Title
Vocal Expression in Schizophrenia: Less Than Meets the Ear

Journal
Journal of Psychopathology and Clinical Science, 125(2)

ISSN
2769-7541

Authors
Cohen, Alex S
Mitchell, Kyle R
Docherty, Nancy M
et al.

Publication Date
2016-02-01

DOI
10.1037/abn0000136

Peer reviewed

HHS Public Access

Author manuscript

J Abnorm Psychol. Author manuscript; available in PMC 2016 July 30.

Published in final edited form as:

J Abnorm Psychol. 2016 February ; 125(2): 299–309. doi:10.1037/abn0000136.

Vocal expression in schizophrenia: Less than meets the ear

Alex S. Cohen[a], **Kyle R. Mitchell**[a], **Nancy M. Docherty**[b], and **William P. Horan**[c]

[a]Louisiana State University, Department of Psychology

[b]Kent State University, Department of Psychology

[c]VA Greater Los Angeles Healthcare System, University of California, MIRECC

Abstract

Abnormalities in nonverbal communication are a hallmark of schizophrenia. Results from studies using symptom rating scales suggest that these abnormalities are profound (i.e., 3–5 standard deviations) and occur across virtually every channel of vocal expression. Computerized acoustic analytic technologies, employed to overcome practical and psychometric limitations with symptom rating scales, have found much more benign and isolated abnormalities. In order to better understand vocal deficits in schizophrenia and to advance acoustic analytic technologies for clinical and research applications, we examined archived speech samples from five separate studies, each employing different speaking tasks (patient $N = 309$; control $N = 117$). We sought to: a) employ Principal Component Analysis (PCA) to identify independent vocal expression measures from a large set of variables, b) quantify how patients with schizophrenia are abnormal with respect to these variables, c) evaluate the impact of demographic and contextual factors (e.g., study site, speaking task), and d) examine the relationship between clinically–rated psychiatric symptoms and vocal variables. PCA identified seven independent markers of vocal expression. Most of these vocal variables varied considerably as a function of context and many were associated with demographic factors. After controlling for context and demographics, there were no meaningful differences in vocal expression between patients and controls. Within patients, vocal variables were associated with a range of psychiatric symptoms – though only pause length was significantly associated with clinically-rated negative symptoms. The discussion centers on explaining the apparent discordance between clinical and computerized speech measures.

Keywords

schizophrenia; negative; alogia; blunted affect; computer

Abnormalities in vocal expression, defined in terms of decreased/increased vocal production (i.e., alogia, pressured speech) and intonation/emphasis (i.e., blunted affect; affective lability) are a staple of schizophrenia and of serious mental illness more generally (Cohen, Najolia, Kim, & Dinzeo, 2012; Galynker, Cohen, & Cai, 2000; Kulhara & Chadda, 1987; Mueser et al., 2010; Tremeau et al., 2005). Despite vocal expression reflecting an important

Send correspondence to: Alex S. Cohen, Ph.D., Louisiana State University, Department of Psychology, 236 Audubon Hall, Baton Rouge, LA 70803, Phone: (225) 578-7017, Fax: (225) 578-4125, acohen@lsu.edu.

Author note: David B. Watson served as the guest editor for this manuscript.

Research Domain Criteria (RDoC; "Production of Non-Facial Communication") our understanding of its nature is poor. An obstacle in understanding and measuring vocal abnormalities is a reliance on interviewer-based rating scales (e.g., Andreasen, 1989; Kirkpatrick et al., 2011; Kring, Gur, Blanchard, Horan, & Reise, 2013; Lukoff, Nuechterlein, & Ventura, 1986). These scales are: relatively insensitive to change in a treatment context given their limited range of response options and ambiguous operational definitions, produce ordinal data that often cover wide temporal swaths, and are imprecise for isolating specific behaviors (Alpert, Shaw, Pouget, & Lim, 2002; Cohen, Alpert, Nienow, Dinzeo, & Docherty, 2008; Cohen & Elvevåg, 2014; Cohen, Mitchell, & Elvevåg, 2014; Mueser, Sayers, Schooler, Mance, & Haas, 1994; Stahl & Buckley, 2007). Emerging computerized technologies allow for assessment of vocal symptoms with near perfect "inter-rater" reliability and greater sensitivity and specificity for isolating specific behaviors than clinical rating scales, and thus can provide a more fine grained, and objective analysis compared to clinical rating scales assessing vocal psychopathology (Cohen & Elvevåg, 2014). These technologies are particularly attractive in that they can be used in mobile assessment and telemedicine applications, and can be automated such that they are efficient and inexpensive for clinical use (e.g., Satt, Hoory, König, Aalten, & Robert, 2014). Moreover, vocal variables have been associated with clinical measures of functioning in at least some studies (Cohen, Najolia, Kim and Dinzeo, 2012), and the behaviors they tap are part of the operational definition of social competence (e.g., Bellack, Morrison, Mueser, Wade & Sayers, 1990). The present study applied acoustic-analytic technologies to understand vocal deficits in schizophrenia using archived natural speech samples from five independent studies.

A recent meta-analysis (Cohen, Mitchell, et al., 2014) highlights potential limitations with clinical ratings scales and our poor understanding of vocal abnormalities in schizophrenia more generally. With respect to vocal expression ratings from the Scale for the Assessment of Negative Symptoms (e.g., blunted vocal affect; Andreasen, 1989), patients with schizophrenia versus nonpatient controls showed profound deficits on the order of three to five standard deviations (i.e., Cohen's $d = 3.42 – 4.39$). In contrast, computer-based measures of vocal expression were associated with more modest, and variable, abnormalities in patients. For example, patients showed normal vocal patterns in some variables (e.g., latency of response; $d = -.21$, k = 2) and medium to large effects in variables related to vocal production (e.g., number of words expressed; $d = -.60$, k = 6; average pause length, $d = -1.10$; k = 4). These findings suggest that, in contrast to results from symptom-rating based studies, only isolated aspects of vocal expression are abnormal in schizophrenia. Consider further a recent meta-analysis of six studies employing behavioral coding of patients while engaged in structured laboratory tasks. The range of effect sizes was fairly broad (range of d's = .61 to 1.95; $d_{mean} = 1.10$; Hoekert & Kahn, 2007), perhaps reflecting differences in coding systems and tasks across studies. Moreover, ratings were fairly holistic in their appraisal of vocal expression; a potential issue in that vocal expression is multi-faceted. Regardless, even behavioral ratings fail to approximate the magnitude of deficits associated with symptom rating scales. Localizing and clarifying the vocal abnormalities associated with schizophrenia is an important step towards understanding the nature and

pathophysiology of negative symptoms, and for identifying biomarkers for their objective assessment. This was the first aim of this study.

When evaluating vocal deficits in patients, it is important to consider that vocal expression is not static. In the general population, vocal expression is highly dynamic across contexts and is modulated by a number of factors, such as presence/absence of social stimuli (Wells & MacFarlane, 1998), cognitive resources (Cohen, Dinzeo, Donovan, Brown, & Morrison, 2014), and affective/arousal state (e.g., angry, happy), (Cohen, Minor, Najolia, & Hong, 2009; Johnstone et al., 2007; Sobin & Alpert, 1999) to name a few. In schizophrenia, vocal deficits are conceptualized as being temporally stable and occurring across contexts (e.g., Kirkpatrick & Galderisi, 2008) – a phenomenon that has been referred to as "Affective Nonresponsivity" (Andreasen, 1989). Indeed, symptom rating scales tapping vocal deficits show modest to high levels of stability over extended periods of time (Long & Brekke, 1999; Strauss, Harrow, Grossman, & Rosen, 2010). However, several studies employing vocal/speech analysis while manipulating speaking context - by increasing the cognitive demands of the speaking task – found that aspects of vocal expression worsen as a function of context (i.e., word count in Barch & Berenbaum, 1994; pause length in Cohen, McGovern, Dinzeo, & Covington, 2014). This suggests that vocal deficits are exacerbated by certain contextual demands, and are not as static as indicated by symptom rating scale data. Insofar as cognitive demands play a role, a wide range of contextual variables (presumably related to cognitive resources when speaking), such as complexity of speech topic, competing tasks (e.g., driving, walking) and familiarity with speaker may be important to understanding vocal deficits. This topic has received very little attention to date. Answering the question "Does vocal expression in schizophrenia vary as a function of contextual variables" was our second aim.

