Phonetic Measurement of Reduced Facial Muscle Movement among Young Adults with First-Episode Schizophrenia-Spectrum Disorders

Michael A. Covington, Anya Lunden (The University of Georgia)

Sarah Cristofaro, Stephanie Johnson, Claire Ramsay (Emory University)

Beth Broussard (The George Washington University)

Shayi Zhang, C. Thomas Bailey, Robert Fogarty (The University of Georgia)

Michael T. Compton (The George Washington University)

Background

Disturbances of language are prominent symptoms of schizophrenia. This project is part of a larger research agenda to develop computer-based measurements of language impairment in the context of schizophrenia.

One of the "negative symptoms" of schizophrenia is monotonous, indistinct speech. Another is reduced movement of facial muscles. Cannizzarro et al. (2005) conclude that facial bradykinesia (slowing of muscle movement) contributes to altered speech quality in schizophrenia.

Approach

Many of the muscle movements involved in producing speech can be measured phonetically as variation of formant frequencies. Indeed, formants are how humans recognize many speech sounds, especially vowels.

Formants F1 and F2 indicate, roughly, tongue height or mouth opening and tongue front-to-back position, respectively. The absolute value of these quantities depends on the speech sound being made and the speaker; what interests us is their variability.

Although it is not a formant, the pitch at which the vocal folds are vibrating is traditionally designated F0. If speech in schizophrenia is indeed monotonous, reduced variation of F0 should correlate with negative symptom severity in schizophrenia.

Recruitment

This study is part of the ACES (Allied Cohort on Early course of Schizophrenia) project at Emory and George Washington University. Data were collected at Emory University as part of a larger study of the early course of psychotic disorders in a predominantly African-American, low-income, socially disadvantaged population.

First-episode patients aged 18-40 years were eligible to participate. Exclusion criteria included inability to speak English, known mental retardation, MMSE score < 24, or medical conditions that could compromise ability to participate.

Each patient was presenting with schizophrenia for the first time, and virtually all were assessed within a week of initiating antipsychotic treatment.

Data collection

All patients were recorded on video during clinical interviews. The purpose of the video recordings was to assess gross movement abnormalities, such as dyskinesias and stereotypies.

Diagnoses of psychotic disorders and substance use disorders were made using the *Structured Clinical Interview for DSM-IV Axis I Disorders* (SCID). Symptom severity was assessed by clinical researchers (blinded to the ratings of speech abnormality, which had not been made yet) with the widely used *Positive and Negative Syndrome Scale* (PANSS).

Of 56 recordings attempted, 47 yielded video quality adequate for the original purpose. However, many of these had sound quality insufficient for speech analysis.

