



Published in final edited form as:

*J Psychiatr Res.* 2023 January ; 157: 43–49. doi:10.1016/j.jpsychires.2022.11.013.

## Mild Vs. Moderate: How Behavioral Speech Measures Predict Metacognitive Capacity Across Different Levels of Formal Thought Disorder

Evan J. Myers<sup>a,\*</sup>, Danielle B. Abel<sup>a</sup>, Kathryn L. Hardin<sup>a</sup>, Robert J. Bettis<sup>a</sup>, Ashlynn M. Beard<sup>a</sup>, Michelle P. Salyers<sup>a</sup>, Paul H. Lysaker<sup>b,c</sup>, Kyle S. Minor<sup>a</sup>

<sup>a</sup>Department of Psychology, Indiana University – Purdue University Indianapolis, Indianapolis, IN, United States

<sup>b</sup>Richard L. Roudebush VA Medical Center, Department of Psychiatry, Indianapolis, IN, United States

<sup>c</sup>Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, United States

### Abstract

Disorganized speech is a key component of formal thought disorder (FTD) in schizophrenia. Recent work has tied disorganized speech to deficits in metacognition, or one's ability to integrate experiences to form complex mental representations. The level of FTD at which differences in metacognitive capacity emerge remains unclear. Across two studies, using different cut scores to form FTD groups, we aimed to 1) explore the relationship between disorganized speech and metacognition and 2) compare trained rater and automated analysis methods. Clinical interviews were coded for disorganized speech and metacognition using the Communication Disturbances Index (CDI), Coh-Metrix multidimensional indices, and Metacognition Assessment Scale. In Study 1, we examined CDI and Coh-Metrix's ability to predict metacognition in FTD ( $n = 16$ ) and non-FTD ( $n = 29$ ) groups. We hypothesized the FTD group would have lower metacognition and that both CDI and Coh-Metrix would account for significant variance in metacognition. In Study 2, we conducted the same analyses with an independent sample using more stringent FTD cut scores (FTD:  $n = 23$ ; non-FTD:  $n = 23$ ). Analyses indicated that at a moderate but not mild cutoff: 1) automated methods differentiated FTD and non-FTD groups, 2) differences in metacognition emerged, and 3) behavioral measures accounted for significant variance (34%) in metacognition. Results emphasize the importance of setting the FTD cutoff at a moderate level and using samples

\*Corresponding author: Evan J. Myers, IUPUI School of Science, 402 N. Blackford Street, LD 124, Indianapolis, IN 46202, USA, evjmyers@iupui.edu.

**Evan J. Myers** conducted the literature search, coded study data, performed data analysis, and wrote or co-wrote all drafts of the manuscript. **Danielle B. Abel**, **Kathryn L. Hardin**, **Robert J. Bettis**, and **Ashlynn M. Beard** coded study data and co-wrote all drafts of the manuscript. **Michelle P. Salyers** co-wrote all drafts of the manuscript. **Paul H. Lysaker** was the Principal Investigator of one of the parent studies and co-wrote all drafts of the manuscript. **Kyle S. Minor** was a Principal Investigator of one of the parent studies, co-designed the study, played an integral role in data interpretation, and co-wrote all drafts of the manuscript.

Declarations of interest: none

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

that contain high levels of FTD. Findings extend research linking FTD and metacognition and demonstrate the benefit of pairing trained rater and automated speech measures.

## Keywords

Schizophrenia; Formal thought disorder; Disorganized speech; Metacognition

---

## 1. Introduction

Effective communication is disrupted in schizophrenia by disorganized speech, a symptom that has been linked to poor functional outcomes such as social disengagement, impaired friendships, and poor work performance (Bowie & Harvey, 2008; Racenstein et al., 1999). Metacognition, or the processes by which a person forms an integrated sense of self and others from their experiences, has substantial overlap with disorganized speech (Lysaker et al., 2020). At a fundamental level, people with schizophrenia experience a fragmentation of thought and language that makes connecting ideas difficult (Bleuler, 1911). This inability to integrate information may underlie disturbances in metacognition (i.e., the ability to form complex mental representations of self and others) and communication (i.e., the ability to form cohesive and coherent speech).

An important question concerning disorganized speech is the threshold at which a person is considered to have a formal thought disorder (FTD). Studies linking FTD to deficits in cognitive functioning have found differences between FTD and non-FTD groups when using a cutoff of a moderate level of FTD (Barrera et al., 2005; Dwyer et al., 2019; Mackinley et al., 2020). Recent research on metacognition suggests that connections between neurocognition and metacognition break apart at minimal to mild levels of FTD (Minor et al., 2015), hinting that mild levels of FTD may have an important relationship with metacognition. To further examine this relationship, we conducted two independent studies using different levels of FTD (mild, moderate) to form groups.

FTD is a multifaceted construct, and different measurement approaches capture unique aspects (e.g., semantic coherence, communication failure). Research has shown that combining measures can help account for greater variance in cognitive domains (Minor et al., 2019). Examining measures in combination but also evaluating the individual contributions of each method adds to our understanding of the utility of different types of measures. The present studies examined the relationship between trained rater and automated analysis approaches for identifying FTD, and whether they could account for greater variance in metacognition when integrated.

