PREDICTING ACCENTEDNESS: ACOUSTIC MEASUREMENTS OF CHINESE-ACCENTED ENGLISH

Vincent Porretta[1] and Benjamin V. Tucker[2]
[1]Dept. of Linguistics, University of Alberta, Alberta, Canada, T6G 2E7, porretta@ualberta.ca
[2]Dept. of Linguistics, University of Alberta, Alberta, Canada, T6G 2E7, bvtucker@ualberta.ca

1. INTRODUCTION

Listeners perceive foreign-accented speech as different from native speech because it deviates from native speaker acoustic targets. These deviations may occur across many acoustic dimensions such as word duration, vowel duration, vowel formant values, and voice onset time. Researchers have shown that non-native speakers produce longer and less variable word durations than native speakers and these measures are correlated with accentedness ratings [3]. Listeners are sensitive to these types of deviations and are capable of detecting accentedness in as little as a single 30 ms burst release [6]. Two previous studies employed distance measures to quantify these deviations in an attempt to model foreign-accented speech. For Arabic speakers of English, distance measurements for vowel duration, first and second formant values, and formant movement predict accentedness ratings of Arabic speakers' English vowels [9]. Similarly, for English speakers of Thai, distance measures of F0 valley and F2 values in diphthongs predict accentedness [12]. These findings suggest that native listeners compare non-native vowel tokens to the distributional properties of their learned native language.

The present study investigates the following questions: 1) How does Chinese-accented English differ from American English across gross acoustic measures such as word duration, vowel duration, and vowel formant values? 2) Which of these acoustic measures are most predictive of accentedness rating?

2. METHODS

The speech of nine male native Chinese speakers and one male native English speaker (low mean accentedness rating = 1.04) was analyzed for this study. Recordings and accentedness ratings were retrieved from the Wildcat Corpus of native- and foreign-accented English [11]. Forty-one monosyllabic words, each with three repetitions, were examined, totaling 123 tokens per talker. Measurements of word duration, vowel duration (of both monophthongs and diphthongs) and format values for F1-F3 (measured at the midpoint) were hand-measured in PRAAT [3] and extracted for analysis.

3. ANALYSIS AND RESULTS

Accentedness ratings were modeled using ordinary least squares linear regression performed in R using the *rms* package [8]. Prior to analysis, formant values were log transformed and plotted revealing considerable variation in vowel space. Additionally accented speakers produce word durations longer and shorter than the native speaker; however, they produce vowel durations longer than the native speaker. The vowel-to-word ratio was calculated by dividing the vowel duration of a word by the total duration of that word. If accentedness is a result of non-native productions approaching (to varying degrees) native-like acoustic targets, quantifying the distance from the native speaker norm allows examination of how variation along different variables affects perceived accentedness. For each numeric variable the token value of each non-native speaker was subtracted from that of the native speaker. The absolute value of that difference yielded a positive number representing the distance between the non-native production and a typical, native-like production. Collinearity between variables was verified and all numeric variables were centered.

These variables were used as input for the model. Restricted cubic splines (df = 5) were used for all three formant variables to account for nonlinearity of the variables. Interactions between variables were checked and included if they were significant. For this model two significant interactions resulted; the first between *Word Duration Distance* and *Vowel Duration Distance*, and the second between *log F1 Distance* and *log F2 Distance*. Outliers (less than 5% of the data) were identified and removed from the model following the procedures of Baayen [1] and the model was refitted. A significant effect resulted for Vowel-to-Word Duration Ratio (β = 1.9444, t (1127) = 2.75, p = .0061). As the ratio increases the predicted accentedness rating increases indicating that accentedness is in part predicted by the amount of a monosyllabic word that consists of a vowel. There is a similar result for the interaction between Word Duration Distance and Vowel Duration Distance (β = -38.3949, t (1127) = -2.38, p = .0174). As illustrated in Figure 1, high accentedness ratings result when speakers produce word durations far from the norm of a native speaker, but vowel durations that are similar to the norm. When both distances are high (i.e., lower ratio) accentedness ratings decrease, indicating that listeners may be sensitive to the ratio between vowel length and word length thus influencing their evaluation of accentedness. The interaction between log F1 Distance and log F2 Distance (Fig. 2) indicates that when the distances are equally high, accentedness rating increases. When F1 distance is low, but F2 distance is high, accentedness rating decreases. However, when F2 distance is low and F1 distance is high, accentedness rating increases dramatically. This indicates that there is a strong effect for F1 values that are far from native speaker norms which is perhaps a strong cue to listeners that the speaker has a foreign accent.

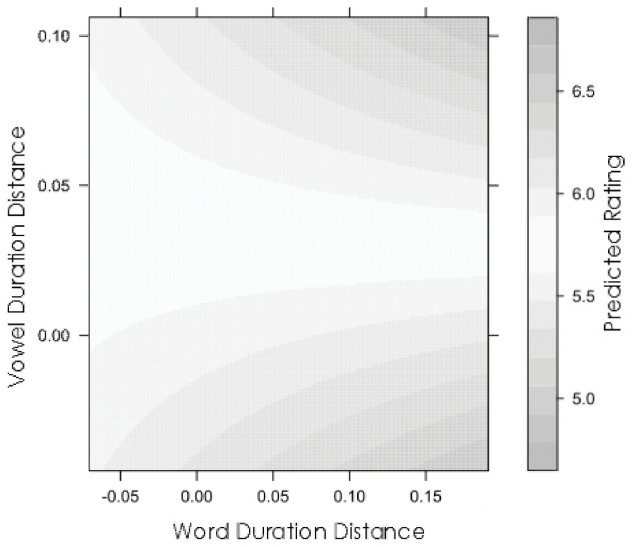

Figure 1: Contour plot of interaction between Vowel and Word Duration Distances for Predicted Rating.

