Experiment design

The listening experiment was written and carried out using the *E-prime* experiment management software (Schneider, Eschman, & Zuccolotto, 2002). It contained a total of 64 sentences. There were four different but equivalent versions of the experiment, and all 36 listeners were randomly assigned an experiment version. Each version was designed to be completely balanced across talkers, sentences, and type of acoustic manipulation. For example, each experiment was designed to contain all 64 sentences present in lists 7-10 of the BKB-R protocol. The speaker and the type of acoustic manipulation varied across experiment versions. Thus, every listener heard all 64 sentences, but from different speakers and under different manipulated conditions depending upon the experiment version. However, because some of the manipulated sound files had to be discarded, as described above, adjustments were made. In the end, each experiment version contained two to three sentences that participants heard twice. When sentences were repeated, it was always with a different speaker and under a different manipulated condition. Sentences were presented to listeners in a completely randomized order.

All sentences were normalized for loudness such that the peak amplitude of each sentence was the same. The sentences were presented binaurally at 70 dB HL. Sentences were presented to listeners in background noise, with a S/N ratio of 0 dB. Wide-band white noise was used as background noise.

Participants completed the experiment in a sound-treated booth. They wore Sennheiser 280 HD headphones for the duration of the experiment. Participants were

seated in front of a 17" computer monitor and were presented with instructions for the experiment in 36-point courier font, which they read off the computer screen. The onscreen instructions informed participants that they would be hearing sentence spoken by people with a foreign accent. Participants were instructed to listen carefully, as each sentence would be presented only once. Participants pressed any key to begin the experiment. Before presentation of a sentence in background noise, participants saw a screen with the words "listen carefully" in 36-point courier font. Upon hearing a sentence in background noise, they were prompted via a pop-up box onscreen to type back what they had heard using a standard Qwerty computer keyboard. Participants had as much time to type as they liked, but were encouraged not to deliberate too long and to move on to the next sentence if they had no idea what was said. Participants were in control of the timing of the experiment, as the next sentence was not presented until they pressed a key indicating they were ready. The experiment took participants anywhere from 15 to 30 minutes. The short duration of the experiment prevented fatigue on the part of listeners, and allowed for their performance to remain relatively consistent throughout the experiment rather than potentially waning at the end due to inattention.

3.4 Scoring

The BKB-R sentences contain three to four keywords per sentence, for a total of fifty keywords per list. When administering the BKB-R protocol in a standardized way as an audiometric tool, only these keywords are scored. Past research on foreign-accented speech using the BKB-R sentences as stimuli has also adhered to this convention (Bent & Bradlow, 2003). For the present study, a different procedure was

used. Rather than looking only at the keywords, the entire sentence, including articles, was scored. The decision to score the entire sentence was based on the fact that while some of the sets of BKB-R keywords convey the basic meaning of the sentence, others do not. For example, in the sentence, "She looked in her mirror," the keywords "she," "looked," "in," and "mirror" seem to appropriately capture the essence of the utterance. However, the same cannot necessarily be said for a sentence like, "The coat is on a chair," in which "coat," "is," and "chair" are the only keywords. Regardless of how well the keywords represent the sentence, in many cases altering or omitting the non-key words changes the meaning of the utterance. This study examines speech intelligibility where two widely spoken global languages are concerned, and is germane to auditory-only communicative interactions that take place over long distances. Thus, requiring a higher level of intelligibility from the listening participants, one in which the full, intended meaning of the sentence is extracted, seems justified.

The 2,304 intelligibility responses provided by the listeners were automatically stored in an *E-prime* database. Responses were then scored with the help of a data spreadsheet and thoroughly inspected by hand to ensure accuracy. Obvious misspellings (which were fairly frequent) and homophones (which were rare) were counted as correct. Sentences for which the listener's only response was "the" or "a" (which were rare) were counted as incorrect, because of the high likelihood that this response represented a blind guess or had simply been replicated based on previously heard sentences. The total number of possible correct words differed across the 64 BKB-R sentences, so raw intelligibility scores were converted to percent correct scores. These percent correct

scores were used for data analysis and are represented in the remainder of this paper.

4. Results

Intelligibility data were submitted to single-factor, within-subjects analyses of variance (ANOVAs) examining intelligibility scores. A secondary correlation analysis was conducted to examine the association between intelligibility scores and various descriptive characteristics of the participants, including age, sex, and previous experience with foreign languages. All analyses were performed using SPSS statistical software using a 95% confidence interval.

A within-subjects ANOVA showed a significant main effect of sentence type on intelligibility (F [3, 105] = 69.542, $p < 0.001$, partial $\eta^2 = 0.67$). The most intelligible sentences were the unmanipulated ones (M = 78.4%, SE = 2.4%), followed by those with manipulated duration (M = 64.6%, SE = 2.5%). Sentences with manipulated F0 (M = 58.4%, SE = 2.2%) and those where both duration and F0 had been manipulated (M = 55.3%, SE = 2.3%) were the least intelligible. These results can be observed in Figure 8 below:

