Capturing the onset of psychosis with automated speech markers

Julianna Olah[1]

[1]King's College London, Institute of Psychiatry, Psychology and Neuroscience
julianna.olah@kcl.ac.uk

Abstract

Psychotic disorders are characterized by key abnormalities in language and speech. Due to advances in technology, such abnormalities can be objectively and automatedly measured via speech analysis and Natural Language Processing techniques. Here, I present initial results on acoustic, prosodic, structural and semantic markers of speech that might be able to detect psychosis and its early signs from few-minutes long, remotely collected speech samples.

Index Terms: psychosis, semantic speech graph, latent semantic analysis, prosody, remote assessment

Introduction

Psychosis is a severe mental health condition and key characteristic of schizophrenia and bipolar disorder. Besides the enormous personal suffering, these disorders are associated with huge socioeconomic costs and healthcare burden. A key factor behind the huge burden of psychotic disorders is their delayed detection and intervention which results in increased social, personal, cognitive, and neurodevelopmental impact, worse prognosis, and poorer response to treatment [1].

Systematically and automatedly assessed speech markers represent promising digital phenotypes to objectively signal the presence or the later onset of psychosis as subtle changes in speech coherence and prosody appear in the early developmental stage of psychotic disorders [2-5]. Beyond diagnostic possibilities, studies have shown that speech can be used to identify individuals who will develop psychotic disorders, between 6 to 12 months before their clinical onset [6]. These findings have brought increased research interest from both clinical and engineering perspectives into the speech-based assessment and prediction of psychosis.

However, many former studies often used limited sample sizes and laboratory-based speech recordings that limits the testing of complex feature sets. Also, to date, the complex relationship between alterations in speech, related to psychosis and those related to often presented comorbid conditions such as depression and anxiety has not been investigated. It is, therefore, an open question whether speech markers can effectively discriminate between psychosis and other conditions or are just differ between healthy individuals and those with mental health conditions.

My PhD aims to investigate the connection between speech alterations and psychosis in a context of the challenges of translation into clinical application. I am investigating whether remotely collected, self-recorded speech samples can be used to identify early psychotic symptoms and whether speech alterations overlap in psychosis, depression and anxiety.

1. Research objectives

Psychotic disorders can be seen as extreme manifestations of unique, quantitative characteristics and symptoms that are continuously distributed in the general population [7], creating a spectrum. I am investigating acoustic, prosodic, semantic (coherence) and structural markers of speech collected from people at different stages from this spectrum to identify distinctive features that are associated with either clinical symptoms (from the extreme end of the spectrum), or subclinical psychotic-like experiences such as schizotypy and delusionary ideation [8] that are connected to the increased risk of developing psychotic. After the identification of distinctive speech markers, the final aim of the investigations is to build machine learning models to investigate whether these combined markers are informative for the detection of individuals who are having psychosis or are at elevated risk of having it in the future.

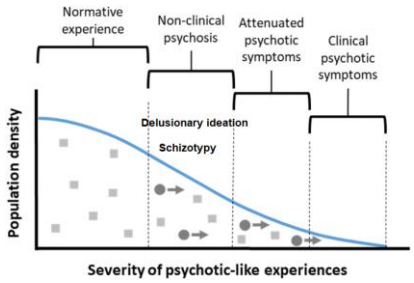

Figure 1: *Schizotypy and delusionary ideation on the psychosis spectrum.*

My concrete research goals are firstly, to test whether (i) depression, generalized anxiety and high schizotypy have similar voice characteristics; (ii), which acoustic markers of online collected speech are the strongest predictors of schizotypy; and (iii) whether controlling for generalized anxiety and depression symptoms can improve the prediction of it.

Secondly, to explore (i), the relation of speech coherence (semantic speech markers) and speech connectivity (structural speech markers) with schizotypy and delusionary ideation.

My third goal is to investigate whether the combination of semantic and structural speech markers can discriminate individuals with high schizotypy.

Finally, my fourth study will investigate (i) the relation of structural, semantic and acoustic speech markers in clinical sample; and (ii), test whether the combination of these features in machine learning models can discriminate between healthy and affected individuals.

2. Progress report

To date, I have completed the first and third study, while for the second study, preliminary results are available on a subset (around 70%) of the target sample. Study four will be completed in the next year, depending upon data availability.

3. First results

Study 1: Exploring the relationship between overlapping acoustic markers of schizotypy, depression and anxiety.

Data & Design: We collected cross-sectional (N = 441), online-recorded speech data from the general population, assessing demographics, symptoms of depression and generalized anxiety and schizotypy. Participants were asked to describe 8 ambiguous pictures for 1 minute each for the speech task, resulting in 8-minutes long sample per participants. Acoustic and prosodic features were extracted using the openSMILE software using the extended Geneva Acoustic Minimalistic Parameter Set (eGeMAPS).

Results: Online collected speech could predict schizotypy, depression, and anxiety symptoms, however, most influential features of these models largely overlapped. Only loudness measures were uniquely represented in schizotypy prediction. The predictive power of speech marker-based models of schizotypy significantly improved after controlling for symptoms of depression and generalized anxiety (from R2 = 0.296 to R2 = 0. 436).

Conclusions: Acoustic features of online collected speech are predictive of psychometric schizotypy as well generalized anxiety and depression symptoms. The acoustic characteristics of schizotypy, depression and anxiety symptoms significantly overlap, to the best of my knowledge, this is the first time such a result has been shown. Speech models that are designed to predict schizotypy or symptoms of the psychosis spectrum might therefore benefit from controlling for symptoms of depression and anxiety.

