Acoustic characteristics of vowel productions and vowel space area in individuals with depression

Erin Miley
Queen Margaret University

Abstract

Background: The current study examined the effects of depression on vowel production and vowel space area (VSA). Depressed speech was compared to that of healthy controls to determine whether vowels differed quantitatively according to the first (F1) and second (F2) formants. VSA measures were also analysed in relative to depression severity to determine whether these were clinically meaningful indicators of affective state.

Research design and methods: Vowels were extracted from the read speech of nine depressed and nine healthy participants over four recording time points. Raw formant values were normalised and plotted to compare group means. VSA measures were examined in relation to depression severity using Kendall's tau. The predictive value of diagnostic group for formant measurements (F1 and F2) were assessed using linear mixed effect models.

Results: Variable differences emerged for vowel types between depressed and healthy participants. Significant interactions emerged for group and vowel type, with groupwise mean comparisons indicating that low vowels /æ/ and /ɑ/ are disproportionately affected by depressive status. Global VSA measures were not significantly correlated with depression severity scores due to these asymmetric effects.

Discussion and implications: Specific vowels may be linked to depressive diagnosis based on a combination of motor incoordination and facial tension, particularly with respect to jaw and tongue height. Global measures of VSA were not sensitive indicators of depressive symptom severity.

1. Introduction

Depression is consistently proving to be one of the most common mental health conditions in our lifetime, with prevalence rates reaching as high as 27% in a recent meta-analysis (Wang et al., 2017). In particular, a significant portion of these symptoms were prevalent in clinical outpatients as compared to healthy control, with rates peaking around individuals aged 30–40 years old, and again at 80–90 years of age. Given the fact that such a high number of individuals will experience a depressive episode in their lifetime, it is becoming ever more apparent that accurate diagnosis of depression is necessary to initiate earlier interventions and management strategies.

The voice has long been considered a prime candidate for insight into mental health status (Hall et al., 1995; Sobin & Sackeim, 1997), effectively providing information in a low-cost manner both in terms of equipment and effort on the part of the speaker. Previous research into the depressed voice has centred around investigations of pitch (e.g. Alpert et al., 2001; Cannizzaro et al., 2004; Darby et al., 1984; Kuny & Stassen, 1993; Mundt et al., 2007; Nilsonne et al., 1988; Quatieri & Malyska, 2012; Tolkmitt et al., 1982; Yang et al., 2012), loudness (Alpert et al., 2001; Darby et al., 1984; Kuny & Stassen, 1993; Stassen et al., 1991) and pausing behaviours (Alpert et

AL., 2001; CANNIZZARO ET AL., 2004; MUNDT ET AL., 2007, 2012; STASSEN ET AL,. 1998; TREVINO ET AL., 2011). Little consensus has been achieved with regards to any of the aforementioned parameters, as individual variability and heterogeneous research designs make for difficult comparisons of the effect of depression on speech between studies. For example, an early study by Darby and colleagues (1984) reported pre- and post-intervention differences in loudness variation, with levels significantly increasing with response to treatment. However, these results were only found to a mild extent (KUNY & STASSEN, 1993), or non-significantly (ALPERT ET AL., 2001) in more recent investigations. A lack of consistency in inclusion criteria diagnostic cut-offs, speech tasks and recording conditions all may be impacting the generalisability of these results for future research. The most promising feature does appear to be speech pausing time (cf. CUMMINS ET AL., 2015), with depressed speakers often engaging in more frequent and longer pauses as compared to healthy controls; this behaviour also changes in response to treatment (MUNDT ET AL., 2007, 2012).

More recently, fine-grained analysis into phoneme production and articulation have emerged as potentially promising variables in investigations of depressive speech, linking in with psychomotor symptoms commonly seen in depression and other mental health diagnoses. Imaging studies have demonstrated that motor coordination may be impaired in depression due to a decrease of dopamine in the nigrostriatal pathway and basal ganglia (e.g. DAILLY ET AL., 2004), causing downstream impairments in initiating movement, maintaining motor control and abnormal muscle tension (CALIGIURI & ELLWANGER, 2000; WALTHER ET AL., 2012). These changes have compounding effects on speech, resulting in decreased articulator and laryngeal control, and can be exhibited in findings such as increased perturbation measures (i.e. jitter and shimmer) in depressed as compared to healthy speakers (OZDAS ET AL., 2004; QUATIERI & MALYSKA, 2012), as well as articulatory undershoot evidenced through decreased F1 measurements in treatment-resistant depressed individuals (MUNDT ET AL., 2012)