Additionally, symptom heterogeneity and comorbidity in schizophrenia may be relevant. Vocal abnormalities in schizophrenia have primarily been examined in the context of negative symptoms, for which increased pauses, decreased signal variability (e.g., change in frequency, intensity), and tonal abnormalities have been associated with negative symptom severity in at least some studies (Alpert et al., 2002; Cohen et al., 2008; Cohen, Mitchell, et al., 2014). However, these same vocal anomalies have been associated with other clinical disorders and symptoms, such as those involving depression and anxiety (Cannizzaro, Harel, Reilly, Chappell, & Snyder, 2004; Cohen, Kim, & Najolia, 2013; Mundt, Vogel, Feltner, & Lenderking, 2012); symptoms that are common in schizophrenia and conceptually related to negative symptoms. Moreover, vocal exaggerations, as opposed to deficits, may be observed in some patients with schizophrenia in that hostility, mania, and even thought disorder are conceptually related to increases in vocal production (e.g., pressured speech), intonation and emphasis (Sobin & Alpert, 1999). The third aim of this study was to evaluate relationships between vocal expression and a broad range of psychiatric symptoms.

The present study leveraged acoustic analytic technologies to understand vocal expression in schizophrenia and nonpsychiatric controls from a large database of speech samples, collected from five different studies – each employing different speaking tasks and procedures. Given the relative novelty of this endeavor, and the considerable inconsistency in vocal features employed in prior studies, (Cohen, Mitchell, et al., 2014; Cohen, Renshaw,

Mitchell, & Kim, 2015) our first step was to subject a relatively large number of commonly-used vocal variables to data reduction procedures (e.g., Principal Component Analysis [PCA]). Following this, we evaluated the degree to which vocal expression varied across contextual factors (represented in the five separate speaking tasks/study sites) and whether, after controlling for these contextual factors, patients differed from controls. Finally, we evaluated the links between vocal expression and various psychiatric symptoms in the patient group to explore potential heterogeneity across patients.

METHODS

Participants

Data were aggregated from five separate archived studies (Cohen et al., 2008; Cohen et al., 2013; Cohen, McGovern, et al., 2014; Docherty, 2012; Horan et al., 2009) conducted with outpatients and long-term forensic inpatients meeting criteria for *Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition* (DSM-IV; American Psychiatric Association, 1994) schizophrenia or schizoaffective disorders. Diagnoses were established using structured clinical interviews (First, Gibbon, Spitzer, & Williams, 1996). All patients were deemed clinically stable at the time of testing and were receiving treatment from multidisciplinary teams. Nonpatient controls from these studies were also examined, for whom lack of psychosis and mood-spectrum diagnoses was established by structured clinical interviews (First et al., 1996). Data for 309 patients and 117 nonpatient controls were available. Table 1 contains demographic and clinical information. Each of the studies was approved by the appropriate Institutional Review Boards and all subjects provided written informed consent prior to beginning the study.

Diagnostic and symptom ratings

Psychiatric symptoms were measured using either the 18-item (Overall & Gorham, 1962) or the Expanded Brief Psychiatric Rating Scale (BPRS; Lukoff et al., 1986). Unfortunately, only factor scores (per Guy, 1976) were available for all five studies. Cronbach Alpha values were generally adequate (i.e., > .60), though two scales (i.e., Activation and Hostility) were combined to achieve this. Factors included: Thought Disturbance (α = .72), Anergia (α = . 73), Anxiety-Depression (α = .74.) and Activation/Hostility (α = .62).

Speaking tasks (i.e., context)

Across studies, participants were asked to produce speech as part of one of five different tasks that varied in topic, cognitive demands, interviewer participation and length (see Table 1). For Studies 1 (Cohen et al., 2008) and 5 (Docherty, 2012), research assistants engaged the subject in a semi-structured conversation about daily activities, hobbies and interests. The speech sample for study 1 was 300 seconds and 600 seconds for study 5; though due to computing demands of processing extended speech, only the first 300 seconds of the latter samples were analyzed. Studies 2 (Cohen et al., 2013) and 3 (Cohen, McGovern, et al., 2014) involved having subjects provide monologues on daily activities/hobbies (180 seconds) and neutrally-valenced autobiographical memories (90 seconds) respectively, for which the interviewer was absent the subjects' field of vision and was silent for the task. Although Studies 2 and 3 were conducted in similar geographic regions and similar

treatment settings and with similar research assistants, we opted to consider them separately given differences in speaking time and content, which in a prior study of these tasks in nonpatients was associated with significant differences in vocal variables (Cohen, Renshaw, Mitchell, & Kim, 2015). Study 4 (Horan et al., 2009) involved verbal responses from a social cognition task, involving responses to 12 different videos containing animated shapes (range of responses = 42 – 75 seconds; Castelli, Frith, Happe & Frith, 2002). All interviewer speech was digitally removed from the speech samples before analysis. Age did not significantly vary across the five studies (p's > .68). Study 5 had a significantly higher percentage of women than the other studies ($x^2[4] = 13.38$, $p = .01$). With respect to ethnicity, Study 3 had significantly more Caucasians than the other studies ($x^2[4] = 12.64$, $p = .01$), and Studies 3 and 4 had significantly fewer African-Americans than the other studies ($x^2[4] = 37.97$, $p < .001$). Some significant differences in Thought Disturbance (Study 5 < all other studies), Anxiety Depression (Study 1 < all other studies) and Activation-Hostility (Study 5 > than Studies 1, 3 & 4) were observed as well (Scheffe test p's < .01).

Computerized acoustic analysis of natural speech

The Computerized assessment of Affect from Natural Speech (CANS; Cohen, Hong, & Guevara, 2010; Cohen et al., In Press), developed by our lab to assess vocal expression from natural speech, was employed here. The CANS is based, in part, on Praat (Boersma & Weenink, 2013), a program that has been used extensively in speech pathology and linguistic studies. Digital audio files were organized into "frames" for analysis (i.e., 100 per second). During each frame, four basic speech properties were quantified: Fundamental Frequency (F0), 1st Formant (F1), 2nd Formant (F2) and Intensity (see Figure 1). To limit potential background noise, we employed optimization filters for measuring F0 (Vogel, Maruff, Snyder, & Mundt, 2009): Low: 75 Hertz; High: 300 Hertz. Due to the nonlinear nature of the hertz frequency scale, F0, F1 and F2 values were converted to semi-tones– a linear scale employed for parametric statistics of hertz-scale data.

There is no consensus on how to quantify vocal signals, and there is considerable variability in summary statistics used (e.g., means, standard deviations) and across variable time scales (e.g., computed across the whole versus portions of the speech sample). To address this issue, we identified 21 systematically-defined and commonly-reported variables, based in part on a recent psychometric evaluation of 1500 young adults (Cohen et al., In Press) and a broader literature review (Huttunen, Keranen, Vayrynen, Paakkonen, & Leino, 2011; Scherer, 2003; Sobin & Alpert, 1999; Table 2). Three of these variables involved "Vocal Production", defined in terms of pauses (i.e., 50 milliseconds or more of signal absence) bounded by an epoch of F0 signal greater than 150 milliseconds in length with no contiguous pause greater than 50 milliseconds. The remaining measures were based on variability in F0, F1, F2 and intensity signals in terms of three different epochs: "local", "global" and "instability". "Local" refers to variability within utterances, whereas "global" refers to variability across utterances (i.e., across the speech sample). We computed local and global measures for both "Range" (i.e., the difference between the highest and lowest values within utterances) and for standard deviation (SD) scores. For F0 and intensity signals, we also computed "instability", which refers to subtle signal fluctuations (i.e., change in consecutive frames; see Table 2).

