

Depression Diagnosis and Treatment Through Speech (DDaTTS)

*Carolina De Pasquale*¹

¹Dublin Institute of Technology, Ireland

carolina.depasquale@dit.ie

1. Motivation and Related Work

Clinical depression is a widespread illness with a variety of symptoms likely to have a negative impact on several aspects of a patient's life. According to the World Health Organization (WHO), 25% of the European population suffers from depression or anxiety each year, and these conditions account for 50% of chronic sick leaves [1]. Depression is considered a leading cause of disability in the United States amongst people aged 15-44 and it is estimated that the burden of individuals with depression has increased from \$173.2 billion in 2005 to \$210.5 billion in 2010 [2], while the European division of the WHO estimates a cost of about 170 billion per year for the combined cost of mood disorders and depression [1], with a cost for depression alone estimated at 118 billion for 2004 [3].

It is an affective disorder, characterised by a number of different factors that may be present in a number of different combinations and severities: low mood, anhedonia (loss of interest in pleasure in hobbies), anergia, feelings of guilt, insomnia (early, middle and late), psychomotor retardation (slowness of thoughts and action), agitation, anxiety, somatic symptoms (gastrointestinal and more general symptoms), loss of libido, hypochondriasis, weight loss and de-personalisation (feelings of unreality). Overall, these aspects lead to a social withdrawal and general social impairment of the patient. Moreover, depression often co-occurs with other psychiatric illnesses such as anxiety disorders, schizophrenia, substance misuse, and with pervasive developmental disorders such as autistic spectrum disorder.

Despite this, it is estimated that about 50% of major depressive disorder goes untreated, and it is believed that one of the factors impacting on the treatment of depression is the ability of medical staff to recognise symptoms associated with the mood disorder and perform an accurate diagnosis [4]: there is a lack of an objective diagnosis method for depression, and at this time diagnosis depends on subjective assessments by expert practitioners, who have to rely on patients' self-reports and self-assessments or perform interviews to assess the impact and prevalence of symptoms in the individual [5].

People engaged in conversation tend to adapt their communicative behaviour to that of their interlocutor, with numerous terms being used to describe this phenomenon, particularly with regards to speech: accommodation, synchrony, mimicry, convergence, alignment and entrainment [6]. Vocal accommodation occurs between speakers when changes in their prosodic parameters move in synchronous alignment or when they converge towards a common point [7]; it is a particularly important aspect of social interaction as it facilitates comprehension and understanding between interlocutors [8].

In an effort to determine an objective measure for depression, recent studies have explored the impact that the disorder has on a number of prosodic variables, which are affected by a variety of symptoms that occur in depressed individuals, such as psychomotor retardation, muscle tension and cog-

nitive impairments [9]; indeed, changes to the prosodic behaviour of individuals with mental health issues are well known [5, 10, 11, 12, 13, 14, 15, 16, 17, 18]. These studies have found that depression induces measurable, manifest changes in a person's speech that can be indicative of depression severity or even suicidal ideation. The DDaTTS project will investigate the development of a rule-set that codifies the link between depression and speech, whilst utilising the findings to develop applications to aid in the diagnosis and treatment of depression by clinical practitioners.

2. Aims and Research Questions

The goal of this research is to examine how depression affects a person's speech in order to develop tools and methods for the detection and treatment of depressive illness. The research will build upon existing research on depressed speech and will examine important socio-communicative non-verbal speech characteristics, such as prosodic accommodation, that can be used to measure rapport and involvement between speakers. This has the potential to provide advanced insights into the depressive state of a patient, including a determination of depression severity. The outcomes of the research will further be used to develop mobile applications to augment current clinical practices, and to aid and monitor depressed patients in a non-clinical setting.

The research will examine whether the amount and type of non-verbal prosodic accommodation can be used as a measure of depression severity in conjunction with more standard speech analysis methods. The research outputs will be used to inform the development of desktop and mobile software for use by clinical practitioners in a clinical setting and patients in a non-clinical setting.

The DDaTTS project will investigate the development of a rule-set that codifies the link between depression and speech, whilst utilising the findings to develop applications to aid in the diagnosis and treatment of depression by clinical practitioners.

3. Research Methodology

The literature suggests that there is a need for real audio recorded during actual psychiatric interviews: currently, a number of studies have obtained audio through various elicitation methods and mood induction procedures that can colour the data [9].

For the purpose of answering to this need, the participants will be recorded during live clinical interviews with high quality recording devices. Audio from non depressed sources will also be captured to provide a control group.

Due to the stringent ethical guidelines that are placed upon experiments with depressed individuals, audio will also be collected from non-clinical populations and used to examine analysis methods and equipment, as well as testing experiment set-ups if needed. The non-clinical audio will be used for the

improvement of analysis workflow, and for the development of automatic speech analysis methods.

4. The results from completed work

A pilot study has been carried out with the participation of a psychiatric consultant. This examined a small number of patient interactions to get preliminary data to ascertain the validity of the investigation and test the best method for recording the speech of the interaction and how feasible it was to get patients to participate. Multiple logistical issues in relation to the collection of data have been clarified, with the result that a less obtrusive recording method will be put in place for future experiments. Moreover, there were some interesting results obtained that point to the potential benefits, insights, and impact of the research.

A summary of the results is given below:

- Clear moments of high accommodation were observed in the speech data.
- The data suggested that clinicians made most of the effort to adapt, despite speaking less than patients.
- Prosodic features typical of depressed individuals were found (see [9, 18]).
- Some of those prosodic features were present in the clinicians' speech too: this indicates a dynamic aspect to the conversations, and there is clear evidence of clinical adaptation in some of the conversations.

