REDUCED VOWEL SPACE IS A ROBUST INDICATOR OF PSYCHOLOGICAL DISTRESS: A CROSS-CORPUS ANALYSIS

Stefan Scherer[†], *Louis-Philippe Morency*[†], *Jonathan Gratch*[†], *John Pestian*[‡]

[†]University of Southern California
Institute for Creative Technologies
Playa Vista, CA

[‡]Cincinnati Children's Hospital Medical Center
Department of Pediatrics, Biomedical Informatics
Cincinnati, OH

ABSTRACT

Reduced frequency range in vowel production is a well documented speech characteristic of individuals' with psychological and neurological disorders. Depression is known to influence motor control and in particular speech production. The assessment and documentation of reduced vowel space and associated perceived hypoarticulation and reduced expressivity often rely on subjective assessments. Within this work, we investigate an automatic unsupervised machine learning approach to assess a speaker's vowel space within three distinct speech corpora and compare observed vowel space measures of subjects with and without psychological conditions associated with psychological distress, namely depression, post-traumatic stress disorder (PTSD), and suicidality. Our experiments are based on recordings of over 300 individuals. The experiments show a significantly reduced vowel space in conversational speech for depression, PTSD, and suicidality. We further observe a similar trend of reduced vowel space for read speech. A possible explanation for a reduced vowel space is psychomotor retardation, a common symptom of depression that influences motor control and speech production.

Index Terms— Depression, PTSD, Suicide, Vowel space, Conversational speech

1. INTRODUCTION

Hypoarticulated voice or reduced frequency range in vowel production are well documented speech characteristics of individuals' suffering from psychological and neurological disorders, including but not limited to depression [1, 2], cerebral palsy [3], amyotrophic lateral sclerosis [4], and Parkinson's disease [5]. The assessment and documentation of hypoarticulation and reduced vowel space often either rely on subjective assessments or on analysis of speech under constrained laboratory conditions (e.g. sustained vowel production), rendering analysis impractical and expensive [6]. Within this work we seek to confirm prior findings of hypoarticulation in depression automatically within conversational speech and across a wider spectrum of *psychological distress*, comprising post-traumatic stress disorder (PTSD) as well as suicidality. We seek to support clinicians and healthcare providers with much needed complementary, quantified, and objective measures of nonverbal behavior and in particular robust voice characteristics to allow for a more informed and objective assessment of an individual's health status [7].

The work depicted here is sponsored by the U.S. Army Research Laboratory (ARL) under contract number W911NF-14-D-0005. Statement and opinions expressed and content included do not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. Corresponding Author: Stefan Scherer; scherer@ict.usc.edu

The present work extends findings on the relationship between depression and vowel space in conversational speech [8]. The proposed automatic unsupervised machine learning approach to assess a speaker's vowel space - defined as the frequency range spanned by the first and second formant of the vowels /i/ (as in *heed*), /a/ (as in *hod*), and /u/ (as in *who'd*) with respect to the reference population - is evaluated across three distinct datasets as well as analyzed with respect to three psychological conditions, namely depression, PTSD, as well as suicidality. We relate the measure of a speaker's vowel space to psychomotor retardation a common symptom and observed hypoarticulation in individuals with depression. Specifically, our approach is based on an accurate voiced-speech detector, a robust formant tracker, and a subsequent vector quantization step using the k-means algorithm.

We set out to answer three research questions and investigate the robustness of automatic vowel space assessment across multiple datasets and recording conditions. In particular, we address the following research questions within this work:

Q1: Based on subjective, manually assessed observations from prior work, we aim to confirm if automatically assessed vowel space of subjects with depression is significantly reduced when compared to subjects without depression. In particular, psychomotor retardation - a common symptom in depression - is hypothesized to have an impact on an individual's vowel space due to its effect on motor control and speech production.

Q2: We investigate if a reduced vowel space can be observed in data recorded in different recording conditions. We seek to confirm that vowel space is reduced in depression for both conversational as well as read speech.