Analyses

The analyses were conducted in four steps. Our first step focused on data reduction of the 21 vocal variables using PCA with Oblimin rotation and Kaiser normalization. A fixed factor number was determined using parallel analysis (O'Conner, 2000). Our primary goal was to find a common factor solution for patients and controls to facilitate comparison between them; so PCA were first conducted on all participants simultaneously and then verified in patients and controls separately. Component factor scores were computed and used in subsequent analyses. Second, hierarchical linear regressions were employed to evaluate the relative relationships between the seven vocal variables (dependent variables) and diagnostic group (Step 3) above and beyond those associated with demographic variables (age, sex, ethnicity; Step 1) and study (Step 2). The ethnicity (2 separate variables) and study (5 separate variables) were "dummy-coded". To complement these analyses, bivariate correlations between all predictor and independent variables were computed. Third, we evaluated the relationships between psychiatric symptoms and vocal variables using both Pearson and partial (i.e., controlling for age, sex and study) correlations. Finally, we complemented the analyses in step three by comparing vocal variables between patients with varying levels of clinically-rated negative symptoms, including: No (i.e., all Anergia symptoms rated as "absent"; 90), Mild (i.e., average of Anergia ratings in the "mild" range [1 to 3]; 179), and High (i.e., average of Anergia ratings greater than "mild" [greater than 3]; 23). For normalization purposes, "extreme" vocal scores (i.e., <> 3.5 SDs) were "winsorized" to ± 3.5.

Results

Data reduction of vocal variables

Parallel analysis of the 21 vocal variables in the entire sample suggested a five-factor solution. The five-factor PCA solution of these variables explained 79.15% of the variance, though the factor structure did not replicate separately in the patient (80.73% of the variance) and control groups (79.41% of the variance). In the patient sample, the instability items cohered heavily together (i.e., component scores > .60) whereas in the control sample they did not; instead loading heavily with three of the four Vertical Tongue Movement items (i.e., component scores > .51). In the patient sample, Pause N did not load heavily with any of the factors (i.e., component scores < .32) whereas in controls, Pause N heavily loaded with other Pause items and a single F1 item (i.e., component scores > .60). Excluding Pause N, Pitch Instability and Volume Instability from the PCA resulted in a coherent four-factor solution (as indicated by Parallel analysis) that was consistent across both patient and control samples. The PCA solution explained 85.16% of the variance. The KMO (i.e., .61) and Bartlett's ($p < .0001$) test statistics were both adequate for this analysis. The pattern matrix for patients and controls are presented in Table 3. A component correlation matrix of the four factors suggested they were not redundant with each other (e.g., mean r values were .13 for controls and .11 for patients; r's < .36; Table 4), nor were they redundant with the three variables excluded from the PCA. The component scores for each factor for the patient and control groups were highly correlated with each other (r's > .92). Given the potential independence of the Instability Scores in the control groups, their relatively low correlation with each other (e.g., $r = -.27$ for patients), and the fact that a prior study of these

variables revealed they are independent with respect to some psychopathological processes (Cohen et al., In Press), we decided to examine these variables separately. Additionally, we examined Pause N separately. Consequent analyses focused on seven variables.

Vocal variables: relative contributions of demographic, context and diagnostic group variables

Bivariate correlations (Table 5) and hierarchical linear regressions (Table 6) between predictor and independent variables were computed. Regressions indicated that significant contributions were made by demographic (i.e., three of seven regressions) and contextual (seven of seven regressions) variables. Diagnostic group was negligible for all seven analyses and in all but one analysis, were not statistically significant. The contribution of context was quite pronounced, accounting for over half of the variance in some variables. Inspection of the regression coefficients suggested that women had shorter Pauses ($\beta = .10$, $t = 2.20$, $p = .03$) and greater Intonation ($\beta = -.24$, $t = 5.10$, $p < .001$) and Tongue Movement ($\beta = .11$, $t = 2.25$, $p = .03$). Age was associated with shorter Pauses ($\beta = -.11$, $t = 2.39$, $p = .02$) and less Pitch Instability ($\beta = -.25$, $t = 5.33$, $p < .001$). Caucasians versus non-Caucasians had shorter Pauses ($\beta = -.28$, $t = 3.69$, $p = .001$) and greater Tongue Movement ($\beta = -.26$, $t = 3.18$, $p = .002$).

To illustrate the relationship between vocal variables and context, standardized means (both patients and controls; controlling for demographic variables) for each vocal variable were computed and plotted in Figure 2. There was no obvious pattern of differences across studies, though study five was relatively low in Tongue Movement and Emphasis whereas study 1 was notably high in Intonation and Emphasis.

Patients and controls were then compared in vocal variables for each study independently. Of the potential 28 analyses (seven variables for four studies; Study 2 contained no controls), only four significant speech deficits were observed. In Study 1, patients versus controls showed significantly fewer Pauses ($d = -.50$, 95% CI[−.72, −.29]) and less Volume Instability ($d = -.52$, 95% CI[−.73, −.30]). In Study 3, patients versus controls showed less Intonation ($d = -.59$, 95% CI[−.81, −.37]), and in Study 4, patients showed longer Pauses ($d = -.47$, 95% CI[−.68, −.25]) (all p's < .05).

Vocal variables and psychiatric symptoms

Table 7 contains the correlations between symptom ratings and vocal variables. Despite some significant correlations, the magnitudes were all, with one exception, in the small to negligible range (i.e., r's < .30). That being said, there were some statistically significant findings even after controlling for demographic and contextual variables. Anergia, but not any other variables, was significantly associated with increased Pause lengths. Depression/anxiety symptoms were associated with fewer Pauses, and less Pitch Instability. Thought disorder was also associated with fewer Pauses, and also greater Intonation and Pitch Instability. Activation/Hostility was associated with greater Tongue Movement.

Clinically-defined negative symptoms

The groups differed significantly only in Pause Length (F[3, 405] = 3.30, p = .02), which was driven by the Non-Negative group showing shorter pauses than the other groups (p's < .03). The High-Negative group significantly differed from the Non-Negative but not the Low-Negative group. Trend-level group differences were observed with Tongue Movement (F[3, 405] = 2.31, p = .08); which reflected the High-Negative group showing significantly more Tongue Movement than the Control and Mild-Negative symptom group (p's < .03) and at a trend level compared to the No-Negative symptom group (p = .08). Standardized scores are reported in Figure 3.

Discussion

This study is the largest investigation of vocal abnormalities in schizophrenia using computerized analytic technologies to date. There were four main findings. First, our data reduction analysis identified seven independent vocal factors/variables. Given the overwhelming number of vocal variables available for analysis and the considerable variability in scores reported across studies (Cohen, Mitchell, et al., 2014), these findings provide a psychometrically-informed framework for future vocal analysis in clinical populations. Second, vocal expression varied considerably across the studies examined in both patients and controls, indicating that contextual variables (e.g., speaking task, setting) should be considered when evaluating vocal abnormalities in schizophrenia. Third, when controlling for demographic and contextual variables, there were no meaningful differences between controls and patients in any of the variables examined. This finding is in stark contrast to the dramatic differences in symptom rating scales reported between patients, as a group, and controls (Cohen, Mitchell, et al. 2014). Even in patients with high levels of clinically-rated negative symptoms, most facets of vocal expression were not abnormal (with the possible exception of pause length; discussed below). Fourth, vocal variables were significantly correlated (albeit at a small effect size level) with psychiatric symptoms in the patient group.