Figure 8

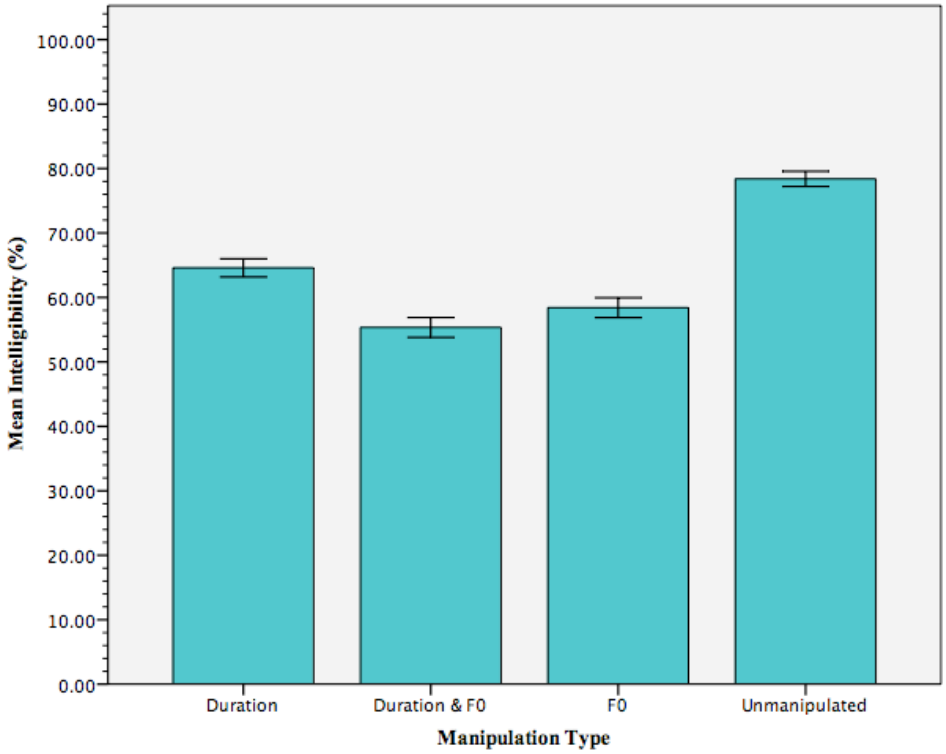

Figure 8. Mean intelligibility of Hindi-accented sentences by manipulation type across all listeners, shown as percent correct scores. Error bars reflect a standard error of +/- 1 *SE*.

Post-hoc pairwise comparisons with a Bonferroni-corrected α level of 0.05 examined the relationship between sentence types. This analysis revealed that the only pairwise comparison not achieving statistical significance was between sentences where F0 only was manipulated (*M* = 58.4%, *SE* = 2.2%) and those where both duration and F0 were manipulated (*M* = 55.3%, *SE* = 2.3%). Significant differences in intelligibility were found for all other pairwise comparisons. These results are summarized in Table 1 below:

Table 1
Pairwise comparisons of sentence types

Sentence type	Comparison	Mean	Standard error	Significance
Duration/F0	Duration	-0.092	0.017	<0.001
	F0	*-0.031*	*0.015*	*0.256**
	Unmanipulated	-0.230	0.017	<0.001
Duration	Duration/F0	0.092	0.017	<0.001
	F0	0.062	0.018	0.009
	Unmanipulated	-0.138	0.017	<0.001
F0	*Duration/F0*	*0.031*	*0.015*	*0.256**
	Duration	-0.062	0.018	0.009
	Unmanipulated	-0.200	0.020	<0.001
Unmanipulated	Duration/F0	0.230	0.017	<0.001
	Duration	0.138	0.017	<0.001
	F0	0.200	0.020	<0.001

**Not statistically significant*

A within-subjects ANOVA was conducted and showed that listener sex and prior experience with other languages did not have a significant effect on intelligibility performance, nor did either of these variables interact with sentence type. It is not entirely surprising that experience with foreign languages was not a significant factor affecting intelligibility performance. Research remains inconclusive as to whether familiarity with a particular foreign language or accent results in better comprehensibility or intelligibility on the part of native listeners (Adank, 2009), and none of the listeners reported familiarity with Hindi or Hindi-accented English in the first place. Correlation analyses revealed a significant relationship between listener age and intelligibility

performance for two of the sentence types. Age was correlated negatively with performance for sentences where both duration and F0 had been manipulated (r = -0.36, N = 36, p (2-tailed) = 0.031). For unmanipulated sentences, the negative correlation between age and intelligibility performance approached, but did not fully achieve, statistical significance (r = -0.328, N = 36, p (2-tailed) = 0.051). This last finding is perhaps not unexpected. Past research has demonstrated that listeners' perception of non-accented speech in noise declines with age, and that this decline is not accounted for solely by the presence of peripheral hearing loss (Cooper & Gates, 1991; Walton, Simon, & Frisina, 2002; Helfer & Freyman, 2008; Wong et al., 2009). The same is true for listeners' perception of foreign-accented speech in noise (Burda et al., 2003; Ferguson et al., 2010).

The mean intelligibility of each of the four Hindi-accented talkers was calculated, and is illustrated in Figure 9 below. Three of the talkers had a nearly identical mean intelligibility, with the fourth (a male speaker) being slightly less intelligible across all sentence types than the other talkers. This did not result in a statistically significant difference between talkers, nor did it account for the difference in intelligibility across sentence types. The mean task performance of each of the 36 listeners is illustrated below in Figure 10. This figure shows that there was one listener who performed significantly below the others in this experiment. This listener reported no hearing impairments and no history of speech or language disorders, and also reported no difficulty in hearing the stimuli at the time of the experiment. However, this listener was one of several in the over-40 age bracket. This may account for some of the negative

correlation between age and performance in this experiment, described above, and is perhaps partly explained by the general relationship between age and speech-in-noise performance, also described above.

