INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

Acoustic characteristics of depression in older adults' speech: the role of covariates

Carmen Mijnders[1], *Esther Janse*[1], *Paul Naarding*[2], *Khiet P. Truong*[3]

[1]Radboud University, The Netherlands
[2]GGNet Mental Health, The Netherlands
[3]University of Twente, The Netherlands

carmenmijnders@hotmail.com, esther.janse@ru.nl, p.naarding@ggnet.nl, k.p.truong@utwente.nl

Abstract

Depression in older adults is often associated with various physical conditions and is hence different from depression at a younger age. Ageing may come with cognitive decline, medication use, and frailty, which are known to be predictors of late-life depression. One common symptom of depression is psychomotor retardation, that may also affect speech production. Most speech studies on depression so far have focused on younger or middle-aged adults. In this study, we used speech data from a large longitudinal Dutch study on late-life depression and its comorbid symptoms to compare speech acoustics in persons with depression (PWD) and controls. We investigated whether groups differed by taking several covariates into account (e.g., frailty, slowness, and medication use). Group differences were found in within-vowel F2 range, speech rate and mean pause duration. These data indicate that speech acoustics can be used to differentiate PWDs and controls, even with low-quality speech data.

Index Terms: late life depression, speech characteristics

1. Introduction

According to the Global Burden of Disease (GBD) study, mental disorders are a significant burden globally, with depression being a major component [1]. Depression can occur at different stages in life, however, depressive symptoms become more frequent in later life and seem to increase with ageing [2]. The high prevalence and severity in combination with the ageing population make research on depression in the older adults relevant [3]. The ageing population will require more resources from society in terms of money, time, and manpower for mental health care for older people.

Depression in older adults is different from depression at a younger age, as depression in older adults is often associated with other physical conditions. Changes in the brain, such as dementia and the use of various medications, play a greater role in later life. This makes depression in older people a complex issue and difficult to diagnose. In addition, the diagnosis of depression is still largely based on the subjective judgment of a psychiatrist. A more objective measure for diagnosing depression is desirable.

Such a measure of depression could be based on speech features. Studies have indeed identified acoustic characteristics that are indicative of the presence or severity of depression, e.g., [4, 5]. However, studies on the impact of depression on speech often focus on (young) adults and not specifically on older adults [6]. It is known that psychomotor retardation (PMR) is a major symptom of depression that also affects speech production [7]. PMR includes cognitive and motor symptoms that are more prominent in older people with a depression. It is represented in various motor productions, such as speech [7, 8]. The role of these age-related characteristics (including co-morbid conditions), is often overlooked when studying the relation between speech characteristics and depression, which could be partly due to the availability of (speech) data.

Finding available speech data that allows for a study into the relation between late-life depression and speech characteristics while taking age-related covariates into account is challenging. In this study, we used the NESDO database: a large longitudinal Dutch study on late-life depression and its comorbid symptoms [3]. More than 500 older adults aged 60–93 years with and without depression were followed for a period of six years. Participants were assessed using questionnaires, interviews, medical exams, and cognitive tests as part of a fixed protocol. Many of these sessions were audio-recorded with a single channel device. Since speech analysis was not the purpose of recording, the audio quality however is relatively low, there are no timestamped transcriptions available, and there is no specific speech sample designed for speech analysis that we can use.

Our study contributes to a gap in depressive speech research by focusing on depression in older adults while taking participant characteristics (such as age-related and PMR-related variables, and information about smoking and medication use) into account. We re-purposed existing audio recordings initially not processed or made for speech analyses aims, and applied (linear mixed) regression modelling to investigate the effect of depression and several person-characteristic variables on certain speech acoustics.

2. Related work

2.1. Speech indicators of depression

Numerous studies have shown that there are significant relations between depression and acoustic characteristics in speech, and that it is possible to detect depression from speech automatically. Readers interested in extensive reviews of the literature are referred to [5, 4, 7]. Here, we give a brief summary of the observed relations that are most relevant for the current work.

