Published in final edited form as:

Curr Opin Psychiatry. 2014 May ; 27(3): 203–209. doi:10.1097/YCO.0000000000000056.

Automated Computerized Analysis of Speech in Psychiatric Disorders

Alex S. Cohen and

Department of Psychology, Louisiana State University

Brita Elvevåg

Psychiatry Research Group, Department of Clinical Medicine, University of Tromsø & The Norwegian Centre for Integrated Care and Telemedicine (NST), University Hospital of North Norway, Tromsø, Norway

Abstract

Purpose of Review—Disturbances in communication are a hallmark of severe mental illnesses. Recent technological advances have paved the way for objectifying communication using automated computerized linguistic and acoustic analysis. We review recent studies applying various computer-based assessments to the natural language produced by adult patients with severe mental illness.

Recent Findings—Automated computerized methods afford tools with which it is possible to objectively evaluate patients in a reliable, valid and efficient manner that complements human ratings. Crucially, these measures correlate with important clinical measures. The clinical relevance of these novel metrics has been demonstrated by showing their relationship to functional outcome measures, their in vivo link to classic ‘language’ regions in the brain, and, in the case of linguistic analysis, their relationship to candidate genes for severe mental illness.

Summary—Computer based assessments of natural language afford a framework with which to measure communication disturbances in adults with SMI. Emerging evidence suggests that they can be reliable and valid, and overcome many practical limitations of more traditional assessment methods. The advancement of these technologies offers unprecedented potential for measuring and understanding some of the most crippling symptoms of some of the most debilitating illnesses known to humankind.

Keywords

natural language; latent semantic analysis; computational linguistics; acoustic analysis; speech disturbances

Introduction

Disturbances in natural speech are defining characteristics of many adult severe mental illnesses (SMI), notably schizophrenia, depression, bipolar, autism and even some

personality disorders such as histrionic, borderline, schizotypal and schizoid personality disorders (1). Across these disorders, and even within them, communication can be affected in many different ways. For example, in verbal communication the content can be vague, confusing, illogical, overly sparse or otherwise ineffectual. Moreover, nonverbal channels of speech can be flat or constricted, such that speech production and modulation are reduced or absent altogether. Conversely, language disturbances can present as exaggerated speaking styles in terms of pressured, loud, dramatic and overly circumstantial content. Despite the importance and prevalence of these symptoms in SMI, surprisingly little is known about them in terms of underlying etiology, or even how they manifest similarly or differently across various disorders. The latter point is particularly important in light of recent efforts, most notably by the National Institute of Mental Health (USA), to develop taxonomies of mental illness based on functional processes, mechanisms and biomarkers as opposed to traditional diagnostic categories (2). Here we discuss the recent literature regarding measurement of speech disturbances, focusing primarily on a set of particularly promising methods employing computerized and automated objective analysis of natural speech.

Measures of Communication Disturbances in Severe Mental Illness

Historically speaking, three general approaches to measuring speech disturbances in psychiatric disorders have been employed, involving interviewer-based symptom rating scales, self-report scales and behavioural based measures. While a comprehensive list of the pros and cons for each approach is beyond the scope of this paper, it is worth briefly discussing their applications and major limitations. The vast majority of studies employ symptom rating scales, which involve a trained professional evaluating various aspects of communication based on observation during a clinical interview or interaction. A plethora of such measures exist (e.g., 3, 4), and most if not all employ an ordinal based rating system with a limited response set (often three to seven options) to denote symptom severity. While these scales have been used extensively in psychiatry research, they suffer from serious limitations (see Table 1). Most notably, they have shown poor specificity in isolating the specific facet of communication they are attempting to measure, and are generally insensitive for measuring all but gross changes in severity of symptoms and clinical functioning over time or between individuals (5). Alternative approaches, involving self-report scales and behaviour-based coding systems, have also been advanced. Self-report approaches are limited in that many patients suffer from poor insight and are thus inaccurate in evaluating and reporting their symptom severity. Behaviour-based coding strategies are advantageous in that they quantify communication disturbances using ratio scales and are based directly on behaviour. However, they are time consuming and require intensive levels of training for reliable administration and are impractical for most clinical and research settings.