Figure 9

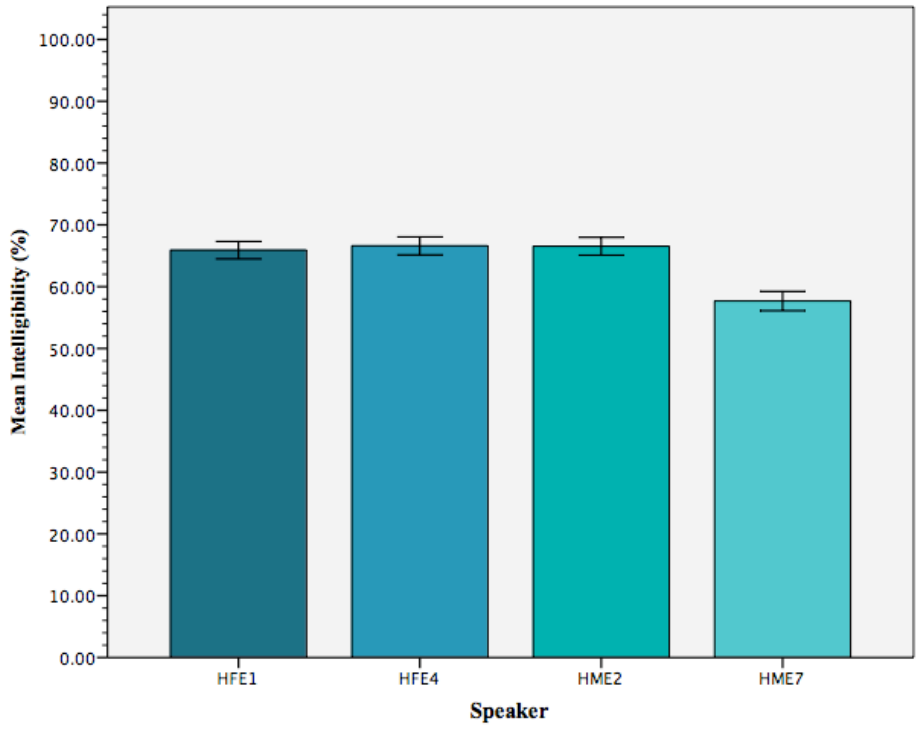

Figure 9. Mean intelligibility of each of the Hindi-accented talkers across all sentence types, shown as percent correct scores. Talkers HFE1 and HFE4 are female; talkers HME2 and HME7 are male. Error bars represent +/- 1 *SE*.

Figure 10

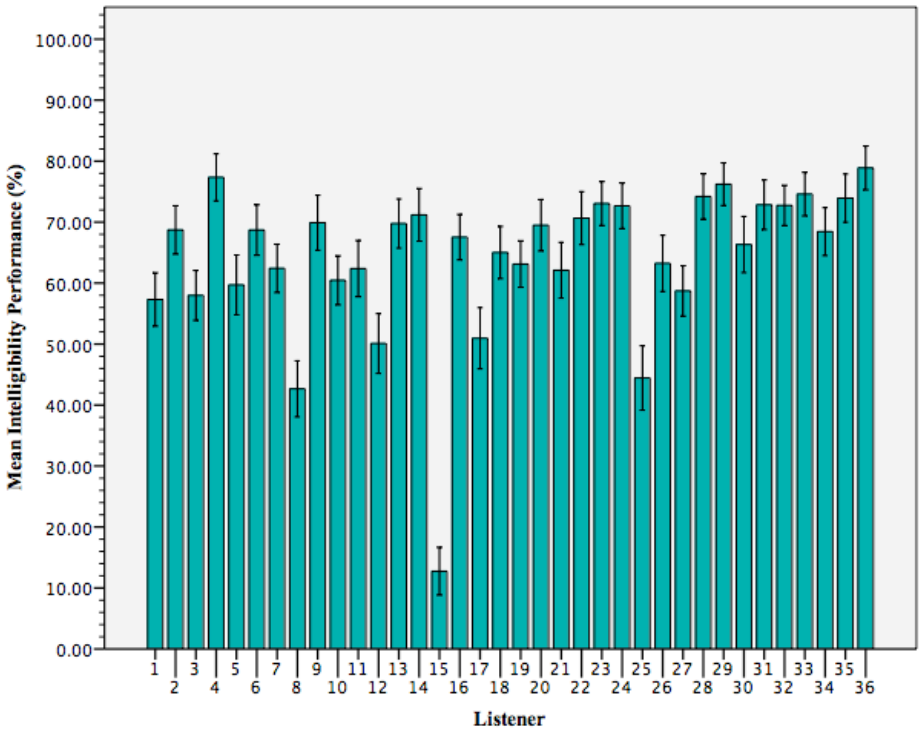

Figure 10. Mean performance on the Hindi-accented speech intelligibility task by listener, shown as percent correct scores. Error bars represent +/- 1 *SE*.

In summary, the results of this experiment show that the intelligibility of unmanipulated Hindi-accented English in background noise was 78.4%. Acoustically manipulating the speech by replacing the prosodic elements of duration and F0 present in Hindi-accented English with those present in American English led to a decrease in intelligibility. When only duration was manipulated, the speech intelligibility decreased to 64.6%. Replacing either the F0 or both the duration and F0 of Hindi-accented sentences with American English prosody resulted in the lowest intelligibility of all, 58.4% and 55.3%, respectively. Listener age was negatively correlated with performance

for some of these sentence types, but not for others. Other demographic variables, such as prior experience with foreign languages and listener sex, did not account for or interact with the statistically significant differences in intelligibility across sentence types.