Q3: Last, we investigate if vowel space has implications on often co-morbid psychological conditions associated with depression. In particular, we investigate vowel space with respect to post-traumatic stress disorder (PTSD) as well as suicidality in adolescents.

The remainder of the paper is organized as follows: Section 2 discusses prior investigations on speech characteristics related to neurological and psychological disorders. Section 3 then describes the automatic approach and utilized datasets in detail. Finally, Section 4discusses and summarizes our findings and concludes the paper.

2. RELATED WORK

The analysis of acoustic characteristics of speech in depression received considerable attention in the past: Investigations revealed reduced speech variability and monotonicity in loudness and pitch [9, 10], reduced speech [11], increased pause duration [12], and varied switching pause duration [13]. Further, depressed speech was found to show increased tension in the vocal tract and the vocal folds [1, 14], and speech characteristics related to psychomotor retardation such as speech articulation were investigated [2].

Within this paper we are interested in vowel space as a measure of overall expressiveness of speech and articulation. Vowel space measures have not been directly investigated for depression. However, some researchers previously assessed vowel space to characterize other clinical conditions including Parkinson's disease [5] and cerebral palsy [3].

In particular, the vowel space of speakers with Parkinson's disease was compared to that of healthy controls in reading tasks [5]. Thirteen subjects and controls read a passage out loud at three different rates, i.e. habitual, fast, and slow rates. The acoustic characteristics of the vowels /i/, /a/, /u/, and /æ/ were investigated along with those of two fricatives /s/ and /ʃ/. The tokens for each of the investigated vowels and fricatives were manually selected from the recordings and spectrally analyzed. The observed average vowel space for subjects with Parkinson's disease was significantly smaller than that of healthy controls (p = .019).

The reduced vowel space of young adults with cerebral palsy, for example, was investigated with respect to the intelligibility of Mandarin [3]. The researchers found that vowel space has been significantly reduced for subjects with cerebral palsy when compared to healthy controls ($p < .001$). The researchers defined the investigated vowel space as the frequency range triangle of the first and second formant, i.e. the resonance frequencies of the vocal tract, spanned by the vowels /i/, /a/, and /u/. Within the present work, we adopt the same definition for consistency. However, here we opt to evaluate vowel space as a ratio between an individual's vowel space and that of a reference population rather than the actual area as measured in Hz^2 in order to render the method gender and age independent.

3. METHODS AND MATERIALS

3.1. Vowel Space Assessment

3.1.1. Speech Processing and Formant Tracking

For the processing of the speech signals, we use the freely available COVAREP toolbox (v1.1.0), a collaborative speech analysis repository available for Matlab and Octave [15][1]. COVAREP provides an extensive selection of open-source robust and tested speech processing algorithms enabling comparative and cooperative research within the speech community[2]. An overview of the algorithm is shown in Figure 1.

Voicing Detection: In [16], a method for fundamental frequency f_0 tracking and simultaneous voicing detection based on residual harmonics is introduced. The method is especially suitable in noisy and unconstrained conditions. In this work we are only interested in the voicing detection part of the algorithm, which is controlled by a basic threshold $\theta = 0.07$ that is applied to the sum of the residual harmonics. In this manner, the unvoiced frames can be discarded as described in [16].

Formant Tracking: The formant tracker used in this approach is introduced in detail in [17]. In particular, we are interested in the first and second formants F1 and F2. In order to remove small

[1] `http://covarep.github.io/covarep/`

[2] The vowel space assessment algorithm presented within this work will be made publicly available within COVAREP after publication.

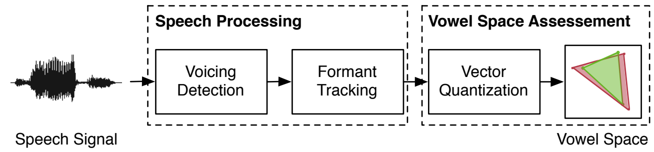

Fig. 1. **Algorithm overview figure.** The automatic assessment of the vowel space ratio is separated into two major steps including speech processing (i.e. voicing detection and vowel tracking) and the vowel space assessment (i.e. vector quantization using k-means clustering and vowel space ratio calculation).

fluctuations we apply a median filter with a filter length n = 15 to the tracked formants. Formants are tracked for all voiced regions, i.e. not only vowels.