Study 2: Exploring the relation of subclinical psychotic-like experiences with semantic and structural speech markers

Data & Design: We collected cross-sectional (N = 500), online-recorded speech data from the general population, assessing demographics, schizotypy, delusional ideation with the picture-description task described in Study 1. Transcripts were used to produce non-semantic speech graphs [9] which generate measures of speech connectivity. We calculated semantic coherence measures using latent semantic analysis by the methodology of Iter and colleagues [10].

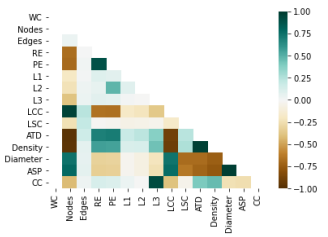

Figure 2: *Correlation of speech connectivity measures with each other*

Results: From the connectivity measures, the *number of parallel edges*, and the *loop of two nodes* were positively and significantly related to delusionary ideation. From coherence measures, *On Topic* score had significant negative relationships with schizotypy. Speech measures showed moderate correlation with each other.

Conclusions: Further analyses are needed, but initial results indicate that speech features appear to be complementary to each other and different features show unique and distinct relation to given psychotic like symptoms.

Study 3: Testing whether semantic and structural speech markers can discriminate between low and high schizotypy

Data & Design: Online recorded speech samples from a picture description test were collected from the general population (N = 482). Schizotypy was measured and recoded into binary high-low groups. According to the methodology described in Study 2, non-semantic speech network analysis and latent semantic analysis were conducted resulting in 20 features of semantic coherence and connectivity. An independent sample T-test was conducted to test for group differences in these measures. Two supervised machine learning approaches, *logistic linear regression* and *Extra Trees classifier* were applied to test the classification performance of the features.

***Results*:** Tangentiality, on topic, number of sentences, number of words, a median of the number of edges, a median of the largest connected component of the speech measures showed significant group differences. Logistic linear regression model could classify individuals with 0. 74 accuracy (precision = 0.94/0.36; recall= 0.74/0.77) while Extra Trees classifier could classify individuals with 0.82 accuracy (precision = 0.90/0.45; recall= 0.88/0.50).

Conclusions: Some speech patterns that characterize psychotic disorders like less coherent and less wordy speech appear to be present in individuals with high schizotypy. Automatedly assessable speech features can discriminate between people with high and low schizotypy. Although further improvements are needed in the classification power of speech feature-based models, initial findings suggest that by further refinement and wider involvement of coherence and connectedness-related markers, there is a potential to identify symptoms of the psychosis continuum via automated, online, few-minutes long speech assessment.

4. Future work

Future work will investigate whether the observed relation between speech markers is present in a clinical population and test how the predictive power of speech alterations change when we discriminate healthy and clinical cases. Given their complimentary nature, I will build both separate and combined models of semantic, structural and acoustic features to predict clinical status and investigate whether given combinations can improve performance. Finally, I would like to investigate whether controlling for anxiety and depression symptoms within machine learning models improves predictive performance on clinical data. By taking part in the doctorial consortium, I hope to get valuable feedback regarding the machine learning and feature combination aspects of my future work.

5. Acknowledgements

Above many others, I sincerely thank Dr. Nicholas Cummins, Dr. Thomas Spencer and Dr. Kelly Diederen for their invaluable guidance and support. My PhD is being supported by the KCL Centre for Doctoral Training in Data-Driven Health.

6. References

McGorry, P. D., Nelson, B., Goldstone, S., and Yung, A. R., "Clinical staging: A heuristic and practical strategy for new research and better health and social outcomes for psychotic and related mood disorders." *The Canadian Journal of Psychiatry*, vol. 55, no. 8, pp. 486-497., 2010

[1] Corcoran, C. M., and Cecchi, G. A, "Using language processing and speech analysis for the identification of psychosis and other disorders." *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 5, no. 8, pp. 770-779, 2020.

[2] Palaniyappan, L., Mota, N. B., Oowise, S., Balain, V., Copelli, M., Ribeiro, S., & Liddle, P. F., "Speech structure links the neural and socio-behavioural correlates of psychotic disorders." *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 88, pp. 112-120, 2019

[3] Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., Bearden, C. E., and Cecchi, G. A., " Prediction of psychosis across protocols and risk cohorts using automated language analysis". *World Psychiatry*, vol. 17,no.), pp. 67-75, 2018

[4] Agurto, C., Pietrowicz, M., Norel, R., Eyigoz, E. K., Stanislawski, E., Cecchi, G., and Corcoran, C.. "Analyzing acoustic and prosodic fluctuations in free speech to predict psychosis onset in high-risk youths." *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC).*, 2020

[5] Mota, N. B., Copelli, M., and Ribeiro, S. "Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance." *npj Schizophrenia*, vol. 3, no. 1, 2017

[6] DeRosse, P., and Karlsgodt, K. H., "Examining the psychosis continuum." *Current Behavioral Neuroscience Reports*, vol. 2, no. 2, pp. 80-89, 2015

[7] Debbane, M., Eliez, S., Badoud, D., Conus, P., Fluckiger, R., and Schultze-Lutter, F, "Developing psychosis and its risk states through the lens of schizotypy." *Schizophrenia Bulletin*, vol. 41, pp. S396-S407, 2014

[8] Mota, N. B., Vasconcelos, N. A., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., Copelli, M., and Ribeiro, S., "Speech graphs provide a quantitative measure of thought disorder in psychosis." *PLoS ONE*, vol. 7, no. 4, 2014

[9] Iter, D., Yoon, J., and Jurafsky, D. "Automatic detection of incoherent speech for diagnosing schizophrenia*." Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*., 2018