Kang, O. (2013). Relative impact of pronunciation features on ratings of non-native speakers' oral proficiency. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference.* Aug. 2012. (pp. 10-15). Ames , IA: Iowa State University.

RELATIVE IMPACT OF PRONUNCIATION FEATURES ON RATINGS OF NON-NATIVE SPEAKERS' ORAL PROFICIENCY

Okim Kang, Northern Arizona University

In the context of second language (L2) assessment, pronunciation has proved to be a problematic and consequently not widely investigated area. Although some effects of fluency, suprasegmental, or individual segmental errors have been identified in L2 oral assessment, the relative impact of a wide array of pronunciation features on non-native speakers' (NNSs) oral proficiency is still uncertain. The current study has investigated to what extent errors in various pronunciation features predict the rated oral proficiency of NNSs' speech. Speech samples of 120 candidates, 1-2 minute long, taken from Individual Long Turn sections in Cambridge ESOL General English Examinations, were analyzed acoustically for measures of speech rate, pauses, stress, intonation, and segmental errors. Results showed that there was a hierarchical priority in the effect of the pronunciation features on ratings of NNSs' oral proficiency. The finding has important implications for the field of L2 speaking assessment and pronunciation instruction.

THE STUDY

Various aspects of non-native speakers' (NNSs) pronunciation can be considered in listeners' assessments of speaker proficiency and both segmentals and suprasegmentals have been examined for their roles in judgments of accented speech. Some studies have investigated the impact of suprasegmental features on listeners' ratings of NNSs' oral performance (Kang, Rubin, & Pickering, 2010) or the correlations between temporal measures and listeners' rating scores (Munro & Derwing, 2001). Other studies have examined the effect of segmentals on listeners' judgments of NNSs' speech (e.g., segmental errors show great effects on listeners' comprehension in Fayer & Krasinski, 1987; high functional-load segmental errors have larger effects on perceptual scales than low functional-load segmental errors in Munro & Derwing, 2006).

However, no consensus has been reached regarding to what extent different pronunciation features contribute to the overall ratings of speaking assessment. Using candidates' spoken responses on the Cambridge ESOL General English Examinations; this study examined the relative contribution of each of the pronunciation categories to the overall oral proficiency judgments.

Speech samples

Cambridge ESOL General English Examinations provided one hundred twenty speech files; these were candidates' responses to their prompts in speaking test. From each speech file, a one-minute-long, monologic section of the response was extracted and transcribed. The current study included the following four proficiency levels using the Common European Framework of Reference for Languages (CEFR) from B1-C2: the

Preliminary English Test (PET, B1), the First Certificate in English (FCE, B2), the Certificate in Advanced English (CAE, C1), the Certificate of Proficiency in English (CPE, C2). Speakers analyzed in this study included 32 in PET, 32 in FCE, 34 in CAE, and 22 in CPE. There were 21 different first languages (L1s) represented.

Speech data coding and analysis

This study auditorily and instrumentally analyzed segmental and suprasegmental features of candidates' output. The instrumental analysis was conducted on two versions of speech analysis software: the Computerized Speech Laboratory (CSL) for stress and intonation analysis and PRAAT for fluency analysis. For the auditory analysis, two analysts participated with reliability reaching .81.

Segmental (vowels and consonants) errors included vowel/consonant substitutions and fourteen other types of errors (e.g., simplification of consonant clusters, linking errors, vowel or consonant epenthesis, vowel or consonant elision). The segmental errors were calculated as the total number of segmental errors divided by the total number of syllables articulated.

The suprasegmental analysis employed the methodology from Kang et al. (2010), i.e., it consisted of fluency, stress and pitch, and tone choices. In fluency measures, the syllables per second, mean length of run (e.g., utterances between pauses of 0.1 seconds and above), the number of silent pauses, mean length of silent pauses, number of hesitation markers, and mean length of hesitation pauses were counted. To measure stress and pitch, space (the proportion of prominent words to the total number of words), pace (number of prominent syllables per run), primary stress (number of misplaced lexical stresses), and overall pitch range were examined.

In addition, nine tone choices were identified following Brazil's (1997) framework for the measures of pitch movement and level: high-rising, high-level, high-falling, mid-rising, mid-level, mid-falling, low-rising, low-level, and low-falling. In other words, the pitch measures included pitch levels of prominent syllables (high, medium, or low) and pitch movement within tone units (rising, level, or falling). Previous research has demonstrated the importance of tone choices on the focus word and pitch level as they can affect perceived information structure in L2 discourse (Kang et al., 2010; Pickering, 2001; Wennerstrom, 1994).