3.1.2. Vowel Space Assessment

Based on the tracked formants F1 and F2 for the voiced regions of speech we compute the vowel space for each recorded subject individually. Figure 2 shows two examples of the assessed vowel space for a depressed and a non-depressed subject. In particular, the observed formant frequency pairs (gray dots), the reference vowel space (red triangle), and the subject's vowel space (green triangle) are seen. We define the vowel space as the frequency region covered by the triangle in the two dimensional frequency space spanned by F1 and F2 for the vowels /i/ (as in *heed*), /a/ (as in *hod*), and /u/ (as in *who'd*) following [3]. These three vowels represent the vowels with the most extreme positions of the tongue and are therefore located in the extremes of this triangularly shaped two-dimensional frequency space [18].

As we do not precisely know when the recorded subjects produced these vowels, we propose to apply a vector quantization approach, namely k-means clustering, to identify the prototypical locations of /i/, /a/, and /u/ for each speaker to automatically assess the individual's vowel space [19] (cf. Figure 1). We closely follow a recently proposed approach to automatically identify the vowel space in speech, that has been validated to highly correlate with manual measures of vowel space ($\rho > .7$) for both male and female speakers [20].

In detail the approach comprises the following steps: (1) To assess an individual's vowel space using k-means, we first initialize the k = 12 cluster centers c_i with $i = 1, ..., 12$ with the prototypical formant frequencies of F1 and F2 for the investigated individual's gender and age as proposed in [21]. (2) We adapt the cluster centers c_i based on the observed formant frequencies $x_m \in \mathbb{R}^2$ for the investigated individual using the basic k-means algorithm. First, each observation x_m is *assigned* to the closest cluster center c_i based on the squared Euclidean distance. This assignment step forms k sets $S_j = \{x_m : ||x_m - c_j||^2 \leq ||x_m - c_i||^2, \forall i = 1, ..., k\}$. Based on these sets S_j the cluster centers c_i are *updated* following the adaptation step $c_i = \frac{1}{|S_i|} \sum_{x_m \in S_i} x_m$, where $|S_i|$ denotes the number of elements in S_i. The iterative application of these two assignment and update steps minimizes the within cluster sum of squares and yields prototypical locations for all k cluster centers. (3) Lastly, after optimization we identify the three cluster centers $c_{/i/}$, $c_{/a/}$, and $c_{/u/}$ closest to the average formant locations of the vowels /i/, /a/, and /u/, as in [21]. At this point we would like to note that the three cluster centers $c_{/i/}$, $c_{/a/}$, and $c_{/u/}$ are not necessarily located near the formant locations of the vowels /i/, /a/, and /u/. After identifying the cluster centers, we compute the area A of the spanned triangle using Heron's formula $A = \sqrt{s(s-a)(s-b)(s-c)}$ with