The most important finding from this study concerns the lack of patient vocal deficits from computerized measures; remarkable because schizophrenia is defined, in part, based on negative symptoms (i.e., alogia, blunted affect) and is associated with profound symptom rating scale abnormalities (Cohen, Mitchell, et al., 2014). We see three possible explanations for the discrepancy between clinical and computer measures. First, it could be that patient's vocal deficits are actually much more benign in magnitude and limited in breadth than clinician's ratings and diagnoses would suggest. Supporting this notion, evidence from our recent meta-analysis (Cohen, Mitchell, et al., 2014) suggests that only variables related to speech production (e.g., Pauses) are consistently and notably (i.e., d > .80) abnormal in schizophrenia. Evidence of clinicians imprecision in evaluating patient speech has been reported in a line of research conducted by Murray Alpert and colleagues. Of particular note, Alpert, Pouget and Silva (1995) manipulated the pause times of pre-recorded patient speech, but no other acoustic characteristics, and asked clinicians to evaluate the recordings using symptom rating scales. Ratings of blunted vocal affect were inappropriately inflated, raising the question of whether measures of vocal expression (and perhaps negative symptoms more

generally) are unduly anchored in patient's pause times. Studies by Alpert and colleagues and our lab have found that clinical ratings of flat affect are significantly correlated with acoustically-derived pause measures (but not measures of vocal affect) as well as global negative symptom scores, neuropsychological functioning and extrapyramidal side effects (Allan, Sison, Alpert, Connolly, & Crichton, 1998; Alpert, Kotsaftis, & Pouget, 1997; Alpert, Pouget, Welkowitz, & Cohen, 1993; Alpert, Rosen, Welkowitz, Sobin, & Borod, 1989; Alpert et al., 2002; Cohen et al., 2008; Cohen et al., 2013). Accordingly, Alpert and colleagues have proposed that clinical ratings of vocal deficits are conflated by perceptions of "global" impressions of patient behavior, based on inferences from highly salient and potentially comorbid or confounding influences – a sort of "halo" effect.

That being said, pauses were not abnormal in the present patient group. This leads us to a second potential explanation for the discordance between computerized and symptom-rating measures – namely that vocal deficits are dynamic and emerge only in certain contexts. Several lines of evidence support this claim. First, in the present study, nearly all vocal variables differed as a function of speaking task, setting or other contextual factors. Second, patients showed abnormal pauses in several studies examined here suggesting that, in some but not all conditions, pauses may become abnormally exaggerated or increased in number. Third, there is evidence that cognition is a "contextual" variable that may underlie vocal deficits. Vocal characteristics are determined, in part, by the involvement and integration of a range of basic cognitive resources. Even in nonpatient populations, restricting these resources leads to more sparse speech and less variability (Barch & Berenbaum, 1994; Cohen, Dinzeo et al., 2014). Given that schizophrenia is associated with a broad range of cognitive deficits, (Cohen et al., 2007; Heinrichs & Zakzanis, 1998) it stands to reason that even minor strains on cognitive resources may have dramatic effects on vocal communication. Note that one of the speaking tasks where patients were abnormal in pause time involved a social cognition task that was at least somewhat demanding in cognitive resources. Finally, Alpert et al. (1997) found that certain pauses, namely those occurring between clauses and when "switching turns" during dyadic exchanges, were particularly pathognomonic in schizophrenia. In sum, the present findings are not inconsistent with the notion that abnormal pause times (a.k.a. alogia) may be limited to contexts for which cognitive resources are limited. Beyond cognitive variables, it is also possible that acoustic differences would have been more pronounced if the speech content was emotionally charged in some manner. With respect to symptom rating scales, it may be that symptom raters may be implicitly weighting (or over-weighting) these incidences when making their ratings.

A third possibility is that symptom measures accurately capture vocal deficits that are somehow missed using the vocal measures employed in the present study. In defense of the present measures, they are conceptually linked to clinical definitions of vocal deficits in schizophrenia and are commonly used in studies of clinical populations and in communication sciences more generally (Cohen et al., 2009; Cohen, Dinzeo et al., 2014; Johnstone et al., 2007; Sobin, 1999; Sobin & Alpert, 1999; Wells & MacFarlane, 1998). Thus, they have shown convergent validity in understanding a broad range of important communicative functions. From a vocal perspective, it is hard to believe, for example, that "blunted vocal affect" would not involve less variability of F0 and Intensity signals within

utterances. That being said, there are literally thousands of ways to process, organize and quantify the human voice, and it is possible that an as-yet unidentified vocal characteristic will meaningfully differentiate schizophrenia from nonpatients. It may also be the case that vocal deficits can only be understood within the confines of dyadic exchanges. This is a particularly salient point in light of a recent study that patients with schizophrenia fail to match their speaking partners in vocal "synergy" – defined in terms of vocal modulations one normally makes during dyadic interaction (Dombrowski, McCleery, Gregory, & Docherty, 2014). It may be the case that vocal abnormalities require grounding within the contextual demands of their conversational partner, and aren't abnormal in-and-of themselves.

The Pitch Instability measure warrants further consideration, as it was correlated (albeit at a small effect size level) with non-negative symptom factors (e.g., thought disturbance, anxiety-depression). In terms of psychiatric correlates, Pitch Instability (vis a vis "jitter") has been associated with stressor-provoked anxiety (Mendoza & Carballo, 1998), but has received little attention in the schizophrenia literature. Pitch Instability reflects "microinstability" in vocal cord vibrations, or random modulation of the source signal. The psychological underpinnings of it are not well known, but there is evidence that they are influenced by arousal levels (*vis a vis* cardiovascular functions). Indeed, a recent double-blind placebo controlled study found that propranolol increased both cardiovascular functions and vocal jitter levels in individuals undergoing a laboratory stressor task (Giddens, Barron, Clark, & Warde, 2010). In the present study, greater F0 instability was associated with symptoms that likely reflect increased arousal states in some capacity. While the small correlational magnitudes across patients may lead some to discount this finding for between-group comparison, it is important to note that computerized vocal analysis can be used for monitoring symptoms within patients in a way symptom-ratings can't. Due, in part, to their near-perfect reliability and high sensitivity, changes in symptom states within individuals can be evaluated using computerized vocal analysis in ways unattainable using likert-scale based symptom ratings. The modest convergence in symptoms and vocal expression across patients in these measures may belie much stronger convergence within patients, and Pitch Instability may be useful as an index of anxiety, depression and psychosis fluctuations within an individual – symptoms which are critical to monitor and track.

Some limitations of the current project warrant mention. Foremost is our inability to control for individual differences in pharmacotherapy or psychosocial interventions. Notably, there were differences, both across and within studies, in how these interventions were prescribed, how individuals complied with these interventions, and in their effects. While this potential confound reflects noise for the present study, it seems unlikely that treatment variables are responsible for either the null or positive findings in this study. Second, we were unable to disentangle the effects of "context" across the five studies examined here. Variables such as speaking task, research setting, research assistant, compensation, clinical setting and geographic region may have contributed, or interacted to contribute to the profound differences in vocal expression across studies. Regardless, our conclusions - that vocal expression varies as a function of context (broadly defined), is appropriate. Third, the reliance on the BPRS is a limitation – as it lacks the specificity of other symptom-rating scale based negative symptom measures (Kirkpatrick et al., 2011; Kring et al., 2013).

Fourth, the speech topics examined here were relatively neutral in emotional tone, and it is possible that more highly charged topics may have resulted in more pronounced group differences. Finally, the overall sample was relatively low in negative symptom severity, though the general null findings comparing patients with clinically-rated negative symptoms versus those without were not dissimilar to the findings from the larger sample.

The present findings reflect an important step in applying computerized vocal analytic technologies for understanding and measuring psychiatric symptoms. From a practical perspective, acoustic analysis is uninvasive, inexpensive, objective and based on natural behavior that is abundant and closely tied to real-world functioning. Moreover, vocal signal can be procured outside traditional clinical or research office settings using telemedicine and mobile technologies; thus greatly expanding the scope and potential sensitivity of assessment. The present data suggest that, in contrast to clinician ratings, vocal expression in schizophrenia are largely normal and, at best, occur only in certain contexts. The present data highlight the use of vocal analysis for precisely quantifying negative symptoms and other symptoms as well. At the same time, vocal expression is a dynamic process and what is “abnormal” in one context might be normal in another context. Going forward, it will be critical to consider speaking context when evaluating vocal expression.

Acknowledgments

The authors wish to thank the patients and subjects who volunteered their time to participate in the study. We are also grateful to Matt Calamia, Ph.D. for his statistical advice and consultation.