To select usable recordings fairly, the audio tracks were extracted and sent to the University of Georgia team, who were blinded to all information about the patients. Two research team members independently rated the files for intelligibility and audio quality and selected the 27 files with the best audio quality (25 after exclusion of two ineligible patients who had inadvertently been included).
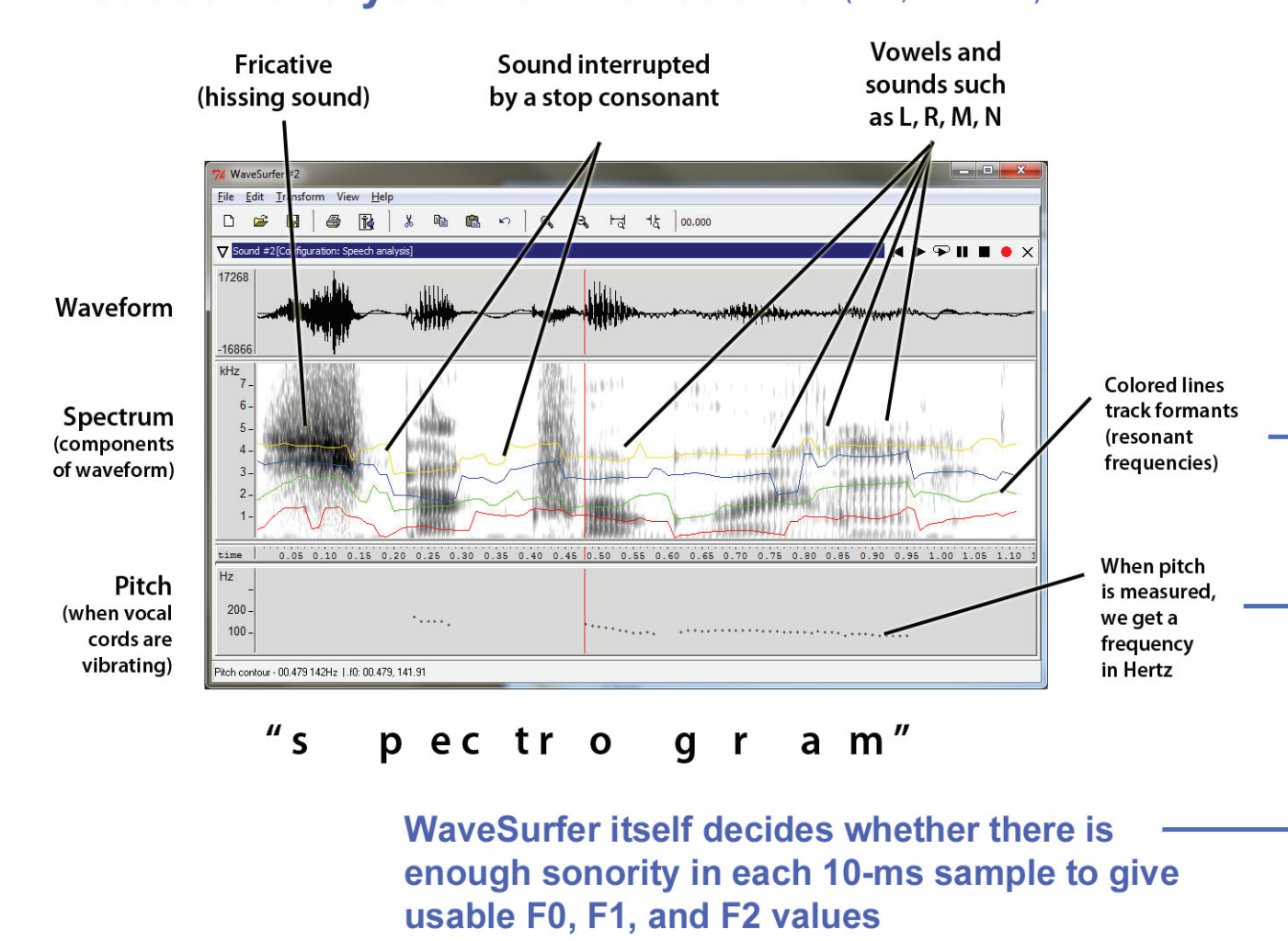
The files were denoised using GoldWave 5.58. Then a trained phonologist extracted a 1-minute sample of each patient's speech, starting 5 minutes into the recording and excluding speech by the interviewer, parts of the recording with noises in the background, and segments when the patient was saying "um-hmm" or similar very short responses.

Analysis

From each sample we created a file of speech data using WaveSurfer 1.8.5. Each file contains F0, F1, and F2 for each 10-millisecond segment, coding F0 as 0 if, according to WaveSurfer's spectral analysis, the vocal folds do not appear to be vibrating or there is for some other reason insufficient sonority for measuring formants.

Using a computer program written locally in Microsoft C#, we computed the standard deviation of log F0 (pitch variability) and the standard deviations of F1 and F2, indicating muscle movements.