Trained rater approaches have shown great utility in evaluating speech samples. Raters count discrete instances of disorganization and calculate a ratio of disorganization per standard unit of speech. Studies using this approach have linked disorganized speech to both neurocognitive (Tan & Rossell, 2019; Bora et al., 2019) and social cognitive deficits (Docherty et al., 2013); however, no studies to date have examined metacognitive deficits. Automated analysis (Elvevåg et al., 2007; Tausczik & Pennebaker, 2010) has been used to evaluate speech samples and have shown great promise for predicting conversion to

psychosis in high-risk samples (Bedi et al., 2015; Corcoran et al., 2018) and predicting schizophrenia diagnosis beyond clinician-rated scales (Mota et al., 2012). A primary advantage of these tools is their efficiency. Following speech transcription, automated analysis of speech occurs rapidly, whereas trained rater approaches require hand coding of transcripts by multiple trained personnel.

One automated analytical tool, Coh-Metrix, has been used to show how linguistic cohesion differs between people with schizophrenia and controls (Willits et al., 2018) and has directly linked disorganized speech to deficits in metacognition (Minor et al., 2019; Lundin et al., 2020). To further examine its clinical application as an assessment instrument, this study will compare Coh-Metrix indices with a well-validated behavior-based measure of disorganized speech and determine whether it can distinguish between FTD and non-FTD groups.

Across separate studies with discrete participants, we tested two primary aims. First, we compared metacognitive capacity in people with and without FTD using mild (Study 1) and moderate (Study 2) cutoffs. We expected people with FTD would exhibit poorer metacognition. Second, we tested the extent to which automated analysis and trained rater measures accounted for variance in metacognition. We expected both measures would account for significant variance. This study adds to a growing literature examining the relationship between disorganized speech and metacognition. It may provide more evidence for using time-efficient automated measures—either by themselves or in tandem with trained rater approaches. The study could also provide preliminary evidence that there may be a differential relationship between disorganized speech and metacognition based upon the presence or relative absence of FTD, which could have important implications for treatment of FTD.

## Study 1: Mild FTD as Cutoff

### 2. Material and Methods

**2.1 Participants**—Study 1 involves secondary data analysis from two completed studies; baseline data from a longitudinal study and a cross-sectional study (see Abel & Minor, 2021; Minor et al., 2022). Participants were adult outpatients primarily recruited from community mental health center and a midwestern VA Medical Center. Inclusion criteria were: a) primary diagnosis of schizophrenia, schizoaffective disorder, or psychotic disorder NOS; b) age 18–60; c) English fluency; d) no change in medication or outpatient status in the 30 days before testing; e) ability to give informed consent; f) no active substance dependence; g) no documented intellectual disability; and h) no history of a neurological illness or traumatic brain injury that resulted in loss of consciousness greater than five minutes. Participants were only included if they completed the Indiana Psychiatric Illness Interview ( $n$  excluded = 6) (see Measures; Lysaker et al., 2002). All study procedures for Studies 1 and 2 were approved by the local institutional review board and carried out in accordance with the latest version of the Declaration of Helsinki. All participants in both studies provided written consent prior to participation.

Participants were divided based on their level of conceptual disorganization, an item on the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). Conceptual disorganization is a clinician-rated item that measures cognitive-verbal processes such as loose associations and circumstantiality on a seven-point Likert scale. Those who were rated < 3 (“Absent” or “Minimal”) formed the non-FTD group ( $n = 29$ ); those  $\geq 3$  (“Mild” or greater) formed the FTD group ( $n = 16$ ). In prior studies, staff achieved good reliability on the PANSS (intraclass correlation = .80).

## 2.2 Measures

**2.2.1 Indiana Psychiatric Illness Interview (IPII; Lysaker et al., 2002):** The IPII is a semi-structured clinical interview that served as the basis for rating metacognitive capacity and disorganized speech. In contrast to most interviews, participants spontaneously generate content with few interviewer prompts. Interviewer prompts are open-ended, allowing participants to speak for as long as they like. Interviews typically last 30–60 minutes and were administered by trained graduate students. Interviews were audio-recorded and later transcribed. For most participants, only speech generated in the first section of the interview, which focuses on the participants’ life story, was analyzed for disorganized speech. However, for participants who generated less than 400 words ( $n = 10$ ), the second part of the IPII (focused on the participants’ illness narrative) was also analyzed. Interviewer speech was excluded from analyses.

**2.2.2 Communication Disturbances Index (CDI; Docherty et al., 1996):** The CDI is a behavior-based measure that identifies specific instances of disorganized speech. Instances are rated based on a loss or ambiguity of meaning across six categories (see Supplementary Material). To calculate a total score, instances are summed, divided by adjusted word count, and multiplied by 100. In calculating adjusted word count, each disturbance is counted as a single word regardless of length. The score is generated as a ratio which accounts for differences in the amount of speech generated by participants. The CDI has shown strong evidence of interrater reliability (Docherty et al., 2003) and construct validity, showing significant correlations with other measures of FTD (Docherty et al., 1997; Docherty et al., 2012). Study authors (E.J.M., D.B.A., K.L.H.) and two undergraduate students rated IPII transcripts independently for CDI, blind to PANSS conceptual disorganization rating, and then met to finalize consensus ratings. Good inter-rater reliability was achieved for 30 randomly selected narratives analyzed prior to consensus meetings, intraclass correlation = .90.

**2.2.3 Coh-Metrix (McNamara et al., 2014):** Coh-Metrix 3.0 is a software program that analyzes texts in terms of coherence and cohesion. Coh-Metrix computes over 100 language indices ranging from basic to complex indices that rate cohesion across words, sentences, and paragraphs. This project examined five complex indices: narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion (see Table 1 for descriptions). These indices were chosen based upon a principal components analysis which showed that these five indices accounted for 54% of the variability in a corpus of language arts, social studies, and science texts (Graesser et al., 2011). Deep cohesion and referential cohesion have previously been linked to metacognition (Lundin et al., 2020; Minor et al., 2019). In

addition to these indices, a Coh-Metrix composite score was computed by averaging the standardized scores of the five indices. All indices were coded such that positive scores indicate more organized speech, and negative scores indicate more disorganized speech.

