

Speech Characteristics of Japanese Speakers Affecting American and Japanese Listener Evaluations

Atsuko Kashiwagi¹ and Michael Snyder²
Showa Women's University

ABSTRACT

The study examines what pronunciation features (i.e., segmental and suprasegmental features, as well as other acoustic properties such as speech rate, intensity, pitch, and pitch range) affect the intelligibility of Japanese learners of English, when judged by native-speaking (NS; American) and non-native-speaking (NNS; Japanese) listeners. Kashiwagi and Snyder (2008) concluded, based on both statistical and interview data, that intelligibility and accentedness judgments were quasi-independent of each other, and that segmental features in the speech samples of Japanese learners were perceived both by NS (American) and NNS (Japanese) listeners to be more problematic than suprasegmentals. The present study is designed to verify these conclusions by re-examining the data with further statistical procedures. Additional data on speech rate, intensity, pitch, and pitch range were also added to the statistical analyses to explore the issue of intelligibility. The resulting data suggest that non-native pronunciation of segmentals, especially of certain vowels, affect the judgments of intelligibility most strongly, and that intelligibility scores and accentedness ratings are controlled by different sets of factors. Speech rate, intensity, pitch and pitch range are also found to have some effects. The data also suggest that NS (American) and NNS (Japanese) listeners are affected differently by certain pronunciation features.

INTRODUCTION

Before the 1960's, the attainment of native-like pronunciation was widely considered the ultimate goal of pronunciation instruction. As more research findings (e.g., Lenneberg, 1967; Scovel, 1988) showed that this was an unrealistic, if not impossible, goal, an emerging consensus (e.g., Celce-Murcia, Brinton, & Goodwin, 1996; Derwing & Munro, 2005; Pennington & Richards, 1986) has been that pronunciation instruction should aim for comfortable intelligibility rather than an imitation of native speech patterns. The concept of intelligible pronunciation has become even more relevant now, as English is spoken by more nonnative speakers (NNSs) than by native speakers (NSs), and instances of NNS-NNS interaction have increased (Graddol, 1997).

¹ Atsuko Kashiwagi teaches English, interpretation/translation and applied linguistics at Showa Women's University. Her research interests include teaching of pronunciation, learning styles, and bilingual education.

² Michael Snyder has taught EFL in Taiwan, China, and Japan. Besides pronunciation, he is interested in developing students' vocabulary and reading ability through content-based language teaching. Correspondence should be sent to Michael Snyder, Department of International Studies, Showa Women's University, 1-7 Taishido, Setagaya-ku, Tokyo, 154-8533, Japan. E-mail: mksnyder@swu.ac.jp

When native-like pronunciation was the goal of instruction, removal of any foreign accent was of primary importance. The notion of *intelligible pronunciation*, however, is based on the concept that accented speech does not automatically reduce intelligibility and that different types of pronunciation errors may affect comprehension in different ways. It follows, therefore, that instruction should focus on features of pronunciation which most affect intelligibility.

Another relevant finding is that intelligibility and accentedness are partially independent of each other (Derwing & Munro, 1997; Kashiwagi, Snyder, & Craig, 2006; Munro & Derwing, 1995). *Intelligibility* is a highly complex phenomenon which depends on a myriad of factors, but empirical data have consistently pointed to the importance of pronunciation in affecting intelligibility (Derwing & Rossiter, 2003; Fayer & Krasinski, 1987; Suenobu, Kanzaki & Yamane, 1992). There have been numerous studies which have investigated which aspects of pronunciation most affect intelligibility, but their findings are mixed at best, possibly due to differences in methodology as well as how the notion of intelligibility is defined. Gimson (1970) argued that correct pronunciation of consonants is more important to comprehension than accurate vowel production, but the findings of Schairer (1992) were the exact opposite. Anderson-Hsieh, Johnson, and Koehler (1992) investigated the relative importance of segmental and suprasegmental features as they relate to intelligibility, and found that the prosodic variable is most strongly associated with the intelligibility scores. While other authors have also argued that prosodic errors are more serious than segmental errors (Johansson, 1978; Palmer, 1976), the results of other studies have been either inconclusive (Derwing & Munro, 1997; Munro & Derwing, 1995) or have found opposite results (Fayer & Krasinski, 1987; Koster & Koet, 1993). More recent studies, instead of trying to debate the primacy of either segmental or suprasegmentals, are more concerned with defining what specific pronunciation features decrease intelligibility. Hahn (2004) reported that misplaced or missing nuclear stress reduced intelligibility significantly. Field (2005) found that misplaced word stress resulted in decreased intelligibility.

