NIH Public Access
Author Manuscript
J Psychiatr Res. Author manuscript; available in PMC 2009 August 1.

Published in final edited form as:
J Psychiatr Res. 2008 August ; 42(10): 827–836.

Computerized Measurement of Negative Symptoms in Schizophrenia

Alex S. Cohen,
Department of Psychology, University of Maryland, College Park

Murray Alpert,
Department of Psychiatry, New York University Medical Center

Tasha M. Nienow,
Department of Psychiatry and Biobehavioral Sciences, UCLA

Thomas J. Dinzeo, and
Department of Psychology, Kent State University

Nancy M. Docherty
Department of Psychology, Kent State University

Abstract

Accurate measurement of negative symptoms is crucial for understanding and treating schizophrenia. However, current measurement strategies are reliant on subjective symptom rating scales which often have psychometric and practical limitations. Computerized analysis of patients' speech offers a sophisticated and objective means of evaluating negative symptoms. The present study examined the feasibility and validity of using widely-available acoustic and lexical-analytic software to measure flat affect, alogia and anhedonia (via positive emotion). These measures were examined in their relationships to clinically-rated negative symptoms and social functioning. Natural speech samples were collected and analyzed for 14 patients with clinically-rated flat affect, 46 patients without flat affect and 19 healthy controls. The computer-based inflection and speech rate measures significantly discriminated patients with flat affect from controls, and the computer-based measure of alogia and negative emotion significantly discriminated the flat and non-flat patients. Both the computer and clinical measures of positive emotion/anhedonia corresponded to functioning impairments. The computerized method of assessing negative symptoms offered a number of advantages over the symptom scale-based approach.

Keywords

negative; symptoms; schizophrenia; functioning; anhedonia

Send correspondence to Alex S. Cohen, Ph.D., Louisiana State University, Department of Psychology, 236 Audubon Hall, Baton Rouge, LA 70803, Email: acohen@lsu.edu.

[3]There was no evidence to suggest that the relationships between the clinical and computerized anhedonia measures and social dysfunction were unduly influenced by individual differences in intellectual functioning across patients. First, intelligence and social functioning were not significantly correlated ($r[58] = .17$, $p > .20$) within the patient group. Second, regression analyses revealed that the contribution of the computerized and clinical anhedonia measures to social dysfunction (entered together; $r^2 = .15$, F change = 5.01, $p < .05$) remained significant after accounting for the contribution made by IQ scores ($r^2 = .03$, F change = 1.98, ns).

1. Introduction

The negative symptoms of schizophrenia are a focus of considerable empirical and clinical attention. A crucial step in understanding and treating these symptoms involves developing reliable and valid methods to measure them. Presently, the most common measurement strategy employs likert-type symptom rating scales completed by trained interviewers (e.g., Andreasen, 1982; Lukoff et al., 1986). While the use of these scales has been instrumental to schizophrenia research (see Earnst & Kring, 1997, Blanchard and Cohen, 2006), their utility is constrained by a number of factors. For example, symptom-rating scales require considerable training to establish inter-rater reliability (Kobak, et al., 2004), can be time-consuming to administer and score, and often require multiple sources of information to make accurate ratings (Ho et al., 2004). From a psychometric perspective, symptom-ratings scales are not ideal because they often employ vague rating systems (e.g., “mild”, “moderate” and “severe“ categories from Andreasen, 1982 and Lukoff et al., 1986) that may be insensitive to subtle changes in symptom severity that occur over time (see Eckert et al., 1996). Moreover, they yield ordinal data that are less than ideal for parametric statistical analysis. To address these issues, some researchers have advocated measuring negative symptoms through computer analysis of behavior (IsHak et al., 2002). These approaches are advantageous because they employ a relatively straightforward and brief administration procedure, they are not constrained by “inter-rater” reliability, they yield ratio data that tends to be normal in distribution, and they are highly sensitive to fluctuations in symptom severity.

Computerized acoustic analysis of speech is a promising approach for measuring two cardinal negative symptoms of schizophrenia: blunt affect – which involves a diminished expression of emotion characterized by reduced vocal inflection and amplitude variability, and alogia – which involves a reduction in the quantity of speech. One acoustic analytic protocol in particular, the VOXCOM system (Alpert et al., 1986), has demonstrated utility as a negative symptom measure. Using VOXCOM, patients with clinically-rated “flat affect” have shown less vocal inflection and less overall speech production than nonflat patients with schizophrenia (Andreasen et al., 1981) and nonpatient controls (Alpert et al., 2000). On the downside, the VOXCOM system employs obsolete technology that limits its potential application to research and clinical settings. However, recent technological advances in psycholinguistics have paved the way for acoustic analysis using freely available software. The present project is meant to examine the feasibility of using a novel acoustic analytic protocol for measuring blunt affect and alogia by replicating the findings of Alpert et al. (2000). Specifically, we predicted that patients with clinically-rated flat affect would show less vocal inflection and slower speech rate than nonflat patients and nonpatient control subjects.