A complimentary and potentially more sophisticated approach, which is the focus of this paper, involves automated, objective, computer-based assessment of natural speech. This approach has been employed extensively in other disciplines, (e.g., communication sciences, military and aeronautics, education), but the focus on SMI has been relatively limited. Due in large part to technological and methodological advances, the ability to employ computer-based assessment has become inexpensive and efficient (see Table 1) and a number of

commercially available programs exist that can be adapted for their use (6). The major limitations regarding their use are that they have been employed in research and, for a variety of reasons, have not been normed in terms of sensitivity and specificity. Moreover, these technologies have almost exclusively been applied to a few SMI disorders (e.g., schizophrenia-spectrum disorders, depression, autism), to the exclusion of other illnesses within the SMI spectrum (e.g., “cluster b and c” personality, bipolar-spectrum, obsessive compulsive, anxiety and post-traumatic stress disorders). In the next section, we review recent studies employing automated computer analysis to understand communication disturbances in schizophrenia-spectrum disorders and depression. This section will focus on two distinct types of speech analysis, that involving semantics and linguistics and another focusing on speech production, variability and acoustic characteristics of speech.

Automated Computer Analysis of Semantic and Linguistics in Severe Mental Illness

In SMI, numerous dimensions of communication can be affected, and of particular importance is semantic communication which involves the meaning of linguistic expression. In schizophrenia, for example, the incoherent discourse - a disjointed flow of ideas, loose associations between words or digressions from the topic - frequently appears during spontaneous speech, and is important for diagnosis, treatment monitoring and prognostic considerations (1). Semantic disturbances are also prominent in individuals with bipolar and histrionic personality disorders, wherein patients often communicate in wildly circumstantial styles. Measuring this incoherence is complex. However, recent proof of concept studies using automated computational linguistic analysis methods show it is possible to evaluate patients based on open-ended verbalizations (7, 8) using speech samples of several minutes as well as using responses that constitute only a few words. Importantly, these novel measures have been validated in that they correlate with traditional clinical measures of thought disorder. This technology employed a number of computational analyses of natural language including latent semantic analysis (LSA), a computational model of meaning that closely mimics human understanding of the contextual use of language (9). The key idea behind these models is that people are sensitive to weak statistical regularities in the linguistic environment, such as the co-occurrence of words in a sentence. Using text corpora, LSA can learn the meaning of a word by estimating the relatedness of any arbitrary set of words as a function of the contexts in which they co-occur. It is important that the input is a large database of word occurrences (e.g., very large text corpora of say 100,000 unique words), or else the solutions are likely unreliable (10). Indeed large text corpora used to construct LSA spaces have been validated and related to a variety of cognitive processes in healthy individuals such as sorting and category judgments, similarity judgments, and lexical priming tasks (11). One of the key advantages of this approach to semantics is that representations for all kinds of words, including abstract or low frequency words can be derived, even if those words never co-occur in the same text or sentence through the use of singular value decomposition or probabilistic inference.

At present these computational linguistic approaches to operationalizing illness severity or symptom ratings are primarily employed within research and a recent special issue of the

journal *Cortex* showcases a variety of scenarios in which such computational and automated approaches to semantics and language are useful in psychiatry and neurology research (for an overview, see (12)). Importantly such studies have shown that these new metrics are clinically useful in that they are related to disconnected speech and functional (outcome) behaviour in a large sample of elderly patients with schizophrenia (13). Moreover, such approaches can complement cognitive (neuropsychological) assessment via automated scoring of the verbal output (i.e., recall) generated by participants on traditional verbal memory task such as the Logical Memory subtest of the Wechsler Memory Scale (14.), and moreover can provide additional insights into the sources of language disruption and underlying mechanisms of prose recall (15). Indeed, such a framework to assessment naturally eliminates potential problems of inter-rater reliability, and in cases that use computational semantic approaches there is also some clinical validation in terms of clinically rated thought disorder and/ or speech coherence in schizophrenia (16–18). Computational language methods have also been used to detect changes in language use (pertaining to representations of self) after treatment for depression (19), as well as to detect personality changes such as in Hubris syndrome as assayed in the speech of senior politicians (20).