**2.2.4 Metacognition Assessment Scale - Abbreviated (MAS-A; Lysaker et al., 2005).** The MAS-A consists of four subscales (self-reflectivity, understanding the mind of the other, decentration, and mastery) that measure participants' understanding of their own and others' mental states and apply that understanding to cope with problems or challenges in their life. Total scores range from 0 to 28, with higher scores indicating increased metacognitive capacity. Transcripts were scored by raters who had demonstrated good interrater reliability (intraclass correlation = .80). No rater was also involved in CDI ratings. Interrater reliability for the total score has consistently been in the good to excellent range within schizophrenia samples (Lysaker et al., 2013; Lysaker et al., 2014).

**2.3 Analyses**—Analyses were conducted using IBM SPSS Statistics 26 and R version 4.0.3. All analyses used the significance threshold of  $p < 0.05$ . First, FTD and non-FTD groups were compared on possible confounding variables. Second, as a construct validity check, groups were compared on CDI and Coh-Metrix composite scores using independent samples t-tests. Third, an independent samples t-test was conducted comparing FTD and non-FTD groups on metacognitive capacity. Finally, two stepwise regressions predicting metacognitive capacity were conducted across the entire sample; one with CDI and Coh-Metrix composite scores as predictors, and the other using CDI total score and Coh-Metrix indices that significantly correlated with metacognition.

### 3. Results

**3.1 Demographic Characteristics**—Participants were mostly middle-aged ( $M = 46.9$ ,  $SD = 10.1$ ), male (58%), Black (56%), non-Hispanic (87%), and had at least a high school education or GED (73%). FTD ( $n = 16$ ) and non-FTD groups ( $n = 29$ ) did not significantly differ on demographic variables, length of studied speech segment, or general symptom severity (see Table 2).

**3.2 Grouping Validity Check**—On disorganized speech measures, groups significantly differed on the CDI, with the FTD group showing greater speech disorganization (Table 2). Groups did not significantly differ on Coh-Metrix. Thus, only the CDI was able to adequately distinguish the groups based on disorganized speech. In Study 1, CDI and Coh-Metrix composite scores were minimally correlated ( $r = .04$ ).

**3.3 Aim 1: Comparing Metacognitive Capacity in FTD and Non-FTD Groups**—FTD and non-FTD groups did not significantly differ in metacognitive capacity or in metacognitive domains (Table 2). Results do not support the hypothesis that those with FTD exhibit greater metacognitive deficits than those without FTD when a mild cutoff is used to distinguish groups.

**3.4 Aim 2: Predicting Metacognition Using Automated Analysis and Trained Rater Measures**—To compare disorganized speech measures, stepwise regression models

were conducted across the entire sample (Table 3). In two different models, CDI and Coh-Metrix composite scores were entered individually as a predictor in step one to examine their relationship with metacognition in isolation. Composites were entered together at step two. Neither individual model was significant at step one, CDI:  $F(1, 43) = .54, p = .47$ ; Coh-Metrix:  $F(1, 43) = 1.12, p = .30$ . At step two, the overall model was not significant,  $F(2, 42) = .75, p = .48$  (Adj.  $R^2 = -.01$ ).

Post-hoc correlations were performed between Coh-Metrix indices that comprised the composite score and metacognition (Table 4). Syntactic simplicity ( $r = -.38$ ), narrativity ( $r = .37$ ), and deep cohesion ( $r = .33$ ) showed statistically significant medium size correlations with metacognition without correction for multiple comparisons. A stepwise regression was conducted with these three indices entered at step one and CDI total score added at step two. The model was significant at step one,  $F(3, 41) = 3.44, p = .03$ . At step two, the overall model reached the level of significance,  $F(4, 40) = 2.77, p = .04$  (Adj.  $R^2 = .14$ ). However, no individual predictors were significantly related to metacognition in either model. Taken together, these results do not support the hypothesis that the CDI would outperform Coh-Metrix in predicting metacognition. Neither instrument was a significant predictor.

**3.5 Study 1 Discussion**—In Study 1, three findings of interest were observed. First, groups did not differ in metacognition. Second, neither speech measure accounted for significant variance in metacognition. Finally, FTD and non-FTD groups differed on only one of the two disorganized speech measures. Only the trained rater measure differentiated groups, which may reflect its ability to detect subtle instances of disorganized speech (Docherty et al., 1996; Kerns & Becker, 2008). This may be particularly relevant for the observed sample, which generally had minimal to mild levels of disorganization. This finding also suggests that setting a higher FTD threshold may be necessary to observe group differences.

## Study 2: Moderate FTD as Cutoff

### 4. Material and Methods

**4.1 Participants**—Study 2 involved secondary analysis of a subset of data from a cross-sectional study (see Lysaker et al., 2013) and used a higher threshold for FTD. Participants who were rated  $\geq 4$  (“Moderate” or greater) formed the FTD group ( $n = 23$ ). Those rated  $< 4$  formed the non-FTD group ( $n = 23$ ) and were selected to demographically match the FTD group on age, race, sex, and education. Participants were recruited from a midwestern VA Medical Center following the same eligibility criteria but were distinct from Study 1 participants.

**4.2 Measures**—To approximate the methods used in Study 1, we again analyzed IPII transcripts using the CDI, Coh-Metrix, and MAS-A. CDI ratings were conducted by study authors (E.J.M., R.J.B., A.M.B.) and two additional undergraduate students.