5. Discussion

The goal of this study was to investigate the intelligibility Hindi-accented English, a language background that is underrepresented in the foreign accent literature. It looked at the prosodic variables of duration and intonation, because past research has been inconclusive in regards to the role each of these plays in the intelligibility of foreign-accented speech. This study found that manipulating F0 leads to a greater detriment in intelligibility than does manipulating duration. Even when F0 was manipulated to be native-like and to therefore enhance intelligibility, this manipulation led to worse performance on the part of native listeners. Acoustically manipulating the duration of Hindi-accented speech also resulted in significantly reduced intelligibility, as the duration-manipulated speech was harder for listeners to understand than the unmanipulated speech. Furthermore, the lowest intelligibility performance of all resulted from sentences where both duration and F0 had been modified.

These findings are surprising. The easiest speech to understand was the speech that retained all prosodic characteristics of the original Hindi-accented utterances. The most difficult speech to understand was that which had been manipulated to have the most characteristics of American English. Can it be concluded that intact patterns of F0

are most important thing for native listeners to have in order to understand Hindi-accented speakers? Or, that modifying Hindi-accented speech in any way renders it less intelligible?

As part of their 2008 investigation, Bradlow and Bent determined intelligibility for one Hindi-accented talker whose unmodified speech was presented to native English listeners in noise at a S/N ratio of +5 dB. This talker's speech received an intelligibility score of above 90 measured on the RAU scale. Overall intelligibility of talkers from multiple native language backgrounds ranged from 53 to 93 RAU. Both of these findings represent higher intelligibility scores than were obtained in the present study. Bent and Bradlow (2003) also obtained mean intelligibility scores for different groups of foreign-accented talkers (not including Hindi) and found that the mean intelligibility score across groups for native English listeners was approximately 67 RAU. This finding indicates lower intelligibility of unmodified accented speech than was obtained in the current study. Tajima et al. (2007) found that, for unmodified Chinese-accented utterances, intelligibility ranged from approximately 35% to 45% depending on the level of background noise. Clearly, this is lower than the 78.4% intelligibility found for unmodified Hindi-accented utterances in the present study. The reasons for these differences are not completely clear, but native language background of the talkers and study methodologies almost certainly played a role. For example, unlike the studies by Bradlow and Bent (2008) and Bent and Bradlow (2003), the current study scored the intelligibility of entire sentences rather than just keywords. Unlike Tajima et al. (2007), the current study used an open-set response instead of a forced-choice alternative. Both

of these choices were amply justified in terms of ecological validity, but also made the scoring criteria extremely strict.

Many of the studies reviewed in this paper's introduction acknowledge the artificiality of the speech stimuli that often results from prosody-swapping algorithms. Perhaps one finding of the current study is that the prosody-swapping algorithm used is not sophisticated enough to produce speech that is as intelligible after modification as it is before. Indeed, Yoon (2006) describes how the algorithm functions at the segment level, but not at the sub-segment level (i.e., formant transitions of vowels or diphthongs). It may be that this level of processing results in digital artifacts that render the modified speech less intelligible. Although in this experiment, as is surely the case in other studies, the author listened carefully to the modified stimuli and rejected any that sounded obviously unnatural, such a barrier perhaps cannot be avoided until more sophisticated methods for speech manipulation are developed. To further illustrate this point, below is a series of figures representing a case in which the prosody-swapping algorithm produced a very unnatural and unintelligible utterance. They represent a female native Hindi speaker's production of the sentence, "The dog came back," and the subsequent prosodic manipulations. The American English female talker's utterance, which provided the reference prosody, is also shown. None of the manipulated tokens was included in the listening experiment, due to the poor quality of the speech signal.

Figure 11

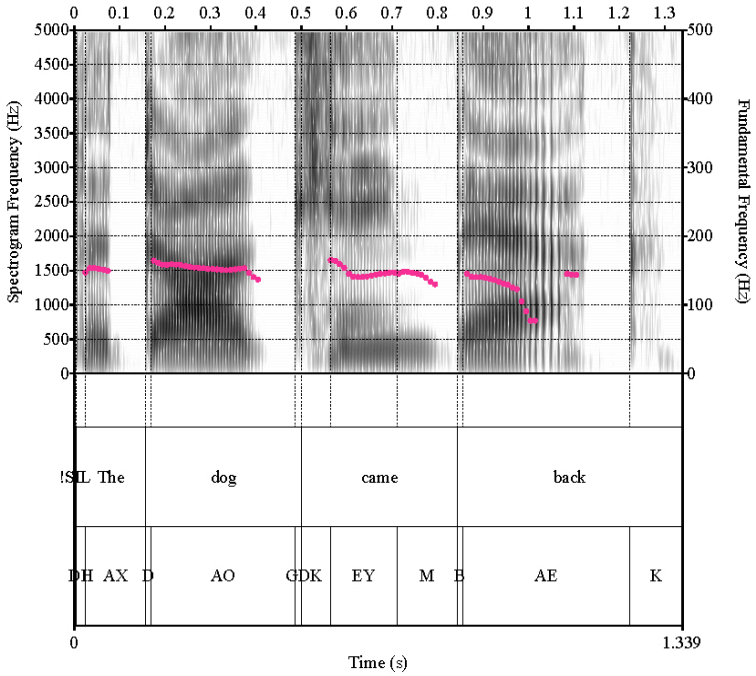

Figure 11. Spectrogram showing the segmental duration and F0 contour of the original American English reference utterance spoken by a female talker.