Moreover, such a computational framework provides an objective way to index coherence in speech that can be usefully related to underlying neurophysiological processes as revealed by fMRI (i.e., LSA scores of speech samples correlate with classic ‘language’ regions in the brain, thus suggesting LSA scores are reflective of underlying neurophysiological mechanisms) and are consistent with existing evidence for a role of the superior temporal cortex in thought disorder in psychosis (21). Indeed, the biological relevance of such automated approaches is further illustrated by a candidate gene approach in schizophrenia combined with a computational analysis of the linguistic data that found that novel LSA-derived phenotypes were associated with single-nucleotide polymorphisms (SNPs) previously associated with candidate genes for major depression, schizophrenia and reading-related traits and risk for dyslexia (22). Put differently, the actual meaning of participants’ responses provides neurocognitively useful metrics and ones that relate to underlying neurobiology, and as such has both clinical and research value. A quite different way of uncovering the phenotypic signature of SMI’s for clinical and genomic studies is to text mine electronic medical records, a process that can be done on a very large scale (e.g., thousands of patients). Adopting such an approach, Lyalina et al (23) explored the boundaries between various neuropsychiatric disorders in terms of how the expert-defined disease boundaries compared to the realities of patient care. The authors focused on autism, bipolar disorder, and schizophrenia, and after analyzing the records of 7000 patients found considerable overlap between schizophrenia and bipolar disorder, but a clear distinction with autism. Indeed there are many potential biomedical applications of text-based mining approaches that extract clinical concepts from patient reports or even psychiatric narrative (5), for example as an aid for clinicians who are not yet experts and thus not especially good at detecting important patterns in clinical narratives.

Automated Analysis of Nonverbal Vocal Expression in Severe Mental Illness

Although automated analysis of nonverbal speech features (a.k.a. “acoustic analysis”) has existed for decades, its application to SMI research has been modest, at best. Historically speaking, acoustic analysis has been used somewhat in children with autism spectrum research (24), and in adult populations with autism-spectrum (25) schizophrenia-spectrum (5) and mood (26) disorders, and, much of this work occurred over a decade ago. To our knowledge, other psychiatric disorders, notably dysthymia, bipolar, borderline, antisocial, histrionic and other personality disorders, have been given little attention. Acoustic analysis yields a multitude of diverse variables that can be organized into three conceptual categories, including speech production – the amount and rate of speech articulated, speech variability – concerning modulation of fundamental frequency (i.e., “pitch) and volume, and speech quality – tonal, spectral or other articulatory factors (e.g., tongue or lip movement). In terms of the psychometric properties of acoustic analysis measures in psychiatric populations, speech variables have shown relatively high temporal stability over hour and week-long epochs (27, 28) and shown modest (or better) correlations to symptom rating scale measures (27–29). Given that the procedure is automated, near perfect test-retest reliability can be expected for the same speech sample. In sum, automated acoustic analysis of speech can be considered reliable and valid.

Looking globally at the psychiatry literature, there is evidence that speech production and speech variability is abnormal in adults with autism-spectrum disorders (30), though the nature of this relationship appears to be influenced by functioning level such that high functioning autism is associated with greater intonation/inflection (31) and low functioning autism is associated with decreased intonation/inflection (32). There is also evidence that schizophrenia (5), schizotypal personality (33) and depressive disorders (26), as a group, are characterized by reduced speech production and/or variability in at least some domains compared to nonpsychiatric control subjects. However, the magnitude of these differences varies considerably across studies. Some differences have been reported to be in the large range in patients with schizophrenia (e.g., Cohen’s $d = 4.62$ for average pause length in (26)). More modest results have been reported in recent studies. For example, small to medium effect sizes were reported for speech production and variability between schizophrenia and controls (5). Moreover, small effects were noted in two recent studies of individuals with schizotypal personality characteristics (33, 34). As yet, evaluation of this literature fails to identify a single speech characteristic that is consistently deficient at a clinical level across studies and within disorder groups. In large part, this reflects individual studies examining and reporting on different aspects of speech (e.g., pause duration, percent time talking), making the findings difficult to compare with each other. Thus, conclusions regarding the magnitude of vocal deficits in SMI are premature at this time.

A related issue with the literature of nonverbal expression in SMI is that many studies fail to employ a nonpsychiatric control group. In this regard, it is difficult to evaluate the degree to which patient speech is statistically deviant. It is worth noting that a recent study found negligible differences in speech production and speech variability in patients with

schizophrenia, depression and bipolar disorders (27), suggesting that speech may not differ across these disorders. Unfortunately, control data were not available for this study, so it is unclear the degree to which patients were anomalous more generally. It is also worth noting a modest, but interesting, literature focusing on how computer-based measures of speech production and speech variability are associated with various aspects of illness. There is evidence that non-language neurocognitive abilities (e.g., attention) and acoustic measures of speech production and variability are correlated with each other in patients with schizophrenia, schizotypal personality, depression and bipolar disorders (35) (36). Such findings have fueled theories that reductions in speech production, and possibly speech variability, fundamentally reflect neurocognitive failures. Experimental investigation of this issue has been limited, though evidence that speech is unduly affected in some, but not all, individuals with schizotypal personality characteristics has been reported (34). Recent studies examining neurobiological and treatment correlates have also recently been conducted (33)(37). In sum, examining neurocognitive, neurobiological and treatment response correlates of speech characteristics measured using acoustic analysis reflect important avenues for future research.