Computerized analysis of speech is also a potential measure of “experiential-based” negative symptoms through the use of lexical analysis. Lexical analysis focuses on the semantic level of communication (i.e., word choice) and can be used to measure psychological processes such as mood (e.g., Pennebaker et al., 2003). For example, lexical analysis of emotion words has been used as a measure of state emotion during mood induction procedures to investigate symptoms of numbing in patients with Post-Traumatic Stress Disorder (Orsillo et al., 2004). Experiential deficits are relevant to schizophrenia research because a reduction in the ability to experience emotion is considered to be an important negative symptom (American Psychiatric Association, 1994, Horan et al., 2004, Kirkpatrick et al., 1989). However, current strategies of measuring anhedonia are problematic in that they require clinicians to make judgments based on patient self-report. Schizophrenia is associated with impairments in insight (Kirkpatrick et al., 2000; Mintz et al., 2003), communication deficits (Docherty, 1996) and alexithymia (Cedro, et al., 2001), so, at least in some cases, patients ability to appraise and communicate their emotional states may be compromised. Given that lexical analysis is not dependent on patients’ conscious appraisals of their emotional states and is not constrained by subjective clinician judgments, lexical analysis offers a potentially powerful tool for understanding the emotional deficits in schizophrenia.

There has been limited application of lexical analysis for measuring anhedonia in schizophrenia. A recent study by Alpert et al., (2000) compared emotion words from natural speech between patients with flat affect, those without and nonpatient controls. Patients as a group did not significantly differ from controls in the number of overall emotion words used. Moreover, patients with flat affect did not significantly differ from nonflat patients in their use of emotion words, supporting the widely held notion that flat affect does not necessarily reflect an attenuation in emotional experience (e.g., see Kring, 1999). What is not clear from this study is the extent to which lexical analysis of emotion words can be used as a clinical measure of anhedonia. A secondary aim of this project was to conduct a preliminary examination using lexical analysis to measure anhedonia/diminished positive emotion. This was accomplished by examining the convergence between computerized and clinical-based measures of positive emotion.

Finally, it is important to consider how negative symptoms might be related to real world functioning. Clinically-rated negative symptoms have been associated with more severe functional impairments (e.g., Milev et al., 2005), but the functional correlates of computer-based negative symptom measures have yet to be explored. Interestingly, there is suggestion that certain clinically-rated negative symptoms are related to functional impairments more than others. A large scale factor analysis of the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1982) found that a symptom factor containing anhedonia symptoms was significantly related to functional impairment, whereas factors containing diminished expression and alogia symptoms were generally not (Sayers et al., 1996). In evaluating the computerized negative symptom measures, it would be important to determine the extent to which they are related to functioning. A final aim of this study was to examine the functional correlates of the acoustic, lexical and clinically-based negative symptoms. Given the findings of Sayers et al. (1996), we were particularly interested in the relationship between the measures of positive emotional experience and social functioning.

2. Methods

2.1 Subjects

The patient group was comprised of 44 male and 16 female volunteers who were recruited from inpatient psychiatric state hospitals within the Northcoast Behavioral Healthcare system which serves northern Ohio, USA. Each of these individuals met Diagnostic & Statistical Manual of Mental Disorders 4th edition (DSM-IV; APA, 1994) criteria for schizophrenia or schizoaffective disorder based on information obtained from the patients' medical records and a semi-structured interview (Schedule for Affective Disorders and Schizophrenia; Endicott & Spitzer, 1978) that was adapted for use with the DSM-IV. Twenty of these patients were Caucasian, 39 were African-American and 1 was Hispanic. The majority of these patients were committed under a forensic status (i.e., forensic evaluation, restoration of competency to stand trial, or a criminal finding of not guilty by reason of insanity). To provide assurance that patients could offer informed consent, we consulted with their treating psychiatrist and obtained their written permission prior to approaching the potential subjects. Further, we excluded patients with severe levels of psychosis (defined as a Global Assessment of Functioning rating (GAF; APA, 1994) below 30) or with documented evidence of mental retardation with the idea that these conditions would interfere with individuals ability to understand the study protocol. Other exclusion criteria included having a DSM-IV current alcohol or drug abuse or a history of alcohol or drug dependence suggestive of severe physiological symptoms (e.g., delirium tremens, repeated loss of consciousness), and a history of significant head trauma (requiring

NIH-PA Author Manuscript

NIH-PA Author Manuscript

overnight hospitalization). All patients were clinically stable at the time of testing and were receiving pharmacotherapy under the supervision of a multi-disciplinary team.

The control group was comprised of 13 male and 6 female university support staff. Ten of these subjects were Caucasian, and 9 were African-American. The control group was matched to the patient group in age, gender, and parental combined Socio-Economic Index (SEI; Hauser & Warren, 1996) variables. Controls were excluded if they met DSM-IV criteria for any substance-abuse or lifetime psychotic disorder, or if they had a history suggestive of the possibility of organic damage, as above. All subjects in this study provided written informed-consent prior to participating in the study after the goals and methods of the study were explained. This study is part of a larger investigation into the emotional and neuropsychological concomitants of social functioning (see Cohen et al., 2005; Nienow et al., 2006). Human Subjects Review Board approval was granted by Kent State University and by the Northcoast Behavioral Healthcare system.

2.2 Clinically-rated symptom measures

The Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984) and the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1982) were used to measure patients' symptom severity. As in Alpert et al., (2000), the "flat affect" group included all subjects with a two (defined as "mild") or higher score on at least three of the following subscales: unchanging facial expression, decreased spontaneous movements, paucity of expressive gestures, affective nonresponsivity and lack of vocal inflections. Cronbach's alpha was .90 for these five items suggesting that these measures were tapping into a common underlying construct. Fourteen patients were included in the "Flat Affect" group and 46 patients were included in the "Non-flat Affect" group. Ratings were made by one of three graduate level researchers who demonstrated adequate reliability based on independent reviews of 10 videotaped SADS interviews (Intra-class Correlation Coefficient values for each of the individual SAPS/SANS global ratings were > .70). Higher scores reflect more severe symptoms. To measure patients' overall positive and negative symptom severity, the global scores from the SAPS and SANS were summed.

