

# Lexical predictability in schizophrenia: a computational approach to quantifying and understanding thought disorder

Victoria Sharpe<sup>1</sup>, Sabrina Ford<sup>2</sup>, Samer Nour-Eddine<sup>1</sup>, Lena Palaniyappan<sup>2,3</sup>, and Gina Kuperberg<sup>1,4</sup>

<sup>1</sup>Tufts University, Department of Psychology, Medford, Massachusetts, United States

<sup>2</sup>Robarts Research Institute, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

<sup>3</sup>Douglas Mental Health University Institute, Department of Psychiatry, McGill University, Montreal, Quebec, Canada

<sup>4</sup>Department of Psychiatry and the Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, United States

## Background:

Language has long been considered a window into thought in schizophrenia. Computational language models allow us to extract objective, quantifiable data from natural speech that can potentially serve as a clinical biomarker. However, it is unclear how these measures relate to clinical symptoms (particularly positive thought disorder) and neurocognitive processes. In healthy adults, lexical predictability, the likelihood that a given word will be produced based on its prior context, is among the best predictors of behavioral processing and neural activity. There is also a large body of evidence that the predictability of language output plays an important role in communication. We therefore used a predictive language model, GPT-2, to quantify word-by-word predictability in natural speech from people with schizophrenia. We asked whether, relative to healthy adults, (a) lexical predictability is reduced in the language output of patients; (b) patients are relatively more impaired in using global versus local context to produce upcoming words; and (c) whether these abnormalities are linked to clinical ratings of thought disorder.

## Methods:

We asked 74 first-episode psychosis patients (FEP), 16 chronic schizophrenia patients (CS), and 36 healthy controls to describe three pictures for one minute each. For each word in the transcribed speech, we used GPT-2 to extract measures of lexical predictability given different context lengths, ranging from all available context (very global) to only the previous word (very local). We then used mixed effects regression, controlling for demographic factors and item-level variables (e.g., transcript length), to

test our hypotheses. Within the patient groups (FEP + CS combined), we also probed correlations with thought disorder, as measured using the Thought and Language Index (TLI).

### **Results:**

We found a main effect of Group on lexical predictability, such that the speech produced by FEPs was significantly less predictable than that produced by healthy controls (Est. = -0.09,  $p = 0.00$ ); the difference between the CS group and controls was not significant (Est. = -0.06,  $p = 0.20$ ), possibly due to the small sample size of the CS group. Moreover, there was a significant interaction between Group (FEP vs. controls) and Context Length (which ranged from 1 – 30 words), such that the effect of Context Length on predictability was smaller in the FEP group (Est. = -0.03,  $p = 0.00$ ). Follow-up analyses confirmed that, as predicted, this interaction was driven by lower global predictability in the FEP group than in controls (Est. = -0.05,  $p = 0.03$ ); in contrast, differences in local predictability were numerically smaller and non-significant (Est. = -0.01,  $p = 0.64$ ).

Among patients, the effect of Context Length on lexical predictability was smaller in patients with more severe thought disorder (a significant interaction between Context Length and TLI score; Est. = -0.01,  $p = 0.02$ ). This effect was again driven by a reduced sensitivity to global context in patients with thought disorder.

### **Discussion:**

These findings suggest that the incoherent language output often seen in schizophrenia may relate to impairments in using global (vs. local) context. This is in line with previous neural evidence showing that patients with schizophrenia have difficulties using global (vs. local) context to predict upcoming words during language comprehension. At a mechanistic level, this work connects to a large body of research documenting abnormalities in predictive processing in schizophrenia across multiple domains. We suggest that lexical predictability may provide a useful metric that is easily quantified by computational models, has face validity with thought disorder, and may provide insights into neurocognitive mechanism.