Specifically with regard to articulatory control in vowel productions, studies have shown that individuals with depression exhibit formant differences from healthy controls. An early study by FLINT and colleagues (1993) demonstrated a reduction in F2 transitions on the vowels in 'light' and 'dial' ([aɪ] and [aɪe]) by depressed speakers. A more recent study (France et al., 2000) analysed existing audiotape recordings of interviews being unmedicated depressed women and their therapists, the latter of which were used as healthy controls. Results suggest that women with depression produced elevated first formant frequencies as compared to this healthy control group, an articulatory pattern linked to less opening of the jaw and/or tongue (LINDBLOM & SUNDBERG, 1971). The authors posited these changes may have been due to an increase in muscle tone (which would particularly affect the jaw) and rigidity in the depressed women's vocal tracts, potentially affecting the ability to meet acoustic targets along the first formant axis and resulting in the demonstrated findings of centralised vowel productions

Vowel Space Area (VSA) measures have also emerged in recent years as a promising estimation of speech motor control. Measured using the triangular (LIU ET AL., 2005) or quadrilateral corner vowels (GOBERMAN & ELMER, 2005), VSA can be utilised to observe compression or expansion of the vowel space within—as well as between—speakers to track systematic variation using the Euclidean distance between coordinates in the first formant (F1) and second formant (F2) space. Vowel space compression has been linked to reduced intelligibility or casual speech (BRADLOW ET AL., 1996; FOURAKIS, 1991). To date, one previous study has investigated VSA in depressed speech (SCHERER ET AL., 2015), with results suggesting that depressed speech in both conversational and read conditions demonstrating decreased overall space. However, this study utilised only 3 corner vowels to develop the measure (/i/, /æ/ and

/u/), neglecting to include other vowels in the formation of the overall space, and only analysed the speech of North American English speakers.

1.1. Current exploratory study

The current study aims to assess the clinical utility of vowel formant and VSA measures as sensitive indicators of depression in speakers of British English, and will answer the questions:

- Do individuals with depression produce vowels that are significantly different (with respect to F1 and F2) as compared to healthy controls?
- Is there evidence to suggest that VSA measures are clinically meaningful indicators of depression and depressive severity?

2. Participants and Methods

2.1. Participants

Recordings were utilised from a pilot acoustic analysis study investigating depressed speech at Newcastle University. As such, there was no control for dialect in recruitment of participants; however, the majority involved had spent the majority of their lives in the Newcastle or Sunderland area, and had a Tyneside, Northwestern or Southeast Scottish accent. The wide variety of accents was balanced amongst groups and genders. All speakers were also native monolingual English speakers. In this study, nine depressed (five females, four males) and nine non-depressed controls (five females, four males) were recruited for participation. The control group consisted of adults aged 18–65 with no history of speech disorders or mental illness, as assessed by a psychiatric interview tool (MINI; Sheehan et al., 1998). The depressed group consisted of adults aged 18–65 with a diagnosis of depression without history of bipolar disorder. Descriptive information concerning age, baseline depression severity score and dialect are described in Table 1. The Hamilton Depression Rating Scale (HAMD-17; HAMILTON, 1960) was used to assess baseline severity; scores can range from 0–52, with higher scores denoting worse symptomatology.

2.2. Experimental procedure

This exploratory study was a longitudinal, comparison-based study, with four time points of data collection spread out over 26 weeks. At baseline, written consent from each participant was obtained, and a HAMD-17 (HAMILTON, 1960) and QIDS-SR16 (RUSH ET AL., 2006) were completed. The QIDS-SR16 is a 16-item self-report survey asking about symptoms such as appetite, feelings of guilt and sadness and psychomotor activity; larger scores represent more severe symptomatology. This was followed by a short five minute speech recording. At 8-, 16- and 26-weeks post baseline, participants were asked to revisit the clinic to complete the QIDS-SR16 and another five-minute speech recording. It should be noted that several participants completed less than four recordings for the final analysis due to missed appointments and/or an unacceptable amount of background in the recording.

	Depressed group	Healthy group
No. of subjects	9	9
No. of males/females (M/F)	4M/5F	4M/5F
Age (years)	46.3 ($\pm$ 9.8)	42.6 ($\pm$ 12.5)
HAMD-17 baseline score	15.6 ($\pm$ 5.2)	1.7 ($\pm$ 1.7)
Dialects	2 Tyneside English (1M/1F)	2 Tyneside English (2M)
	6 General Northern English (3M/3F)	6 General Northern English (4M/2F)
	1 Southeastern Scottish (1F)	1 Southeastern Scottish (1F)
Ethnicity	8 White British	9 White British
	1 British 'Other'	
Minimum educational attainment	All college educated or further education	

Table 1: Characteristics of recruited participants.