Figure 12

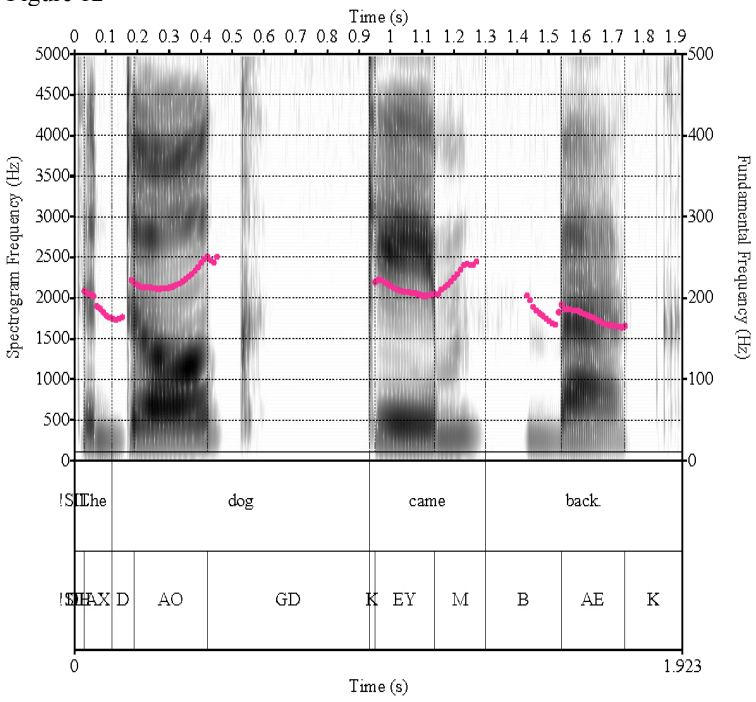

Figure 12. Spectrogram showing the segmental duration and F0 contour of an unmanipulated Hindi-accented utterance spoken by a female talker.

Figure 13

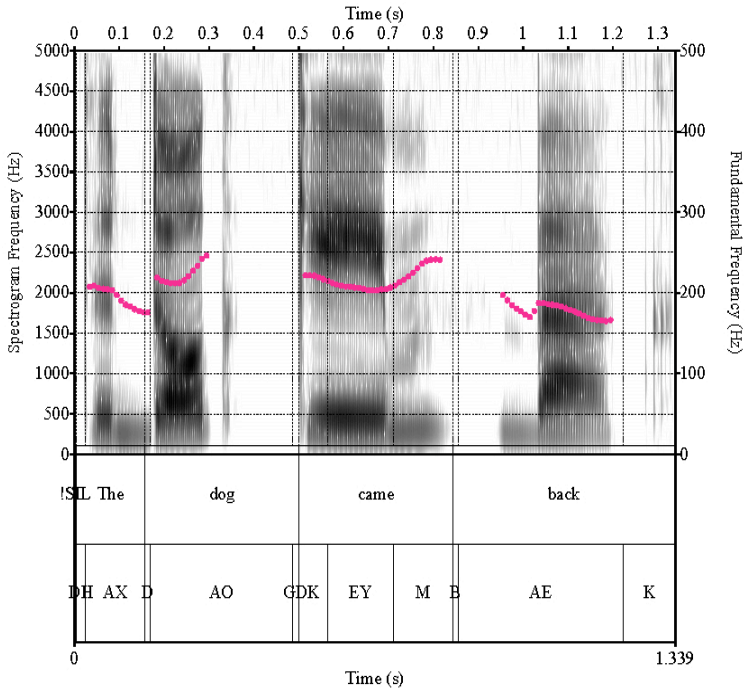

Figure 13. Spectrogram showing the unsuccessful duration manipulation of a Hindi-accented utterance spoken by a female talker.

Figure 14

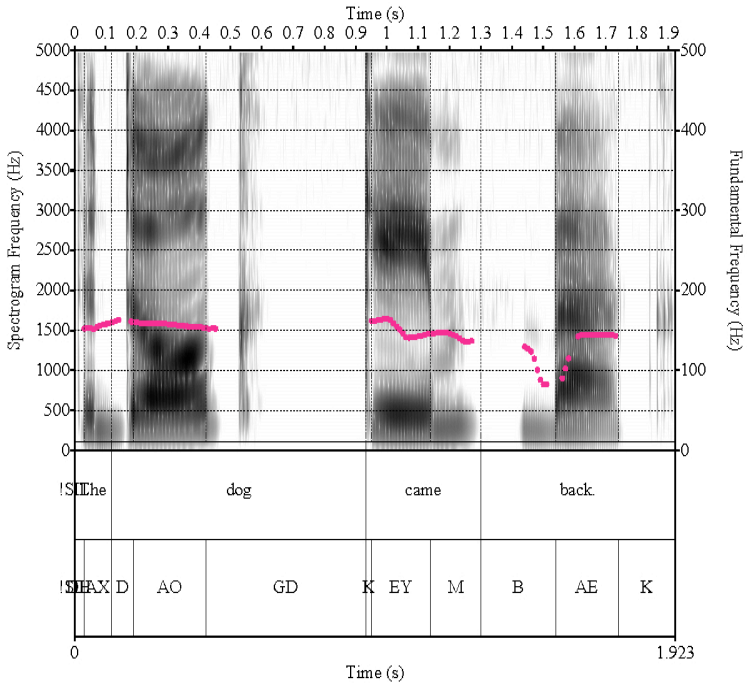

Figure 14. Spectrogram showing the unsuccessful F0 manipulation of a Hindi-accented utterance spoken by a female talker.