Depression appears to be associated with higher jitter and shimmer values [4, 10, 11, 9], indicating lower quality voices that are perceived as more breathy or hoarse. This is in concordance with findings of lower harmonics-to-noise ratios (HNRs) in depressive speech [10], possibly caused by PMR. For formants, the findings are less conclusive: F1 and F2 (averages and ranges) either increases or decreases with severity of depression [5]. It is possible that the articulatory effort associated with the formants is affected in different ways by PMR or medication causing these mixed results. On the contrary, reduced variability in F0, associated with monotonousness, has been found rather consistently in various studies, e.g., [5, 13, 14],

10.21437/Interspeech.2023-839

Table 1: *Speech indicators of depression with arrows indicating the direction of the found effect for* PWD. *Checkmarks indicate whether the current study includes the feature for S(troop) or I(nterview) data.*

		Previous work Effect	Used S	Used I
voice quality	jitter	▲ [4, 9, 10, 11]	✓	
	shimmer	▲ [4, 10, 11]	✓	
	HNR	▼ [10]	✓	
	spectral tilt	▼ [11, 12]		
formants	mean F1	▼ [13, 4]	✓	
	mean F2	▼ [13] ▲ [4]	✓	
	range F1	▲ [12, 4]	✓	
	range F2	▼ [12, 13]	✓	
prosody	mean F0	▼ [12] ▲ [10]	✓	✓
	stand. dev. F0	▼ [14, 15, 13, 12, 10, 5]		✓
	mean intensity	▼ [16]		
	range intensity	▼ [16]		
	mean pause duration	▲ [16, 17, 18, 19, 20]		✓
	pause ratio	▲ [17, 18, 19]		✓
	speech rate	▼ [15, 18, 20]		✓
interaction	gap duration	▲ [20]		✓
	utterance duration	▲ [21, 22, 17, 18]		

as well as reduced intensity levels and speech rate. However, for mean F0, findings have been inconsistent. Group differences have also been reported in pausing behaviours: people with depression produce more and longer pauses. These differences in speech rate and pausing behaviour are most likely caused by slowness in cognitive and motor aspects of spoken language. Finally, utterance and gap length have been reported to be larger in people with depression. A concise overview of acoustic features affected by depression is given in Table 1.

In many of the studies discussed, confounding variables such as sex, age, native language are controlled for [4]. However, medication that is known to affect speech characteristics is less controlled for [4]. In some studies, subjects who use medication are excluded [13, 4], while in other studies, no differences were found between medicated and nonmedicated participants [23, 24]. Since the effects of medication on speech characteristics in depression are not conclusive, it is recommended to include this variable in the analysis, see Section 4.

2.2. Late life depression

Late life depression differs from depression in younger adults, and affects speech characteristics differently. Ageing affects speech production, particularly vocal quality, speech rate and pausing, and articulatory precision. Moreover, older adults with depression experience more frequent comorbid physical and/or cognitive symptoms which may impact speech production. Psychomotor and cognitive slowing and a decreased gait speed are more prominent in older adults [25]. Frailty, which is associated with ageing, can also lead to a poorer prognosis when combined with depression [26]. Frailty reflects a decline in the functional and physiological reserves of various bodily processes, making individuals with frailty at risk of negative health effects even with mild stressors [27]. Motor symptoms such as slowness and weakness may manifest as a result of frailty [28].

Psychomotor retardation (PMR) is an important characteristic of depression (American Psychiatric Association, 2013) and is defined by motor, cognitive, and verbal symptoms [29, 7]. PMR is not unique to depression and can also occur in neuropsychiatric disorders such as schizophrenia and Parkinson's disease, and includes symptoms such as cognitive and motor slowing. PMR in depression can be observed in speech, facial expressions, body posture, self-touching, eye movements, and speed and presence of movements [7, 8].