Conclusion

Merely a century ago, diagnosis and symptom assessment for virtually all medical illnesses was almost exclusively made based on observation – now technologies have complimented this process and allowed for unprecedented efficiency, accuracy and potentially improved outcomes. One can imagine a future when mental illnesses are diagnosed and measured using, in large part, automated, objective bio-behavioural assessments. Indeed, greater emphasis is being placed on objective assessment by a number of funding agencies (2). However, the complexity of the central nervous system does not allow for such advances to be easy, and for this reason, objective measures of psychiatric symptoms have been slow to develop. Assessment of speech disturbances, important to many SMIs, reflects one obvious target for such technological advancement. The fact that communication can be assessed across a range of domains, such as over the internet, phone and “smart-phone”, expands its practical application well beyond the traditional boundaries of the therapy office and laboratory. Moreover, there is potential for combining data across centers and languages with consortium-level effort, much in the same way genetic research has been conducted. As can be seen from this review, progress in developing these technologies has been slow. Nevertheless, applying the technological and methodological advances from the fields of linguistic and acoustic sciences offers the potential to revolutionize the way we assess and understand mental illness.

Acknowledgments

Alex Cohen was supported, in part, by a National Institute of Mental Health (R03 MH092622) grant.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

* of special interest

** of outstanding interest

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed.. Arlington, VA: American Psychiatric Publishing; 2013.
2. Cuthbert BN, Insel TR. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC medicine*. 2013; 11:126. Epub 2013/05/16. [PubMed: 23672542]
3. Kring AM, Gur RE, Blanchard JJ, Horan WP, Reise SP. The Clinical Assessment Interview for Negative Symptoms (CAINS): final development and validation. *Am J Psychiatry*. 2013; 170(2): 165–172. Epub 2013/02/05. [PubMed: 23377637] * This article reports on a new interviewer-based measure of negative symptoms in schizophrenia. Included in this measure are ordinal based evaluations of vocal prosody and speech. This measure is comprehensive, empirically derived and applicable to research contexts.
4. Strauss GP, Keller WR, Buchanan RW, Gold JM, Fischer BA, McMahon RP, et al. Next-generation negative symptom assessment for clinical trials: validation of the Brief Negative Symptom Scale. *Schizophr Res*. 2012; 142(1–3):88–92. Epub 2012/11/07. [PubMed: 23127378] * This article reports on another new interviewer-based measure of negative symptoms in schizophrenia. Included in this measure are ordinal based evaluations of blunted vocal expression and speech production (i.e., alogia). This measure is relatively comprehensive, empirically derived and brief.
5. Cohen AS, Alpert M, Nienow TM, Dinzeo TJ, Docherty NM. Computerized measurement of negative symptoms in schizophrenia. *J Psychiatr Res*. 2008; 42(10):827–836. Epub 2007/10/09. [PubMed: 17920078]
6. Boersma P, Weenink D. Praat: doing phonetics by computer (Version 5.3.59). 2013
7. Elvevåg B, Foltz PF, Rosenstein M, DeLisi L. An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. *J Neurolinguistics*. 2010; 23:270–284. [PubMed: 20383310]
8. Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res*. 2007; 93(1–3):304–316. Epub 2007/04/17. [PubMed: 17433866]
9. Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*. 1997; 104:211–240.
10. Voorspoels W, Storms G, Longenecker J, Verheyen S, Weinberger DR, Elvevåg B. Deriving semantic structure from category fluency: clustering techniques and their pitfalls. *Cortex*. 2013 [Epub ahead of print].
11. Dumais ST. Latent semantic analysis. *Annual Review of Information Science and Technology*. 2005; 38:188–230.
12. Elvevåg B, Garrard P. Language, computers and cognitive neuroscience. *Cortex*. 2014 Special issue forthcoming. ** Although beyond the scope of this paper, automated computerized measures of language are also of enormous research use within neurology. This special issue of cortex features articles devoted to this topic. Articles focus on the development of an extremely large corpus (near 20 million words) for use with LSA to explore the association to cue words in Broca's aphasia. Articles reviewed in this special edition also focus on automated approaches that have been useful in understanding the nature of semantic dementia from transcripts, as well as in MRI by using machine learning approaches to diagnosis and detecting laterality effects in semantic dementia discourse. The focus of this special edition also includes the viability for using narrative speech transcripts to automatically classify primary progressive aphasia subtypes, predict those at risk for developing dementia, as well as those who do develop dementia.
13. Holshausen K, Harvey PD, Elvevåg B, Foltz P, Bowie CR. Latent semantic variables predict formal thought disorder and adaptive behavior in older inpatients with schizophrenia. *Cortex*. 2013 [Epub ahead of print]. * This study examined correlations between LSA-based measures of semantic coherence and measures of verbal fluency, functioning and clinically-rated speech disorder in 165 older inpatients with schizophrenia. The authors found that a computational measure of word unusualness differentially predicted disorganized speech, overall verbal fluency,

and adaptive functioning, and this was the case even when accounting for demographic, clinical, and pre-morbid cognition.