RESULTS

A step-wise multiple regression was performed with the four proficiency levels as a dependent variable and each of the four categories of pronunciation features as independent variables (segmental errors, fluency, stress and pitch, and tone choices). Table 1 illustrates the proportion of variance (R^2) for each of the four pronunciation categories.

Table 1

Proportion of Variance (R^2) Explained by Pronunciation Features

Pronunciation Features	R Square (R^2) Change
Fluency	.267
Stress and pitch	.309
Tone choice	.045
Segmental errors (high FL vowels and consonants, Low FL vowels and consonants, and 13 other segmentals)	.080
Total	.701

The results showed that 70% of the variance in Cambridge ESOL four proficiency levels was attributed to pronunciation errors. Among those features, fluency took up 26.7%, stress and pitch, 30.9%, tone choice, 4.5%, and segmental errors, 8%. Consequently, from the linguistic analysis perspectives in speaking assessment, we can hypothesize that there is a hierarchical priority in the effect of these pronunciation features (Figure 1) on L2 oral proficiency. That is, in assessing NNSs' oral performance, various pronunciation criterion features contribute to raters' judgments differently. Stress and pitch played a more important role than any other features, followed by fluency, segmental errors, and intonation tone choices.

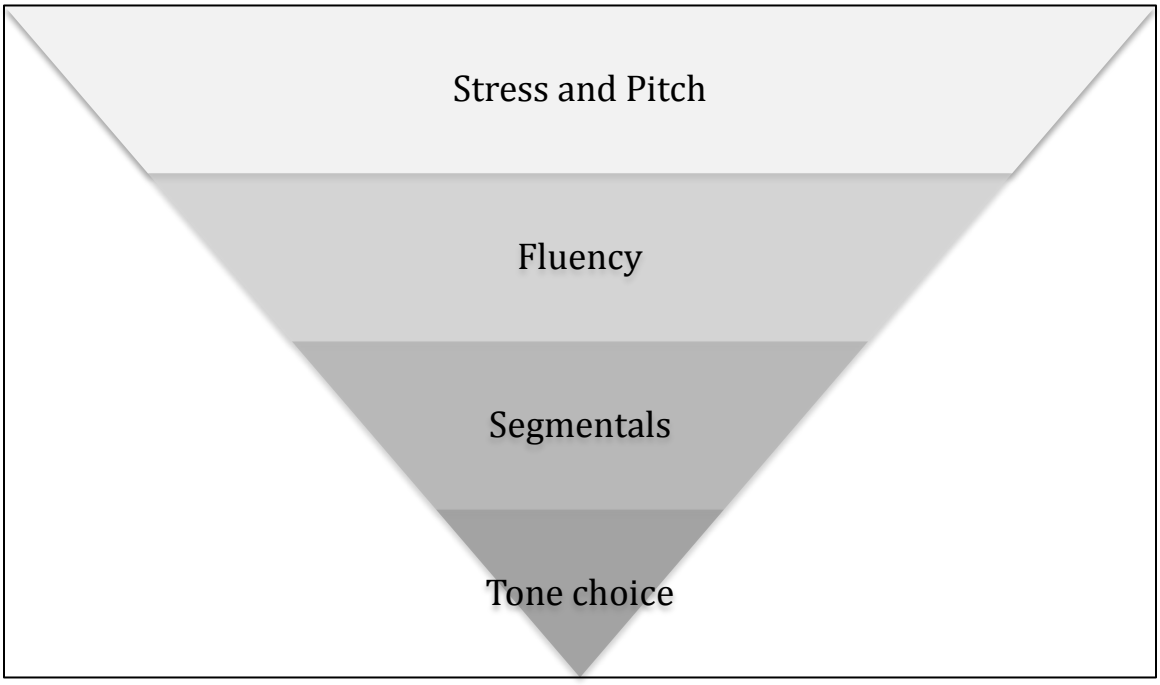

Figure 1. Hierarchical Priority in Pronunciation Features

DISCUSSION AND CONCLUSION

The results show the contribution of each of the pronunciation criterion features to the oral proficiency assessment. Note that most previous research has been limited in terms of the scope of the pronunciation features examined (e.g., Kang et al., 2010). In this study,

a variety of pronunciation variables (segmental errors, fluency, stress and pitch, and tone choices) were examined together and explained 70% of the variance. Fluency and stress/pitch features were strong contributors to this judgment (combined 56%) whereas segmental errors and intonation contributed more weakly to proficiency judgments.