Late life depression is a clinically complex illness interacting with variables involved in the (normal) ageing process that can affect speech production. PMR is an important feature in (late life) depression, and can be characterized by neuromotor and cognitive slowness [23, 30]. Slowness is also a symptom of frailty [28]. Frailty in turn is related to PMR and is highly comorbid in adults with late life depression [31]. The clinical complexity of late life depression introduced by interactions among these symptoms such as slowness and frailty ask for an elaborate analysis that takes these factors into account.

3. Data

3.1. Participants

We randomly included 19 participants from the NESDO database, of which 9 belonged to the experimental group PWD (=people with depression) and 10 belonged to the healthy control group CG. Inclusion criteria for PWD were a primary diagnosis of a major depressive disorder, dysthymia (persistent depressive disorder), minor depression or adjustment disorder with depressed mood. Excluded were people with an (expected) diagnosis of schizophrenia, dementia, addiction problems or bipolar disorder, and people with severe cognitive problems (Mini-Mental State Examination score lower than 18). People in the control group CG do not have a (prior) diagnosis of depression or dementia. As will be explained in Section 3.3, we will carry out acoustic measurements on segmental and suprasegmental levels that ask for different types of speech materials. From the whole recording, we selected the verbal Stroop test [32] and part of the interview about demographic information. Since in some of the participant recordings these selected speech materials are missing, the participant groups for both types of speech materials are not exactly the same but overlap largely. The overlap is 17 participants (9 controls and 6 persons with depression) whose data are available for the Stroop materials as well as for the interview. Demographics and participant-related information for both participant groups are shown in Table 2.

3.2. Procedure

The participants participated in a longitudinal study into late-life depression in the Netherlands, NESDO [3]. In this study, researchers followed the participants for 6 years long and assessed and interviewed them at 3 different moments in time: the first assessment was the baseline assessment, the second and third assessments took place 2 and 6 years after the baseline assessment respectively. We used the third assessments which were recorded on digital recording devices. Interviewers were instructed to put the microphone and the digital recorder in between themselves and the participants. The same recording equipment was used throughout, but room acoustics could not be kept constant. Each assessment followed a strict experimen-

Table 2: *Demographic and other participant-related information (numbers in brackets are standard deviations). Numbers are counts unless indicated otherwise.*

		Stroop		Interview	
		PWD	CG	PWD	CG
Group		9	10	8	9
Sex	F/M	5/4	5/5	6/2	4/5
Age (yrs)		73.4 (7.0)	77.9 (7.3)	71.6 (4.5)	77.3 (7.5)
Smoking	yes/no	2/7	0/10	1/7	0/9
Medication	yes/no	6/3	1/9	5/3	1/8
MMSE (score)		28.0 (2.2)	28.2 (1.9)	28.0 (2.5)	28.4 (1.8)
Slowness	yes/no	5/4	4/6	3/5	4/5
Frailty	yes/no	3/6	0/10	2/6	0/9

tal protocol that consisted of blood sampling, interviews about the participant's medication use and demographics, the Stroop test [32], MMSE (Mini Mental State Examination), and reading aloud sequences of numbers among others. As our speech samples, we selected the Stroop test and the interview about demographics. Participants included in the NESDO database all gave written consent at the start of the baseline assessment but additional consent was requested from the NESDO database for specific use of the audio files (as our goal was not part of the primary goals of the database).

3.3. Selected material

For each participant, 40 vowels from the colour words in the verbal Stroop test (balanced over congruency conditions and colours) were manually extracted (/oː,u,eː,ɑu/ in "rood", "groen", "geel", and "blauw"). T-tests showed no significant differences between CG and PWD with respect to MMSE and the Interference Stroop scores. In addition, for each participant, speech samples without crosstalk or any other hampering noise conditions of at least 6.5 seconds long were selected from the demographic interview to study suprasegmental and interactionals characteristics. To facilitate this process and to identify speech segments, we used automatic speech recognition[1] to transcribe the audio recordings.