14. Rosenstein M, Diaz-Asper C, Foltz PW, Ellevåg B. A computational language approach to modeling prose recall in schizophrenia. *Cortex*. 2014 ** This study examined in detail two methods for the automatic scoring of prose recall (from the Wechsler Memory Scale) so as to assess the time-varying properties of recall at various retrieval intervals (immediate, 30 minutes and 24 hours later) in patients with schizophrenia, unaffected siblings and healthy unrelated control participants. The authors found that a syntax-based (n-gram text categorization) technique can reproduce the scores from the traditional (human) ratings, but that a more semantics-based (LSA) automated approach predicts group membership (i.e. schizophrenia vs. not) even more accurately than either the human scores or an automated approach based on n-grams. A model that combines the syntactic and semantic features improved accuracy of group membership prediction further.
15. Hoffman R, Grasemann U, Gueorguieva R, Quinlan D, Lane D, Mikkilainen R. Using computational patients to evaluate illness mechanisms in schizophrenia. *Biol Psychiatry*. 2010; 69:997–1005. [PubMed: 21397213]
16. Cabana A, Valle-Lisboa JC, Ellevåg B, Mizraji E. Detecting order-disorder transitions in discourse: implications for schizophrenia. *Schizophr Res*. 2011; 131(1–3):157–164. Epub 2011/06/07. [PubMed: 21640558]
17. Valle-Lisboa JC, Pomi A, Cabana A, Ellevåg B, Mizraji E. A modular approach to language production: Models and facts. *Cortex*. 2013 Epub 2013/03/23. ** This paper adopted a neurocomputational framework to explore possible connections between language and underlying substrates. The authors assume that many aspects of producing and comprehending language are reducible to dynamics within a modular network of neural associative memories. Based upon this assumption, they connect three theoretical and empirical domains, namely (i) models of language processing and production, (ii) statistical methods to construct functional brain images, and (iii) corpus linguistic tools (including LSA). The authors use these models to explore potential mechanisms that could be responsible for the alterations in discourse displayed by patients with schizophrenia.
18. Mota NB, Vasconcelos NA, Lemos N, Pieretti AC, Kinouchi O, Cecchi GA, et al. Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS One*. 2012; 7(4):e34928. Epub 2012/04/17. [PubMed: 22506057] * The authors quantified natural speech coherency from clinical interviews for eight patients with schizophrenia, eight with mania and eight controls. The authors employed a quantitative network model derived from graph theory. Classifiers based on these measures sorted the patient groups with 93.8% sensitivity and 93.7% specificity, compared to standard symptom rating scales with achieved only 62.5% sensitivity/specificity.
19. Arvidsson D, Sikström S, Werbart A. Changes in self and object representations following psychotherapy measured by a theory-free, computational, semantic space method. *Psychother Res*. 2011; 21(4):430–446. Epub 2011/05/31. [PubMed: 21623547]
20. Garrard P, Rentoumi V, Lambert C, Owen D. Linguistic biomarkers of Hubris syndrome. *Cortex*. 2013 Epub 2013/10/01. ** This paper reports on a new quantitative approach to characterize the linguistic patterns associated with the development of a personality disorder known as Hubris Syndrome. Previously it has been suggested that a number of prime ministers (including Margaret Thatcher and Tony Blair) may have developed this syndrome while in power. Hubris syndrome is often associated with an overestimation of one's own competence, accomplishments or capabilities. The paper focuses on the analyses of transcribed discourse samples from three British prime ministers, namely Margaret Thatcher, Tony Blair and John Major. The authors employed a variety of mathematical and computational methods and derived three main measures from the language samples, namely (i) entropy, (ii) the time path of ngrams (1, 2, and 3 word sequences) to index relative frequency changes over time, and (iii) word 'keyness' to reflect how specific words varied against a reference corpus. The authors argued for the development of Hubris Syndrome in Margaret Thatcher and Tony Blair as evident in the increase in the use of the 'royal we', the use of keywords indicating excessive confidence such as 'certain', 'sure', 'confident', 'succeed/success', and references to God or history. They also found that language became more complex and less predictable during hubristic periods. These linguistic patterns were not reflected in the language of John Major.