In order to examine the convergence between the clinically-rated and computer-based measures of negative symptoms, we matched SANS items to computer-based variables based on overlap in coverage. These measures included: a) the SANS lack of vocal inflections which we expected to be inversely related to the computer-based inflection variable, b) SANS poverty of speech, increased latency of speech and global alogia, which was expected to be inversely related to the computer-based speech rate variable, and c) SANS global anhedonia; expected to be inversely corresponded to the computer-based positive emotion variable. We also include the SANS global affective flattening score as a "multi-modal" measure of affect, including facial, vocal and other physical expressiveness.

2.3 Computer-based negative symptom measures

Subjects each provided a five-minute narrative during which they responded to a set of predetermined open-ended probes that were designed to be emotionally-ambiguous in valence and content (e.g., "tell me about yourself and what kind of person you are"? "What do you like to do with your time"?). Subjects were expected to do most of the talking, but interviewers elicited elaboration from the subject when indicated. The narratives were recorded onto analogue tape using a portable tape recorder and a stereo microphone that was placed in front of the subject.

Acoustic analysis was conducted using PRAAT (Boersma & Weenink, 2006), a program that has been used extensively in speech pathology and linguistic studies. To prepare the narratives for acoustic analysis, the audiotapes were digitized and the interviewers' speech was removed from the sound file using sound edit tools. The resulting narratives ranged in length from 229 to 300 seconds. In order to remove background noise, each sound file was processed using low frequency noise reduction software (Wavhum Inc. v1.9, 2005). The PRAAT system organizes the sound file into "frames" for analysis, which for the present study was set at a rate of 20 per second. The entire speech sample, as opposed to select clips, was subject to analysis. We measured blunt affect in terms of inflection, which was computed as the standard deviation of the fundamental frequency [1]. Alogia was measured as speech rate in words per second. This variable was computed as the total word count divided by the total length of the speech sample with the interviewer's voice removed. Due to variability in recording conditions across subjects we were unable to measure vocal emphasis, a component of blunt affect.

The audio-taped narratives were transcribed by laboratory assistants and carefully proofread to ensure correct spelling and accurate transcription. Each narrative was reviewed so that exactly five minutes of speech were represented in the transcript, and then the interviewer speech was deleted. The Linguistic Inquiry and Word Count program (LIWC; Pennebaker, 2001) was used for lexical analysis of the speech samples. The LIWC program processes text files one word at a time, attempting to match the base form of the word to a "dictionary" of over 2290 words stems. These word stems are organized into 83 different categories. LIWC analysis yields a frequency count of the total instances of target words from each category. These numbers are then divided by the total number of words in the text to control for individual differences in verbosity. Thus, scores reflect a percentage of word matches in that category. LIWC contains separate positive and negative emotional categories, which are comprised of words related to emotional states (e.g., angry, happy, friendly).

2.4 Functioning

Social functioning was measured using the Social Functioning Scale (SFS; Birchwood, Smith et al., 1990). The SFS is a 79-item questionnaire that assesses a broad range of social behavior. The SFS was completed using information from patients' self-report, interviewer behavioral observation and from patients' medical records. A global social functioning score was computed by summing the individual subscale scores, transformed using standard scores (norms provided in Birchwood, 1990). The SFS has been shown to have acceptable coefficient alphas (.80), sensitivity and construct validity in patients with schizophrenia and non-psychiatric controls (Birchwood et al., 1990). Factor analysis of the SFS suggested the presence of a single factor, suggesting that each of the subscales reflect a single latent construct (Birchwood, 1990). For this reason, and to reduce the overall number of analyses, total SFS scores are presented in this study. Higher SFS scores reflect better social functioning. Intellectual functioning was measured using estimated IQ scores derived from the Shipley Institute of Living scales (Zachary et al., 1985), which is based on separate tests of vocabulary and abstract reasoning abilities. This scale has been found to be highly correlated with full scale WAIS-R IQ scores (>.79).

[1]The standard deviation of the fundamental frequency mean is commonly employed as a measure of inflection in studies of vocal prosody (e.g., Hagenaars & van Minnen, 2005), although it is important to contrast this measure from the VOXCOM output variable that has been used in studies by Alpert and colleagues. The measure used in the present study takes into account all fundamental frequency values in the sample that exceed the silence threshold and are within limits of the bandpass filter. Conversely, the VOXCOM measure considers only the fundamental frequency value for each syllable of spoken word at the point of maximum amplitude. To determine the convergence between these variables, we computed inflection values using both approaches for ten patient speech samples based on the first thirty seconds of speech. Reliability statistics between these measures were excellent for both fundamental frequency mean (single Intra-class Correlation Co-efficient value (ICC) = .99) and standard deviation (ICC = .90) values, suggesting that they are equivalent measures.