This project was funded by grant R03MH092622 from the National Institute of Mental Health to Alex Cohen

References

Allan ER, Sison CE, Alpert M, Connolly B, Crichton J. The relationship between negative symptoms of schizophrenia and extrapyramidal side effects with haloperidol and olanzapine. Psychopharmacological Bulletin. 1998; 34(1):71–74.

Alpert M, Kotsaftis A, Pouget ER. At issue: speech fluency and schizophrenic negative signs. Schizophrenia Bulletin. 1997; 23(2):171–177. DOI: 10.1093/schbul/23.2.171 [PubMed: 9165627]

Alpert M, Pouget ER, Silva R. Cues to the assessment of affects and moods: speech fluency and pausing. Psychopharmacological Bulletin. 1995; 31(2):421–424.

Alpert M, Pouget ER, Welkowitz J, Cohen J. Mapping schizophrenic negative symptoms onto measures of the patient’s speech: set correlational analysis. Psychiatry Research. 1993; 48(3):181–190. DOI: 10.1016/0165-1781(93)90070-W [PubMed: 8272441]

Alpert M, Rosen A, Welkowitz J, Sobin C, Borod JC. Vocal acoustic correlates of flat affect in schizophrenia. Similarity to Parkinson’s disease and right hemisphere disease and contrast with depression. British Journal of Psychiatry. 1989; (Suppl 4):51–56. [PubMed: 2605011]

Alpert M, Shaw RJ, Pouget ER, Lim KO. A comparison of clinical ratings with vocal acoustic measures of flat affect and alogia. Journal of Psychiatric Research. 2002; 36(5):347–353. DOI: 10.1016/S0022-3956(02)00016-X [PubMed: 12127603]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. IVth. Arlington, VA: American Psychiatric Publishing; 2013.

Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. British Journal of Psychiatry. 1989; (Suppl 7):49–58. [PubMed: 2695141]

Barch D, Berenbaum H. The relationship between information processing and language production. Journal of Abnormal Psychology. 1994; 103(2):241–251. Retrieved from http://ccpweb.wustl.edu/pdfs/1994JAbPsych241.pdf. [PubMed: 8040493]

Bellack AS, Morrison RL, Mueser KT, Wade JH, Sayers SL. Role play for assessing the social competence of psychiatric patients. Psychological Assessment. 1990; 2:248–255. DOI: 10.1037/1040-3590.2.3.248

Boersma, P.; Weenink, D (Producer). Praat: doing phonetics by computer (Version 5.3.59). 2013. Retrieved from http://www.praat.org/

Cannizzaro M, Harel B, Reilly N, Chappell P, Snyder PJ. Voice acoustical measurement of the severity of major depression. Brain and Cognition. 2004; 56(1):30–35. DOI: 10.1016/j.bandc.2004.05.003 [PubMed: 15380873]

Castelli F, Frith C, Happe F, Frith U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. Brain. 2002; 125(Pt 8):1839–1849. [PubMed: 12135974]

Cohen AS, Alpert M, Nienow TM, Dinzeo TJ, Docherty NM. Computerized measurement of negative symptoms in schizophrenia. Journal of Psychiatric Research. 2008; 42(10):827–836. DOI: 10.1016/j.jpsychires.2007.08.008 [PubMed: 17920078]

Cohen AS, Dinzeo TJ, Donovan NJ, Brown CE, Morrison SC. Vocal acoustic analysis as a biometric indicator of information processing: Implications for neurological and psychiatric disorders. Psychiatry Research. 2014; 226(1):235–241. DOI: 10.1016/j.psychres.2014.12.054 [PubMed: 25656172]

Cohen AS, Elvevåg B. Automated computerized analysis of speech in psychiatric disorders. Current Opinion in Psychiatry. 2014; 27(3):203–209. DOI: 10.1097/YCO.0000000000000056 [PubMed: 24613984]

Cohen AS, Hong SL, Guevara A. Understanding emotional expression using prosodic analysis of natural speech: refining the methodology. Journal of Behavioral Therapy and Experimental Psychiatry. 2010; 41(2):150–157. DOI: 10.1016/j.jbtep.2009.11.008

Cohen AS, Kim Y, Najolia GM. Psychiatric symptom versus neurocognitive correlates of diminished expressivity in schizophrenia and mood disorders. Schizophrenia Research. 2013; 146(1–3):249–253. DOI: 10.1016/j.schres.2013.02.002 [PubMed: 23481582]

Cohen AS, McGovern JE, Dinzeo TJ, Covington MA. Speech deficits in serious mental illness: a cognitive resource issue? Schizophrenia Research. 2014; 160(1–3):173–179. DOI: 10.1016/j.schres.2014.10.032 [PubMed: 25464920]

Cohen AS, Minor KS, Najolia GM, Lee Hong S. A laboratory-based procedure for measuring emotional expression from natural speech. Behavior Research Methods. 2009; 41(1):204–212. DOI: 10.3758/BRM.41.1.204 [PubMed: 19182141]

Cohen AS, Mitchell KR, Elvevag B. What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. Schizophrenia Research. 2014; 159(2–3):533–538. DOI: 10.1016/j.schres.2014.09.013 [PubMed: 25261880]

Cohen AS, Najolia GM, Kim Y, Dinzeo TJ. On the boundaries of blunt affect/alogia across severe mental illness: implications for Research Domain Criteria. Schizophrenia Research. 2012; 140(1–3):41–45. DOI: 10.1016/j.schres.2012.07.001 [PubMed: 22831770]

Cohen AS, Renshaw T, Mitchell K, Kim Y. A psychometric investigation of “Macroscopic” speech measures for clinical and psychological science. Behavior Research Methods. 2015; :1–12. DOI: 10.3758/s13428-015-0584-1 [PubMed: 24683129]

Cohen AS, Saperstein AM, Gold JM, Kirkpatrick B, Carpenter WT Jr, Buchanan RW. Neuropsychology of the deficit syndrome: new data and meta-analysis of findings to date. Schizophrenia Bulletin. 2007; 33(5):1201–1212. DOI: 10.1093/schbul/sbl066 [PubMed: 17159230]

Costello AB, Osborne JW. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Practical Assessment Research & Evaluation. 10(7) Available online: http://pareonline.net/getvn.asp?v=10&n=7.

Docherty NM. On identifying the processes underlying schizophrenic speech disorder. Schizophrenia Bulletin. 2012; 38(6):1327–1335. DOI: 10.1093/schbul/sbr048 [PubMed: 21562141]

Dombrowski M, McCleery A, Gregory SW Jr, Docherty NM. Stress reactivity of emotional and verbal speech content in schizophrenia. Journal of Nervous and Mental Disease. 2014; 202(8):608–612. DOI: 10.1097/NMD.0000000000000169 [PubMed: 25010108]

First, MB.; Gibbon, M.; Spitzer, RL.; Williams, JBW. Structured Clinical Interview for DSM-IV Axis I Disorders Patient Edition. New York: Biometrics Research; 1996.

Galynker II, Cohen LJ, Cai J. Negative symptoms in patients with major depressive disorder: a preliminary report. Neuropsychiatry, Neuropsychology, and Behavioral Neurology. 2000; 13(3): 171–176.

Giddens CL, Barron KW, Clark KF, Warde WD. Beta-adrenergic blockade and voice: a double-blind, placebo-controlled trial. Journal of Voice. 2010; 24(4):477–489. DOI: 10.1016/j.jvoice.2008.12.002 [PubMed: 19846273]

Guy, W. ECDEU Assessment Manual for Psychopharacology: Publication ADM 76–338. Washington DC: US Department of Health, Education, and Welfare; 1976.