While the above research is mainly concerned with NSs listening to NNS speech, recent researchers have brought up another very important perspective in the issue of intelligibility, namely *intelligibility as judged by NNS listeners*. Jenkins (2000; 2002) studied NNS-NNS interaction data collected in naturalistic settings and reported that segmental errors were clearly the most frequent causes of miscommunication in these interactions. Most suprasegmentals such as features of connected speech, pitch movements to signal attitude or grammatical meaning, placement of word stress, and stress-timed rhythm, did not lead to any serious intelligibility problems. Kashiwagi and Snyder (2008) studied how American and Japanese judges evaluated speech samples of Japanese students. Based on the extensive interviews with the judges, they reported that most of the mis-hearings were caused by segmental, not suprasegmental mistakes, and that there were no significant differences among the judges which were attributable to their language backgrounds. Munro, Derwing, and Morton (2006) examined how native speakers of Cantonese, Japanese, Mandarin and English responded to English utterances from native speakers of Cantonese, Japanese, Polish and Spanish and found striking similarities in the judgments of intelligibility, comprehensibility and accentedness. Riney, Takagi and Inutsuka (2005) reported that American listeners and Japanese listeners responded to NNS speech differently, but they found that American listeners primarily used segmental cues in their perceptions of accent, and Japanese listeners primarily used non-segmental parameters (intonation, fluency, and speech rate). Their study does not directly address the issue of intelligibility, but nonetheless provides an interesting perspective.

Many studies have also examined the effects of other acoustic characteristics of speech on intelligibility. Munro and Derwing (1998) studied how varying speech rates of NNS speech affected comprehension and found that, in general, NSs preferred to listen to NNSs speaking at slower rates. There is a vast body of research looking into the effects of such acoustic properties as voice quality, speaking rates, loudness, pitch, and pitch range on intelligibility in the fields of speech therapy and synthesized voice (Doyle, Danhauer, & Reed, 1988; Horga & Liker, 2006; Qi & Weinberg, 1991; Tjaden & Wilding, 2004; Tomokiyo, Black, & Lenzo, 2005). Their results suggest that these acoustic properties significantly affect intelligibility. However, there is a paucity of similar research with special focus on nonnative English speech.

The increasing body of research into intelligibility of NNS speech has shed considerable light on the issue. The emerging picture, while still far from complete, seems to suggest that intelligibility and accentedness must be understood as two quasi-independent phenomena, and that not all pronunciation features contribute equally to intelligibility. Errors in the production of certain segmental features do cause misunderstandings, just as non-native suprasegmental features were responsible for reduced intelligibility. There is still limited research on the effects of other acoustic dimensions of NNS speech on intelligibility and listener variables (e.g., NS listeners vs. NNS listeners), and they need to be explored further.

Purpose of the Study and Research Questions

The purpose of the present study is to gain more understanding of the issue of intelligibility and accentedness with a special focus on English as an international language (EIL) spoken by Japanese speakers in contexts where they need to be understood both by native and nonnative speakers of English. The study builds on the data obtained in the authors' 2008 research, and is designed to answer the following questions:

1. Do statistical analyses support the authors' previous conclusions?

The original study used extensive interview data to draw conclusions that most intelligibility problems were caused by segmental mistakes, and that NS (American) and NNS (Japanese) judges were affected by non-native pronunciation features in similar ways. The present study aims to verify the conclusions with statistical analyses.