For the purposes of this study, vowel formant measurements were averaged over the trimmed 80% central portion of each labelled segment. Recordings were run through a forced alignment program (pyalign), an implementation of the Penn Forced Aligner (YUAN & LIBERMAN, 2008) and visually checked in Praat (BOERSMA & WEENINK, 2019). A custom script in Praat provided the values for each vowel produced in a given speech task, per participant. Each vowel segment was perceptually evaluated and checked against the alignment program's assigned ARPABET label to assess accuracy. A total of 2,128 tokens were extracted from the depressed participants' speech, and 2,172 from the healthy participants' speech (51% and 49% of the total vowel tokens analysed, respectively).

2.3. Statistical analyses

All statistical analyses were completed in R. After removing outliers which lay outwith 2 standard deviations from the individual's vowel means, a linear mixed effects model was used to assess difference in each of the formants (e.g. F1 and F2). Fixed effects included diagnostic group, vowel type, vowel duration and one interaction term between Group and Vowel. Speaker and word context were included as random effects.

Correlational analyses were also completed using Kendall's τ (KENDALL & GIBBONS, 1990) to assess correlations between Polygon and Hull Area measures and QIDS-SR16 scores within the read speech samples. Although less powerful of a test than Pearson's *r*, Kendall's τ is less sensitive to non-parametric distributions.

2.4. Acoustic variables

To negate coarticulation effects, F1 and F2 values were extracted from the central 80% of the vowels /i æ u o ɒ ɔ/ (cf. KENT & VORPERIAN, 2018). All F1 and F2 values were normalised according to the Lobanov procedure (LOBANOV, 1971) which directly transforms raw Hz into

z-scores.

VSA was measured using the polygon area produced by the mean values for each vowel (e.g. FOX & JACIEWICZ, 2008) as well as the area of the convex hull defined as the area encompassing all vowel tokens (SANDOVAL ET AL., 2013).

2.5. Recording procedure

All recordings were made using a Neumann i89 condenser microphone with an omnidirectional polar pattern in a small clinical interview room. The microphone to mouth distance was maintained at 20 cm. Participants were asked to read the 'Dog and Duck' story (BROWN & DOCHERTY, 1995), read out nine phonetically balanced sentences and speak for 30 seconds about an emotionally neutral topic at each data collection timepoint.

3. Results

3.1. Vowel formant differences

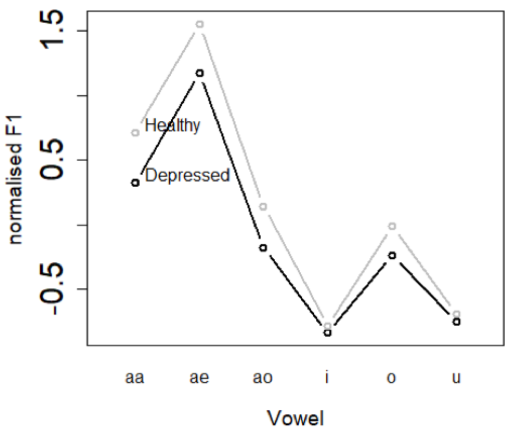
Fixed factors were assessed separately to check impact on the overall goodness-of-fit to the dataset. Having a diagnosis of depression significantly predicted F1 (b = -0.378, $t(21.67)$ = -3.408, $p < .001$), with the coefficient denoting a negative relationship between diagnosis and this formant; depressed speakers produced lower F1 values as compared to healthy counterparts.

A significant interaction predictive of F1 was Group $\times$ Vowel. A groupwise comparison of means (using the Tukey method; TUKEY, 1949) provided insight into which particular vowels were affected in depression, given the fact that asymmetrical differences were ascertained from visual inspection of the vowel plots. As demonstrated in Figure 1, low vowels /æ ɒ/ were significantly different between depressed and healthy speakers, with depressed speakers producing these with lower F1 values (i.e. raised productions).

A secondary linear mixed effect model for F2 exhibited no predictive value of including Group as a predictive factor (b = -0.02 , $t(19.95)$ = -0.187, $p < .05$).

