Umeå University

This is a submitted version of a paper published in *Journal of Voice*.

Citation for the published paper:

Karlsson, F., Doorn, J. (2012)

"Letter to the Editor"

Journal of Voice, 26(5): 676.e1

URL: <http://dx.doi.org/10.1016/j.jvoice.2011.08.010>

Access to the published version may require subscription.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-48349>

<http://umu.diva-portal.org>

Vowel articulation in Parkinson's disease: Erroneous calculation of vowel space area

Fredrik Karlsson^{1*} and Jan van Doorn¹

¹ Department of Clinical Sciences, Division of Speech and Language Pathology, Umeå University, Sweden

In a recent article, Skodda et al. (1) report the results from measurements of the acoustic working space of patients with Parkinson's disease (PD), as well as normal controls. While we commend the authors on their efforts to establish the most effective measures to demonstrate differences between controls and patients with PD, we wish to point out that the paper is marred by an error in the formula for calculation of the triangular vowel space area. As Skodda et al. have pointed out, the most established method for evaluating vowel space is the area of the vowel triangle formed by the vowels /i/, /u/ and /a/ in the F2-F1 plane. This area can be found using a standard geometric calculation of the area of a triangle with coordinates (F2i, F1i), (F2u, F1u) and (F2a, F1a) for its three vertices. Various equivalent formulae for this area can be found in coordinate geometry textbooks. The formula used by Sapir et al. (2) is probably the most common, i.e.

$$\text{Area} = |0.5 \cdot (F1i(F2a - F2u) + F1a(F2u - F2i) + F1u(F2i - F2a))|$$

---

* Corresponding author
Fredrik Karlsson
PhD
Department of Clinical Sciences/Speech and Language Pathology
Umeå University
SE-90185 Umeå
Sweden
Tel. +46 90 785 89 70
Email. fredrik.karlsson@ling.umu.se

There are, however, various other forms that can be proved to be mathematically equivalent to the formula above. It is one of these alternate forms that Skodda et al. (2011) have attempted to use. They have written the formula as (p 467)

$$tVSA = \frac{1}{2} \times [(F2u+F2i) \times (F1u-F1i) - (F2a +F2u) \times (\mathbf{F1a-F1u}) - (F2a+F2i) \times (F1a-F1i)] \quad (1)$$

where tVSA is the area of the triangular vowel space.

However, formula (1) is incorrect because F1a and F1u have been transposed in the middle term. (F1a-F1u) should be (F1u-F1a), to give

$$tVSA = \frac{1}{2} \times [(F2u+F2i) \times (F1u-F1i) - (F2a +F2u) \times (\mathbf{F1u-F1a}) - (F2a+F2i) \times (F1a-F1i)] \quad (2)$$

The proof to show that the revised formula is equivalent to that used by Sapir et al (2) is given in Appendix 1.

It should also be noted that this formula would have been more mathematically complete if it included absolute value symbols to indicate that the calculated area should always be positive.

We believe that the error in the formula has been carried through to the actual calculations in the Skodda et al. study (1), based on the magnitudes of the vowel space areas (tVSA) for the control subjects. The tVSA results for control subjects calculated by Skodda et al. deviate to a large extent in magnitude from what has been reported previously in the literature. Skodda et al. report the mean tVSA estimates to be 602,363 Hz² (SD 225,753 Hz²) for male controls and 1,089,856 Hz² (SD 258,019 Hz²) for female controls. Results from previous studies are 300,000 Hz² or less for both male and female speakers (2-4). Thus, the results presented by Skodda et al. deviate significantly from what has been found previously for normal, healthy speakers. We suspect that the reason for the deviation is the formula error rather than some type of population sampling anomaly.

In order to investigate effective difference between Skodda et al.'s formula for tVSA (Sk_tVSA), and the formulas provided elsewhere in the literature, a data set of 257 vowel spaces were constructed from /i/, /u/ and /a/ measurements obtained from the 181 normal controls (5-7) available in table format within the Praat computer program (8). Using data provided from Peterson et al. (5), separate vowel spaces were constructed for the first and second vowel listed with the same intended target. The data set included both male and female speakers, as well as children with both genders, and therefore covers a large range of possible vowel space areas obtainable in acoustic measurements.

Using this data set, it was first established that the algorithms used in the literature generally do not differ in their area estimates for the constructed data set. Comparisons of the Sapir et al. (2)'s formula and Heron's formula showed a maximum difference in results of $< 9.895 \times 10^{-10} \text{ Hz}^2$ most likely introduced by rounding errors. Thus, Sapir et al.'s formula is therefore used as a proxy for both of these formulas in the following treatment.

Although not identical to the normal control data used by Skodda et al., our data set appears to be similar to what has been obtained in the literature. As seen in Figure 1, tVSA values for the vowel spaces are comparable to what was found in previous studies (2-4) when applying Sapir et al.'s formula, and comparable to Skodda et al.'s results when applying their formula. Thus, we conclude that the vowel spaces that we have extracted fit reasonably well within the results obtained previously, and may therefore be used for the purpose of the present evaluation.