4. Method

4.1. Acoustic features

Following previous work (Section 2), we selected voice quality, formants, prosody and interaction-related features, see Table 1. We did not select intensity however since the distance between microphone and participant was not well controlled for. The acoustic features were extracted using a Praat script [33]. Jitter, shimmer and HNR (harmonics-to-noise ratio) were extracted over the extracted vowels, as well as mean F0, F1 and F2, and (within-vowel) range (defined as the difference between the 95th percentile and 5th percentile values) of F1 and F2. In the interview recordings, mean and standard deviation of F0 were measured, as well as mean pause duration and ratio (defined as the duration of pauses divided by the duration of speech), response RT (defined as the gap duration between the question and answer given by the participant), and speech rate (number of syllables/second, manually determined).

[1] https://webservices.cls.ru.nl/asr_nl

4.2. Covariates

As explained in Section 2, late life depression is clinically complex and associated with several interconnected (comorbid) factors. Hence, as factors related to PMR, we include the MMSE score [34], slowness (assessed through a walking test, where slowness was considered to be present if walking speed corresponded to the lowest quintile of stratified gender and height groups) and observed frailty (considered to be present if participants presented with ≥3 out of the following 5 criteria: weight loss, weakness, slowness, exhaustion, and low physical activity level) as covariates. The use of medication, i.e., antidepressants, antipsychotics and/or benzodiazepine, is included as well (coded as yes/no). Finally, sex, age, and smoking are considered, see Table 2 for all covariates included.

4.3. Analysis

All acoustic features (derived from the Stroop data and from the interview data) were analysed in two steps. In step 1 we fitted simple models (separate models for each acoustic feature) to investigate whether groups (PWD, CG) differed (without inclusion of covariates). For each of the acoustic features based on the Stroop experiment, there were repeated measurements of four different colour words expressed several times. Step 1 statistical models for the Stroop acoustic features were therefore linear mixed-effect regression (lmer) models testing for a fixed effect of Group on the acoustic feature at hand, including a random participant intercept, and a random colour word intercept. For the acoustic measures derived on the basis of the interview, all acoustic measurements (cf. Table 3) were averaged data (yielding only one observation per participant). Step 1 statistical models for the interview data were therefore simple linear regression (lm) models, testing for a Group effect on the acoustic feature at hand, without any random effect structure.

In step 2, we fitted more elaborate models to investigate whether groups differed in specific acoustic features, taking multiple personal characteristics into account. More specifically, we set up models to test for group differences (PWD, CG) in speech acoustics, taking the following seven covariates into account (cf. Table 2): age, gender, cognition (MMSE score), frailty (yes/no), smoking (yes/no), use of psychotropic medication (yes/no), and slowness symptoms (yes/no). Sum coding was used for all binary variables. Ratio variables were centred. Again, as the acoustic feature data derived from the Stroop experiment represented 40 observations per participant, step 2 statistical models for Stroop acoustic features were lmer models, in which we tested for a fixed effect of Group, as well as of the seven covariates on the acoustic feature at hand, including a random participant and a random colour word intercept. Step 2 statistical models for the interview data were linear models, testing for a Group effect on the acoustic feature at hand, next to any potential effects of the seven covariates. All step 2 starting models thus had Group and the seven covariates in the fixed structure. These starting models were simplified by step-wise backstripping to arrive at the most parsimonious models, taking out insignificant effects one by one, always starting from the effect with the lowest t-value in the previous model.

5. Results

Descriptive data for the acoustic features and step 1 analysis results are provided in Table 3. As can be seen in Table 3, the Step

1 analyses showed no group differences between PWD and CG in any of the acoustic features, except for speech rate. Speech rate of the controls during the interviews was faster than that of the PWD (β = 1.034, SE = 0.465, t = 2.222, p < 0.05).