Heinrichs RW, Zakzanis KK. Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. Neuropsychology. 1998; 12:426–445. DOI: 10.1037/0894-4105.12.3.426 [PubMed: 9673998]

Hoekert M, Kahn RS, Pijnenborg M, Aleman A. Impaired recognition and expression of emotional prosody in schizophrenia: review and meta-analysis. Schizophrenia Research. 2007; 96(1):135–145. DOI: 10.1016/j.schres.2007.07.023 [PubMed: 17766089]

Horan WP, Nuechterlein KH, Wynn JK, Lee J, Castelli F, Green MF. Disturbances in the spontaneous attribution of social meaning in schizophrenia. Psychological Medicine. 2009; 39(4):635–643. DOI: 10.1017/S0033291708003838 [PubMed: 18606048]

Huttunen K, Keranen H, Vayrynen E, Paakkonen R, Leino T. Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. Applied Ergonomics. 2011; 42(2): 348–357. DOI: 10.1016/j.apergo.2010.08.005 [PubMed: 20832770]

Johnstone T, van Reekum CM, Bänziger T, Hird K, Kirsner K, Scherer KR. The effects of difficulty and gain versus loss on vocal physiology and acoustics. Psychophysiology. 2007; 44(5):827–837. DOI: 10.1111/j.1469-8986.2007.00552.x [PubMed: 17608801]

Kirkpatrick B, Galderisi S. Deficit schizophrenia: an update. World Psychiatry. 2008; 7(3):143–147. Retrieved from http://www.ncbi.nlm.nih.gov.libezp.lib.lsu.edu/pmc/articles/PMC2559917/. [PubMed: 18836581]

Kirkpatrick B, Strauss GP, Nguyen L, Fischer BA, Daniel DG, Cienfuegos A, Marder SR. The brief negative symptom scale: psychometric properties. Schizophrenia Bulletin. 2011; 37(2):300–305. DOI: 10.1093/schbul/sbq059 [PubMed: 20558531]

Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. American Journal of Psychiatry. 2013; 170(2):165–172. DOI: 10.1176/appi.ajp.2012.12010109 [PubMed: 23377637]

Kulhara P, Chadda R. A study of negative symptoms in schizophrenia and depression. Comprehensive Psychiatry. 1987; 28(3)(87):229–235. 90029–0. DOI: 10.1016/0010-440X [PubMed: 3595110]

Long JD, Brekke JS. Longitudinal factor structure of the Brief Psychiatric Rating Scale in schizophrenia. Psychological Assessment. 1999; 11(4):498–506. DOI: 10.1037/1040-3590.11.4.498

Lukoff D, Nuechterlein KH, Ventura J. Manual for the Expanded Brief Psychiatric Rating Scale (BPRS). Schizophrenia Bulletin. 1986; 12:594–602.

Mendoza E, Carballo G. Acoustic analysis of induced vocal stress by means of cognitive workload tasks. Journal of Voice. 1998; 12(3)(98):263–273. 80017–9. DOI: 10.1016/S0892-1997 [PubMed: 9763177]

Mueser KT, Pratt SI, Bartels SJ, Forester B, Wolfe R, Cather C. Neurocognition and social skill in older persons with schizophrenia and major mood disorders: An analysis of gender and diagnosis effects. Journal of Neurolinguistics. 2010; 23(3):297–317. DOI: 10.1016/j.jneuroling.2009.08.007 [PubMed: 21113403]

Mueser KT, Sayers SL, Schooler NR, Mance RM, Haas GL. A multisite investigation of the reliability of the Scale for the Assessment of Negative Symptoms. American Journal of Psychiatry. 1994; 151(10):1453–1462. DOI: 10.1176/ajp.151.10.1453 [PubMed: 7916540]

Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. Biological Psychiatry. 2012; 72(7):580–587. DOI: 10.1016/j.biopsych.2012.03.015 [PubMed: 22541039]

Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. Psychological Reports. 1962; 10:799–812.

Satt A, Hoory R, König A, Aalten P, Robert PH. Speech-Based Automatic and Robust Detection of Very Early Dementia. Interspeech. 2014; 3:6.doi: 10.13140/2.1.1258.8805

Scherer KR. Vocal communication of emotion: A review of research paradigms. Speech Communication. 2003; 40(1):227–256. DOI: 10.1016/S0167-6393(02)00084-5

Sobin C, Alpert M. Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. Journal of Psycholinguistic Research. 1999; 28(4):167–365. DOI: 10.1023/A:1023237014909

Stahl SM, Buckley PF. Negative symptoms of schizophrenia: A problem that will not go away. Acta Psychiatrica Scandinavica. 2007; 115(1):4–11. DOI: 10.1111/j.1600-0447.2006.00947.x [PubMed: 17201860]

Strauss GP, Harrow M, Grossman LS, Rosen C. Periods of recovery in deficit syndrome schizophrenia: a 20-year multi–follow-up longitudinal study. Schizophrenia Bulletin. 2010; 36(4):788–799. DOI: 10.1093/schbul/sbn167 [PubMed: 19095758]

Tremeau F, Malaspina D, Duval F, Correa H, Hager-Budny M, Coin-Bariou L, Gorman JM. Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects. American Journal of Psychiatry. 2005; 162(1):92–101. DOI: 10.1176/appi.ajp.162.1.92 [PubMed: 15625206]

Vogel AP, Maruff P, Snyder PJ, Mundt JC. Standardization of pitch-range settings in voice acoustic analysis. Behavior Research Methods. 2009; 41(2):318–324. DOI: 10.3758/BRM.41.2.318 [PubMed: 19363172]

Wells B, MacFarlane S. Prosody as an interactional resource: Turn-projection and overlap. Language and Speech. 1998; 41(3–4):265–294. DOI: 10.1177/002383099804100403 [PubMed: 10746359]

General Scientific Summary

Patients with schizophrenia are rated by clinicians as having profound and chronic vocal expressive deficits. Data from this study, using computerized vocal analysis of patients, suggests that vocal expression in patients is similar to nonpatients. Vocal deficits, if they occur in schizophrenia, are much more subtle than clinical ratings suggest, and may be contextually dependent.