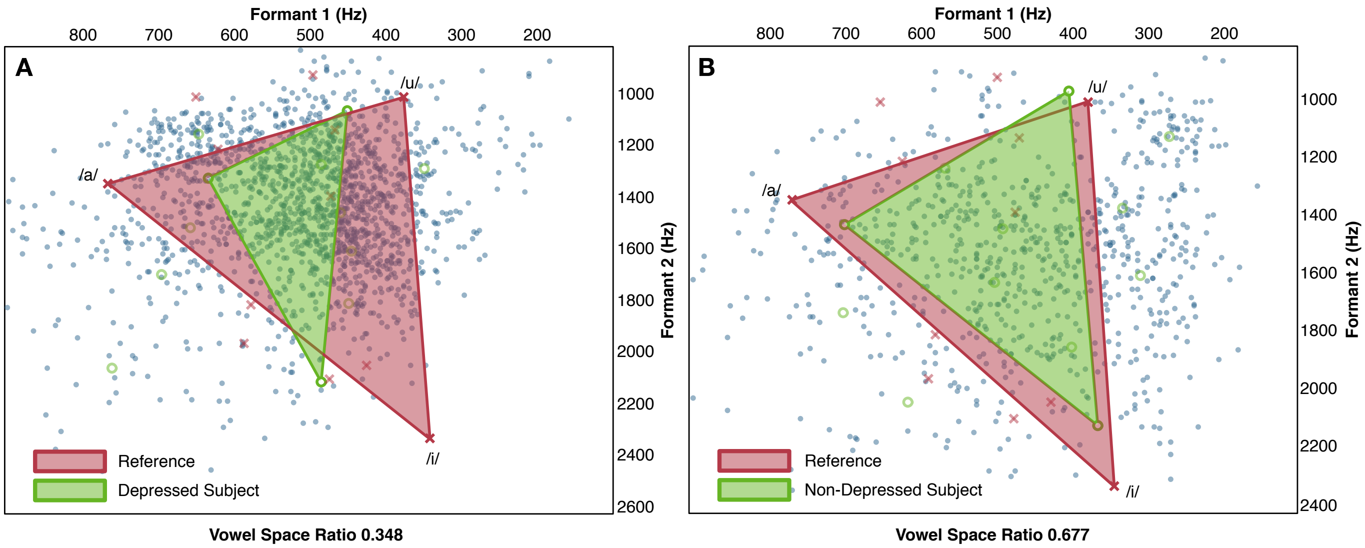

Fig. 2. **Example vowel space assessment for two male subjects of DAIC.** The male reference population vowel space (i.e. /i/, /a/, /u/) depicted in red is compared to the subjects' vowel spaces depicted in green, for a depressed subject (A) and a non-depressed subject (B) respectively. The vowel spaces are visualized on a two-dimensional plot with Formant 1 on the x-axis and Formant 2 on the y-axis (both in Hz). Additional two-dimensional vowel centers are displayed for both the male reference population (red x-symbolds) and the investigated subjects' vowel space cluster centroids (green circles). The grey dots depict all observations of the first two formants across an entire interview. The depressed subject's vowel space (A) is visibly smaller than the non-depressed subject's vowel space (B).

$s = \frac{a+b+c}{2}$ and a, b, c the lengths of the triangle's three sides. We then compute the ratio $vowelspace = \frac{A_{ind}}{A_{ref}}$ of the individual's vowel space area A_{ind} and the reference vowel space area A_{ref} to characterize how large the individual's vowel space is to the reference population vowel space with respect to the individual's corresponding gender or age.

3.2. Datasets

In order to assess the robustness of the vowel space assessment (cf. Section 3.1) we compare results on three different datasets, namely the Distress Assessment Interview Corpus (DAIC) [22], the AVEC 2013 audio-visual depression corpus (AVEC) [23], and an interview dataset of suicidal and non-suicidal adolescents recorded at the Cincinnati Children's Hospital Medical Center [24], here we refer to the dataset as Cincinnati Children's Interview Corpus (CCIC).

Distress Assessment Interview Corpus (DAIC) - We utilize the Distress Analysis Interview Corpus (DAIC; IRB #UP-11-00342), a large multimodal collection of semi-structured clinical interviews [22]. Within DAIC we use a virtual human as an interviewer[3]. The interviews were collected as part of a larger effort named *SimSensei* to create a virtual agent that interviews people and identifies verbal and nonverbal indicators of mental illness [25]. Participants are coded for depression and PTSD based on accepted psychiatric questionnaires: The PTSD Checklist-Civilian version (PCL-C) [26] is a self-report measure that evaluates all 17 PTSD criteria using a 5-point Likert scale as well as the Patient Health Questionnaire-Depression 9 (PHQ-9). The PHQ-9 is typically used as a screening tool for assisting clinicians in diagnosing depression as well as selecting and monitoring treatment.