**4.3 Analyses**—Study 1 analyses were replicated with this independent sample.

## 5. Results

**5.1 Demographic Characteristics**—Participants were mostly middle-aged ( $M = 46.6$ ,  $SD = 9.3$ ), male (100%), nearly equally Black (52%) and White (48%), and most had at least a high school education or GED (85%). FTD and non-FTD groups did not significantly differ on speech segment length, but did differ on general symptom severity, with the FTD group showing greater symptoms (see Table 2).

**5.2 Grouping Validity Check**—Groups significantly differed on CDI and Coh-Metrix composite scores, with the FTD group showing greater disorganized speech and less linguistic cohesion (Table 2). This suggests that both instruments differentiated groups when moderate threshold was used for FTD. In Study 2, CDI and Coh-Metrix composite scores were significantly inversely correlated,  $r(44) = -.33$ ,  $p = .025$ .

**5.3 Aim 1: Comparing Metacognitive Capacity in FTD and Non-FTD Groups**—Overall metacognitive capacity did not significantly differ between FTD and non-FTD groups. However, self-reflectivity was significantly lower in the FTD group ( $t(44) = 2.03$ ,  $p = .024$ ). No other domains significantly differed. These results partially support the hypothesis that those with FTD would show greater metacognitive impairment.

**5.4 Aim 2: Predicting Metacognition from Automated Analysis & Trained Rater Measures**—To compare disorganized speech measures, similar stepwise regression models were run as in Study 1 (Table 5). The CDI model was significant at step one,  $F(1, 44) = 12.24$ ,  $p = .01$ , with the CDI total score accounting for 20% of the variance in metacognition. The Coh-Metrix model was not significant at step one,  $F(1, 44) = .51$ ,  $p = .48$ . At step two, the overall model reached the level of significance,  $F(2, 43) = 6.07$ ,  $p = .01$  (Adj.  $R^2 = .18$ ), with the CDI total score being the only significant individual predictor of metacognition.

Post-hoc correlations were performed between Coh-Metrix indices and metacognition (Table 4). In this study, only syntactic simplicity ( $r = -.42$ ) and referential cohesion ( $r = .29$ ) showed significant medium size correlations ( $p < .05$ ) without correction for multiple comparisons. A stepwise regression was conducted, entering these indices at step one and adding CDI total at step two. The model was significant at step one,  $F(2, 43) = 6.04$ ,  $p = .01$  (Adj.  $R^2 = .18$ ). At step two, the overall model reached the level of significance,  $F(3, 42) = 8.78$ ,  $p = .04$  (Adj.  $R^2 = .34$ ). Both syntactic simplicity and CDI total score were significant individual predictors of metacognition ( $p < .01$ ). Taken together, these results support the hypothesis that the CDI would outperform Coh-Metrix in predicting metacognition. In addition, the speech measures had minimal overlap and accounted for unique variance in metacognition.

**5.5 Study 2 Discussion**—In Study 2, three key findings emerged. First, both disorganized speech measures helped differentiate FTD groups. This provides additional evidence of the clinical utility of automated analysis to detect moderate FTD. Second, although our small sample was unable to detect statistically significant differences, there was a medium effect size group difference in overall metacognition. Moreover, the FTD group

was significantly lower in self-reflectivity. Third, behavioral speech measures accounted for significant unique variance in metacognition. Disorganized and discohesive speech were inversely related to metacognition, replicating and extending prior findings (Lundin et al., 2020).

## 6. General Discussion

Our studies built on prior research identifying links between disorganized speech and metacognition. Based on past studies, the threshold of FTD at which we might expect differences in metacognition to emerge was unclear; our findings suggest that differences are only noticeable at moderate or greater levels of FTD. There were three noteworthy findings when FTD was set at a moderate versus mild threshold. First, at moderate but not mild levels, automated methods differentiated FTD and non-FTD groups. Second, self-reflectivity was lower in the FTD group and differences in overall metacognitive capacity began emerging at a moderate threshold. Third, at a moderate threshold, trained rater and automated methods accounted for 34% of the variance in metacognition. At a mild threshold, they only accounted for 14%. Therefore, studies examining disorganized speech should include sufficiently high levels of FTD and use a more conservative cut score (i.e., moderate) to form groups.

Both trained rater and automated speech measures may have utility in differentiating FTD. Groups differed on automated methods only when the threshold for FTD was set at moderate levels (Study 2), while the trained rater method differentiated groups in both studies, showing raters are sensitive even at low levels of disorganized speech. The automated analysis tool was unable to differentiate groups at lower levels of FTD, signaling it may only measure gross disturbances in disorganized speech and may not capture more subtle signs that human raters can identify. It is important to note that the CDI was designed to measure FTD whereas Coh-Metrix was developed to assess cohesion—which is similar but distinct from FTD. However, Coh-Metrix has the advantage of being more efficient, which is especially important in clinical settings. One explanation for the discrepant performance of these measures is that the trained rater measure may be more similar to the clinician-rated instrument used to differentiate FTD groups in this study: both measures rate disorganized speech according to the “naked eye.” Further research is needed to clarify which automated indices are most relevant to the study of disorganized speech and to better understand the utility and limitations of automated analytical tools in measuring speech within psychiatric samples. Understanding the limitations of automated tools can help in translating their use to complement a traditional psychiatric evaluation. Research utilizing other promising automated tools such as speech graphs, latent semantic analysis, and acoustic analysis is also needed (Mota et al., 2017; Holshausen et al., 2014; de Boer et al., 2021).