Figure 2 shows the difference between the absolute value of the Sk_tVSA compared to tVSA (Sapir's formula is used as a proxy). Three main observations are made from Figure 2. First, for all vowel space areas tested, the difference between Sk_tVSA and tVSA is substantially bigger than zero. Second, the difference between the two results sets increases for vowel spaces with larger tVSA according to the established algorithms. Third, the Sk_tVSA estimate introduces variability where there should be none. Together, the result

comparisons presented in Figure 2 suggest that Skodda et al.'s formula is an inaccurate estimate of tVSA, and that it introduces systematic errors and it may serve as a potent additional source of variance in data sets containing vowel space estimates, possibly depending on the shape of the vowel space. An increase in variance that is very likely to induce a Type II error of analysis.

While we support the view that tVSA may not be a good metric for estimating changes in vowel space, and that the vowel articulation index (VAI) may be preferable, we conclude that the data published in Skodda et al.'s article is flawed because the vowel space area metric is based on an incorrect formula. Thus the data need to be re-evaluated before any of the conclusions drawn by Skodda et al. could be considered to be backed by empirical evidence.

We would also like to point out that the metric formant centralization ratio (FCR) described by Skodda et al. (1) in their discussion as a 'further surrogate parameter'..... 'based on F1 and F2 measurements, but with different mathematical transformations' p. 471 is somewhat misleading. FCR is simply the inverse of VAI, and can thus not really be considered as fundamentally different from VAI.

Figure 1 Vowel space area estimates in the test data set, when the two different formulas are applied.

Figure 2 Absolute value of the difference between estimates of tVSA using Skodda et al.'s formula (Sk_tVSA) and Sapir et al.'s formula, plotted against the Vowel space area of the triangle (Sapir et al.'s estimate is used as a proxy for the concurring results).

References

1. Skodda S, Visser W, Schlegel U. Vowel Articulation in Parkinson's Disease. *Journal of Voice*. 2011 Jul.;25(4):467–472.
2. Sapir S, Ramig LO, Spielman JL, Fox C. Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech. *Journal of Speech, Language, and Hearing Research*. 2010;53(1):114–114.
3. Turner GS, Tjaden K, Weismer G. The Influence of Speaking Rate on Vowel Space and Speech Intelligibility for Individuals With Amyotrophic Lateral Sclerosis. *Journal of Speech and Hearing Research*. 1995;38(5):1001–1013.
4. Roy N, Nissen SL, Dromey C, Sapir S. Articulatory changes in muscle tension dysphonia: evidence of vowel space expansion following manual circumlaryngeal therapy. *Journal of communication disorders*. 2009;42(2):124–135.

5. Peterson GE, Barney HL. Control methods used in a study of the vowels. The Journal of the Acoustical Society of America. 1952;24(2):175–184.
6. Pols L, Tromp H. Frequency analysis of Dutch vowels from 50 male speakers. The Journal of the Acoustical Society of America. 1973; 53:1093-1101.
7. Weenink DJM. Formant analysis of Dutch vowels from 10 children. In: Proc. Inst. of Phonetic Sciences University of Amsterdam. Amsterdam: 1985. p. 45–52.
8. Boersma P, Weenink D. Praat, a system for doing phonetics by computer. Glot international. 2001;5(9/10):341–345.

Appendix 1

Please note that the equations are written using standard mathematical notation, where the areas are denoted A_{Sapir} , A_{Skodda} , A_{Corr} .

Proof

$$A_{\text{Sapir}} = |0.5 * (F1i(F2a-F2u) + F1a(F2u-F2i) + F1u(F2i-F2a))| \quad \text{Equation 1}$$

$$A_{\text{Skodda}} = 0.5 * (F2u + F2i) (F1u - F1i) - (F2a + F2u) (F1a - F1u) - (F2a + F2i) (F1a - F1i) \quad \text{Equation 2}$$

$$A_{\text{Corr}} = |0.5 * (F2u + F2i) (F1u - F1i) - (F2u + F2a) (F1u - F1a) - (F2a + F2i) (F1a - F1i)| \quad \text{Equation 3}$$

Expanding terms in Equation 3

$$\begin{aligned} A_{\text{Corr}} &= |0.5 * [(F2uF1u - F2uF1i + F2iF1u - F2iF1i) - (F2uF1u - F2uF1a + F2aF1u - F2aF1a) - (F2aF1a - F2aF1i + F2iF1a - F2iF1i)]| \\ &= |0.5 * (F2uF1u - F2uF1i + F2iF1u - F2iF1i - F2uF1u + F2uF1a - F2aF1u + F2aF1a - F2aF1a + F2aF1i - F2iF1a + F2iF1i)| \\ &= |0.5 * (-F2uF1i + F2iF1u + F2uF1a - F2aF1u + F2aF1i - F2iF1a)| \end{aligned}$$

Collecting terms

$$\begin{aligned} &= |0.5 * (F1i(F2a-F2u) + F1a(F2u-F2i) + F1u(F2i-F2a))| \\ &= A_{\text{Sapir}} \quad \text{QED} \end{aligned}$$