In total 253 subjects interacted with the automatic SimSensei system. Overall, 186 male subjects and 67 female subjects with an average age of 44.7 (σ = 12.37) years were recorded. On average each conversation lasted for 18.76 minutes. Following the self-assessment questionnaires, 18.6% scored positive for depression (N = 47) and 34.6% for PTSD (N = 88). The self-reported conditions for PTSD and depression are significantly correlated for both the categorical (i.e. positive vs. negative) as well as the score assessments (i.e. assessed severity scores) [27].

AVEC 2013 audio-visual depression corpus (AVEC) - We further utilized a subset of the AVEC 2013 audio-visual depression corpus (AVEC) in our analysis [23]. In particular, we focus our analysis on a read speech passage of the Max Frisch novel *Homo Faber* to complement our interview data available in the DAIC corpus. Each recording has an indicated Beck Depression Inventory (BDI) score, a self-reported measure of depression [28]. We utilized a cut-off score of 20 for moderate to severe depression to split the dataset into depressed and non-depressed recordings as suggested in [28].

We excluded recordings of subjects with an identifier of 300 and above, as these recordings have a wrong sample rate and it is unclear what effect this can cause for the assessment of vowel space. In total, a remaining 68 recordings were analyzed of which 29 recordings were classified as depressed and the remaining 39 as not-depressed. The average length of the recordings for depressed subjects is 239.1 seconds (σ = 75.7) and not-depressed subjects 205.2 seconds (σ = 37.2).

Cincinnati Children's Interview Corpus (CCIC) - From March 2011 through August 2011, 60 patients (30 control and 30 subjects; average age of 15.47, σ = 1.5) were enrolled in a prospective, controlled trial at the Cincinnati Children's Hospital Medical Center (CCHMC) Emergency Department (ED; IRB #2008-1421). Eligible patients were between the ages of 13 and 17 and had come to the ED with suicidal ideation, gestures, and attempts. Patients with orthopedic injuries were enrolled as controls, because they are seen as having the fewest biological and neurological perturbations

[3]Sample interaction between the virtual agent and a human actor can be seen here: `http://youtu.be/ejczMs6b1Q4`

of all of the ED patients. Data were collected by a trained social worker. One control patient was excluded from the study due to a severe interruption in the ED and recording difficulties during the interview. Due to a large variation in interview length between suicidal and non-suicidal control patients only the first 60 seconds of speech for each participant is utilized for the analysis [29].

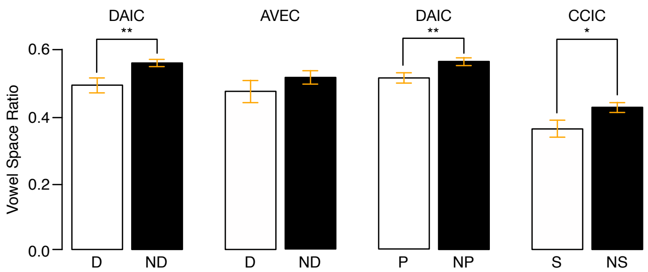

Fig. 3. **Vowel space ratio across conditions.** Observed mean vowel space ratios across conditions depression (D) vs. no-depression (ND) for interview data as well as read speech, PTSD (P) vs. no-PTSD (NP) and suicidal (S) vs. non-suicidal (NS) for interview data. The displayed whiskers signify standard errors and the brackets show significant results with * ... p $<$ .05 and ** ... p $<$ .01.

Dataset	μ (σ)	μ (σ)	Hedges' g
	Depression	**No Depression**	
DAIC	0.49 (0.15)	0.55 (0.15)	**-0.43****
AVEC	0.47 (0.18)	0.51 (0.12)	-0.27
	PTSD	**No PTSD**	
DAIC	0.51 (0.14)	0.56 (0.15)	**-0.34****
	Suicidal	**Non Suicidal**	
CCIC	0.36 (0.14)	0.42 (0.08)	**-0.55***

Table 1. **Observed vowel space ratios.** Results indicate reduced vowel space for distressed subjects across datasets. The arithmetic mean μ and the standard deviations σ (in brackets) are shown along with Hedges' g a measure for effect size. ** ... indicate significant difference with p-values $<$ .01 and * ... p-values $<$ .05 respectively.