2. Are intelligibility and accentedness quasi-independent phenomena, as the authors' previous study suggested?

The original study found that strong accentedness did not automatically presuppose reduction of intelligibility. The present study employs statistical analyses to further explore whether different factors affect intelligibility and accentedness.

3. Do other acoustic dimensions of speech such as pitch, intensity, pitch range, and speech rate affect the judgments of accentedness and intelligibility?

The present study is designed to investigate this question with statistical analyses.

RESEARCH METHOD

The Original Study – Kashiwagi and Snyder (2008)

In Kashiwagi and Snyder (2008), speech samples from 20 Japanese female college students were evaluated by three NS (American) and three NNS (Japanese) judges for both their intelligibility and accentedness. We were unable to use NS judges who represent other NS English varieties, nor could we use NNS judges who represent other language backgrounds than Japanese, due to the difficulties in finding a sufficient number of qualified judges from various nationalities. All of the American judges were college English teachers from the U.S. with more than 5 years' experience in teaching Japanese speakers, and had a good understanding of English phonology. All three Japanese judges were also college English teachers, who had obtained master's degrees in the U.S. and had an equally good understanding of English phonology. Japanese Judge A had also lived in the US as a child. As for the students, their English proficiency was false-beginning to intermediate with their Test of English for International Communication (TOEIC) scores ranging from the high 200's to the low 600's.

Following Munro and Derwing (1995), *intelligibility* was broadly defined in Kashiwagi and Snyder (2008) as the extent to which a speaker's utterance is actually understood by a listener and calculated in percentage points by exact word matches between intended messages and transcriptions. *Accentedness*, on the other hand, was defined as the extent to which a speaker's pronunciation is perceived to differ from a NS version, and was measured impressionistically on a scale of seven with 1 indicating "very strongly accented" and 7 "no accent." The definition of accentedness is also based on the study by Munro and Derwing (1995).

The 20 students read two short passages each, and a total of 40 different passages were recorded. Prepared passages, rather than spontaneous speech, were used to avoid judgments of pronunciation from being influenced by any syntactic and semantic errors made by the speaker (Briere, 1967; Varonis & Gass, 1982). To avoid unnatural pausing or intonation (which would automatically lead to reduced intelligibility) when the students read the passages for recording, we took time to talk to each student in Japanese before the recording to make sure their understanding of the passages were accurate. The recorded passages were played to the American and Japanese judges to be transcribed in standard orthography. After the completion of the transcription, the judges were given a break of about 5 minutes and listened to the passages a second time to rate the accentedness of each utterance impressionistically on a scale of 7, with 1 indicating "very strongly accented" and 7 "no accent." The transcriptions and the accentedness ratings were done on two separate listenings so that the judgments of accentedness would not be affected by the process of transcription.

The researchers held an interview with each of the six judges several days later to discuss what pronunciation features they perceived to be the cause of their misunderstandings. Each judge was presented with the original passages along with their transcriptions with mismatches highlighted in red, when listening to the recordings once more, and were asked to freely discuss possible causes of their misunderstandings. The suggested causes were written down by the researchers, who later compared their notes and tabulated the results. The authors decided to use these self-reports as their core data, as in any discussion of intelligibility, which is a heavily listener-dependent phenomenon, listener perceptions must be given an important role.

General American pronunciation (GA) was used as "a point of reference" in our descriptions of how the subjects' pronunciation differs from the NS version, as all the NS judges in this study are native speakers of American English and all the NNS judges studied in the United States. The purpose of this study was to observe, as objectively as possible, how departures from GA affect American as well as Japanese judges, and it is not in any way implied that GA is the only target variety of English for Japanese students to emulate. The IPA-modified

system used in *Longman Dictionary of Contemporary English* (2003) was used for phonetic transcriptions.

The Present Study

The original study investigated the judges' perceptions of problematic pronunciation features by means of extensive interviews. The authors are aware, however, that collecting data in interviews has certain limitations, as there is the definite possibility of perceptual data being clouded by listener bias (Derwing & Munro, 2005). Based on the understanding that interview data may not constitute sufficient evidence in themselves, the present study was conducted to examine whether statistical analyses bore out the judges' self-reports. Instrumentally measured acoustic data on speech properties were also added to the overall picture to explore how they relate to intelligibility and accentedness.