Acoustic analysis with WaveSurfer (KTH, Stockholm)

WaveSurfer itself decides whether there is enough sonority in each 10-ms sample to give usable F0, F1, and F2 values

Results

Speech measurements were tested for correlation with the PANSS negative symptom subscale score, with the following results:

| PANSS negative total vs:	Pearson correlation	P (2-tailed)
$SD_N \log_{10} F0$	-0.107	0.611
SD_N F1	-0.339	0.098 (suggestive)
SD_N F2	-0.446	0.026 (significant)

Increased negative symptoms correlated:

**Significantly** with reduced standard deviation of F2, indicating reduced front-to-back movement.

Suggestively with reduced standard deviation of F1, indicating reduced movement in height and/or mouth opening. (A one-tailed test could have been justified since the hypothesis was one-sided, and then this result would have come out significant.)

Not at all with reduced standard deviation of F0, which would indicate flattened intonation.

Movement in tongue height and/or mouth opening
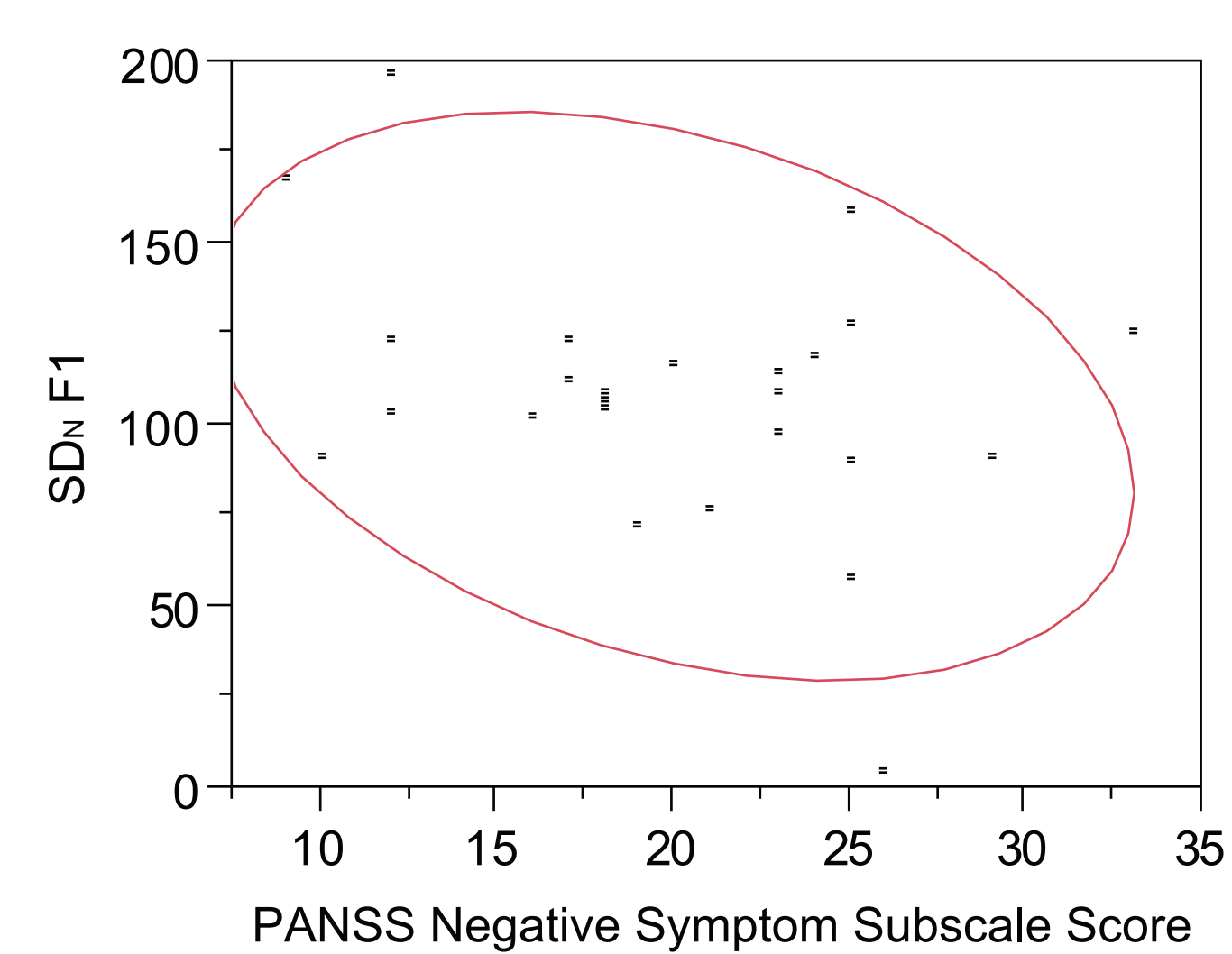
Movement in tongue front-to-back position
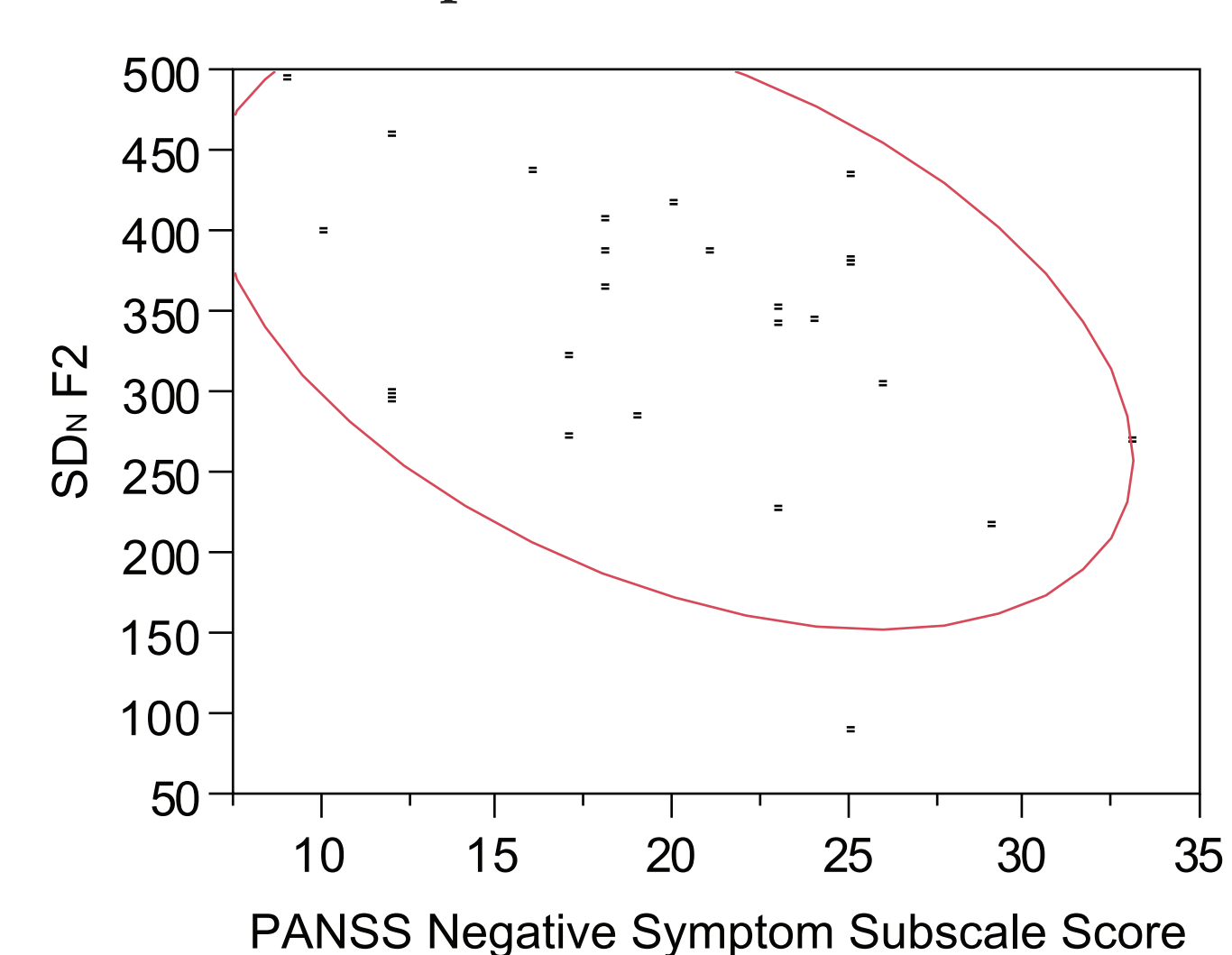
This scale measures "negative symptoms" such as reduced activity, avolition, and other impairments, as distinct from "positive symptoms" such as hallucinations, delusions, or suspiciousness.