21. Tagamets MA, Cortes CR, Griego JA, Ellevåg B. Neural correlates of the relationship between discourse coherence and sensory monitoring in schizophrenia. *Cortex*. 2013 [Epub ahead of print]. ** This study examined fMRI correlates of semantic coherence in 11 patients with schizophrenia and 11 healthy controls. Semantic coherence (calculated using computational semantic analysis which is a variant of LSA) was used to evaluate neural correlates of coherent speech and sensory monitoring. While undergoing fMRI, patients and controls had to decide whether a presented spoken or written word was the same as one previously presented immediately beforehand. There were three conditions of the task, namely (i) with the spoken and written words matching, (ii) the stimuli comprising homophones (e.g., sale/sail), and (iii) and homographs (e.g., bow, pronounced in two different ways). Correlations were examined between task activations (versus a baseline condition) and semantic coherence scores in each of the three conditions in the patient and control group. No significant correlations were seen in the controls, but there were significant correlations in the patient group in all three task conditions; in particular, a cluster in the left middle/superior temporal cortex. The authors take these findings to suggest that computational models of coherence offer a useful method to probe the neurophysiological underpinnings of thought disorder and how these are related to sensory processing.
22. Nicodemus KK, Ellevåg B, Foltz PW, Rosenstein M, Diaz-Asper C, Weinberger DR. Category fluency, latent semantic analysis and schizophrenia: a candidate gene approach. *Cortex*. 2014 ** This study provides a novel attempt to delineate the underlying genomic architecture of the widely used category fluency task using phenotypes obtained via computational linguistics. The category fluency test requires the naming of as many words belonging to a certain category (e.g., animals) as possible within a short period of time. In this paper, in addition to the standard metric of overall word count, the authors applied LSA to analyze the clustering pattern of the categories generated. Then they employed a candidate gene approach focusing on single-nucleotide polymorphisms (SNPs) with known function and available from a recent genome-wide association study of schizophrenia. The authors selected candidate genes that have been associated with language and speech, episodic memory, and processing speed, and this resulted in 39 coding SNPs in a sample of 665 participants which included patients with schizophrenia, their unaffected siblings and unrelated healthy controls. The authors argue that the findings, albeit they preliminary, warrant larger studies such as genome-wide association studies of similar phenotypes, so as to provide the needed statistical power to chart how common variation contributes to these novel phenotypes.
23. Lyalina S, Percha B, Lependu P, Iyer SV, Altman RB, Shah NH. Identifying phenotypic signatures of neuropsychiatric disorders from electronic medical records. *J Am Med Inform Assoc*. 2013; 20(e2):e297–e305. Epub 2013/08/21. [PubMed: 23956017] * Medical records were examined for over 7,000 patients with autism, bipolar and schizophrenia disorders to determine the validity of a novel automated text-processing program useful for electronic medical record mining. Results suggest that the novel software successfully distinguished between patient groups (e.g., autism and the other disorders). Data mining of electronic medical records using automated procedures appears to be a promising method for understanding mental illnesses.
24. Diehl JJ, Paul R. Acoustic Differences In The Imitation Of Prosodic Patterns In Children With Autism Spectrum Disorders. *Research in autism spectrum disorders*. 2012; 6(1):123–134. Epub 2011/11/30. [PubMed: 22125576]
25. Hupp JM, Jungers MK. Beyond words: Comprehension and production of pragmatic prosody in adults and children. *Journal of experimental child psychology*. 2013; 115(3):536–551. Epub 2013/02/21. [PubMed: 23422927]
26. 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. [PubMed: 15380873]
27. Cohen AS, Najolia GM, Kim Y, Dinzeo TJ. On the boundaries of blunt affect/alogia across severe mental illness: implications for Research Domain Criteria. *Schizophr Res*. 2012; 140(1–3):41–45. Epub 2012/07/27. [PubMed: 22831770] ** This study employed computerized acoustic analysis of natural speech produced in reaction to a range of visual stimuli in 48 stable outpatients with schizophrenia and mood disorders. Speaking assessments were administered 1 week-apart to examine how temporal stability might vary as a function of clinical diagnosis and symptom severity. Speech characteristics generally did not differ between groups and were similarly, and for the most part, highly stable over time. Aspects of speech were significantly associated with severity of psychosis and negative symptoms, but not with clinical depression/anxiety severity.