2.5 Analyses

The analyses were conducted in four steps. First, patients with clinically-rated flat affect, those without, and nonpatient controls were compared in descriptive, clinical and functioning variables. Second, we compared the three groups on the computerized acoustic and lexical variables with the expectation that flat patients would show significantly less inflection and slower speech rate than the other groups. We also included the computer-based emotion variable in this analysis although we had no a priori hypotheses regarding their outcome. Third, we examined the relationship between the computerized and clinical measures of blunt affect, alogia and emotional experience. We expected that the matching computer and clinical measures would be significantly intercorrelated [2]. Finally, we examined the relationship between the computerized and clinical negative symptom measures and social and intellectual functioning. We hypothesized that decreasing amounts of computer-based positive emotion and clinical-based anhedonia would be associated with poorer social functioning. Intellectual functioning was not expected to be significantly associated with the computer-based negative symptom measures. All results reported are two-tailed.

3. Results

3.1 Sample characteristics

Means and standard deviations were computed for the descriptive, clinical and functioning variables for the flat, nonflat and control groups. These data are presented in Table 1. Group-wise comparisons, using a series of ANOVA and Chi-square analyses suggested that the three groups were not statistically dissimilar in age, parental SEI or gender or ethnic composition. Controls had significantly more education, higher IQ scores, and better social functioning than both patient groups. The patient groups were similar in general and social functioning, and medication variables. Flat patients had significantly less education, more severe negative symptoms, and a trend for less severe positive symptoms compared to nonflat patients. Although the three groups were comparable in gender composition, it is important to discuss the potential impact this variable could have on the negative symptom measures. Females, as a group, typically show higher levels of inflection and are generally considered to be more expressive than males. With respect to the computerized negative symptom measures, patient ($M \pm sd = 23.25 \pm 6.12$) females had more inflection than patient males ($M \pm sd = 15.60 \pm 6.84$; $t[58] = 3.93$, $p. < .01$), and control females ($M \pm sd = 25.73 \pm 3.65$) showed a trend for more inflection than control males ($M \pm sd = 19.63 \pm 7.12$; $t[77] = 1.96$, $p. < .10$), but there were no significant gender differences on any of the other variables (all p. values > .10). On the clinically-rated negative symptom measures, there were no significant gender differences (all p. values > .10). It is also worth considering whether there were any differences in computer-based variables in the “minority” vs. Caucasian groups because the flat group had a non-significantly higher proportion of non-Caucasian participants. There were no significant differences between these groups in negative symptom measures for either the patient or control groups (all p. values > .10).

3.2 Computer-based negative symptoms: between group analyses

Table 2. contains the means and standard deviations for the flat, nonflat and control groups on the computer-based negative symptom measures. Each of these variables was roughly normally

[2]Given the aforementioned limitations of clinical measures for assessing negative symptoms, one could question the utility of evaluating the computerized measure based on its association with a clinical measure. In the present study, clinical measures were viewed as a “fallible indicator” of symptoms rather than a “gold-standard”. A moderate relationship between the clinical and computerized measures would be suggestive of convergent validity while a lack of relationship could indicate a serious validity problem for one or both of the instruments. In this sense, the association between the two measures seems a meaningful metric for determining the computerized measure’s validity.

distributed suggesting that parametric statistics were appropriate for group comparisons. These variables were compared across the three groups using a Multiple ANalysis Of Variance (MANOVA). The overall main effect for group was significant (Wilk's Lambda F [2, 152] = 3.27, p. < .01). Significant between group differences were observed for the inflection (F = 3.67, p. < .05) and speech rate (F = 7.56, p. < .01) variables, and a trend group difference for the negative emotion (F = 2.90, p. < .10) variable. Scheffe' post-hoc tests were then conducted. These findings suggested that flat patients showed significantly less inflection than controls, and significantly slower speech rate than both nonflat patients and controls. To evaluate the magnitude of differences, effect sizes were computed. As seen in Table 2, flat patients had less inflection and slower speech rate than controls at a large effect size, and at a medium or above effect size level compared to nonflat patients. The flat group expressed less negative emotion compared to the nonflat group at a medium effect size level, while effect sizes for positive emotion expression were small or negligible.

Given that females had significantly more inflection than males, and the flat group had slightly fewer females than the other groups, it is possible that the group differences in inflection reflect unequal gender distributions across groups. Group comparisons were recomputed separately for males and females. For males, the overall group main effect was reduced to a trend (F[2, 54] = 2.63, p < .10) while for females, the group effect was not significant (F[2, 19] = 1.19, p > .10). These analyses were underpowered, so we computed effect sizes to help gauge the magnitude of differences. For males, flat patients (M ± SD = 13.20 ± 3.96) had less inflection than controls (M ± SD = 19.63 ± 7.12; effect size = 1.12) and nonflat patients (M ± SD = 16.41 ± 7.44; effect size = .54) at a large and medium effect size respectively. For females, flat patients (M ± SD = 20.47 ± 9.75) had less inflection than controls (M ± SD = 25.73 ± 3.65; effect size = .72) and nonflat patients (M ± SD = 23.89 ± 5.35; effect size = .44) at a medium and small effect size level. Thus, females showed a pattern that was consistent with that seen in males, however, the strength of the effect was attenuated.

3.3 Convergence between computerized and clinical-based negative symptoms

Zero order correlations were computed between the acoustic, lexical and clinical based negative symptoms for the patients. This data is presented in Table 3. In order to compensate for excessive skew in the clinical measures, (range of skew values = 1.7 – 2.5) we employed non-parametric statistics (Spearmans' rho correlation coefficient). It is worth noting that clinically rated negative symptoms were not highly prevalent in this sample (defined as a score of "2 - mild" or greater). Twenty-three percent of the patients were rated as having "lack of inflection", 17% of them having "poverty of speech", 15% having "latency of response" and 27% having "global alogia". The global anhedonia scores were roughly normally distributed (skew value < 1.0) but were still subjected to non-parametric statistics to maintain continuity across these analyses.