Discussion

Phonetic analysis shows a clear correlation between reduced front-back tongue movement and severity of negative symptoms. There is a suggestive correlation between tongue height movement and severity of negative symptoms. We conclude that reduced facial movement in schizophrenia can be demonstrated phonetically.

Surprisingly, we found no correlation between negative symptoms and reduced variation in F0 (measured logarithmically as subjective pitch). Two explanations occur to us. One is that "aprosody" in schizophrenia is not actually pitch flattening, but rather lack of clear articulation; the other is that our patients were all above some minimum threshold of severity, so they all had aprosody.

Future directions

Traditionally, clinicians have noted the quality of patients' speech subjectively and anecdotally. We introduce precise computer-aided measurements which could also be relevant to the assessment of other conditions, such as depression and mania.

This study was limited by the small subject population and the low quality of the recordings, which were originally made for a different purpose. We are planning a further study with a larger population (including healthy controls, not just patients of varying symptom severity), higher recording quality, and a variety of additional language-related cognitive tasks.

Because these first-episode patients had been on medication only a short time, we do not think the abnormalities in their speech were induced by medication; but even if they were, the technique for measuring the abnormalities is useful.

References

Alpert, M., Rosen, A., Welkowitz, J., Sobin, C., Borod, J., 1989. Vocal acoustic correlates of flat affect in schizophrenia: similarity to Parkinson's disease and right hemisphere disease and contrast to depression. Br. J. Psychiatry 154, 51—56.

Alpert, M., Rosenberg, S. D., Pouget, E. R., Shaw, R. J., 2000. Prosody and lexical affect in flat affect schizophrenia. Psychiatry Res. 97, 107—118.

Cannizzarro, M. S., Cohen, H., Rappard, F., Snyder, P. J., 2005. Bradyphrenia and bradykinesia both contribute to altered speech in schizophrenia: a quantitative acoustic study. Cogn. Behav. Neurol. 18, 206—210.

Covington, M. A., He, C., Brown, C., Naçi, L., McClain, J. T., Fjordbak, B. S., Semple, J., Brown, J., 2005. Schizophrenia and the structure of language: the linguist's view. Schiz. Res. 77, 85—98.

Cutting, J., 1985. The psychology of schizophrenia. Churchill Livingstone, Edinburgh.

Schneider, F., Heimann, H., Himer, W., Huss, D., Mattes, R., Adam, B., 1990. Computer-based analysis of facial action in schizophrenic and depressed patients. Eur. Arch. Psychiatry Clin. Neurosci. 240, 67—76.

Spoerri, T., 1966. Speaking voice of the schizophrenic patient. Arch. Gen. Psychiatry 14, 581—585.

Trémeau, F., Malaspina, D., Duval, F., Corréa, H., Hager-Budny, M., Coin-Bariou, L., Macher, J. P., Gorman, J. M., 2005. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. Am. J. Psychiatry 162, 92—101.

Georgetown University Round Table (GURT 2012)