Moreover, stability of speech characteristics generally did not vary as a function of diagnostic group or clinical severity. Speech production was associated with social functioning deficits. These data suggest that speech variables tap a stable and clinically important facet of psychopathology that cut across diagnostic categories. The results of this paper support the use of automated computerized analysis for three practical reasons: First, given the vast number of SMI patients there are not enough health professionals available to monitor patients as frequently as necessary. Second, psychiatric patients frequently require longitudinal monitoring on the scale of years which is cognitively challenging for clinicians since patients have different clinical baselines against which they need to be compared and human memory is simply not accurate enough for such a gigantic task. Third, often patients have met with clinicians a few days prior to a relapse or suicide, thus highlighting how enormously difficult the clinical evaluation is. Combined with telemedicine/ e-health applications could lay the foundation for a longitudinal 'personalized medicine' approach such that each patient becomes their own baseline, and results from other patients with 'similar sorts of illnesses' are used as guides.

28. Yoder PJ, Oller DK, Richards JA, Gray S, Gilkerson J. Stability and validity of an automated measure of vocal development from day-long samples in children with and without autism spectrum disorder. *Autism research: official journal of the International Society for Autism Research*. 2013; 6(2):103–107. Epub 2013/02/26. [PubMed: 23436778] ** Acoustic analysis was employed to understand spoken language in children with autism spectrum disorder (ASD). A unique feature of this study was the use of day-long vocal samples, as opposed to briefer speech samples typically employed in other studies. Results suggest that acoustic analysis variables provide a stable estimate for a measure of vocal development that is highly related to expressive spoken language in a group of young children with ASD and in a group that is typically developing.
29. Covington MA, Lunden SL, Cristofaro SL, Wan CR, Bailey CT, Broussard B, et al. Phonetic measures of reduced tongue movement correlate with negative symptom severity in hospitalized patients with first-episode schizophrenia-spectrum disorders. *Schizophr Res*. 2012; 142(1–3):93–95. Epub 2012/10/30. [PubMed: 23102940] ** Speech quality and variability were examined from natural speech procured from a semi-structured clinical research interview in 25 first-episode psychosis patients. Formant values (F0, F1, F2) were analyzed. Variability in these indices was correlated with severity of interview rated negative and positive symptoms. Of particular note, a measure of variability of tongue front-to-back position (F2) was significantly correlated with the severity of interview-based negative symptoms. This is the only study, to our knowledge, to investigate variability in formant values beyond F0 in patients with SMI.
30. Duffy C, Healy O. Spontaneous communication in autism spectrum disorder: A review of topographies and interventions. *Research in autism spectrum disorders*. 2011; 5(3):977–983.
31. Nadig A, Shaw H. Acoustic and perceptual measurement of expressive prosody in high-functioning autism: increased pitch range and what it means to listeners. *J Autism Dev Disord*. 2012; 42(4):499–511. Epub 2011/04/30. [PubMed: 21528425]
32. DePape AM, Hall GB, Tillmann B, Trainor LJ. Auditory processing in high-functioning adolescents with Autism Spectrum Disorder. *PLoS One*. 2012; 7(9):e44084. Epub 2012/09/18. [PubMed: 22984462]
33. Dickey CC, Vu MA, Voglmaier MM, Niznikiewicz MA, McCarley RW, Panych LP. Prosodic abnormalities in schizotypal personality disorder. *Schizophr Res*. 2012; 142(1–3):20–30. Epub 2012/10/17. [PubMed: 23068317] ** This study employed acoustic analysis of brief speech samples in twenty-eight antipsychotic-naïve volunteers with schizotypal personality disorder and 27 healthy comparison participants. Included in the assessment were magnetic resonance imaging based measures of relevant brain structures. Participants read aloud short sentences and responded to probes delivered by an interviewer. Acoustic variables included both objective measures (i.e., pause proportion, duration, and pitch variability) as well as interview-based measures (e.g., degree of emotion portrayed, and how much they wanted to hear more from the subjects). SPD participants' speech showed significantly more pauses, was slower and had less pitch variability than controls. Moreover, smaller left pars opercularis volumes in SPD participants was correlated with more interview-rated pauses and less emotion. SPD participants displayed statistically significant prosodic deficits that were measurable in speech samples, though these were at a small effect size (as a group).