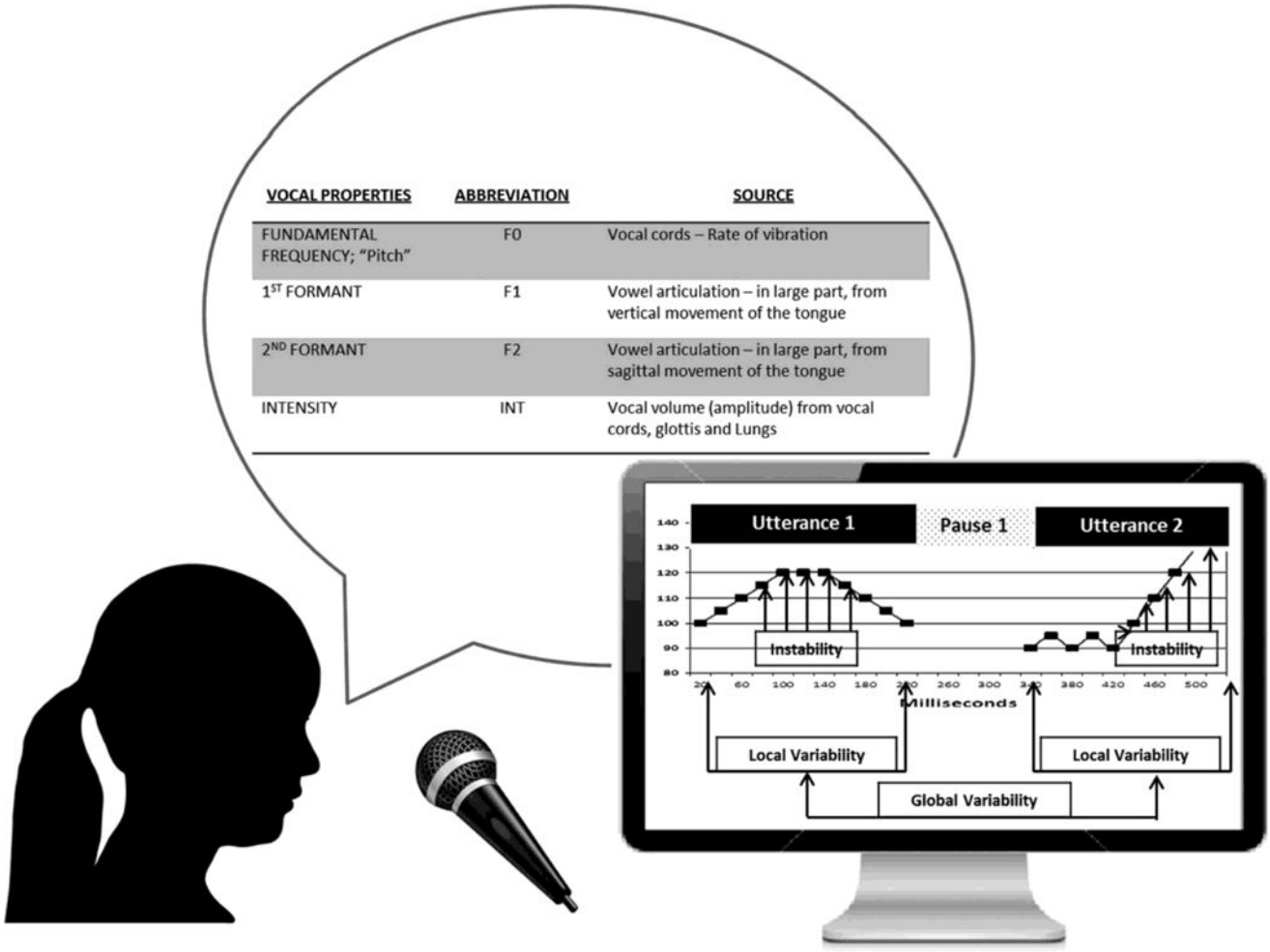

Figure 1.
Vocal properties examined in this study.

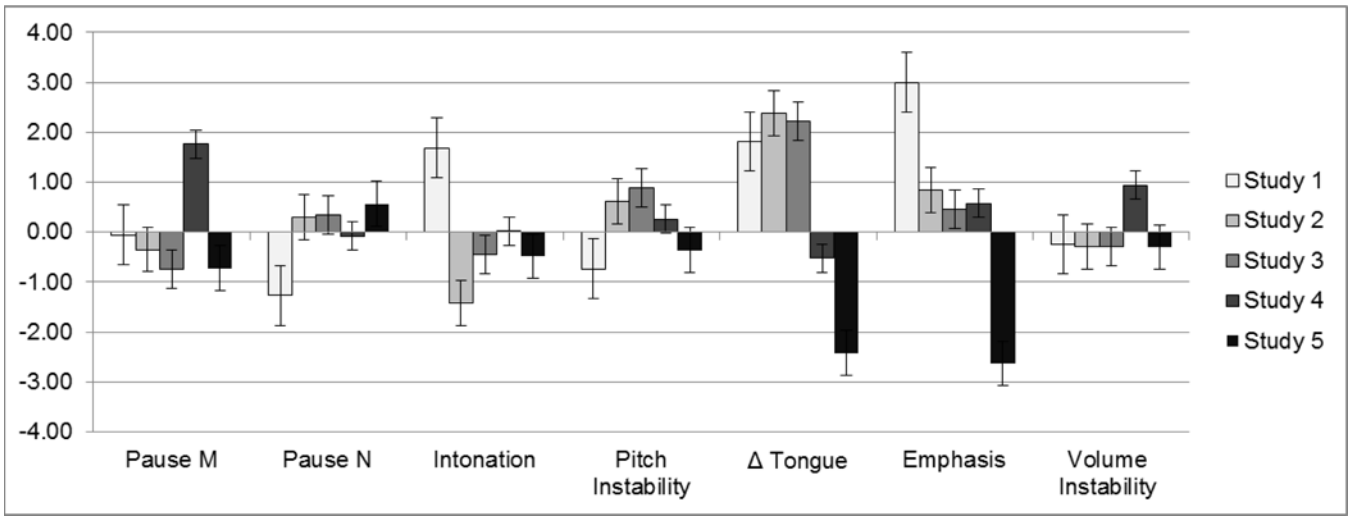

Figure 2.
Standardized means of vocal variables plotted as a function of study (combined patients and controls).

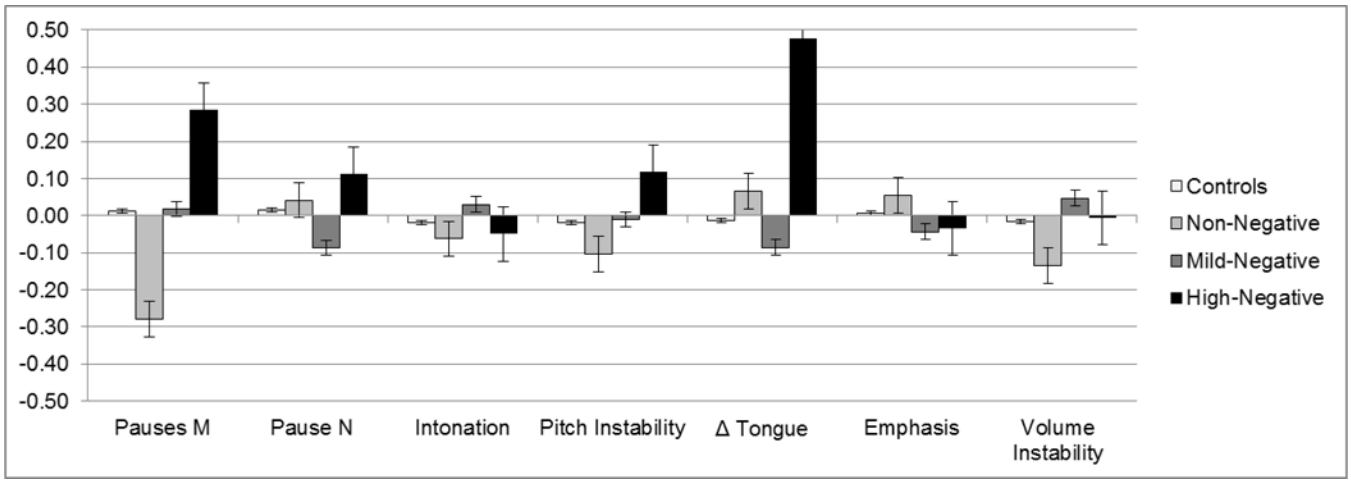

Figure 3.

Standardized means of vocal variables plotted as a function of patients with No (Non), Mild and High clinically-rated negative symptoms and controls.

Table 1

Descriptive and study information.

	Study1[a]	Study 2[b]	Study 3[c]	Study 4[d]	Study 5[e]
N Controls/Patients	21 – 65	0 – 31	33 – 42	40 – 53	23 – 118
Patient Setting	Forensic-Inpatient	Outpatient	Outpatient	Outpatient	Outpatient
% Female (Con – Pts)	33% – 28%	0% – 36%	52% – 31%	25% – 25%	48% – 44%
Age Controls	38.33 ± 10.75	n/a	40.88 ± 12.24	39.73 ± 9.23	38.26 ± 9.17
Patients	42.06 ± 8.44	39.68 ± 10.76	42.33 ± 12.02	40.30 ± 10.71	40.13 ± 7.36
% Cauc (Con – Pts)	57% – 35%	0% – 32%	49% – 52%	40% – 32%	39% – 33%
% AA (Con – Pts)	43% – 65%	0% – 68%	42% – 43%	30% – 32%	61% – 64%
BPRS Factor Scores					
Thought Disturbance	2.80 ± 1.29	2.60 ± 1.15	2.79 ± 1.37	2.54 ± 1.19	1.79 ± 0.74
Anxiety/Depression	1.50 ± 0.46	2.67 ± 1.16	2.25 ± 1.16	2.38 ± 0.82	2.62 ± 1.14
Anergia	1.63 ± 0.70	1.98 ± 0.98	1.63 ± 0.84	1.67 ± 0.65	1.56 ± 0.58
Activation-Hostility	2.80 ± 1.29	2.60 ± 1.15	2.79 ± 1.37	2.54 ± 1.19	1.79 ± 0.74
Speaking Task	Conversation	Monologue	Monologue	Cognitive Task	Conversation
Speaking Task Length	300 sec.	180 sec.	90 sec.	42 – 75 sec.	300 sec.