At a moderate threshold, people with FTD showed reduced metacognitive self-reflectivity. This suggests that people with greater FTD have more difficulty thinking about their own mental states. For example, distinguishing between different cognitive operations (e.g., remembering and deciding) and between different emotions (e.g., resentment, frustration) may be especially challenging. Building these capacities could be a particularly important treatment target for people with significant FTD. Although not statistically significant, the

FTD group also showed reduced overall metacognitive capacity ( $d = 0.46$ ), indicating that people with FTD may have increased difficulty utilizing metacognitive knowledge to cope with psychological problems. Previous work (Minor et al., 2015) suggested that differences in metacognition may emerge at mild levels of FTD, but our findings show that a greater than moderate level may be needed to detect differences. This finding is generally in line with prior work showing an inverse relationship between disorganized symptoms and metacognition (Minor & Lysaker, 2014).

Integrating trained rater and automated approaches helped account for greater variance in overall metacognition, but only when the threshold of FTD was moderate (Study 2). Together, they explained about a third of the variance. There was little overlap in variance explained, indicating that both speech content (e.g., referential errors) and structure (e.g., syntactic complexity) may be independently relevant to metacognition, and that automated speech methods hold promise for gaining further insight into FTD and cognitive processes. The significant time burden of trained rater approaches makes it impractical clinically. However, efforts to reduce rating time or automate ratings of communication disturbances could make it useful to implement in tandem with other automated approaches to account for greater variance and characterize deficits.

We expected the trained rater measure and linguistic cohesion to account for significant variance in metacognition, consistent with past work using clinician-rated scales (Minor et al., 2019). The trained rater measure codes instances of speech with a loss of meaning due to vagueness, ambiguity, or insufficient context. In the context of life narratives, people who have difficulty forming complex understandings of themselves and others (i.e., those with lower metacognition) would be expected to have difficulty conveying their life story and inner experiences (i.e., thoughts and emotions) with requisite detail. Similarly, because people who show low metacognitive capacity have difficulty distinguishing between different cognitive operations (e.g., remembering and deciding) and between emotions, they would be expected to use simpler language structures and have fewer logical links between ideas, resulting in lower overall cohesion. Consistent with this conceptual link, referential cohesion has shown relationships with metacognition (Minor et al., 2019). Similarly, Lundin and colleagues (2020) found that deep cohesion plays an important role in linking different aspects of cognition, mediating relationships between executive functioning and metacognition in schizophrenia. Our finding that syntactic simplicity is inversely related to metacognition extends this work, demonstrating that those with greater metacognitive deficits use simpler syntactic structures. Syntactic simplicity was also significantly correlated with metacognition in an independent sample in Study 1 so this relationship may exist across levels of disorganization.

Strengths of this study include its two-study design examining constructs at different FTD thresholds and use of multiple behavioral measures of FTD. A primary study limitation, however, is that neurocognition was not assessed. Neurocognition is impaired in schizophrenia and has been tied to both disorganized symptoms and metacognition (Lysaker et al., 2005; Ventura et al., 2010). Thus, a neurocognitive “third variable” (e.g., impaired working memory) may underly these variables and be responsible for the associations between them. A second limitation is that the participants in Study 2 had greater

overall symptoms, which might contribute to the significant findings observed. Another limitation is the small sample size across groups, which did not provide sufficient power to test regressions within groups. Finally, although samples were racially diverse, Study 2 contained only men, which may limit generalizability.

In conclusion, these studies highlight the importance of using samples with high levels of FTD when studying disorganized speech in schizophrenia. Results suggest those with FTD may have lower metacognition than those without. The results also demonstrate the utility of adding automated analysis to existing methods in predicting cognitive variables, as together they explained over a third of the variance in metacognition. Future work could extend these findings by examining the real-world implications of the relationship between disorganized speech and metacognition (i.e., how it pertains to functioning and quality of life).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

- Abel DB, Minor KS, 2021. Social functioning in schizophrenia: Comparing laboratory-based assessment with real-world measures. *J. Psychiatr. Res.* 138, 500–506. 10.1016/j.jpsychires.2021.04.039 [PubMed: 33971484]
- Barrera A, McKenna PJ, Berrios GE, 2005. Formal thought disorder in schizophrenia: An executive or a semantic deficit? *Psychol. Med.* 35, 121–132. 10.1017/S003329170400279X [PubMed: 15842035]
- Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM, 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophr.* 10.1038/npjschz.2015.30
- Bleuler E, 1911. *Dementia praecox or the group of schizophrenias*, New York: International Universities Press; 1950.
- Bora E, Yalincetin B, Akdede BB, Alptekin K, 2019. Neurocognitive and linguistic correlates of positive and negative formal thought disorder: A meta-analysis. *Schizophr. Res.* 209, 2–11. 10.1016/j.schres.2019.05.025 [PubMed: 31153670]
- Bowie CR, Harvey PD, 2008. Communication Abnormalities Predict Functional Outcomes in Chronic Schizophrenia: Differential Associations with Social and Adaptive Functions. *Schizophr. Res.* 103, 240–247. 10.1038/jid.2014.371 [PubMed: 18571378]
- Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CE, Cecchi GA, 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 17, 67–75. 10.1002/wps.20491 [PubMed: 29352548]
- De Boer J, Voppel A, Brederoo S, Schnack H, Truong K, Wijnen F, Sommer I, 2021. Acoustic speech markers for schizophrenia-spectrum disorders: A diagnostic and symptom-recognition tool. *Psychol. Med.* 1–11. doi:10.1017/S0033291721002804
- Docherty NM, DeRosa M, Andreasen NC, 1996. Communication Disturbances in Schizophrenia and Mania. *Arch Gen Psychiatry* 53, 358–364. [PubMed: 8634014]
- Docherty NM, Miller TN, Lewis MA, 1997. Communication disturbances in the natural speech of schizophrenic patients and non-schizophrenic parents of patients. *Acta Psychiatr. Scand.* 95, 500–507. 10.1111/j.1600-0447.1997.tb10138.x [PubMed: 9242845]
- Docherty NM, Cohen AS, Nienow TM, Dinzeo TJ, Dangelmaier RE, 2003. Stability of formal thought disorder and referential communication disturbances in schizophrenia. *J. Abnorm. Psychol.* 112, 469–475. 10.1037/0021-843X.112.3.469 [PubMed: 12943025]
- Docherty NM, 2012. On identifying the processes underlying schizophrenic speech disorder. *Schizophr. Bull.* 38, 1327–1335. 10.1093/schbul/sbr048 [PubMed: 21562141]