34. Cohen AS, Morrison SC, Brown LA, Minor KS. Towards a cognitive resource limitations model of diminished expression in schizotypy. *J Abnorm Psychol.* 2012; 121(1):109–118. Epub 2011/05/11. [PubMed: 21553936] ** The authors evaluated a theory that constricted affect in individuals at putative risk for schizophrenia reflects a cognitive resource issue, the notion being that as cognitive resources are depleted, expression becomes diminished in individuals with schizotypy. Acoustic analysis of natural speech was procured during experimentally manipulated baseline and high cognitive-load dual tasks and examined in 38 individuals with psychometrically defined schizotypy and 34 controls. For nearly all participants in the study, expression significantly decreased as a function of increased task demands. Interestingly, there were no group differences in expression or magnitude of change across baseline to high cognitive-load conditions. A subset of schizotypal participants, those with self-reported constricted affect and those with poor performance on the cognitive-load task, showed significant reductions in expression under high-load versus baseline speaking conditions relative to other schizotypal and control participants. These findings suggest that diminished expression occurs as a function of limited cognitive resources in schizotypy.
35. Cohen AS, Kim Y, Najolia GM. Psychiatric symptom versus neurocognitive correlates of diminished expressivity in schizophrenia and mood disorders. *Schizophr Res.* 2013; 146(1–3): 249–253. Epub 2013/03/14. [PubMed: 23481582] ** Computerized acoustic analysis of natural speech was employed to understand the neurocognitive correlates of blunted vocal affect and alogia in patients with schizophrenia and mood disorders. Specifically, the degree to which speech characteristics tapping alogia (i.e., average pause duration) and blunted affect (i.e., prosody computed from fundamental frequency and intensity) was evaluated in its association with psychiatric symptoms and neurocognitive deficits. Subjects included 26 volunteers with schizophrenia and 22 subjects with mood disorders. For both the schizophrenia and mood disorder groups, neurocognitive functioning was significantly correlated with increased pause time (large effect size levels) and, for the schizophrenia group only, reduced prosody (large effect size level). The pattern of neurocognitive associates was more diffuse for patients with mood disorder versus schizophrenia, though still in the medium to large effect size range. Psychiatric symptoms were not significantly associated with speech characteristics for either group (generally, negligible effect sizes). These results suggest that there is a link between expressivity and neurocognitive dysfunctions for both patients with schizophrenia and mood disorders.
36. Cohen AS, Iglesias B, Minor KS. The neurocognitive underpinnings of diminished expressivity in schizotypy: what the voice reveals. *Schizophr Res.* 2009; 109(1–3):38–45. Epub 2009/02/24. [PubMed: 19230620]
37. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry.* 2012; 72(7):580–587. Epub 2012/05/01. [PubMed: 22541039] ** Speech production was evaluated using computerized acoustic analysis of speech in 105 with major depression. All participants were recruited into a 4-week, randomized, double-blind, placebo-controlled research methodology study and natural speech samples were collected at baseline and study conclusion. Speech production changes associated with clinical outcomes in a prior study were found to be reliable and significant predictors of outcome. Depressed patients who did not clinically improve showed smaller vocal acoustic changes compared to treatment responders.

Key Points

- Evidence suggests that automated computerized assessments of natural speech can be reliable and valid for understanding important aspects of severe mental illness, and can overcome many practical limitations of more traditional assessment methods.
- Automated computerized assessments of natural speech are beginning to be applied to understand the neurobiological and genetic underpinnings of severe mental illness. This reflects a particularly promising application of this technology.
- Automated computerized assessments of natural speech are also being applied as an index of clinical functioning, and reflect a particularly promising outcome variable for pharmacological and psychosocial treatment studies.
- Despite the promise of automated computerized assessments of natural speech, much work remains on refining these technologies, particularly so they can be applied in clinical settings.

Table 1

Issues with Rating-Based Measures of Speech Disturbances and How Computer Measures Address These Issues

Issue with Symptom Rating Scales	How Computer Measures Address These Issues
Operational definitions are vague (e.g., mild, moderate)	Human behaviour is directly quantified
Can be biased to ethnicity and culture	Uninfluenced by individual difference factors
Poor at isolating specific facets of communication	Isolates specific aspects of behaviour
Insensitive to change: often covers weeks to months	User defined sensitivity, typically in milliseconds
Inter-rater reliability estimates can be far from perfect	Near perfect reliability provided standard parameters
Data are ordinal and inappropriate for parametric statistics	Data are ratio and continuous
Requires a clinical interview	Can be conducted over the phone or in the field
Population distribution has a strong “right” skew	Data are normally distributed
Does not quantify “normal” behaviour	Can compare patient behaviour to established norms
Interviews and ratings can be time consuming	Automated procedure processes behaviour rapidly
Requires extensive training	Requires basic computer experience