5. Future Work

A series of studies will be carried out to test the results obtained in the pilot and investigate new areas of research.

A larger controlled study with a clinical protocol is underway with the collaboration of St. Patrick's University Hospital Dublin.

A different study will take place to examine correlation between prosodic accommodation and distinct brain states, in addition to providing material for the main investigation; this will provide further insight on the interpersonal dynamics at play during a conversation: EEG data will offer information on the interpersonal dynamics affecting the conversational behaviour (see [19]) of depressed patients in a clinical setting.

6. Acknowledgements

Special thanks to the project supervisors: Dr. Brian Vaughan and Dr. Charlie Cullen.

This study was funded by the Fiosraigh Dean of Graduate Research School Award, from Dublin Institute of Technology, Ireland.

7. References

- [1] WHO Europe, "Depression in Europe: facts and figures," 2016. [Online]. Available: <http://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/news/news/2012/10/depression-in-europe/depression-in-europe-facts-and-figures>
- [2] P. E. Greenberg, A.-A. Fournier, T. Sisitsky, C. T. Pike, and R. C. Kessler, "The Economic Burden of Adults With Major Depressive Disorder in the United States (2005 and 2010)," *The Journal of Clinical Psychiatry*, vol. 76, no. 2, pp. 155–162, feb 2015. [Online]. Available: <http://www.psychiatrist.com/jcp/article/Pages/2015/v76n02/v76n0204.aspx>
<http://www.psychiatrist.com/jcp/article/Pages/2015/v76n02/v76n0204.aspx>
- [3] P. Sobocki, B. Jonsson, J. Angst, and C. Rehnberg, "Cost of depression in Europe," *The journal of mental health policy and economics*, vol. 9, no. 2, pp. 87–98, jun 2006.
- [4] WHO Europe, "Depression in Europe," 2012. [Online]. Available: <http://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/news/news/2012/10/depression-in-europe>
- [5] M. Asgari, I. Shafran, and L. B. Sheeber, "Inferring clinical depression from speech and spoken utterances," in *2014 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, sep 2014, pp. 1–5. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6958856>
- [6] C. De Looze, C. Oertel, S. Rauzy, and N. Campbell, "Measuring Dynamics of Mimicry," *ICPhS*, vol. 1, no. August, pp. 1294–1297, 2011.
- [7] C. De Looze and S. Rauzy, "Measuring speakers' similarity in speech by means of prosodic cues: Methods and potential," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2011, pp. 1393–1396.
- [8] C. De Looze, S. Scherer, B. Vaughan, and N. Campbell, "Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction," *Speech Communication*, vol. 58, pp. 11–34, mar 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167639313001386>
<http://linkinghub.elsevier.com/retrieve/pii/S0167639313001386>
- [9] 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. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0167639315000369>
- [10] N. Cummins, J. Epps, M. Breakspear, and R. Goecke, "An investigation of depressed speech detection: Features and normalization," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2011, pp. 2997–3000.
- [11] E. Moore II, M. A. Clements, J. W. Peifer, and L. Weisser, "Critical Analysis of the Impact of Glottal Features in the Classification of Clinical Depression in Speech," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 96–107, jan 2008. [Online]. Available: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=>
<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4360055>
- [12] J. R. Williamson, T. F. Quatieri, B. S. Helfer, G. Ciccarelli, and D. D. Mehta, "Vocal and Facial Biomarkers of Depression based on Motor Incoordination and Timing," in *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge - AVEC '14*. New York, New York, USA: ACM Press, 2014, pp. 65–72. [Online]. Available: <http://doi.acm.org/10.1145/2661806.2661809>
<http://dl.acm.org/citation.cfm?doi=2661806.2661809>
- [13] L.-S. A. Low, N. C. Maddage, M. Lech, L. Sheeber, and N. Allen, "Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2010, pp. 5154–5157. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5495018>
- [14] D. J. France, R. G. Shiavi, S. Silverman, M. 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, jul 2000.
- [15] S. Alghowinem, R. Göcke, M. Wagner, J. Epps, M. Breakspear, and G. Parker, "From Joyous to Clinically Depressed: Mood Detection Using Spontaneous Speech," in *Twenty-Fifth International FLAIRS Conference*, 2012, pp. 141–146.

- [16] L.-s. A. Low, N. C. Maddage, M. Lech, L. Sheeber, and N. Allen, "Content based clinical depression detection in adolescents," in *17th European Signal Processing Conference (EUSIPCO 2009)*. Glasgow: EURASIP, 2009, pp. 2362–2366.
- [17] L.-S. A. Low, M. C. Maddage, M. Lech, L. B. Sheeber, and N. B. Allen, "Detection of Clinical Depression in Adolescents' Speech During Family Interactions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 574–586, mar 2011. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5629355>
- [18] S. Scherer, L.-P. Morency, J. Gratch, and J. Pestian, "Reduced vowel space is a robust indicator of psychological distress: A cross-corpus analysis," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, apr 2015, pp. 4789–4793. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7178880>
- [19] G. J. Stephens, L. J. Silbert, and U. Hasson, "Speaker-listener neural coupling underlies successful communication," *Proceedings of the National Academy of Sciences*, vol. 107, no. 32, pp. 14 425–14 430, aug 2010. [Online]. Available: <http://www.pnas.org/content/107/32/14425.full.pdf?withds=yes> <http://www.pnas.org/cgi/doi/10.1073/pnas.1008662107>