4. RESULTS AND DISCUSSION

Following the three research questions **Q1** - **Q3** (cf. Section 1), we report and discuss our findings on the relationship between vowel space and depression (**Q1**), the effect of different speaking conditions (**Q2**), and the relevance of vowel space for PTSD and suicidality (**Q3**). In addition to mean values μ, we present the p-values of two-tailed t-tests and Hedges' g values as a measure of the effect size. The g value denotes the estimated difference between the two population means in magnitudes of standard deviations [30]. We summarize our statistical evaluation results in Table 1 and Figure 3.

Q1 - Reduced vowel space characterizes depressed conversational speech: Our investigations reveal that participants categorized as having depression by the PHQ-9 within the interview corpus DAIC exhibited smaller vowel space than those not categorized as having depression (depressed μ = 0.49, non-depressed μ = 0.55, t(251) = 2.69, p = .008, Hedges g = -0.43). This effect is further visualized in two example vowel spaces; Figure 2 A shows the reduced vowel space of a depressed speaker and Figure 2 B the vowel space off a non-depressed speaker respectively. Figure 2 A shows a high density of formant frequency observations of a depressed speaker towards the center and a narrow spread. The measure assesses the longitudinal frequency coverage of the first and second formant for an individual in an unconstrained interaction. Thus, the measure captures the range and extremes of a speaker's vowel articulation and aims to capture assessments of both overall articulation as well as psychomotor retardation, a commonly found symptom of depression [2]. While we expect that psychomotor retardation is correlated with the assessed vowel space measure further investigations are required to draw a direct link. Within the present study, we do not have access to diagnosis and expert assessments of psychomotor retardation, which we plan to accomplish in the near future.

Q2 - Vowel space measure is affected by speaking conditions: We investigate the vowel space measure with a separate dataset of depressed speech. Specifically, we analyze the AVEC dataset read speech portions. We found that the vowel space ratio is reduced for depressed subjects, however, the effect is not significant (depressed μ = 0.47, non-depressed μ = 0.51, t(66) = 1.12, p = .268, Hedges g = -0.27). While the effect is not significant, we still observe a similar trend for read speech. Several factors might have influenced the findings: Read speech is articulated differently from conversational speech and reading proficiency might be a confounding factor. Further investigations are required.

Q3 - Reduced vowel space characterizes distressed conversational speech: We further assess the relation between PTSD and suicidality with the assessed vowel space. We identified that participants categorized as having PTSD in DAIC by the PCL-C had smaller vowel space measurements than those not categorized as having PTSD (PTSD μ = 0.51, non-PTSD μ = 0.56, t(251) = 2.55, p = .01, Hedges g = -0.34). This finding might be explained as a characteristic of PTSD or by the high overlap and correlation between conditions of PTSD and depression within the investigated population (ϕ = .494; p $<$ .001). Similarly, suicidal adolescents in the CCIC showed a reduced vowel space when compared to their non-suicidal peers (suicidal μ = 0.36, non-suicidal μ = 0.42, t(57) = 2.14, p = .037, Hedges g = -0.55). The overall reduced vowel space for adolescents may be due to the increased size of the reference vowel space used in the approach [21]. The present findings indicate shared vowel space characteristics with depression. Indeed the comorbidity between PTSD as well as suicidality and depression has been previously identified in the literature [31, 32] and the observed strong correlation between conditions has been further discussed in our prior work, where we have identified the more generic condition of *psychological distress* as a common denominator [27].

5. CONCLUSION

Our investigations show that reduced vowel space indeed characterizes conversational speech of individuals with depression, PTSD, as well as suicidality, which can be summarized under the more comprehensive term of *psychological distress*. Further, the proposed approach robustly reveals reduced vowel space ratios across three distinct corpora. In the future, we plan to confirm our findings with respect to specific symptoms of depression, such as psychomotor retardation. We are convinced that automatically assessed vowel space from conversational data will be an essential piece for the objective analysis and assessment by healthcare providers for a wide range of psychological or neurologic conditions. Future applications could include psychological distress screening, diagnosis, and symptoms monitoring.