Cauc = Caucasian, AA = African-American,

[a] Cohen et al., 2008;

[b] Cohen et al., 2013;

[c] Cohen, McGovern, et al., 2014;

[d] Horan et al., 2009;

[e] Docherty, 2012

J Abnorm Psychol. Author manuscript; available in PMC 2016 July 30.

Table 2

Summary of vocal variables included in this study.

PAUSES	
Pauses M - Mean of Pause lengths (ms).	Pause SD - SD of pause lengths.
Pause N - Total N of pauses.	
INTONATION	**VERTICLE TONGUE MOVEMENT**
F0 SD Global - SD of all SDs computed within utterances.	F1 SD Global - SD of all SDs computed within utterances.
F0 SD Local - M of all SDs computed within utterances.	F1 SD Local - M of all SDs computed within utterances.
F0 Range Global - SD range computed within utterances.	F1 Range Global - SD range computed within utterances.
F0 Range Local - M range computed within utterances.	F1 Range Local - M range computed within utterances.
SAGITTAL TONGUE MOVEMENT	**EMPHASIS**
F2 SD Global - SD of all SDs computed within utterances.	INT SD Global - SD of all SDs computed within utterances
F2 SD Local - M of all SDs computed within utterances	INT Range Global - SD range computed within utterances.
F2 Range Global - SD range computed within utterances.	INT Range Local - M range computed within utterances.
F2 Range Local - M range computed within utterances.	INT SD Local - M of all SDs computed within utterances
PITCH INSTABILITY	**VOLUME INSTABILITY**
Ptich Instability - Absolute value of F0 change in consecutively voiced frames within utterance (M across utterances).	Volume Instability - Absolute value of INT change in consecutively voiced frames within utterance (M across utterances).

M = Mean, SD = Standard Deviation, INT = Intensity

Table 3

Principal Components Analysis results; Pattern matrix for patients and controls (in parentheses)

	Δ Tongue	Emphasis	Intonation	Pause M
F1 Range Global	**.95 (.81)**	−.01 (−.02)	−.01 (.21)	−.07 (−.01)
F1 Range Local	**.92 (.86)**	.01 (.00)	−.10 (.13)	−.07 (−.07)
F2 Range Local	**.89 (.86)**	.01 (−.01)	.06 (−.01)	.13 (.30)
F1 SD Global	**.89 (.79)**	.01 (−.02)	−.13 (.19)	−.10 (−.13)
F2 SD Local	**.89 (.85)**	.00 (−.03)	.03 (−.09)	.16 (.32)
F2 Range Global	**.87 (.77)**	.02 (.06)	.14 (−.13)	.09 (−.10)
F1 SD Global	**.86 (.64)**	−.02 (−.01)	−.03 (.08)	−.12 (−.36)
F2 SD Global	**.78 (.68)**	.00 (.01)	.16 (−.23)	.12 (.08)
INT SD Local	−.01 (.00)	**1.00 (1.00)**	−.01 (.00)	.00 (.01)
INT SD Global	−.02 (.00)	**1.00 (1.00)**	−.01 (.00)	.00 (.00)
INT Range Global	−.01 (.00)	**1.00 (1.00)**	.00 (.00)	.00 (.00)
INT Range Local	.03 (.02)	**1.00 (1.00)**	.02 (.01)	.00 (.01)
F0 SD Global	−.21 (−.25)	−.02 (.02)	**.92 (.87)**	.01 (−.04)
F0 Range Global	−.02 (.01)	.00 (.01)	**.92 (.87)**	.01 (−.10)
F0 SD Local	.10 (.10)	.00 (−.01)	**.91 (.88)**	−.05 (.20)
F0 Range Local	.23 (.30)	.02 (−.01)	**.83 (.80)**	−.06 (.17)
Pause SD	−.01 (−.02)	.03 (.08)	−.03 (.08)	**.96 (.90)**
Pause M	.03 (.01)	−.02 (.02)	−.04 (.06)	**.95 (.92)**

[a]The instability and Pause N measures are presented independently (see text for elaboration)

Table 4

Zero-order correlation matrix of factor scores for controls (italicized) and patients.

	1.	2.	3.	4.	5.	6.	7.
1. Pause M		−.04	.03	.01	.08	.10	.15**
2. Pause N	*−.28**		−.12*	.14*	.23**	.02	−.14*
3. Intonation	*.14*	*.04*		.06	.20**	.02	.36**
4. Pitch Instability	*.14*	*.05*	*.04*		.18**	.00	−.27**
5. Δ Tongue	*.07*	*.23**	*.15*	*.31***		.04	−.07
6. Emphasis	*.18*	*.05*	*.01*	*.07*	*−.02*		.02
7. Volume Instability	*.25***	*−.07*	*.26***	*−.21**	*−.09*	*.01*	

* = $p < .05$

Table 5

Bivariate correlations between vocal variables and demographic, contextual and group predictor variables.

	Pause M	Pause N	Intonation	Pitch Instability	Δ Tongue	Emphasis	Volume Instability
Demographic Variables							
Age	−.12*	.05	−.04	.24*	.08	−.07	.02
Sex	.11*	.05	−.24**	.11*	.11*	.04	.01
Caucasian Ethnicity [a]	.00	.10*	−.02	−.06	.02	.06	.00
African-American [a]	−.13**	−.10*	.01	−.02	.01	−.09	−.04
Contextual Variables							
Study 1[a]	.02	−.03	.37*	−.36*	.52*	−.05	.46*
Study 2[a]	−.07	.05	−.13*	.17*	.20*	−.04	−.39*
Study 3[a]	−.23*	.38*	−.08	.42*	.28*	−.07	−.48*
Study 4[a]	.55*	.10*	.00	.14*	−.01	.27*	.13*
Study 5[a]	−.28*	−.40*	−.18*	−.25*	−.76*	−.11*	.09
Group Variables							
Group[a]	−.02	−.15*	−.10*	−.07	−.04	−.12**	−.01

[a] Variable is dummy-coded.

* = $p < .05$

Table 6

Relative contributions of demographic (step 1), contextual (step 2) and diagnostic (control, patient) variables to vocal variables.

	Demographic[a]		Contextual[b]		Diagnosis	
Dependent Variables	**ΔR²**	**ΔF**	**ΔR²**	**ΔF**	**ΔR²**	**ΔF**
Pauses M	0.07	7.79*	0.28	44.74*	0.01	5.59*
Pause N	0.01	1.57	0.23	32.63*	0.00	1.62
Intonation	0.06	6.73*	0.18	25.48*	0.00	1.94
Pitch Instability	0.09	10.53*	0.33	58.11*	0.00	0.17
Δ Tongue	0.02	2.16	0.72	282.24*	0.00	1.14
Emphasis	0.01	1.56	0.07	7.97*	0.00	2.19
Volume Instability	0.01	0.61	0.52	112.36*	0.00	0.00

[a] Ethnicity, sex and age,

[b] Variables were dummy-coded and entered in a single step,

* = $p < .05$

Table 7

Correlations between clinically-rated symptoms and vocal variables, computed separately without control and controlling for sex, ethnicity and context (in parentheses) as partial correlations (patients only).

	Thought Disorder	Anxiety-Depression	Anergia	Activation-Hostility
Pauses M	0.11 (0.05)	0.02 (0.07)	0.13* (0.14*)	−0.06 (0.09)
Pause N	0.04 (−0.13*)	−0.15* (−0.15*)	0.05 (0.02)	−0.20* (−0.10)
Intonation	0.17* (0.15*)	−0.10 (0.08)	−0.08 (−0.07)	0.01 (0.09)
Pitch Instability	0.24* (0.20*)	−0.01 (−0.13*)	0.05 (0.02)	0.02 (0.11)
Δ Tongue	0.34* (0.00)	−0.28 (0.02)	0.03 (−0.10)	−0.17* (0.23*)
Emphasis	−0.06 (−0.09)	−0.03 (−0.02)	−0.08 (−0.08)	−0.10 (−0.07)
Volume Instability	−0.04 (0.04)	−0.19* (−0.07)	−0.14* (−0.08)	0.02 (0.00)

* = $p < .05$