Some findings from Table 3 bear mention here. First, as hypothesized, the computer-based speech rate measure was significantly associated with each of the clinical alogia ratings suggesting convergence between these measures. Second, computer-based inflection was significantly and inversely correlated with the clinically-rated lack of vocal inflections variable. It is noteworthy that clinically-rated lack of vocal inflections was also significantly associated with the computer-based speech rate measure, raising the possibility that the inflection ratings were contaminated by other characteristics of speech besides inflection. Third, in contrast to expectations, clinically-rated global anhedonia was not significantly related to computer-based positive emotion suggesting that the clinically-rated anhedonia and computer-based positive emotion variables were capturing different phenomenon. Fourth, the acoustic measures were not correlated much with each other (all r values < .10) suggesting that these measures were each tapping separate constructs. On the other hand, the clinically-rated lack of vocal inflections and alogia measures (i.e., SANS poverty of speech, increased latency and global alogia) were all significantly intercorrelated (r values range = .37 – .63, all p values < .05) suggesting that there was considerable overlap in the scope of these measures. Finally, the clinical ratings of global anhedonia were significantly related to more severe clinically-rated poverty of speech, global alogia, latency of response and global affective flattening indicating that clinically-rated anhedonia tended to co-occur with other clinically-rated negative symptoms.

3.4 Negative symptom measures and social and intellectual functioning

Bivariate correlations were computed between the functioning variables and the computer (r = Pearson's correlational coefficient) and clinical-based (r = Spearman's correlational coefficient) negative symptoms. Computer-based positive emotion (r = .30, p < .05) and clinically-rated global anhedonia (r = −.32, p < .05) were each significantly associated with poorer social functioning, but none of the other clinical or computer-based measures were significantly associated with social functioning (all p. values > .10). None of the computer-based measures were significantly associated with intellectual functioning (all p. values > .10), but the clinical-based lack of vocal inflection (r = −.35, p < .05), global anhedonia (r = −.36, p < .05) and global affective flattening (r = −.33, p < .05) measures each corresponded to lower intellectual functioning. In sum, lower levels of positive emotionality, measured using either the clinical or computerized methods, were associated with poorer social functioning. Unlike the clinical measures, the computer-based negative symptoms were relatively uncorrelated with intellectual functioning.

In order to determine the specific contributions of anhedonia severity to functional impairments beyond that made by other negative symptoms, we computed separate hierarchical regressions for the computer and clinical measures. These were set up so that anhedonia/positive emotion measure was entered in the second step and the other negative symptoms were entered as a block in the first step. Computer-based anhedonia explained 6% (F change = 5.03, p < .05) of the variance in social functioning beyond the negligible contribution made by other computer-based measures of negative symptoms (r^2 = .04, F change = 1.48, ns). Similarly, clinically-rated anhedonia explained 11% (F change = 7.02, p < .05) of the variance above and beyond the contribution made by the other clinically-rated measures of negative symptoms (r^2 = .06, F change = .82, ns). Together, the computer and clinical anhedonia measures explained 19% (F change = 6.18, p < .01) of the social functioning score variance. Thus, anhedonia, more than other negative symptoms, was associated with social functioning impairments.

4. DISCUSSION

The primary aim of this study was to evaluate the feasibility of measuring key negative symptoms using computerized acoustic and lexical analysis of natural speech. The present findings suggest that the computerized measures have important advantages over traditional clinical-based measures. From a psychometric perspective, the computer-based procedures provided normally-distributed, ratio-level data, whereas the data from the clinical rating scales was highly skewed with a restricted range. Moreover, the acoustic analysis was sufficiently sensitive to detect differences in vocal inflection between males and females, whereas the clinical ratings of vocal inflection were not. The computer-based measures also appeared to provide a more precise measure of individual negative symptoms than the clinical measures. The computer-based measures were relatively unrelated to each other suggesting that they were tapping into distinct, separable constructs while the clinical ratings showed a high degree of overlap. These findings add to growing concerns that clinical ratings of negative symptoms tend to inappropriately reflect raters' global impressions about a patient's negative symptoms rather than reflecting a specific operationalized sign or symptom (see Alpert et al., 2002). As an illustration of this point, clinical ratings of inflection were more related to the computer-based alogia scores than the computerized vocal inflection scores. This suggests that the clinical ratings of inflection may have been inappropriately contaminated by aspects of patients' speech other than inflection, such as speech rate.

In support of the validity of our acoustic analytic protocol as a measure of vocal blunt affect and alogia, there was moderate convergence between the clinical and acoustic-based measures. Consistent with the findings of studies using a different acoustic analytic protocol (i.e., VOXCOM; Alpert et al., 2000; Alpert et al., 2002), patients with clinically-rated blunt affect showed low levels of inflection and slow speech rate. The magnitude of the group comparisons between flat and nonflat patients and controls were in the medium to large effect size range, suggesting that these differences were not trivial. Within patients, the computerized and clinical measures of alogia and blunt affect showed statistically significant convergence (at small to medium effect size levels).