Figure 15

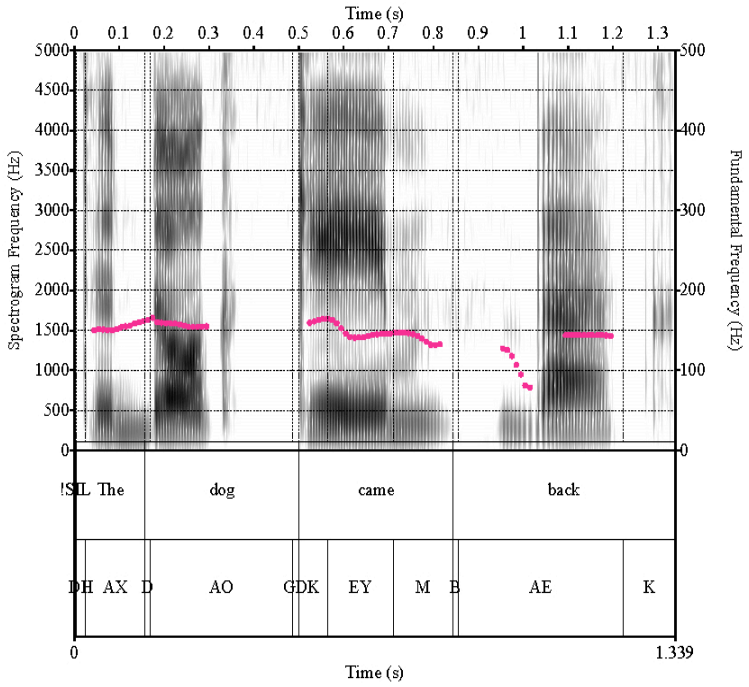

Figure 15. Spectrogram showing the unsuccessful duration-and-F0 manipulation of a Hindi-accented utterance spoken by a female talker.

Lending additional support to the notion that the resynthesis process itself may have reduced the intelligibility of the Hindi-accented speech is a study by Spitzer et al. (2007). In their investigation, the authors manipulated the speech of one American talker along several parameters. To test the quality of the resynthesis process, carried out using the Analysis Synthesis Lab software, they included a condition in which the speech was resynthesized using pitch synchronous, residual-driven LPC methods, but not acoustically manipulated in any way. That is, the speech was taken apart and put back together using its own prosodic features as material, but none of those prosodic features was manipulated. The authors report that untrained listeners were unable to distinguish the resynthesized phrases from the original phrases, though they do not comment on

whether or not they themselves were able to perceive a difference. Nevertheless, Spitzer et al. found that the intelligibility of these resynthesized phrases in noise for 90 English-speaking listeners was slightly lower than that of the natural, but otherwise identical, phrases. While not statistically significant, this finding suggests that resynthesizing speech in any way leads to a decrease in intelligibility. This may account for some of the decrement in speech intelligibility observed for the acoustically manipulated sentences in the present study.

Although they did not investigate intelligibility proper, and although it has already been argued that ratings of perceived comprehensibility are not the same as intelligibility, two studies by Munro and Derwing may shed some light on why it was that manipulating the duration resulted in decreased intelligibility in this study. In both Munro and Derwing (1998) and Munro and Derwing (2001), the effects of speaking rate on the foreign-accented English of non-native speakers from various language backgrounds was investigated. Speaking rate was manipulated synthetically, and native listeners rated the perceived comprehensibility of the speech. When the speaking rate became slower, the perceived comprehensibility dropped. When the speaking rate was moderately speeded, the perceived comprehensibility increased. However, if the non-native speaking rate was speeded to the extent that it became as fast as the native speaking rate, the perceived comprehensibility dropped again. These studies imply that perceived comprehensibility is affected by speaking rate in the following way: a faster rate (compared with the rate of unmodified accented speech) is beneficial, but a speaking rate as fast as the native rate is unfavorable. It could be hypothesized, then, that by altering the duration of Hindi-

accented speech to match the duration of a native American English speaker, a detriment in intelligibility could have resulted. In the absence of detailed acoustic analysis to substantiate this hypothesis, however, it is uncertain if altering the duration of the Hindi-accented speech truly affected intelligibility in such a way.

Again, it is difficult to draw conclusions when past research has examined the speech of talkers from so many different L1 backgrounds, using so many different methodologies. However, the results of this investigation support the work of authors like Bent and Bradow (2003) and Holm (2008), who found that the intelligibility of foreign-accented speech and the respective roles played by prosodic factors like duration and intonation depend very much upon the native language background of both speakers and listeners. And what about a native Chinese speaker communicating in German? Or a native French speaker communicating in Arabic? These speaker/listener pairings would very likely yield different results than what has been concluded by research efforts to date. At this stage, it seems, pointing to concrete evidence and subsequently coming up with practical suggestions about what speakers from various L1 backgrounds can do to maximize their intelligibility is still not attainable.

5.1. Limitations

It is possible that the use of read sentences as opposed to spontaneous speech affected the results of this investigation. It may be that the prosodic characteristics of longer, more natural samples of speech would have been perceived differently by listeners. Another limitation is that no acoustic analyses of the speech of the Hindi-accented and native American talkers were conducted, which may have proved useful

given the fact that Hindi-accented English has not been well examined in past literature. An additional limitation is the fact that no intelligibility data exist for the two native American English speakers who provided the reference prosody for the prosody-swapping procedure. Such data might help to determine whether the decrement in intelligibility observed for sentences whose prosody had been replaced by that of the American talkers could have been related to the overall intelligibility of those speakers. Finally, even though this study had a wider cross section of listeners than is often achieved when recruiting at a university, the listeners were relatively homogenous in that they were all well educated and fairly young, with only a few ethnicities represented. Perhaps a greater diversity in listeners would have resulted in different outcomes.