Table 3: *Descriptive statistics (group means and SDs) for the acoustic features based on the Stroop recordings and interview recordings. T-values testing for Group differences in Step 1 models are also provided in the right-most column (bold print denotes significance level of p < 0.05.)*

	PWD	CG	Group difference t-value
Stroop recordings			
Jitter (%)	0.007 (0.005)	0.006 (0.005)	-0.798
Shimmer (%)	0.171 (0.064)	0.158 (0.058)	-1.432
HNR	10.656 (4.050)	12.569 (4.229)	1.941
Mean F1	441.434 (73.594)	459.108 (88.894)	0.975
Mean F2	1252.820 (463.969)	1298.262 (472.113)	0.771
F1 range	124.952 (122.889)	157.749 (140.882)	1.593
F2 range	399.046 (463.459)	545.142 (545.364)	1.558
Mean F0	166.102 (58.726)	180.770 (46.702)	0.787
Interview recordings			
Mean F0 (Hz)	217.162 (51.629)	182.140 (35.078)	-1.653
SD F0 (Hz)	107.385 (44.232)	79.428 (33.416)	-1.481
Mean pause duration (s)	0.721 (0.352)	0.551 (0.444)	-0.869
Speech rate (syll/s)	3.530 (0.651)	4.563 (1.161)	**2.222**
Response RT (s)	0.370 (0.233)	0.651 (0.447)	1.646

Below, we will only report step 2 models for acoustic features that included group differences. That is, models that only showed effects of covariates (like e.g., age, gender, or cognitive status) in the absence of group differences will not be reported on here. Whenever group differences were observed, we also verified that these were not due to inclusion of outlier values.

For the acoustic features derived from the Stroop recordings, only F2 range showed a significant group effect, as well as an effect of medication. More specifically, F2 range within colour-word vowels was larger for control participants (CG) than for PWD (β = 289, SE = 100, t = 2.877, p < 0.05), and participants using medication showed larger F2 range than those not using medication (β = -255, SE = 104, t = -2.456, p < 0.05). Similar effects of group and medication use were observed for within-vowel F1 range, but these were less stable. For the interview-based acoustic features, step 2 analyses showed group differences in mean pause duration and speech rate. The most parsimonious model for mean pause duration showed that controls had shorter mean pause durations than PWD (β = -0.416, SE = 0.163, t = -2.556, p < 0.05), and that, unexpectedly, those with higher cognitive scores had longer pause durations (β = 0.103, SE = 0.042, t = 2.458, p < 0.05). The pause duration model also showed an effect of sex, such that male participants had longer pauses than female (β = 0.654, SE = 0.184, t = 3.561, p < 0.01). The most parsimonious model for speech rate included effects of group and of sex: controls had higher speech rates than PWD (β = 1.329, SE = 0.438, t = 3.036, p < 0.01), and male participants had slower rates than females (β = -0.967, SE = 0.444, t = -2.177, p < 0.05).

6. Discussion

This study was set up to investigate speech acoustics in older PWD, as compared to controls. We aimed to investigate acoustic features associated with depression in relation to other (comorbid) characteristics, such as symptoms of psychomotor retardation. The results found are largely in line with results found in previous work. When no covariates are added, only speech rate is affected by Group: PWD speak more slowly than CG (also found in [15, 18, 20]). When covariates are added, Group and medication show a significant effect on F2 range (in line with [12, 13]. For mean pause duration, in accordance with previous research [16, 17, 18, 19, 20], we also find that controls have shorter mean pause durations than PWD. However, we also found that those with higher cognitive scores (MMSE) have longer pause duration, which is unexpected and cannot be explained at the moment. Finally, speech rate was found to be higher for controls than PWD (in line with [15, 18, 20]).

An additional aim was to assess whether the NESDO speech materials that had been recorded as part of an existing protocol not aimed at speech analysis could actually be used for speech analysis after all. Given the results, we were sufficiently able to re-use a small portion of the available data for a different purpose. Relatively large manual effort was needed for this re-purposing process, so future research should look into how this process can be scaled up, since one of the limitations of this study is its small sample size.