The interview data in the original study showed that the judges perceived vowel errors, consonant errors and stress errors to have contributed most heavily to the reduction of intelligibility. Of the vowel errors, those of r-colored vowels and five other vowels (/æ/, /ɑ:/, /ʌ/, /ɔ:/, /oʊ/) were most often identified to be problematic; of the consonant errors, /l/, /r/, /θ/, /ð/, /f/ and /v/ were most often reported to have caused unintelligibility. The error rates of these vowels and consonants were calculated for each speech sample. While in the past studies, researchers tended to treat all phonological errors equally when examining their effects on intelligibility (Anderson-Hsieh, Johnson & Koehler, 1992; Munro & Derwing, 1995), we decided to examine the effects of what our interview data found to be the most problematic vowels and consonants, as including other vowels and consonants whose pronunciation errors did not cause many intelligibility problems might have diluted the data for statistical analysis. The researchers, both of whom were trained phoneticians, listened to each speech sample, and counted errors in the pronunciation of the vowels and consonants specified above. The error rates were then calculated by dividing the number of errors by the total number of possible occurrences of error in each sample. The two researchers also identified errors in stress placement. The rates of irregular word stress and phrase stress (i.e., stress on noun compounds, adjective-noun phrases and verb phrases) were calculated by dividing the number of syllables affected by a particular instance of irregular stress by the total number of syllables in the speech sample.

Acoustic dimensions of speech, which are often discussed with regard to speech quality (i.e., pitch, intensity, pitch range, and speech rate), were measured for each of the speech samples by using Praat speech analysis software. Praat is being developed by Boersma and Weenink (2007) at the University of Amsterdam, and is available by free download from <<http://www.praat.org/>>. The average pitch and intensity of each speech sample were computed. Pitch means highness or lowness of the voice and is measured in hertz (Hz), while intensity means loudness or softness of the voice and is measured in decibels (dB). The gross pitch range was computed by subtracting the minimum pitch value from the maximum for the entire sample. The speech rate was computed by dividing the number of syllables in each of the speech samples by its total duration, which is measured by Praat to the nearest .01 second. Voice quality, which is also believed to be a major factor determining speech quality, was not included in this study due to the fact that the term is still not well defined, and that its measurement methods are not established.

Separate stepwise multiple regression analyses were conducted for the American judges and Japanese judges with intelligibility scores as the dependent variable and seven factors (vowel

error rate, consonant error rate, stress error rate, average pitch, average intensity, pitch range and speech rate) as the independent variables to find out which of these can best predict intelligibility, and whether the same results are observed for the two groups of judges. The same analyses were conducted with accentedness ratings as the dependent variable to see if the two dimensions of pronunciation, intelligibility and accentedness, are affected similarly by the seven factors.

RESULTS AND DISCUSSION

The Original Study – Kashiwagi and Snyder (2008)

Brief descriptions of the results of Kashiwagi and Snyder (2008) are reported first to prepare readers for the present study. The original study conducted Pearson’s Correlation analyses between accentedness ratings and intelligibility scores and found wide individual variations among the judges (Table 1). The scatter plot (Figure 1) of all the accentedness ratings and intelligibility scores of the six judges also showed that even though there is a weak correlation between the two sets of data (Pearson’s $r = .0195^{**}$), high intelligibility scores do not necessarily presuppose low accentedness ratings.

TABLE 1
Correlations between Intelligibility and Accentedness (Pearson’s r)

American Judge A ($n = 40$)	American Judge B ($n = 40$)	American Judge C ($n = 40$)	Japanese Judge A ($n = 40$)	Japanese Judge B ($n = 40$)	Japanese Judge C ($n = 40$)
.306	.561**	.355*	.546**	.219	.257

* = $p < .05$ ** = $p < .01$

FIGURE 1
Scatter Plot between Accentedness and Intelligibility



The original interview data showed that for both American and Japanese judges, mispronunciation of a vowel was the most often cited reason for misunderstanding, which accounted for 139 (35.3%) of the total 393 identified reasons. Of the 139 cases, r-colored vowels were the most often cited causes, followed in order by the five vowels, /æ/, /ɑ:/, /ʌ/, /oʊ/ and /ɒ:/. These vowels were responsible for a total of 95 instances of misunderstanding.