- Docherty NM, McCleery A, Divilbiss M, Schumann EB, Moe A, Shakeel MK, 2013. Effects of social cognitive impairment on speech disorder in schizophrenia. *Schizophr. Bull.* 39, 608–616. 10.1093/schbul/sbs039 [PubMed: 22416265]
- Dwyer K, David AS, McCarthy R, McKenna P, Peters E, 2019. Linguistic alignment and theory of mind impairments in schizophrenia patients' dialogic interactions. *Psychol. Med.* 1–9. 10.1017/S0033291719002289
- Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE, 2007. Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia Brita. *Schizophr. Res.* 93, 304–316. [PubMed: 17433866]
- Graesser AC, McNamara DS, Kulikowich JM, 2011. Coh-Metrix : Providing Multilevel Analyses of Text Characteristics. *Educ. Res.* 40, 223–234.
- Holshausen K, Harvey PD, Elvevåg B, Foltz PW, Bowie CR, 2014. Latent semantic variables are associated with formal thought disorder and adaptive behavior in older inpatients with schizophrenia. *Cortex.* 55, 88–96. doi: 10.1016/j.cortex.2013.02.006. [PubMed: 23510635]
- Kay SR, Fiszbein A, Opler LA, 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13, 261–276. 10.1093/schbul/13.2.261 [PubMed: 3616518]
- Kerns JG, Becker TM, 2008. Communication disturbances, working memory, and emotion in people with elevated disorganized schizotypy. *Schizophr. Res.* 100, 172–180. 10.1016/j.schres.2007.11.005 [PubMed: 18068952]
- Lundin NB, Hochheiser J, Minor KS, Hetrick WP, Lysaker PH, 2020. Piecing together fragments: Linguistic cohesion mediates the relationship between executive function and metacognition in schizophrenia. *Schizophr. Res.* 215, 54–60. 10.1016/j.schres.2019.11.032 [PubMed: 31784337]
- Lysaker PH, Clements CA, Plascak-Hallberg CD, Knipscheer SJ, Wright DE, 2002. Insight and personal narratives of illness in schizophrenia. *Psychiatry* 65, 197–206. 10.1521/psyc.65.3.197.20174 [PubMed: 12405078]
- Lysaker PH, Carcione A, Dimaggio G, Johannesen JK, Nicolò G, Procacci M, Semerari A, 2005. Metacognition amidst narratives of self and illness in schizophrenia: Associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatr. Scand.* 112, 64–71. 10.1111/j.1600-0447.2005.00514.x [PubMed: 15952947]
- Lysaker PH, Gumley A, Luedtke B, Buck KD, Ringer JM, Olesek K, Kukla M, Leonhardt BL, Popolo R, Dimaggio G, 2013. Social cognition and metacognition in schizophrenia: evidence of their independence and linkage with outcomes. *Acta Psychiatr. Scand.* 127, 239–247. 10.1111/acps.12012 [PubMed: 22967227]
- Lysaker PH, Vohs J, Hamm JA, Kukla M, Minor KS, de Jong S, van Donkersgoed R, Pijnenborg MHM, Kent JS, Matthews SC, Ringer JM, Leonhardt BL, Francis MM, Buck KD, Dimaggio G, 2014. Deficits in metacognitive capacity distinguish patients with schizophrenia from those with prolonged medical adversity. *J. Psychiatr. Res.* 55, 126–132. 10.1016/j.jpsychires.2014.04.011 [PubMed: 24811777]
- Lysaker PH, Minor KS, Lysaker JT, Hasson-Ohayon I, Bonfils K, Hochheiser J, Vohs JL, 2020. Metacognitive function and fragmentation in schizophrenia: Relationship to cognition, self-experience and developing treatments. *Schizophr. Res. Cogn.* 19, 100142. 10.1016/j.scog.2019.100142 [PubMed: 31828019]
- Mackinley M, Chan J, Ke H, Dempster K, Palaniyappan L, 2021. Linguistic determinants of formal thought disorder in first episode psychosis. *Early Interv. Psychiatry* 15, 344–351. 10.1111/eip.12948 [PubMed: 32129010]
- McNamara DS, Graesser AC, McCarthy PM, Cai Z, 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix.* Cambridge University Press., New York, NY.
- Minor KS, Lysaker PH, 2014. Necessary, but not sufficient: Links between neurocognition, social cognition, and metacognition in schizophrenia are moderated by disorganized symptoms. *Schizophr. Res.* 159, 198–204. 10.1016/j.schres.2014.08.005 [PubMed: 25192756]
- Minor KS, Marggraf MP, Davis BJ, Luther L, Vohs JL, Buck KD, Lysaker PH, 2015. Conceptual disorganization weakens links in cognitive pathways: Disentangling neurocognition, social cognition, and metacognition in schizophrenia. *Schizophr. Res.* 169, 153–158. 10.1016/j.schres.2015.09.026 [PubMed: 26441007]