6. REFERENCES

[1] J. K. Darby, N. Simmons, and P. A. Berger, "Speech and voice parameters of depression: a pilot study," *Journal of Communication Disorders*, vol. 17, no. 2, pp. 75–85, 1984.

[2] A. J. Flint, S. E. Black, I. Campbell-Taylor, G. F. G. Gailey, and C. Levinton, "Abnormal speech articulation, psychomotor retardation, and subcortical dysfunction in major depression," *Journal of Psychiatric Research*, vol. 27, no. 3, pp. 309–319, 1993.

[3] H.-M. Liu, F.-M. Tsao, and P. K. Kuhl, "The effect of reduced vowel working space on speech intelligibility in mandarin-speaking young adults with cerebral palsy," *Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3879–3889, 2005.

[4] G. S. Turner, K. Tjaden, and G. Weismer, "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis," *Journal of Speech and Hearing Research*, vol. 38, no. 5, pp. 1001–1013, 1995.

[5] P. A. McRae, K. Tjaden, and B. Schoonings, "Acoustic and perceptual consequences of articulatory rate change in parkinson disease," *Journal of Speech, Language, and Hearing Research*, vol. 45, no. 1, pp. 35–50, 2002.

[6] M. Alpert, E. R. Pouget, and R. R. Silva, "Reflections of depression in acoustic measures of the patient's speech," *Journal of Affective Disorders*, vol. 66, no. 1, pp. 59–69, 2001.

[7] J. Gratch, L.-P. Morency, S. Scherer, G. Stratou, J. Boberg, S. Koenig, T. Adamson, and A. Rizzo, "User-state sensing for virtual health agents and telehealth applications.," *Studies in health technology and informatics*, vol. 184, pp. 151–157, 2012.

[8] S. Scherer, G. Lucas, J. Gratch, A. Rizzo, and L.-P. Morency, "Reduced vowel space in conversational speech indicates depression," *PlosONE*, under review.

[9] A. Nilsonne, "Speech characteristics as indicators of depressive illness," *Acta Psychiatrica Scandinavica*, vol. 77, no. 3, pp. 253–263, 1988.

[10] D. J. France, R. G. Shiavi, S. E. Silverman, M. K. Silverman, and D. M. Wilkes, "Acoustical properties of speech as indicators of depression and suicidal risk," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 829–837, 2000.

[11] J. A. Hall, J. A. Harrigan, and R. Rosenthal, "Nonverbal behavior in clinician-patient interaction," *Applied and Preventive Psychology*, vol. 4, no. 1, pp. 21–37, 1995.

[12] C. Sobin and H. A. Sackheim, "Psychomotor symptoms of depression," *American Journal of Psychiatry*, vol. 154, no. 1, pp. 4–17, 1997.

[13] Y. Yang, C. Fairbairn, and J. F. Cohn, "Detecting depression severity from vocal prosody," *IEEE Transactions on Affective Computing*, vol. 4, no. 2, pp. 142–150, 2013.

[14] S. Scherer, G. Stratou, J. Gratch, and L.-P. Morency, "Investigating voice quality as a speaker-independent indicator of depression and ptsd," in *Proceedings of Interspeech 2013*. 2013, pp. 847–851, ISCA.

[15] G. Degottex, J. Kane, T. Drugman, T. Raitio, and S. Scherer, "Covarep - a collaborative voice analysis repository for speech technologies," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014)*, 2014, pp. 960–964.

[16] T. Drugman and A. Abeer, "Joint robust voicing detection and pitch estimation based on residual harmonics," in *Proceedings of Interspeech 2011*. 2011, pp. 1973–1976, ISCA.

[17] B. Bozkurt, B. Doval, C. d'Alessandro, and T. Dutoit, "Improved differential phase spectrum processing for formant tracking," *Proceedings of Interspeech - ICSLP*, pp. 2421–2424, 2004.

[18] B. Lindblom, "Explaining phonetic variation: A sketch of the h&h theory," *Speech Production and Speech Modeling*, pp. 403–439, 1990.