Mispronunciation of a consonant was the second most often cited reason for misunderstanding for both groups, and accounted for 94 instances (23.9%). Of the 94 consonant errors cited by the judges, /r/ topped the list, followed in order by /l/, /ð/, /θ/, /f/ and /v/. These six consonants were responsible for a total of 62 citations. Both American and Japanese judges reported suprasegmental errors to be less problematic in their comprehension, and identified a total of only 46 cases in which suprasegmental errors were perceived to be the main cause of misunderstanding. Of the 46 cases, 33 were caused by irregular word stress or phrase stress (stress on noun compounds, adjective-noun phrases and verb phrases), and the remaining 13 were caused by lack of insufficient sentence stress on content words. Intonation, rhythm patterns and features of connected speech were not cited to have caused misunderstandings by either of the two groups of judges.

The results of the interviews are summarized below (Table 2). A Chi-square analysis

compared the answers between the American and Japanese judges across the three main categories of error types (segmental errors, suprasegmental errors, and segmental + suprasegmental errors). The result, $\chi^2(2) = .10, p = .95$, showed that there were no significant differences between the two groups.

TABLE 2
Identified Causes of Misunderstandings by Error Types

Error Types		American Judges			Japanese Judges			Total
		A	B	C	A	B	C	
Segmental Errors	Consonant Error (Including Consonant Cluster)	14	10	13	13	17	27	94
	Vowel Error	24	16	22	18	33	26	139
	Combination of Segmental Errors	10	8	10	10	10	11	59
	Epenthesis (Sound Insertion)	5	5	7	2	4	2	25
	Subtotal	53	39	52	43	64	66	317
Supra-segmental Errors	Word or Phrase Stress Error	3	5	5	6	7	7	33
	Sentence Stress Error	3	2	2	2	2	2	13
	Subtotal	6	7	7	8	9	9	46
Supra-segmental Errors + Segmental Errors	Word or Phrase Stress Error + Segmental Error(s)	3	1	8	5	4	9	30
	Irregular Sentence Stress + Segmental Error(s)	0	0	0	0	0	0	0
	Subtotal	3	1	8	5	4	9	30
Total		62	47	67	56	77	84	393

The Present Study

In order to verify the conclusions of Kashiwagi and Snyder (2008) and to add other acoustic properties, stepwise multiple regression analyses with the intelligibility scores as the dependent variable and seven speech characteristics (vowel error rate, consonant error rate, stress error rate, average pitch, average intensity, pitch range, and speech rate) as the independent variables were conducted (Tables 3 & 4). The results showed the importance of vowels to intelligibility, as was suggested in the original interview data. They also suggested that intelligibility assessments of the Japanese judges were not controlled by the same set of variables as those of the American judges.

The analysis of the American judges showed Vowel Error Rate as the single significant predictor variable. This variable accounted for 16% of the variance. Three variables (Speech Rate, Pitch Range and, Vowel Error Rate) were selected as the predictor variables in the analysis of the Japanese judges, and accounted for 16% of the variance. The results seem to suggest that

vowel errors had effects on the reduction of intelligibility for both American and Japanese judges. Speech rates and pitch ranges, however, had more effect on the Japanese judges than vowel errors; the Japanese judges tended to find speech samples at slower rates and with wider pitch ranges easier to understand. For both American judges and Japanese judges, consonant errors and stress errors were excluded from the regression model. Intensity and pitch were also not found to be significant predictors for either of the groups.