- Minor KS, Willits JA, Marggraf MP, Jones MN, Lysaker PH, 2019. Measuring disorganized speech in schizophrenia: Automated analysis explains variance in cognitive deficits beyond clinician-rated scales. *Psychol. Med.* 49, 440–448. 10.1017/S0033291718001046 [PubMed: 29692287]
- Minor KS, Marggraf MP, Davis BJ, Mickens JL, Abel DB, Robbins ML, Buck KD, Wiehe SE, Lysaker PH, 2021. Personalizing Interventions Using Real-World Interactions: Improving Symptoms and Social Functioning in Schizophrenia With Tailored Metacognitive Therapy. *J. Consult. Clin. Psychol.* 90, 18–28. 10.1037/ccp0000672 [PubMed: 34410749]
- Mota NB, Copelli M, Ribeiro S, 2017. Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *npj Schizophr.* 3, 18. 10.1038/s41537-017-0019-3 [PubMed: 28560264]
- Mota NB, Vasconcelos NAP, Lemos N, Pieretti AC, Kinouchi O, Cecchi GA, Copelli M, Ribeiro S, 2012. Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS One* 7, 1–9. 10.1371/journal.pone.0034928
- Racenstein JM, Penn D, Harrow M, Schleser R, 1999. Thought disorder and psychosocial functioning in schizophrenia: The concurrent and predictive relationships. *J. Nerv. Ment. Dis.* 187, 281–289. 10.1097/00005053-199905000-00003 [PubMed: 10348082]
- Tan EJ, Rossell SL, 2019. Language comprehension and neurocognition independently and concurrently contribute to formal thought disorder severity in schizophrenia. *Schizophr. Res.* 204, 133–137. 10.1016/j.schres.2018.08.019 [PubMed: 30126817]
- Tausczik YR, Pennebaker JW, 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *J. Lang. Soc. Psychol.* 29, 24–54. 10.1177/0261927X09351676
- Ventura J, Thames AD, Wood RC, Guzik LH, Helleman GS, 2010. Disorganization and Reality Distortion in Schizophrenia. *Schizophr. Res.* 121, 1–14. 10.1038/jid.2014.371 [PubMed: 20579855]
- Willits JA, Rubin T, Jones MN, Minor KS, Lysaker PH, 2018. Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophr. Res.* 197, 365–369. 10.1016/j.schres.2017.11.014 [PubMed: 29153448]

### Highlights

- Automated analytical tools are novel methods to assess speech in schizophrenia.
- Group differences emerged at a moderate formal thought disorder threshold.
- Findings suggest using a moderate rather than mild threshold for thought disorder.
- Patients with moderate thought disorder had lower metacognition.
- Together, speech measures predicted metacognition in people with schizophrenia.

**Table 1.**

## Coh-Metrix index descriptions

<b>Index</b>	<b>Index description</b>
Narrativity	How much speech tells a story with previously introduced characters/topics
Syntactic simplicity	The extent to which sentences contain fewer words and use simple structures
Word concreteness	The degree to which words are concrete versus abstract
Referential cohesion	How much words and ideas overlap across sentences and a conversation
Deep cohesion	The degree to which speech contains causal and logical links

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.**

Participant data by study and group; with formal thought disorder (FTD) and without (non-FTD)

<b>Study 1</b>	<b>FTD (n = 16)</b>	<b>non-FTD (n = 29)</b>	<b>Test of Significance</b>	<b>Effect Size (Cohen's d)</b>
Age ( <i>SD</i> )	47.00 (10.28)	46.83 (10.20)	$t(43) = -0.05$	
Gender ( <i>n, %</i> )			$\chi^2(1) = 1.23$	
Male	11, 68.8%	15, 51.7%		
Female	5, 31.3%	14, 48.3%		
Race			$\chi^2(2) = 0.53$	
Black	8, 50.0%	17, 58.6%		
White	6, 37.5%	10, 34.5%		
Multiracial	2, 12.5%	2, 6.9%		
Education			$\chi^2(2) = 0.55$	
<HS	5, 31.3%	7, 24.1%		
HS/GED	4, 25.0%	6, 20.7%		
>HS	7, 43.8%	16, 55.2%		
CDI total score	2.58 (1.65)	1.56 (0.81)	$t(19) = -2.31^{*#}$	-0.86
CM composite	0.70 (0.19)	0.83 (0.37)	$t(43) = 1.29$	0.40
MAS-A total	11.06 (3.67)	10.97 (2.66)	$t(43) = -0.10$	-0.03
Self-reflectivity	4.16 (1.23)	4.08 (1.06)	$t(43) = -0.20$	-0.03
Other	2.84 (1.17)	2.95 (0.91)	$t(43) = 0.33$	0.10
Decentration	0.63 (0.59)	0.66 (0.42)	$t(43) = 0.18$	0.09
Mastery	3.44 (1.42)	3.28 (1.15)	$t(43) = -0.41$	-0.13
PANSS total	66.87 (7.36)	61.44 (14.41)	$t(43) = -1.67^{#}$	-0.44
Concept Disorg	3.50 (0.89)	1.31 (0.47)	$t(20) = -9.12^{***#}$	-3.37
IPII word count	2257.56 (2944.88)	1546.52 (1032.19)	$t(17) = -0.90^{#}$	-0.37
<b>Study 2</b>	<b>FTD (n = 23)</b>	<b>non-FTD (n = 23)</b>	<b>Test of Significance</b>	<b>Effect Size (Cohen's d)</b>
Age ( <i>SD</i> )	46.48 (9.73)	46.74 (9.02)	$t(44) = 0.09$	
Gender ( <i>n, %</i> )				
Male	23, 100%	23, 100%		
Race			$\chi^2(2) = 0.53$	
Black	12, 52.2%	12, 52.2%		
White	11, 47.8%	11, 47.8%		
Education			$\chi^2(2) = 0.55$	
<HS	3, 13.0%	4, 17.4%		
HS/GED	11, 47.8%	13, 56.5%		