[19] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. 1967, pp. 281–297, University of California Press.

[20] S. Sandoval, V. Berisha, R. L. Utianski, J. M. Liss, and A. Spanias, "Automatic assessment of vowel space area," *Journal of the Acoustical Society of America*, vol. 134, no. 5, pp. 477–483, 2013.

[21] J. Hillenbrand, L. A. Getty, M. J. Clark, and K. Wheeler, "Acoustic characteristics of american english vowels," *Journal of the Acoustical Society of America*, vol. 97, no. 5, pp. 3099–3111, 1995.

[22] J. Gratch, R. Artstein, G. Lucas, S. Scherer, A. Nazarian, R. Wood, J. Boberg, D. DeVault, S. Marsella, D. Traum, A. Rizzo, and L.-P. Morency, "The distress analysis interview corpus of human and computer interviews," in *Proceedings of Language Resources and Evaluation Conference (LREC)*, 2014, pp. 3123–3128.

[23] M. Valstar, B. Schuller, K. Smith, F. Eyben, B. Jiang, S. Bilakhia, S. Schnieder, R. Cowie, and M. Pantic, "Avec 2013 - the continuous audio / visual emotion and depression recognition challenge," in *Proceedings of ACM international workshop on Audio/visual emotion challenge*, 2013, pp. 3–10.

[24] S. Scherer, J. P. Pestian, and L.-P. Morency, "Investigating the speech characteristics of suicidal adolescents," in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2013, pp. 709–713, IEEE.

[25] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgilia, J. Gratch, A. Hartholt, M. Lhommet, G. Lucas, S. Marsella, F. Morbini, A. Nazarian, S. Scherer, G. Stratou, A. Suri, D. Traum, R. Wood, Y. Xu, A. Rizzo, and L.-P. Morency, "Simsensei kiosk: A virtual human interviewer for healthcare decision support," in *Proceedings of Autonomous Agents and Multiagent Systems (AAMAS)*, 2014, pp. 1061–1068.

[26] E. B. Blanchard, J. Jones-Alexander, T.C. Buckley, and C.A. Forneris, "Psychometric properties of the ptsd checklist (pcl)," *Behaviour Research and Therapy*, vol. 34, no. 8, pp. 669–673, 1996.

[27] S. Scherer, G. Stratou, G. Lucas, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo, and L.-P. Morency, "Automatic audiovisual behavior descriptors for psychological disorder analysis," *Image and Vision Computing Journal, Special Issue on Best of Face and Gesture 2013*, vol. 32, no. 10, pp. 648–658, 2014.

[28] D. Maust, M. Cristancho, L. Gray, S. Rushing, C. Tjoa, and M. E. Thase, *Psychiatric rating scales*, vol. 106 of *Handbook of Clinical Neurology*, chapter 13, pp. 227–237, Elsevier, 2012.

[29] V. Venek, S. Scherer, L.-P. Morency, A. Rizzo, and J. P. Pestian, "Adolescent suicidal risk assessment in clinician-patient interaction: A study of verbal and acoustic behaviors," in *Proceedings of Spoken Langugae Technologies (SLT)*, 2014.

[30] L. V. Hedges, "Distribution theory for glass's estimator of effect size and related estimators," *Journal of Educational Statistics*, vol. 6, no. 2, pp. 107–128, 1981.

[31] D. Campbell, B. Felker, C.-F. Liu, E. Yano, J. Kirchner, D. Chan, L. Rubenstein, and E. Chaney, "Prevalence of depression-ptsd comorbidity: Implications for clinical practice guidelines and primary care-based interventions," *Journal of General Internal Medicine*, vol. 22, pp. 711–718, 2007, 10.1007/s11606-006-0101-4.

[32] M. K. Nock, I. Hwang, N. A. Sampson, and R. C. Kessler, "Mental disorders, comorbidity and suicidal behavior: Results from the national comorbidity survey replication," *Molecular Psychiatry*, vol. 15, pp. 868–876, 2010.