# Optimizing automated objective speech technologies for measuring negative symptoms

**Submission ID** 3000970

## SUBMISSION DETAILS

**I agree to provide poster pdf for attendee download.** Yes

**Methodological Issue Being Addressed** What speaking tasks and speech features are optimal for measuring negative symptoms using automated speech analysis?

**Introduction** Measurement of negative symptoms is dependent on expert ratings from semi-structured interviews. While important for many applications, these ratings are relatively insensitive to change, have reliability constraints, are time consuming to procure and are burdensome for patients and clinic staff. Automated and objective speech analysis is a promising complementary approach to measuring negative symptoms. With respect to clinical trials, it can potentially be used as exploratory endpoints and can screen participants for study suitability.

The present project evaluated reliability and validity of an automated, objective, face-valid and computationally transparent tool for measuring negative symptoms from natural speech. Archived recordings of patient speech were analyzed from a variety of clinical interview, conversational, mobile recording, cognitive, and laboratory tasks with varying probes, instructions and lengths. This allowed us to evaluate: a) the reliability of key speech features with each other (internal consistency), b) the reliability of speech across different speaking tasks (inter-domain reliability, and c) convergence with clinical ratings (convergent validity).

**Methods** Archived audio recordings from nine separate data collections for 19 different speaking tasks were evaluated using a fully automated acoustic and natural language analytic pipeline. Six features were extracted, including: latencies between responses, phrases and words (i.e., average latency before initiating speech, between phrases, and between words respectively), utterance word count, utterance duration, and articulation rate. These features were selected based on their face validity with negative symptoms, their use in the extant literature, their low computational demands and their transparency in calculation. The latter is important for replication and application across socio-culturally diverse peoples.

A variety of clinical ratings scales were used in these studies. Patients were binarized based on expressive negative symptoms being present (i.e., mild or greater; n = 308 people; k = 1401 recordings for analysis) or absent (i.e., absent or questionable; n = 524 people; k = 2978 recordings for analysis). In total, 114,345 speech utterances available were analyzed for this study.

**Results** Reliability analysis suggested that the six features were modestly inter-correlated (coefficient alpha = 0.52), and were relatively dynamic across speaking task (range of single ICC values = -0.29 to 0.38). There was considerable variability in how these features differed in people with versus without negative symptoms (mean of maximum Cohen’s d values for any task comparing patients with versus without negative symptoms = 0.60; range = 0.27 to 1.06). The strongest effects involved exaggerated speech latencies while patients were interacting with an interviewer. The weakest effects were observed while patients performed a cognitive task using a mobile application.

**Conclusion** Automated speech-analysis has potential as a reliable and valid measure of negative symptoms. It is unlikely there is a “one-size” fits all solution for measuring negative symptoms across all tasks and all seemingly relevant speech measures. From our data, clinically rated negative symptoms were best captured using response latencies from interviews. Reasons for this are discussed, as are implications for applying speech analysis to speech from other tasks. These studies were conducted using existing corpuses using tasks with varying levels of standardization. Future studies would benefit from prospective interviews that are standardized (e.g., in terms of interviewer training, duration of interview, between-visit interval, study entry criteria).

## Co-Authors

* Presenting Author

| First Name | Last Name | Affiliation |
|---|---|---|
| Alex * | Cohen * | Louisiana State University, Department of Psychology |
| Mark | Opler | WCG Clinical, Inc. |
| Brian | Kirkpatrick | Psychiatric Research Institute, University of Arkansas for Medical Sciences |
| Gregory | Strauss | University of Georgia, Department of Psychology |
| Seth | Hopkins | Sumitomo Pharmaceuticals America |


## Keywords

| Keywords |
|---|
| Speech |
| Negative symptoms |
| Acoustics |
| Language |
| Cognition |


**Guidelines** I have read and understand the Poster Guidelines

**Disclosures if applicable** Mark Opler, Brian Kirkpatrick, Gregory Strauss and Alex Cohen are owners of Quantic Innovation, a digital phenotyping company working with Sumitomo

Pharmaceuticals America