TABLE 3
Stepwise Multiple Regression Analysis between Intelligibility Scores and
Seven Speech Characteristics (American Judges)

n = 120

Independent Variables	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	<i>SE</i>	β		
Vowel Error Rate	-19.82	4.16	-.40	-4.76	.00

$R^2 = .16$ Adjusted $R^2 = .15$

TABLE 4
Stepwise Multiple Regression Analysis between Intelligibility Scores and
Seven Speech Characteristics (Japanese Judges)

n = 120

Independent Variables	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	<i>SE</i>	β		
Speech Rate	-.10	.05	-.21	-2.26	.03
Pitch Range	.03	.01	.23	2.70	.01
Vowel Error Rate	-11.69	5.73	-.19	-2.04	.04

$R^2 = .16$ Adjusted $R^2 = .13$

To investigate how accentedness ratings were influenced by the above seven speech characteristics and compare the results with those of intelligibility, stepwise multiple regression analyses with accented ratings as the dependent variable and the seven speech characteristics as the independent variables were conducted (Tables 5 & 6). The results suggest that accentedness ratings were influenced by a different set of variables from those affecting intelligibility scores, and that while the American and Japanese judges were affected by similar variables, there were certain differences between the two groups of judges. The analysis of the American judges showed that four variables (Consonant Error Rate, Speech Rate, Intensity and Vowel Error Rate) together accounted for 30% of the variance. The analysis of the Japanese judges also suggested that four variables (Pitch, Consonant Error Rate, Stress Error Rate, Speech Rate) were significant predictors for accentedness ratings. These variables in total accounted for 34% of the variance.

Consonant errors had effects on accentedness ratings for both groups of judges, while they were not found to have significantly influenced intelligibility scores. Stress errors, which were not selected as a significant predictor for intelligibility, were found to have influenced accentedness ratings of the Japanese judges. Vowel error rates in turn were selected only in the regression model of the American judges. Speech rates, which had effects on intelligibility judgments of the Japanese judges, also affected accentedness ratings of both groups of judges.

The results show that both American and Japanese judges found speech samples more accented when they were delivered at slower rates. The American judges were also affected by the intensity in their accentedness ratings; they found louder speech samples more accented. The Japanese judges in turn seemed to be affected by the pitch; they found speech samples spoken at higher pitches less accented.

TABLE 5
Stepwise Multiple Regression Analysis between Accentedness Ratings and Seven Speech Characteristics (American Judges)

n = 120

Independent Variables	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	<i>SE</i>	β		
Consonant Error Rate	-1.37	.49	-.25	-2.82	.01
Speech Rate	-.01	.00	-.25	-3.04	.00
Intensity	-.15	.06	-.22	-2.70	.01
Vowel Error Rate	-1.23	.54	-.22	-2.29	.02

$R^2 = .30$ Adjusted $R^2 = .27$

TABLE 6
Stepwise Multiple Regression Analysis between Accentedness Ratings and Seven Speech Characteristics (Japanese Judges)

n = 120

Independent Variables	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	<i>B</i>	<i>SE</i>	β		
Pitch	.02	.01	.32	4.03	.00
Consonant Error Rate	-1.70	.45	-.30	-3.78	.00
Stress Error Rate	-4.09	1.34	-.23	-3.05	.00
Speech Rate	-.01	.00	-.20	-2.52	.01

$R^2 = .34$ Adjusted $R^2 = .32$

CONCLUSION

The present study seems to verify at least some of the claims made in Kashiwagi and Snyder (2008) based on the interview data. Stepwise multiple regression analyses on speech characteristics showed that vowel errors, which were reported in the interviews to affect intelligibility most frequently, were found to have influenced intelligibility of both American and Japanese judges. Consonant errors and stress errors, which were also identified in the interviews as frequent causes of misunderstanding, however, did not show significant effects on intelligibility, probably because they were responsible for many fewer intelligibility problems than vowels.

Statistical analyses also provided further evidence that intelligibility and accentedness are quasi-independent of each other. Stepwise multiple regression analyses showed that accentedness

ratings were controlled by a different set of variables than intelligibility scores. Consonant errors, which were excluded from the regression model for intelligibility scores, were selected as significant predictors for accentedness ratings for both groups of judges. The results seem to suggest that non-native production of consonants significantly affects listeners' impressions (e.g., accentedness ratings) of pronunciation. Vowel errors, on the other hand, were more serious than consonant errors in affecting intelligibility.