One of the interesting findings is the strong contribution of stress and pitch variables on oral proficiency assessment. Thirty-one percent of the variance in Cambridge ESOL proficiency ratings was attributable to these stress and pitch factors, which is somewhat higher than the fluency dimension (27%). The significance of stress and pitch in NNSs' speech evaluation is in line with recent research examining pronunciation features on listeners' judgments of accented speech (Kang, 2010; Kang et al., 2010; Trofimovich & Isaacs, 2012). Trofimovich and Isaacs (2012) demonstrated that word stress influenced listeners' judgments (both accent and comprehensibility) in the case of rating the oral production of native French speakers of English. In Kang et al. (2010), stress was part of the acoustic fluency cluster, which was the most potent predictor of listeners' evaluation of NNSs' speech. In addition, pitch range was one of the strongest predictors of undergraduate students' ratings of an international teaching assistant's in-class lecture (Kang, 2010).

The importance of stress and pitch is particularly noteworthy. In speaking assessment, most studies have investigated the effects of temporal/fluency measures (speech rate and pauses) on listeners' judgments of NNSs' accented speech (e.g., Iwashita et al., 2008; Kormos & Denes, 2004; Ginther, Dimova, & Yang, 2010). These temporal measures are also frequently preferred by automatic speech recognition systems as objectively measurable parameters that show a high correlation with L2 fluency judgments (Zechner, Higgins, Xi, & Williamson, 2009). In this study, 27% of the variance in the proficiency ratings was explained by these fluency features. Nevertheless, the finding of the present study suggests that stress and pitch should be weighted more heavily in L2 pronunciation assessment. As Trofimovich and Isaacs (2012) stated, the influential contribution of stress to the oral judgments of accented speech can be related to "the fact that stress is one of the most structural and hierarchical aspects of phonology" (p. 10).

Overall, there appears to be a clear hierarchical structure in the importance of pronunciation features. In this study, stress and pitch were ranked first, followed by fluency measures, segmental errors, and tone choices at the end. Accordingly, ESL teachers may need to prioritize pronunciation features in classroom instruction for the promotion of intelligibility or for the preparation of high-stakes speaking tests. This knowledge can be further applied to develop scoring criteria for L2 oral assessment. Finally, future research in L2 pronunciation may provide new ways to operationalize stress or pitch as target variables to measure L2 speakers' oral production.

NOTE

The current study was presented at Pronunciation in Second Language Learning and Teaching 2012 under the title of *Pronunciation Features Distinguishing Examinees' Oral Performances at Different Proficiency Levels.* It examined salient pronunciation features that could distinguish Common European Framework of References (CEFR) speaking levels (B1-C2) in Cambridge ESOL General English Examinations. The full study appears in Cambridge ESOL *Research Notes 52* (2013). A comprehensive discussion of

the project funded by the Cambridge ESOL Research Program can be found in manuscripts currently under review for publication in other venues. Due to the constrained copyright of the articles, in the current proceedings, the author briefly reports one aspect of the findings, focusing on the relative impact of a wide array of pronunciation features on NNSs' oral proficiency.

ABOUT THE AUTHOR

Okim Kang is assistant professor of applied linguistics at Northern Arizona University. Her research concerns aspects of L2 pronunciation, speech perception and production, automated speech scoring, oral language proficiency assessment, language attitudes, and World Englishes. Email: okim.kang@nau.edu

Department of English
Liberal Arts Building 18
Room 140
Northern Arizona University
Flagstaff, Arizona, 86011-6032

Phone: 928.523.2059
Fax: 928.523.7074

REFERENCES

Brazil, D. (1997). The communicative value of intonation in English. Cambridge, UK: Cambridge University Press.

Fayer, J.M. & Krasinski, E.K. (1987) Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*, 313–26.

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27*, 379-399.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct?. *Applied Linguistics, 29*(1), 24-49.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of l2 comprehensibility and accentedness. *System, 38*, 301-315.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*, 554-566.

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164.

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

Munro, M. J. & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*, 520-531.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*(2), 233-255.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language And Cognition, 15*(4), 905-916.

Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of nonnative speakers. *Applied Linguistics* 15: 399-421.

Zechner, K., Higgins, D., Xi, X., Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication, 51*, 883-895.