3.2. VSA measures

Results from correlational analyses demonstrate no significant correlations between either of the VSA measures and depression severity scores (τ = .056, $p < .05$ and τ = .222, $p < .05$, for polygon area and hull area, respectively).



Vowel	b	SE	*t*	*p*
/ɒ/	0.378	0.111	3.408	0.03
/ae/	0.362	0.109	3.323	0.04
/ɔ/	0.302	0.113	2.683	ns
/i/	0.031	0.110	0.283	ns
/o/	0.201	0.110	1.822	ns
/u/	0.042	0.112	0.376	ns

Figure 1: Normalised differences in F1 by diagnostic group and vowel type, with a summary of vowel type as a fixed effect. ARPABET symbols are utilised on the x-axis, where *aa* refers to /ɒ/ and *ao* refers to /ɔ/.

4. Discussion

4.1. Vowel formant differences

Results support previous findings (e.g. France et al., 2000; Mundt et al., 2012) that the first formant is particularly salient as an indicator of depression in speech. From visual inspection of the vowel plot alone, it is apparent that depression affects vowels asymmetrically, with certain vowels being produced at the same location in the oral cavity and others demonstrating substantial differences. This finding echoes that of a similar study carried out in Parkinson's disease patients (Bang et al., 2013), where centralisation of vowels were seen on some vowels (/a e i/) but not others (/u/). Authors attributed the overall reduction in vowel space area to be due to the effects of these specific vowels, possibly as a side effect of articulatory undershoot due to reduced range of movement.

Muscle tension also may be a causal factor in the decreased first formant of low vowels due to a reduction in jaw opening and corresponding changes to tongue height. Previous research has linked depression with decreased facial muscle activity (Gehricke & Shapiro, 2000), although this was specifically inspected over the brow and cheek region of the face, rather than the jaw. It may be that facial tension extends across regions of the face, and upper facial tension may also coincide with tension in the jaw and articulators.

4.2. VSA measures

Given the findings of asymmetric vowel effects, it is not surprising then to find that Polygon and Hull global VSA measures were not related to any of the depressed participants' QIDS-SR16 scores. Results do align with findings by Scherer and colleagues (2015), whereby VSA was not systematically reduced in individuals with depression as compared to healthy controls in read speech, although that study did produce significant findings between groups under conversational speech conditions. It may be that speech task itself plays an important role in the articulatory behaviour of depressed individuals, with read speech possibly being hyperarticulated or spoken in a clearer voice, thus negating any effects the mental health condition may have on total vowel space (Bradlow et al., 1996).

5. Conclusion

This study was the first known cross-dialectal investigation into assessment of vowel productions within depressed speech as compared to healthy individuals. Findings provide support to the theory that depressive abnormal vowel articulations affect formants asymmetrically, with F1 appearing to be demonstrably different between healthy and depressed individuals. Global measures of VSA and ED may not be applicable in depressive speech investigations given these asymmetric vowel effects. More sensitive measures of vowel productions, possibly based on F1 only, may be more clinically meaningful indicators of worsening depressive symptomatology.

Future research may wish to tease further assess this distinction between males and females in the expression of psychomotor behaviours. Previous research has determined gender-based differences in vowel articulation (e.g. Simpson, 2001, 2002, 2003; Weirich & Simpson, 2014a, 2018), as well as vowel space area (Hillenbrand et al., 1995; Simpson & Ericsdotter, 2007; Weirich & Simpson, 2014b) and vowel duration (Weirich & Simpson, 2018). Studies have linked these findings to variation in pharyngeal constriction, which may be more complete in male speakers, while also pointing to female speakers' tendency to use more of their articulatory space in speech than males (Weirich & Simpson, 2018). The current findings provide further supportive evidence for gender-based differences in articulatory targets between male and female speakers.

This study was limited in the fact that there was no control for dialect within the groups; although each group had fairly even distribution of accents, it would be prudent to revisit these findings in samples with one identifiable dialect.

Acknowledgements

This study was jointly funded by the Academic Psychiatry Unit at the Wolfson Research Centre, Newcastle University and the Clinical Audiology Speech and Language Research Centre, Queen Margaret University. The author wishes to thank the participants involved in the study, as well as the staff at the Ageing Research Laboratories for their continued support.

References

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders*, 66(2), 59–69.

Bang, Y. I., Min, K., Sohn, Y. H., & Cho, S. R. (2013). Acoustic characteristics of vowel sounds in patients with Parkinson's disease. *Neurorehabilitation*, 32, 649–654.

Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer [version 6.0.30]. Retrieved from `http://www.praat.org`

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I:

Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20(3), 255–272.

BROWN, A. & DOCHERTY, G. J. (1995). Phonetic variation in dysarthric speech as a function of sampling task. *European Journal of Disorders of Communication*, 30(1), 17–35.

CALIGIURI, M. P., & ELLWANGER, J. (2000). Motor and cognitive aspects of motor retardation in depression *Journal of Affective Disorders*, 57(1–3), 83–93.

CANNIZZARO, M., HAREL, B., REILLY, N., CHAPPELL, P., & SNYDER, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition*, 56(1), 30–35.

DAILLY, E., CHENU, F., RENARD, C. E., & BOURIN, M. (2004). Dopamine, depression and antidepressants. *Fundamental & Clinical Pharmacology*, 18(6), 601–607.

DARBY, J. K., SIMMONS, N., & BERGER, P. A. (1984). Speech and voice parameters of depression: A pilot study. *Journal of Communication Disorders*, 17(2), 75–85.

FLINT, A. J., BLACK, S. E., CAMPBELL-TAYLOR, I., GAILEY, G. F., & LEVINTON, C. (1993). Abnormal speech articulation, psychomotor retardation and subcortical dysfunction in major depression. *Journal of Psychiatric Research*, 27(3), 309–319.

FOURAKIS, M. (1991). Tempo, stress and vowel reduction in American English. *Journal of the Acoustical Society of America*, 90(4), 1816–1827.

FOX, R. A., & JACEWICZ, E. (2008). Analysis of total vowel space areas in three regional dialects of American English. *Proceedings of Acoustics '08 Paris*, 495–500.

FRANCE, D. J., SHIAVI, R. G., SILVERMAN, S., SILVERMAN, M., & WILKES, M. (2000). Acoustical properties of speech as indicators of depression and suicidal risk. *IEEE Transactions on Biomedical Engineering* , 47(7), 829–837.

GEHRICKE, J., & SHAPIRO, D. (2000). Reduced facial expression and social context in major depression: discrepancies between facial muscle activity and self-reported emotion. *Psychiatry Research*, 95(2), 157–167.

GOBERMAN, A. M., & ELMER, L. W. (2013). Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease. *Journal of Communication Disorders*, 38(3), 215–230.

HALL, J. A., HARRIGAN, J. A., & ROSENTHAL, R. (1995). Nonverbal behavior in clinician-patient interaction. *Applied and Preventive Psychology*, 4(1), 21–37

HAMILTON, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23(1), 56–62.

HILLENBRAND, J., GETTY, L. A., CLARK, M. J., & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *European Journal of Disorders of Communication*, 97(5), 3099–3111.

KENDALL, M, & GIBBONS, J. D. (1990). *Rank Correlation Methods* (5th Ed.). Oxford, UK: Oxford University Press.

KENT, R.D., & VOPERIAN, H. K. (2018). Static measurements of vowel formant frequencies and bandwidths: A review. *Journal of Communication Disorders*, 74, 74–97.

KUNY, S., & STASSEN, H. H. (1993). Speaking behaviour and voice sound characteristics in depressive patients during recovery. *Journal of Psychiatry Research*, 27(3), 289–307.

LINDBLOM, B. E., & SUNDBERG, J. E. (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America*, 50(4), 1166–1179.

LIU, H., TSAO, F., & KUHL, P. (2005). The effect of reduced vowel space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of the Acoustical Society of America*, 117(6), 3879–3889.

LOBANOV, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49(2B), 606–608.

MUNDT, J. C., SNYDER, P. J., CANNIZZARO, M. S., CHAPPIE, K., & GERALTS, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, 20(1), 50–64.

MUNDT, J. C., VOGEL, A. P., FELTNER, D. E., & LENDERKING, W. R. (2012). Vocal acoustic biomarkers of depression severity and treatment response. *Biological Psychiatry*, 72(7), 580–587.

NILSONNE, A., SUNDBERG, J., TERNSTROM, S., & ASKENFELT, A. (1988). Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. *Journal of the Acoustical Society of America*, 83(2), 716–728.

OZDAS, A., SHIAVI, R. G., SILVERMAN, S. E., SILVERMAN, M. K., & WILKES, D. M. (2004). Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk. *IEEE Transactions on Biomedical Engineering*, 51(9), 1530–1540.

QUATIERI, T. F., & MALYSKA, N. (2012). Vocal-source biomarkers for depression: A link to psychomotor activity. *Proceedings of Interspeech 2012*, 1059–1062.