5.2. Future Research

Future research should present native English-speaking listeners with English speech manipulated to have Hindi-accented duration and pitch characteristics, in order to see whether doing so affects intelligibility. It should acoustically analyze the speech of both the native Hindi and English speakers in detail, in order to pinpoint any significant speech acoustic parameters affecting intelligibility. Finding and utilizing a means of prosody-swapping that is as sophisticated as possible is another worthwhile goal for research in this area. When conducting future investigations of accented speech using prosody swapping, matching the median F0 of the speakers in question may prove useful as a first step. Doing so might alleviate some of the unnaturalness and digital artifacts created by resynthesizing the speech. Having an experimental condition in which listeners rate the intelligibility of speech that has been resynthesized using its own

prosodic features (i.e., replacing the prosody of an utterance, whether Hindi-accented or native English, with its own segmental duration and F0) might lend support to past research that has found that even subtly resynthesized speech is less intelligible than natural speech (i.e., Spitzer et al., 2007). Future research should also examine whether deliberate modifications of duration and pitch by a live speaker in real time (as opposed to computer signal manipulation) produce any measurable changes in intelligibility. It may be that subtle properties of voice quality and other fine-tuned aspects of speech, while not sufficiently captured using computerized methods, could indeed prove to be significant. Such a finding might offer realistically achievable solutions for those in the ESL and accent modification fields working with individuals who speak with a foreign accent. It might also prove worthwhile to use the same stimuli from this experiment for an experiment on comprehensibility, and see if the results of the two investigations are correlated in any way, as some previous research suggests (i.e., Burda et al. 2003). The same stimuli from this experiment could also be presented to native Hindi speakers, in order to compare their patterns of intelligibility to those of native English speakers. Another goal of future research should be to standardize methodologies so that research outcomes can be reliably compared. Ideally, future research in the area of foreign-accented speech would achieve the following: pinpointing exactly which prosodic features are the most important for intelligibility on the part of native listeners, describing how this compares with what speakers from differing native language backgrounds tend to produce in spoken language, and from there, devising practical solutions for non-native speakers and professionals working in the fields of ESL/EFL and accent modification.

References

Adank (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 520-529.

Anderson-Hsieh, J., Johanson, R., & Koehler, J. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmental, prosody, and syllable structure. *Language Learning, 42*(4), 529-556.

Bamford, J., and Wilson, I. (1979). Methodological considerations and practical aspects of the BKB sentence lists. In (Eds.), J. Bench & J. Bamford *Speech-hearing Tests and the Spoken Language of Hearing-impaired Children* (pp. 148–187). London: Academic.

Bannert, R. (1995). Intelligibility and acceptability in foreign accented Swedish: the effects of rhythmical and tonal features. *PHONUM, 3*, 7-29.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology, 13*(3), 108-112.

Bent, T. & Bradlow, A. R. (2003). The interlanguage intelligibility benefit. *Journal of the Acoustical Society of America, 114*(3), 1600-1610.

Black, J.W., Singh, S., Tosi, O., et al. (1965). Speech and aural comprehension of foreign students. *Journal of Speech and Hearing Research, 8*(1), 43-48.

Boersma, P., & Weenick, D. (2009). Praat Digital Signal Processing System (Version 5.1.25) [Computer software]. Available from http://www.praat.org/

Boula de Mareuil, P. & Vieru-Dimulescu, B. (2006). The contribution of prosody to the perception of foreign accent. *Phonetica, 63,* 247-267.

Bradlow, A. R. & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Language & Communication, 22(*2), 171-185

Burda, A. N., Sherz, J. A., Hageman, C. F., & Edwards, H. T. (2003). Age and Understanding Speakers with Spanish or Taiwanese Accents. *Perceptual and Motor Skills, 97*(1), 11-20.

Cooper, J. C., & Gates, G. A. (1991). Hearing in the elderly-the Framingham cohort, 1983–1985: Part II: Prevalence of central auditory processing disorders. *Ear and Hearing*, *12*, 304–311.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America, 116*(6), 3668-3678.

Derwing, T. M. & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*, 1-16.

Ferguson, S. H., Jongman, A., Sereno, J. A., et al. (2010). Intelligibility of foreign-accented speech for older adults with and without hearing loss. *Journal of the American Academy of Audiology, 21*(3), 153-162.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233 – 277). Timonium, MD: York.

Gluszek. (2010). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the united states. Journal of language and social psychology, 29(2), 224-234.

Gut, U. (2007). Foreign accent. In C. Müller, (Ed.), *Speaker classification*. (pp. 75-87). Heidelberg: Springer.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223.

Helfer, K. S., & Freyman, R. L. (2008). Aging and speech-on-speech masking. *Ear and Hearing*, *29*(1), 87–98.

Holm, S. (2008). *Intonational and durational contributions to the perception of foreign-accented Norwegian: An experimental phonetic investigation* (Doctoral dissertation). Det Historisk-Filosofiske Fakultet; Doktoravhandlinger ved NTNU, 2008:282. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:no:ntnu:diva-2331/

Huckvale, M. (2006). The new accent technologies: recognition, measurement and manipulation of accented speech. In P. Zhang, T.-W. Xie, S. Lin, J.H. Xie, A.C. Fang, & J. Xu (Eds.), *Research and application of digitized Chinese teaching and learning.* (Pp. 28-37). Beijing: Language and Culture Press.