7. Conclusion

Late-life depression is assocated with various physical and cognitive conditions that can affect speech acoustics, hence, these conditions should be taken into account. When adding age, gender, smoking, medication, frailty, slowness and MMSE as covariates in our analyses, we found significant group differences (PWD vs. CG) in F2 range within-vowel, speech rate, and mean pause duration. More specifically, participants using medication showed larger F2 range, and participants with higher cognitive scores had longer mean pause durations. Slowness and frailty specifically (covariates related to PMR) did not seem to affect acoustic features: not as variables in their own right, nor was it the case that their inclusion brought up group differences. When this demographic and other participant-related information was *not* added, only speech rate showed group differences. Hence, it is important to take these covariates into account when this information is available in order to obtain a more complete understanding of how late-life depression is associated with speech characteristics.

8. References

[1] G. . M. D. Collaborators *et al.*, "Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–

2019: a systematic analysis for the global burden of disease study 2019," *The Lancet Psychiatry*, vol. 9, no. 2, pp. 137–150, 2022.

[2] M. Luppa, C. Sikorski, T. Luck, L. Ehreke, A. Konnopka, B. Wiese, S. Weyerer, H.-H. König, and S. G. Riedel-Heller, "Age-and gender-specific prevalence of depression in latest-life–systematic review and meta-analysis," *Journal of Affective Disorders*, vol. 136, no. 3, pp. 212–221, 2012.

[3] H. C. Comijs, H. W. van Marwijk, R. C. van der Mast, P. Naarding, R. C. Oude Voshaar, A. T. Beekman, M. Boshuisen, J. Dekker, R. Kok, M. W. de Waal *et al.*, "The Netherlands study of depression in older persons (NESDO); a prospective cohort study," *BMC Research Notes*, vol. 4, pp. 1–10, 2011.

[4] D. M. Low, K. H. Bentley, and S. S. Ghosh, "Automated assessment of psychiatric disorders using speech: A systematic review," *Laryngoscope Investigative Otolaryngology*, vol. 5, no. 1, pp. 96–116, 2020.

[5] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. F. Quatieri, "A review of depression and suicide risk assessment using speech analysis," *Speech Communication*, vol. 71, pp. 10–49, 2015.

[6] D. D. DeSouza, J. Robin, M. Gumus, and A. Yeung, "Natural language processing as an emerging tool to detect late-life depression," *Frontiers in Psychiatry*, vol. 12, p. 719125, 2021.

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

[8] D. J. Widlöcher, "Psychomotor retardation: clinical, theoretical, and psychometric aspects." *Psychiatric Clinics of North America*, 1983.

[9] A. Ozdas, R. G. Shiavi, S. E. Silverman, M. K. Silverman, and D. M. Wilkes, "Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk," *IEEE transactions on Biomedical engineering*, vol. 51, no. 9, pp. 1530–1540, 2004.

[10] T. F. Quatieri and N. Malyska, "Vocal-source biomarkers for depression: A link to psychomotor activity," in *Proceedings of Interspeech*, 2012.

[11] W. J. Silva, L. Lopes, M. K. C. Galdino, and A. A. Almeida, "Voice acoustic parameters as predictors of depression," *Journal of Voice*, 2021.

[12] F. Hönig, A. Batliner, E. Nöth, S. Schnieder, and J. Krajewski, "Automatic modelling of depressed speech: relevant features and relevance of gender," in *Proceedings of Interspeech*, 2014, pp. 1248–1252.

[13] D. J. France, R. G. Shiavi, S. Silverman, M. Silverman, and 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.

[14] 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.

[15] M. Cannizzaro, B. Harel, N. Reilly, P. Chappell, and P. J. Snyder, "Voice acoustical measurement of the severity of major depression," *Brain and Cognition*, vol. 56, no. 1, pp. 30–35, 2004.

[16] M. Garcia-Toro, J. A. Talavera, J. Saiz-Ruiz, and A. Gonzalez, "Prosody impairment in depression measured through acoustic analysis," *The Journal of Nervous and Mental Disease*, vol. 188, no. 12, pp. 824–829, 2000.

[17] J. C. Mundt, P. J. Snyder, M. S. Cannizzaro, K. Chappie, and D. S. Geralts, "Voice acoustic measures of depression severity and treatment response collected via interactive voice response (ivr) technology," *Journal of Neurolinguistics*, vol. 20, no. 1, pp. 50–64, 2007.