>HS	9, 39.1%	6, 26.1%		
CDI total score	2.47 (1.12)	1.29 (0.71)	$t(44) = -4.25^{***\#}$	-1.25
CM composite	0.60 (0.29)	0.81 (0.31)	$t(44) = 2.27^*$	0.67
MAS-A total	9.30 (3.58)	11.04 (3.91)	$t(44) = 1.57$	0.46
Self-reflectivity	3.43 (0.84)	4.04 (1.17)	$t(44) = 2.03^*$	0.60
Other	2.96 (0.86)	3.02 (0.97)	$t(44) = 0.24$	0.07
Decentration	0.41 (0.65)	0.61 (0.80)	$t(44) = 0.91$	0.27
Mastery	2.50 (1.80)	3.37 (1.91)	$t(44) = 1.59$	0.47
PANSS total	84.61 (9.03)	71.52 (10.60)	$t(44) = -4.51^{***}$	-1.33
Concept Disorg	4.22 (0.42)	2.04 (0.88)	$t(32) = -10.70^{***\#}$	-3.16
IPII word count	1241.48 (950.18)	1765.13 (1723.92)	$t(34) = 1.28^\#$	

\*\*\*  
 $p < .001$

\*\*  
 $p < .01$

\*  
 $p < .05$

+  
 $p < .10$

#  
equal variances not assumed

CDI (Communication Disturbances Index); CM (Coh-Metrix); NARz (Narrativity z-score); SYNz (Syntactic simplicity z-score); CNCz (Word concreteness z-score); REFz (Referential cohesion z-score); DCz (Deep cohesion z-score); MAS-A (Metacognitive Assessment Scale – Abbreviated); Other (Awareness of the other’s mind); PANSS (Positive and Negative Syndrome Scale); Study 1: FTD 3 on Conceptual Disorganization item from PANSS; Study 2: FTD 4 on Conceptual Disorganization item

**Table 3.**

Study 1 stepwise regressions using behavioral measures of disorganized speech as predictors of metacognition across the full sample (n = 45)

	Adj. $R^2$	$B$	S.E. $B$	$\beta$
<i>Model with CM composite</i>				
Step One	-0.01			
CDI total score		-0.27	0.36	-0.11
Step One	0.00			
CM composite		1.49	1.41	0.16
Step Two	-0.01			
CDI total score		-0.23	0.37	-0.10
CM composite		1.41	1.43	0.15
<i>Model with CM indices</i>				
Step One	0.14*			
SYNz		-0.80	0.65	-0.21
DCz		0.79	0.74	0.17
NARz		0.87	0.90	0.17
Step Two	0.14*			
CDI total score		-0.32	0.35	-0.13
SYNz		-0.65	0.67	-0.17
DCz		0.86	0.74	0.19
NARz		1.02	0.92	0.20

\*  $p < .05$

CDI (Communication Disturbances Index); CM (Coh-Metrix) ; SYNz (Syntactic simplicity z-score); DCz (Deep Cohesion z-score); NARz (Narrativity z-score)

**Table 4.**

Post-hoc correlations between MAS-A total score and Coh-Metrix indices across the full sample

<b>Study 1</b>	CM comp	NAR	SYN	CNC	REF	DC
MAS-A	-.05	.37*	-.38*	.05	.10	.33*
<b>Study 2</b>	CM comp	NAR	SYN	CNC	REF	DC
MAS-A	.11	-.01	-.42**	.27	.29*	.16

\*  $p < .05$

MAS-A (Metacognitive Assessment Scale – Abbreviated) total score; CM comp (Coh-Metrix composite score); NAR (Narrativity); SYN (Syntactic simplicity); CNC (Word concreteness); REF (Referential cohesion); DC (Deep cohesion)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5.**

Study 2 stepwise regressions using behavioral measures of disorganized speech as predictors of metacognition across the full sample (n = 46)

	Adj. $R^2$	$B$	S.E. $B$	$\beta$
<i>Model with CM composite</i>				
Step One	0.20**			
CDI total score		-1.61	0.46	-0.47**
Step One	-0.01			
CM composite		1.29	1.82	0.11
Step Two	0.18**			
CDI total score		-1.67	0.49	-0.48**
CM composite		-0.65	1.73	-0.05
<i>Model with CM indices</i>				
Step One	0.18**			
SYNz		-2.02	0.74	-0.38**
REFz		1.49	0.96	0.21
Step Two	0.34***			
CDI total score		-1.42	0.42	-0.41**
SYNz		-1.85	0.67	-0.35**
REFz		1.18	0.87	0.17

\*\*  
 $p < .01$

\*\*\*  
 $p < .001$

CDI (Communication Disturbances Index); CM (Coh-Metrix); SYNz (Syntactic simplicity z-score); REFz (Referential cohesion z-score)