Kamiyama, T. (2004). Perception of foreign accentedness in L2 prosody and segments: L2 Japanese speakers learning L2 French. In Bernard Bel & Isabelle Marlien (Eds.) *Speech Prosody 2004 International Conference* (pp. 721-724). Nara, Japan:ISCA Archive. Retrieved from http://www.isca-speech.org/archive/sp2004/

Kortekaas, R. W. & Kohlrausch, A. (1999). Psychoacoustical evaluation of PSOLA. II. Double-formant stimuli and the role of vocal perturbation. *The Journal of the Acoustical Society of America, 105*(1), 522-35.

Lewis, M. P. (Ed.) (2009). Ethnologue: Languages of the World, Sixteenth edition. Dallas, TX: SIL International. Online version retrieved from http://www.ethnologue.com/

Lindemann, S. (2005). Who speaks "broken English"? U.S. undergraduates' perceptions of non-native English. *International Journal of Applied Linguistics, 15*(2), 187-212.

Magen, H S. (1998). The perception of foreign-accented speech. *Journal of phonetics, 26*(4), 381-400.

Moulines, E. & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication, 9*(5-6), 453–467.

Moulines, E. & Laroche, J. (1995). Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Communication 16*(2), 175–205.

Munro, M. J. (1995). Nonsegmental factors in foreign accent: ratings of filtered speech. *Studies in Second Language Acquisition, 17,* 17-34.

Munro, M. J. (1998). The Effects of Noise on the Intelligibility of Foreign-Accented Speech. *Studies in Second Language Acquisition, 20*(2), 139-154.

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Munro, M. J. & Derwing, T. M. (1995b). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language & Speech, 38*(3), 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*(2), 159-182.

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

Munro, M. J. et al. (2010). Detection of nonnative speaker status from content-masked speech. Speech communication, 52(7/8), 626-637.

Nábelek, A. K. & Donahue A. M. (1984). Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America, 75(*2), 632-634.

Rogers, C. L., Dalby, J., & Nishi, K. (2004). Effects of noise and proficiency on intelligibility of Chinese-accented English. *Language and Speech, 47*(2), 139-154.

Schnieder, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime User's Guide.* Pittsburgh: Psychology Software Tools, Inc.

Shukla, S. (1990). Syllable structure and word stress in Hindi. *The Georgetown Journal of Languages and Linguistics, 1*(2), 235-247.

Spitzer, S. M., Liss, J. M., & Mattys, S. L. (2007). Acoustic cues to lexical segmentation: A study of resynthesized speech. *Journal of the Acoustical Society of America, 122*(6), 3678-3687.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of Temporal Correction on Intelligibility of Foreign-Accented English. *Journal of Phonetics, 25*(1), 1-24.

Takata, Y. & Nábelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America, 88*(2), 663-666.

Trofimovich, P. (2006). Learning second language suprasegmentals: Effect of l2 experience on prosody and fluency characteristics of l2 speech. *Studies in Second Language Acquisition, 28*(1), 1-30.

van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *The Journal of the Acoustical Society of America, 112*(6), 3004-3013.

Walton, J., Simon, H., & Frisina, R. D. (2002). Age-related alterations in the neural coding of envelope periodicities. *Journal of Neurophysiology*, *88*, 565–578.

Wong, P. C., Jin, J. X., Gunasekera, G. M., et al. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia, 47*(3), 693-703.

Yoon, K. (2006). Imposing native speakers' prosody on non-native speakers' utterances. Unpublished paper presented at The 9th Western Pacific Acoustics Conference, Seoul, South Korea. Praat script available from: http://www.ling.ohiostate.edu/~kyoon/wespac9/prosody-swapping-all-DF.praat/

APPENDIX – Lists 7, 8, 9, and 10 of the BKB-R protocol

List 7

The children dropped the bag
The dog came back
The floor looked clean
She found her purse
The fruit is on the ground
Mother got a saucepan
They washed in cold water
The young people are dancing
The bus left early
They had two empty bottles
The ball is bouncing very high
Father forgot the bread
The girl has a picture book
The orange was very sweet
He is holding his nose
The new road is on the map

List 8

The boy forgot his book
A friend came for lunch
The match boxes are empty
He climbed his ladder
The family bought a house
The jug is on the shelf
The ball broke the window
They are shopping for cheese
The pond water is dirty
They heard a funny noise
The police are clearing the road
The bus stopped suddenly
She writes to her brother
The football player lost a shoe
The three girls are listening
The coat is on a chair

List 9

The book tells a story
The young boy left home
They are climbing the tree
She stood near her window
The table has three legs
A letter fell on the floor
The five men are working
He listened to his father
The shoes were very dirty
They went on a vacation
The baby broke his cup
The lady packed her bag
The dinner plate is hot
The train is moving fast
The child drank some milk
The car hit a wall

List 10

A dish towel is by the sink
The janitor used a broom
She looked in her mirror
The good boy is helping
They followed the path
The kitchen clock was wrong
The dog jumped on the chair
Someone is crossing the road
The mailman brought a letter
They are riding their bicycles
He broke his leg
The milk was by the front door
The shirts are hanging in the closet
The ground was very hard
The buckets hold water
The chicken laid some eggs