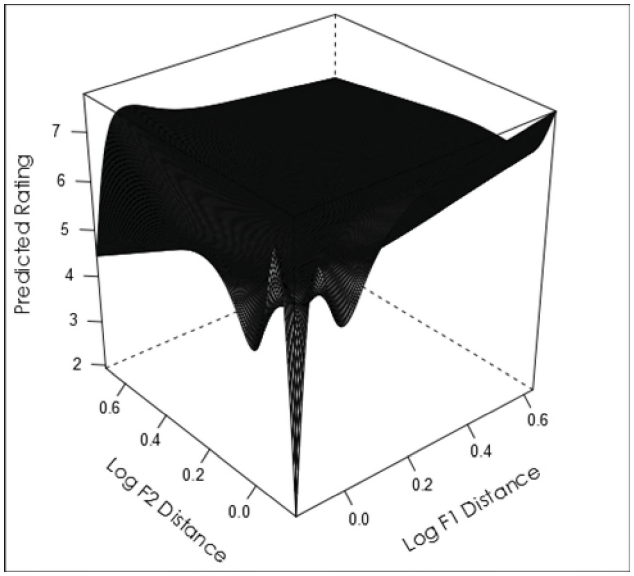

Figure 2: Perspective plot of interaction between log F1 and log F2 Distances for Predicted Rating

4. DISCUSSION AND CONCLUSIONS

This study demonstrates that foreign accentedness ratings of native Chinese speakers can be predicted using measures of acoustic distance. These findings replicate previous research [5, 6] that seek to model non-native vowel productions using measures of distance from typical native speaker values. The present results add to these findings by showing the relationship between word and vowel durations. For Chinese accented speech, vowel duration and word duration (and how they interact) play an important role in how listeners evaluate accentedness. Additionally, deviations of F1 from typical values have a strong effect on the degree of perceived accent. It appears that speakers who are more able to approximate typical native speaker values on these measures are perceived as less accented supporting the idea that, in evaluating degree of foreign accent, listeners compare gross acoustic features like word duration, vowel duration, and formant values to typical native speaker values. This suggests that listeners store distributional information about acceptable native-like productions with which they compare productions from new speakers [10]. By understanding which acoustic variables affect the perception of accentedness and how acoustic information is evaluated by listeners, it is possible to investigate their role in perceptual learning of accented speech. Current research on adaptation to foreign-accented speech [4, 5, 7] does not address the role of standard acoustic measurements, much less acoustic distance measures. Moreover, we plan to further investigate how these variables correlate with markers of cognitive processing when listening to accented speech of varying degrees.

REFERENCES

[1] Baayen, R.H. 2008. Analyzing linguistic data: A practical introduction to Statistics using R. Cambridge: Cambridge University Press.

[2] Baker, R.E., Baese-Berk, M., Bonnasse-Gahot, L., Kim, M., Van Engen, K.J., Bradlow, A.R. 2011. Word durations in non-native English. Journal of Phonetics, 39(1), 1-17.

[3] Boersma, P., Weenink, D. 2011. Praat: doing phonetics by computer. Version 5.3.04, retrieved 12 January 2012 from http://www.praat.org/

[4] Bradlow, A.R., Bent, T. 2008. Perceptual adaptation to non-native speech. Cognition. 106, 707-729.

[5] Clarke, C.M., Garrett, M.F. 2004. Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America, 116,* 3647–3658.

[6] Flege, J.E. 1984. The detection of French accent by American listeners. *The Journal of the Acoustical Society of America*, *76*(3), 692.

[7] Jongman, A., Wade, T., Sereno, J. 2003. On improving the perception of foreign-accented speech. In M. J. Sole, D Recasens, & J Romero (eds). *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, Spain,* 1561-1564. Barcelona, Spain: Universitat Autonoma de Barcelona.

[8] Harrell, F.E., Jr. (2012). rms: Regression Modeling Strategies. R package version 3.5-0. http://CRAN.R-project.org/package=rms

[9] Munro, M.J. 1993. Productions of English vowels native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech, 36(1),* 39–66.

[10] Pierrehumbert, J. (2001) Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds.) *Frequency effects and the emergence of lexical structure*. John Benjamins, Amsterdam. 137-157.

[11] Van Engen, K.J., Baese-Berk, M., Baker, R. E., Choi, A., Kim, M., Bradlow, A.R. 2010. The Wildcat Corpus of Native-and Foreign-accented English: Communicative Efficiency across Conversational Dyads with Varying Language Alignment Profiles. *Language and Speech*, *53*(4), 510 -540.

[12]Wayland, R. 1997. Non-native production of Thai: Acoustic measurements and accentedness ratings. *Applied Linguistics, 18(3),* 345-373.