Distinguishing Dysarthric Speech: Vowel Acoustics and Measurement

Annalise R. Fletcher [1], *Megan J. McAuliffe* [1], *Kaitlin L. Lansford* [2], *Julie M. Liss* [3]

[1] University of Canterbury, NZ, [2] Florida State University, USA, [3] Arizona State University, USA

annalise.fletcher@pg.canterbury.ac.nz

Index Terms: dysarthria, acoustics, vowel space area

1. Introduction

Centralization of vowel formants has been associated with reduced intelligibility in healthy speakers and those with motor speech disorders [1-3]. Hence, estimations of speakers' vowel space area (VSA) provide a promising avenue for assessing speech motor control. However, VSA measurements have traditionally demonstrated limited success in distinguishing healthy and disordered speech [1, 3]. One factor that may play a role is the formant measurement point. Research in speech motor control almost universally takes measurements of formant values from vowels' temporal midpoint. While this provides measurement consistency, the influence of vowels' phonetic context rarely results in symmetrical formant trajectories, and thus midpoint values exhibit considerable variation across word tokens [4].

There is also increasing evidence that a vowel's temporal midpoint reflects a different stage of articulation in faster and slower speech [4]. For people with motor speech disorders, significant reductions in speech rate are common. It has been hypothesized that the extraction of formant values from a static time point could be responsible, in part, for the failure of studies to consistently reveal reductions in the VSA of this population – despite evidence of reduced lip, tongue and jaw movements [1].

This project explores the extent to which different formant measurement points are capable of distinguishing healthy from mild dysarthric speech using common vowel centralization metrics.

2. Method

Forty nine speakers of New Zealand English (NZE) read a standard passage—17 healthy older individuals (6 females, 11 males, mean age: 67), and 32 speakers diagnosed with mild dysarthria (8 females, 24 males, mean age: 69).

Manual phoneme segmentation was completed in Praat using standard criteria and settings. Vowel centralization was calculated using F1 and F2 of the /a/, /i/ and /ɔ/ vowels for modern NZE, with three tokens of each vowel selected for analysis. Vowel formant measures were extracted from: (1) the vowels' temporal midpoint and (2) the vowel target. For the front vowel /i/, the vowel target was selected at peak F2 frequency; for /a/, where F1 reached its maximum; and for /ɔ/, the target point was taken at the lowest value of F2. From these data, the following metrics were calculated: (1) Vowel Space Area (VSA),

$$Hz^2 = 0.5 \times \mathrm{ABS}\,[F1i \times (F2a - F2ɔ) + F1/ɔ/ \times (F2i - F2a) + F1a \times (F2ɔ - F2i)], \quad (1)$$

where ABS = absolute value, F1i = first formant of the /i/ vowel, and so on, and (2) Formant Centralization Ratio (FCR), as adapted from Sapir, et al. [3].

$$(F2ɔ + F2a + F1i + F1ɔ) / (F2i + F1a) \quad (2)$$

3. Results

3.1. VSA

In males, vowel target values revealed a smaller mean VSA in speakers with dysarthria relative to controls, with the difference between groups approaching statistical significance ($t(22.69) = 2.00$, $p = .06$). Smaller VSA scores were also found using midpoint values, but this difference was not statistically significant ($t(22.51) = 0.73$, $p = .47$). For female speakers, vowel target values ($t(8.31) = 1.87$, $p = .10$), and equivalent midpoint values ($t(8.62) = 0.37$, $p = .72$) did not produce statistically significant differences in VSA.

3.2. FCR

In males, vowel target values provided a statistically significant difference in FCR between speaker groups ($t(31.28) = 2.72$, $p = .01$), with increased centralization in the dysarthric speech. Group differences were also observed using midpoint values, but these did not reach statistical significance ($t(28.90) = 0.99$, $p = .33$). For females, vowel target values revealed a statistically significant increase in centralization in speakers with dysarthria ($t(9.82) = 2.77$, $p = .02$). In contrast, equivalent midpoint values did not produce a difference in FCR between the two groups ($t(9.67) = 0.30$, $p = .77$).

4. Discussion

When vowel target values are applied, acoustic metrics demonstrate a greater ability to detect vowel centralization in speakers with dysarthria. This highlights the importance of the formant measurement point in assessing motor speech control.

5. References

[1] G. Weismer, J.-Y. Jeng, J. S. Laures, R. D. Kent, and J. F. Kent, "Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders," *Folia Phoniatrica et Logopaedica,* vol. 53, pp. 1-18, 2000.

[2] S. H. Ferguson and D. Kewley-Port, "Talker differences in clear and conversational speech: Acoustic characteristics of vowels," *Journal of Speech, Language, and Hearing Research,* vol. 50, pp. 1241-1255, 2007.

[3] S. Sapir, L. O. Ramig, J. L. Spielman, and C. Fox, "Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech," *Journal of Speech, Language, and Hearing Research,* vol. 53, pp. 114-125, 2010.

[4] K. Tjaden and G. Weismer, "Speaking-rate-induced variability in F2 trajectories," *Journal of Speech, Language, and Hearing Research,* vol. 41, pp. 976-989, 1998.