R DEVELOPMENT CORE TEAM (2007). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from `http://www.R-project.org`

RUSH, A. J., BERNSTEIN, I. H., TRIVEDI, M. H., CARMODY, T. J., WISNIEWSKI, S., MUNDT, J. C., ... FAVA, M. (2006). An evaluation of the quick inventory of depressive symptomatology and the hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. *Biological Psychiatry*, 59(6), 493–501.

SANDOVAL, S., BERISHA, V., UTIANSKI, R. L., LISS J. M., & SPANIAS, A. (2013). Automatic assessment of vowel space area. *Journal of the Acoustical Society of America*, 134(5), EL477–EL483.

SAPIR, S., POLCZYNSKA, M., & TOBIN, Y. (2009). Why does the vowel space area as an acoustic metric fail to differentiate dysarthric from normal vowel articulation and what can be done about it? *Poznan Studies in Contemporary Linguistics*, 45(2), 301–311.

SCHERER, S., MORENCY, L. P., GRATCH, J., & PESTIAN, J. (2015). Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4789–4793.

SHEEHAN, D. V., LECRUBIER, Y., SHEEHAN, K. H., AMORIM, P., JANAVS, J., WEILLER, E., ... DUNBAR, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, 59(Supplement 20), 22–33.

SIMPSON, A. P., & ERICSDOTTER, C. (2003). Sex-specific durational differences in f0 and vowel space. *Proceedings of the 16th International Congress of Phonetic Sciences*, 933–936.

SIMPSON, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *Journal of the Acoustical Society of America*, 109(5), 2153–2164.

SIMPSON, A. P. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics*, 30(3), 417–435.

SIMPSON, A. P. (2003). Possible articulatory reasons for sex-specific differences in vowel duration. *Proceedings of the 6th International Seminar on Speech Production*, 261–266.

SOBIN, C., & SACKEIM, H. (1997). Psychomotor symptoms of depression. *American Journal of Psychiatry*, 154(1), 4–17.

STASSEN, H. H, KUNY, S., & HELL, D. (1998). The speech analysis approach to determining onset of improvement under antidepressants. *European Neuropsychopharmacology*, 8(4), 303–310.

STASSEN, H. H., BOMBEN, G., & GUNTHER, R. (1991). Speech characteristics in depressio. *Psychopathology* , 24, 88–105.

TJADEN, K., LAM, J., & WILDING, G. (2013). Vowel acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear, loud and slow speaking conditions. *Journal of Speech, Language, and Hearing Research*, 56(5), 1485–1502.

TOLKMITT, F., HELFRICH, H., STANDKE, R., & SCHERER, K.R. (1982). Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *Journal of Communication Disorders*, 15(3), 209–222.

TREVINO, A. C., QUATIERI, T. F., & MALYSKA, N. (2011). Phonologically-based biomarkers for major depressive disorder. *EURASIP Journal on Advances in Signal Processing*, 2011(1), 42.

TUKEY, J. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2), 99–114.

WALTHER, S., HOFLE, O., FEDERSPIEL, A., HORN, H., HUGLI, S., WIEST, R., ... MULLER, T. J. (2012). Neural correlates of disbalanced motor control in major depression. *Journal of Affective Disorders*, 136(1–2), 124–133.

WANG, J., WU, X., LAI, W., LONG, E., ZHANG, X., LI, W., ... LIN, H. (2017). Prevalence of depression and depressive symptoms among outpatients: a systematic review and meta-analysis. *BMJ Open*, 7, e017173.

WEIRICH, M., & SIMPSON, A. P. (2014a). Articulatory vowel spaces of male and female speakers. *Proceedings of the 10th International Seminar on Speech Production (ISSP*, 453–456.

WEIRICH, M., & SIMPSON, A. P. (2014b). Differences in acoustic vowel space and the perception of speech tempo. *Journal of Phonetics*, 43, 1–10.

WEIRICH, M., & SIMPSON, A. P. (2018). Individual differences in acoustic and articulatory undershoot in a German diphthong – Variation between male and female speakers. *Journal of Phonetics*, 71, 35–50.

YANG, Y., FAIRBAIRN, C., & COHN, J. (2012). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing*, 4(2), 142–150.

YUAN, J., & LIBERMAN, M. (2008).Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics '08*, 5687–5690.

Erin Miley
CASL Research Centre
Queen Margaret University
Edinburgh
United Kingdom
EMiley@qmu.ac.uk