Stress errors, while not influencing the accentedness ratings of the American judges, were found to have affected those of the Japanese judges, suggesting that the American and Japanese judges were possibly influenced by different features of pronunciation in their accentedness ratings. The different effects of suprasegmental features on NS and NNS judgments of accent were also reported in the study by Riney, Takagi, and Inutsuka (2005), who found that in their assessments of accent, NNS listeners used non-segmental parameters more while NS listeners were influenced more by segmentals, especially /r/ and /l/.

Other acoustic dimensions of speech were also found to have influenced both intelligibility and accentedness. Slower speech rates were found to have helped the Japanese judges improve their comprehension, but not the American judges, probably because the Japanese judges needed more time to process accented speech than their American counterparts. It should be noted that wider pitch ranges also helped the Japanese judges in their comprehension. Slower speech rates, however, had adverse effects on accentedness ratings; both the American and Japanese judges found slower speech samples more accented. Munro and Derwing (1998) also reported the adverse effects of slower speech rates on accentedness judgments.

Various explanations for the other findings regarding the effects of intensity, pitch, and pitch range are possible, but what we found in the present study is still explorative and there is still a lack of sufficient evidence for valid discussion. We would like to further examine these dimensions of speech in our future research.

When the results of Kashiwagi and Snyder (2008) and the present study are combined, they seem to provide important pedagogical implications. Many researchers as well as teachers now place importance on the instruction of suprasegmentals, on the assumption that they seriously affect intelligibility. Some researchers even argue that suprasegmentals must be given priority in pronunciation instruction (Avery & Ehrlich, 1992; Celce-Murcia, Brinton & Goodwin, 1996). The combined data, however, seem to suggest that the instruction of certain vowels also deserves significant attention if the purpose of the instruction is to increase intelligibility of Japanese speakers.

Apart from the discussion of where instructional priority should be placed, another important issue is the different nature of intelligibility and accentedness, and the need to allow learners to make an informed choice. Learners of EIL, who often equate these two parameters without question, must be informed that a strong accent does not automatically lead to reduction of intelligibility, and that they may still be perfectly intelligible to both NS and NNS listeners even if they retain some non-native pronunciation features. At the same time, however, learners must also be informed that retention of non-native accents may disadvantage them in certain situations. Many language attitude studies have indicated that NNS accents still evoke negative responses compared to "standard" NS accents (Lippi-Green, 1997; Derwing, 2003; Jenkins, 2007). After receiving sufficient information, learners, not teachers or researchers, must decide what pronunciation goal they wish to set for themselves. Jenkins (2002) states, "it will be important not to patronize those learners who, having heard the arguments, still wish to work towards the goal of a native speaker accent, by telling them they have no need to do so" (p. 101).

Suggestions for Future Studies

The present study, while providing some important pedagogical implications, is limited in several ways. First, the speech samples used here were readings of prepared passages, not extemporaneous productions in two-way communicative situations. As Jenkins (2007) indicates, these laboratory settings lack natural communicative behaviors such as accommodation and negotiation for meaning, and the data obtained in these settings may not correctly reflect what actually happens in real life. When extemporaneous speech samples are taken from beginning to intermediate students, their semantic and syntactic errors are likely to affect overall intelligibility judgments, and may make it difficult to correctly assess the effects of non-native pronunciation performance. A study is needed which uses speech samples of advanced learners of English in real communicative situations to truly examine the issues of intelligibility.

Secondly, in the present study, NS speakers were only represented by speakers of GA, and NNS speakers only by Japanese speakers. The fact that these judges also either shared the subjects' native language or had extensive exposure to it may make it difficult to generalize the results. More research is needed to obtain perspectives of NS listeners other than the speakers of GA, as well as of non-Japanese NNS listeners, both of whom are unfamiliar with English spoken with a Japanese accent. In addition, what we found in the present study is only applicable to English spoken by Japanese speakers. Even though most of our results agree with what Jenkins (2000, 2002) reported, further research is needed, which investigates how the pronunciation features of speakers of other NNS variations affect both NS and NNS listeners, before any generalization on NNS speech as a whole is suggested.