We found encouraging, but somewhat mixed support for the use of lexical analysis as a measure of anhedonia in schizophrenia. On one hand, as hypothesized, the computer-based measure of positive emotion was associated with social dysfunction, a finding that is in line with other research on social functioning in schizophrenia (e.g., Blanchard et al., 1998; Sayers et al., 1996). On the other hand, the clinical and computer-based measures of positive emotion were not significantly inter-correlated thus raising questions about the overlap in coverage between these instruments. When attempting to reconcile this null finding, it is important to note that the lexical measure probably reflects state emotions whereas the clinical measure of anhedonia, which covers a month long epoch, is reflective of stable emotional deficits. Thus, a more accurate test of the lexical measure's convergent validity might involve the use of state-emotion self-report or physiological measures. Although symptoms of anhedonia are typically conceptualized in terms of an enduring trait, state instruments will be vital for monitoring subtle fluctuations in symptom severity. Thus, lexical analysis remains an intriguing means of measuring positive emotion in schizophrenia. Efforts to further explore this issue are currently underway by our laboratory.

Of all the negative symptom measures examined in this study, only those pertaining to positive emotion were significantly related to social functioning impairments. Consistent with this finding, anhedonia has been associated with more severe dysfunction in patients (Blanchard et al., 1998; Cohen et al., 2005) and Sayers et al. (1996) reported that a negative symptom factor defined by symptoms of anhedonia were significantly related to functional impairment whereas other negative symptom factors were not. Taken together, these findings suggest that "experiential-based" negative symptoms, such as anhedonia, contribute more to functional outcome than "expression-based" negative symptoms. If true, this has important implications for efforts aimed at improving functional outcome. However, current technologies for measuring emotional experience may be inadequate for discriminating experiential from expressive deficits in schizophrenia (see Horan et al., 2006). Computerized symptom assessment reflects a promising methodology for clarifying the causal mechanism by which negative symptoms compromise functional abilities.

The finding that inflection, measured using acoustic analysis, is somewhat less restricted in flat female patients than flat male patients relative to controls is consistent with evidence that negative symptoms may be sex-linked in some manner. Female patients tend to have less severe negative and deficit symptoms than males (Earnst & Kring, 1997; Kirkpatrick et al., 2001), and first degree relatives of female patients have shown less severe flat affect symptoms compared to first degree relatives of male patients (Goldstein et al., 1990). Similarly, Crow and colleagues have reported evidence that homologous genes on the X and Y chromosomes contribute to different phenotypic expressions in males and females (Crow, 1988; Crow et al., 1994). While interpretation of gender differences in this study is hampered by a lack of statistical power, these results highlight the importance of considering gender when developing norms for negative symptom measures.

Limitations of this study warrant mention here. First, each of the patients in this study was medicated, and our ability to determine the effects of pharmacological agents on the computerized negative symptom measures was constrained. Given that medication type and dosage is not random, it would be difficult to examine this issue without employing an unmedicated sample. Second, the patient sample was somewhat atypical in that it was comprised primarily of forensic cases. It stands to reason that negative symptoms are less frequent in forensic samples, so it could be the case that sample characteristics contributed to a restricted range of negative symptoms. Consequently, some of the analyses in this study were underpowered because there were relatively few "flat" patients in this sample. For the most part, prior findings were replicated suggesting that the present analyses had adequate statistical power. Third, due to technological limitations, we were unable to examine vocal amplitude as a measure of blunt affect. Finally, we were unable to meaningfully examine the impact of neuropsychological deficits or other variables (e.g., premorbid functioning, education level, verbal skill) on acoustic or lexical expression. This remains an important topic for future research.

The present findings have important implications for the study of negative symptoms, namely, that computerized analysis of natural speech is a viable option for measuring negative symptoms that overcomes many of the limitations associated with current measurement methods. An application of this line of research would be to use computerized negative symptom assessment for pharmaceutical trials and clinical settings. For this goal to be realized, an important next step would be to standardize the speech capture protocol so that it is applicable to a wide variety of clinical and laboratory settings. We believe that further refinement of the data processing procedure, for example, through the use of digital recording and speech recognition software, could someday facilitate point of care assessment of negative symptoms.

References

Alpert M, Merewether F, Homel P, Marz J. Voxcom: A system for analyzing natural speech in real time. Behavior Research Methods, Instruments & Computers 1986;18(2):267.

Alpert M, Rosenberg SD, Pouget ER, Shaw RJ. Prosody and lexical accuracy in flat affect schizophrenia. Psychiatry Research 2000;97(2–3):107–118. [PubMed: 11166083]

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. [PubMed: 12127603]

Andreasen NC. Negative symptoms in schizophrenia. Definition and reliability. Archives of General Psychiatry 1982;39(7):784–788. [PubMed: 7165477]

Andreasen, NC. Scale for the Assessment of Positive Symptoms. Iowa City: University of Iowa Press; 1984.

Andreasen NC, Alpert M, Martz MJ. Acoustic analysis: an objective measure of affective flattening. Archives of General Psychiatry 1981;38(3):281–285. [PubMed: 7212958]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington DC: American Psychiatric Association; 1994.

Birchwood, M. The Social Functioning Scale: Unpublished Administration Manual. 1990.

Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. British Journal of Psychiatry 1990;157:853–859. [PubMed: 2289094]

Blanchard JJ, Cohen AS. The structure of negative symptoms within schizophrenia: implications for assessment. Schizophrenia Bulletin 2006;32(2):238–245. [PubMed: 16254064]

Blanchard JJ, Mueser KT, Bellack AS. Anhedonia, positive and negative affect, and social functioning in schizophrenia. Schizophrenia Bulletin 1998;24(3):413–424. [PubMed: 9718633]

Boersma, P.; Weenink, D. Praat: doing phonetics by computer (Version 4.4.05). 2006.