[18] J. C. Mundt, A. P. Vogel, D. E. Feltner, and W. R. Lenderking, "Vocal acoustic biomarkers of depression severity and treatment response," *Biological Psychiatry*, vol. 72, no. 7, pp. 580–587, 2012.

[19] A. C. Trevino, T. F. Quatieri, and N. Malyska, "Phonologically-based biomarkers for major depressive disorder," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, pp. 1–18, 2011.

[20] M. Yamamoto, A. Takamiya, K. Sawada, M. Yoshimura, M. Kitazawa, K.-c. Liang, T. Fujita, M. Mimura, and T. Kishimoto, "Using speech recognition technology to investigate the association between timing-related speech features and depression severity," *PloS one*, vol. 15, no. 9, p. e0238726, 2020.

[21] L. Albuquerque, A. R. S. Valente, A. Teixeira, D. Figueiredo, P. Sa-Couto, and C. Oliveira, "Association between acoustic speech features and non-severe levels of anxiety and depression symptoms across lifespan," *PloS one*, vol. 16, no. 4, p. e0248842, 2021.

[22] A. Esposito, A. M. Esposito, L. Likforman-Sulem, M. N. Maldonato, and A. Vinciarelli, "On the significance of speech pauses in depressive disorders: results on read and spontaneous narratives," *Recent Advances in Nonlinear Speech Processing*, pp. 73–82, 2016.

[23] A. J. Flint, S. E. Black, I. Campbell-Taylor, G. F. 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.

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

[25] J. B. Sanders, M. A. Bremmer, H. C. Comijs, D. J. Deeg, and A. T. Beekman, "Gait speed and the natural course of depressive symptoms in late life; an independent association with chronicity?" *Journal of the American Medical Directors Association*, vol. 17, no. 4, pp. 331–335, 2016.

[26] A. Lugtenburg, M. Zuidersma, K. J. Wardenaar, I. Aprahamian, D. Rhebergen, R. A. Schoevers, and R. C. Oude Voshaar, "Subtypes of late-life depression: a data-driven approach on cognitive domains and physical frailty," *The Journals of Gerontology: Series A*, vol. 76, no. 1, pp. 141–150, 2021.

[27] T. Malmstrom and J. E. Morley, "Frailty and cognition: linking two common syndromes in older persons," *The Journal of Nutrition, Health & Aging*, vol. 17, pp. 723–725, 2013.

[28] L. P. Fried, C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. J. Kop, G. Burke *et al.*, "Frailty in older adults: evidence for a phenotype," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 56, no. 3, pp. M146–M157, 2001.

[29] D. Bennabi, P. Vandel, C. Papaxanthis, T. Pozzo, and E. Haffen, "Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications," *BioMed Research International*, vol. 2013, 2013.

[30] M. P. Caligiuri and J. Ellwanger, "Motor and cognitive aspects of motor retardation in depression," *Journal of Affective Disorders*, vol. 57, no. 1-3, pp. 83–93, 2000.

[31] P. J. Brown, S. P. Roose, K. R. O'Boyle, A. Ciarleglio, B. Maas, K. C. Igwe, S. Chung, S. Gomez, M. Naqvi, A. M. Brickman *et al.*, "Frailty and its correlates in adults with late life depression," *The American Journal of Geriatric Psychiatry*, vol. 28, no. 2, pp. 145–154, 2020.

[32] J. R. Stroop, "Studies of interference in serial verbal reactions." *Journal of Experimental Psychology*, vol. 18, no. 6, p. 643, 1935.

[33] P. Boersma and D. Weenink, "Praat: doing phonetics by computer," http://www.praat.org, 2023, version 6.0.45.

[34] T. Tombaugh, I. McDowell, B. Kristjansson, and A. Hubley, "Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): a Psychometric Comparison and Normative Data." *Psychological Assessment*, vol. 8, no. 1, p. 48, 1996.