Finally, multiple regression analyses in the present study only accounted for small portions of the variance of both intelligibility and accentedness assessments. Even with the understanding that non-acoustic factors (grammar, familiarity with topic, familiarity with speaker's accent, social factors, etc.) are conceivably responsible for some of the variance, more effort to explain the remaining variance is needed.

REFERENCES

- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42, 529-555.
- Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.
- Boersma, P. & Weenink, D. (2007). *Praat: Doing phonetics by computer (Version 4.6.12)*. Retrieved July 27, 2007, from <http://www.praat.org/>
- Briere, E. J. (1967). Phonological testing reconsidered. *Language Learning*, 17, 163-171.
- Bullon, S. (2003). *Longman dictionary of contemporary English*. Harlow, England: Pearson Education.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Derwing, T. M. (2003). What do ESL students say about their accents? *The Canadian Modern*

- Language Review*, 59, 547-566.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 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., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning*, 13, 1-17.
- Doyle, N. P., Danhauer, J. L., & Reed, C. G. (1988). Listeners' perception of consonants produced by esophageal and tracheoesophageal talkers. *Journal of Speech and Hearing Disorders*, 53, 400-407.
- Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37, 313-326.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39, 399-423.
- Gimson, A. C. (1970). *An introduction to the pronunciation of English* (2nd ed.). London: E. Arnold.
- Graddol, D. (1997). *The future of English?* London: British Council.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38, 201-223.
- Horga, D & Liker, M. (2006). Voice and pronunciation of cochlear implant speakers. *Clinical Linguistics & Phonetics*, 20, 211-217.
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23, 83-103.
- Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.
- Johansson, S. (1978). Studies in error gravity: Native reactions to errors produced by Swedish learners of English. *Göteborg Studies in English*, 44, 131-138.
- Kashiwagi, A., & Snyder, M. (2008). American and Japanese Listener Assessment of Japanese EFL Speech: Pronunciation Features Affecting Intelligibility. *The Journal of Asia TEFL*, 5(4), 27-47.
- Kashiwagi, A., Snyder, M., & Craig, J. (2006). Suprasegmentals vs. Segmentals: NNS phonological errors leading to actual miscommunication. *JACET Bulletin*, 43, 43-57.
- Koster, C. J., & Koet, T. (1993). The evaluation of accent in the English of Dutchmen. *Language Learning*, 43, 69-92.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: John Wiley and Sons.
- Lippi-Green, R. (1997). *English with an accent*. London; Routledge.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73-97.
- Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, 48, 159-182.
- Munro, M. J., Derwing, T.M., & S. L. Morton. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28, 111-131.
- Palmer, J. (1976). Linguistic accuracy and intelligibility. *Proceedings of the 4th International*

- Congress of Applied Linguistics, Stuttgart, Germany, 505-513.*
- Pennington, M., & Richards, J. (1986). Pronunciation revisited. *TESOL Quarterly, 20*, 207-225.
- Qi, Y., & Weinberg, B. (1991). Spectral slope of the vowels produced by tracheoesophageal speakers. *Journal of Speech and hearing Research, 34*, 243-247.
- Riney, T.J., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly, 39*, 441-466.
- Schairer, K. (1992). Native speaker reaction to non-native speech. *The Modern Language Journal, 74*, 309-319.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House.
- Suenobu, J., Kanzaki, K., & Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *International Review of Applied Linguistics in Language Teaching, 30*, 146-156.
- Tjaden, K., & Wilding, G. (2004). Rate and loudness manipulations in dysarthria. *Journal of Speech, Language, and Hearing Research, 47*, 766-783.
- Tomokiyo, L. M., Black, A. W. & Lenzo, K. A. (2005). Foreign accents in synthetic speech: Development and evaluation. *Proceedings of Interspeech 2005, Lisbon, Portugal*, 1469-1472.
- Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition, 4*, 114-136.