Cedro A, Kokoszka A, Popiel A, Narkiewicz-Jodko W. Alexithymia in schizophrenia: An exploratory study. Psychological Reports 2001;89(1):95. [PubMed: 11729558]

Cohen AS, Dinzeo TJ, Nienow TM, Smith DA, Singer B, Docherty NM. Diminished emotionality and social functioning in schizophrenia. Journal of Nervous and Mental Diseases 2005;193(12):796–802.

Crow TJ. Sex chromosomes and psychosis. The case for a pseudoautosomal locus. British Journal of Psychiatry 1988;153:675–683. [PubMed: 3255457]

Crow TJ, Delisi LE, Lofthouse R, Poulter M, Lehner T, Bass N, et al. An examination of linkage of schizophrenia and schizoaffective disorder to the pseudoautosomal region (Xp22.3). British Journal of Psychiatry 1994;164(2):159–164. [PubMed: 7818635]

Docherty NM. Affective reactivity of symptoms as a process discriminator in schizophrenia. Journal of Nervous and Mental Diseases 1996;184(9):535–541.

Eckert SL, Diamond PM, Miller AL, Velligan DI, Funderburg LG, True JE. A comparison of instrument sensitivity to negative symptom change. Psychiatry Research 1996;63(1):67–75. [PubMed: 8832775]

Earnst KS, Kring AM. Construct validity of negative symptoms: an empirical and conceptual review. Clinical Psychology Review 1997;17(2):167–189. [PubMed: 9140714]

Endicott J, Spitzer RL. A diagnostic interview: The schedule for affective disorders and schizophrenia. Archives of General Psychiatry 1978;35(7):837. [PubMed: 678037]

Goldstein JM, Faraone SV, Chen WJ, Tolomiczencko GS, Tsuang MT. Sex differences in the familial transmission of schizophrenia. Br J Psychiatry 1990;156:819–26. [PubMed: 2207512]

Hagenaars MA, van Minnen A. The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia. Journal of Anxiety Disorders 2005;19(5):521. [PubMed: 15749571]

Hauser, R.; Warren, J. unpublished manual: Center for Demography and Ecology. 1996. Socioeconomic indexes for occupations: a review, update and critique.

Ho BC, Flaum M, Hubbard W, Arndt S, Andreasen NC. Validity of symptom assessment in psychotic disorders: information variance across different sources of history. Schizophrenia Research 2004;68 (2–3):299–307. [PubMed: 15099611]

Horan WP, Kring AM, Blanchard JJ. Anhedonia in schizophrenia: a review of assessment strategies. Schizophrenia Bulletin 2006;32(2):259–273. [PubMed: 16221997]

IsHak, WW.; Almoguera-Abad, A.; Alpert, M.; IsHak, WW.; Burt, T.; Sederer, LI. Outcome measurement in psychiatry: A critical review. American Psychiatric Publishing, Inc; 2002. Innovative approaches in outcome measurement: A futuristic view; p. 393

Kirkpatrick B, Buchanan RW, Ross DE, Carpenter WT Jr. A separate disease within the syndrome of schizophrenia. Archives of General Psychiatry 2001;58(2):165–171. [PubMed: 11177118]

Kirkpatrick B, Castle D, Murray RM, Carpenter WT Jr. Risk factors for the deficit syndrome of schizophrenia. Schizophrenia Bulletin 2000;26(1):233–242. [PubMed: 10755684]

Kring AM. Emotion in schizophrenia: Old mystery, new understanding. Current Directions in Psychological Science 1999;8(5):160.

Kobak KA, Engelhardt N, Williams JB, Lipsitz JD. Rater training in multicenter clinical trials: issues and recommendations. Journal of Clinical Psychopharmacology 2004;24(2):113. [PubMed: 15206656]

Lader M. Rating scales in schizophrenia: A review of their usefulness for assessing atypical antipsychotics. CNS Drugs 2000;14(1):23.

Lukoff D, Nuechterlein KH, Ventura J. Appendix A: Manual for expanded brief psychiatric rating scale. Schizophrenia Bulletin 1986;12(4):594–602.

Milev P, Ho BC, Arndt S, Andreasen NC. Predictive Values of Neurocognition and Negative Symptoms on Functional Outcome in Schizophrenia: A Longitudinal First-Episode Study With 7-Year Follow-Up. American Journal of Psychiatry 2005;162(3):495. [PubMed: 15741466]

Mintz AR, Dobson KS, Romney DM. Insight in schizophrenia: A meta-analysis. Schizophrenia Research 2003;61(1):75. [PubMed: 12648738]

Nienow, Docherty, Cohen, Dinzeo. Attentional Dysfunction, Social Perception, and Social Competence: What is the Nature of the Relationship? Journal of Abnormal Psychology. 2006

Nkam I, Langlois-Thery S, Dollfus S, Petit M. L'alexithymie chez des schizophranes deficitaires et non deficitaires. Encaphale 1997;23(5):358.

Orsillo SM, Batten SV, Plumb JC, Luterek JA, Roessner BM. An experimental study of emotional responding in women with posttraumatic stress disorder related to interpersonal violence. Journal of Traumatic Stress 2004;17(3):241–248. [PubMed: 15253096]

Pennebaker, JW. Linguistic Inquiry and Word Count. Lawrence Erlbaum Associates; 2001.

Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: Our words, our selves. Annual Review of Psychology 2003;54:547.

Sayers SL, Curran PJ, Mueser KT. Factor structure and construct validity of the Scale for the Assessment of Negative Symptoms. Psychological Assessment 1996;8(3):269.

Tull MT, Medaglia E, Roemer L. An investigation of the construct validity of the 20-Item Toronto Alexithymia Scale through the use of a verbalization task. Journal of Psychosomatic Research 2005;59(2):77–84. [PubMed: 16186002]

Zachary RA, Paulson MJ, Gorsuch RL. Estimating WAIS IQ from the Shipley Institute of Living Scale using continuously adjusted age norms. Journal of Clinical Psychology 1985;41(6):820–831. [PubMed: 4078008]

Table 1

Means and standard deviations for Flat affect (n = 14), Nonflat (n = 46) and Control groups (n = 19) on descriptive, socioeconomic index score (SEI), functioning, clinical and medication data.

MEASURE Descriptive Variables	GROUP Flat (F)	Nonflat (NF)	Controls (C)	Group Comparison
Age	37.86 ± 11.09	42.67 ± 7.34	38.32 ± 10.97	-
Education	10.79 ± 2.39	12.37 ± 1.74	14.21 ± 1.44	F < NF** < C**
Parents combined SEI[1]	59.80 ±15.67	63.32 ± 26.98	59.34 ± 20.69	-
% Male	79%	72%	68%	-
% Caucasian	21%	37%	53%	-
Average hospital stay (wks)[2]	6.64 ± 9.21	5.69 ± 8.57	-	-
Social Functioning Scale[1]	107.35 ± 8.22	107.55 ± 7.70	122.81 ± 4.05	F = NF < C**
GAF,[13]	40.69 ± 6.56	41.30 ± 8.63	-	-
IQ[14]	76.29 ± 9.43	81.51 ± 10.76	101.37 ± 11.82	F = NF < C**
SAPS global totals[5]	4.36 ± 2.79	6.37 ± 4.19	-	F < NF+
SANS global total[5]	8.50 ± 4.26	4.24 ± 3.34	-	NF < F**
Medications				
% on atypical antipsychotics	85%	91%	-	-
% on antidepressants	15%	4%	-	-
% on mood stabilizers	31%	44%	-	-

+ = p value < .10,

** = p. value < .01.

[1] Higher scores reflect better functioning,

[2] One extreme score case (336 weeks) was excluded from Flat group for this analysis only,

[3] GAF = Global Assessment of Functioning, Scale 0–100, (Diagnostic & Statistical Manual of Mental Disorders, American Psychiatric Association, 1994),

[4] In standard scale format (mean = 100, standard deviation = 15),

[5] Higher scores reflect more severe symptomatology.

Table 2

Means and standard deviations and comparison effect sizes for the Flat affect (n = 14), Nonflat (n = 46) and Control groups (n = 19) on computer-based negative symptom measures.

Measures:	Group			Effect size		
	Flat	Nonflat	Controls	Flat vs. Nonflat	Nonflat vs. Controls	Flat vs. Controls
Inflection (Hz.)[1]	14.76 ± 6.02	18.52 ± 7.66	21.56 ± 6.78	.55	.42	1.06
Speech rate (words/sec.)[2]	1.63 ± .39	2.26 ± .60	2.31 ± .59	1.24	.08	1.36
Positive emotion (% total)	3.57 ± 1.98	3.04 ± 1.62	3.19 ± 1.19	.31	.24	.10
Negative emotion (% total) 3	.45 ± .44	.75 ± .48	.55 ± .44	.63	−.22	.42

[1]Flat patients < Controls, $p < .05$,

[2]Flat patients < Nonflat patients = Controls, $p < .01$,

[3]Flat patients < Nonflat patients = Controls, $p < .10$.

Table 3

Zero-order correlations between the computer and clinical-based measures of negative symptoms for patients (n = 60).

COMPUTER MEASURES	COMPUTER MEASURES 1.	2.	3.	4.	5.	6.	CLINICAL RATINGS 7.	8.	9.
1. Inflection (Hz.)	1.00	-	-	-	-	-	-	-	-
2. Speech rate (words/sec.)	.02	1.00	-	-	-	-	-	-	-
3. Positive emotion words (% of total)	−.14	−.14	1.00	-	-	-	-	-	-
4. Negative emotion words (% of total)	.10	.48**	−.35**	1.00	-	-	-	-	-
CLINICAL RATINGS									
5. Lack of Vocal Inflections[1]	−.26*	−.41**	−.04	−.05	1.00	-	-	-	-
6. Poverty of speech[1]	−.06	−.57**	−.09	−.05	.63**	1.00	-	-	-
7. Increased latency[1]	.08	−.40**	.06	−.03	.37**	.62**	1.00	-	-
8. SANS global alogia[1]	.08	−.36**	.07	.08	.53**	.70**	.70**	1.00	-
9. SANS global anhedonia[1]	.30*	−.09	−.02	−.03	.17	.37**	.23+	.28*	1.00
10. SANS global affective flattening	−.15	−.49**	.12	−.28*	.69**	.57**	.56**	.50**	1.00

+ = p value < .10,

* = p. value < .05,

** = p. value < .01.

[1] Spearman's Rho was computed to minimize the impact of skew in the clinical rating scores.

Higher scores reflect more severe symptomatology.