1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Linguistic Determinants of Formal Thought Disorder in First Episode Psychosis.

Michael Mackinley[1,2], Jenny Chan[1], Hannah Ke[2], Kara Dempster[3], Lena Palaniyappan[1,2,3]

[1]Robarts Research Institute, University of Western Ontario, London, Ontario, Canada.

[2] Lawon Health Research Institute, London, Ontario, Canada

[3]Department of Psychiatry, University of Western Ontario, London, Ontario, Canada.

Corresponding Author: Michael MacKinley, mmacki4@uwo.ca

Word Count: 2732

Abstract

Aim: Thought disorder is a core feature of schizophrenia but assessment of disordered thinking is challenging, which may contribute to the paucity of mechanistic understanding of disorganization in early psychosis. We studied the use of linguistic connectives in relation to clinically quantified dimensions of thought disorder using automated speech analysis in untreated, first episode psychosis (FEPs) and healthy controls (HCs).

Methods: 39 treatment-naïve, actively psychotic FEPs and 23 group matched HCs were recruited. Three one-minute speech samples were induced in response to photographs from the Thematic Apperception Test and speech was analyzed using COH-METRIX software. Five connectives variables from the Coh-Metrix software were reduced using principle component analysis, resulting in two linguistic connectives factors. Thought disorder was assessed using the Thought Language Index (TLI) and the PANSS-8.

Results: Connective factors predicted disorganization, but not impoverishment suggesting aberrant use of connectives is specific to positive thought disorder. An independent t-test comparing low and high disorganization FEPs showed higher load of acausal temporal connectives in high disorganization FEPs compared to low disorganization FEPs (mean [SD] in high vs. low disorganization FEPs = 0.64 (1.1) vs. -0.37 (1.02) ; t=2.91, p=0.006). Factor 2 was not correlated with disease severity or cognition suggesting connective use is a specific index of disorganized thinking rather than overall illness status.

Conclusions: Clinical assessment of disorganization in psychosis is likely linked to the aberrant use of connectives resulting in an intuitive sense of incoherence. In early psychosis, thought disorder may be reliably quantifiable using automated syntax analysis.

1.0 Introduction

Schizophrenia is an illness of perceptual and thought disturbances. While research on disordered thinking in schizophrenia is typically focused on unusual thought content (e.g., delusions), disturbances in the formation and expression of thought are core features of the illness. Formal thought disorder (FTD) is defined by substantial disturbances in a patient's ability to express cogent, complex thoughts (Bleuler, 1950). While not necessary for a diagnosis of schizophrenia, the presence of FTD predicts onset of psychosis in individuals who are at clinical high risk of developing psychosis (Dominguez et al., 2010; Ziermans et al., 2014), is present in over 55% of patients in their first episode of psychosis (Roche, Lyne, et al., 2015), and its persistence is linked to worsened social/community functioning (Bowie, Gupta, & Holshausen, 2011; Bowie & Harvey, 2008; Roche et al., 2016), cognitive deficits and poorer clinical outcomes (Cuesta, Peralta, & De Leon, 1994; Roche, Creed, MacMahon, Brennan, & Clarke, 2015). Despite the recognized importance of this syndrome, no effective interventions are available to reduce the burden of FTD in schizophrenia when antipsychotics are ineffective. In part, this therapeutic gap can be attributed to our lack of a clear conceptual understanding of this syndrome and its physiological basis.

FTD captures a number of non-specific thought/language anomalies (Andreasen, 1986) that are evident across a wide range of mental health and cognitive disorders including psychosis, bipolar disorder (Morgan et al., 2017), depression, and Autism Spectrum Disorder (Eussen et al., 2015). Further, the presence of FTD among healthy controls (albeit typically in subtler forms) has been identified (Kircher, Bröhl, Meier, & Engelen, 2018; Kuperberg, 2010) though most existing clinical scales cannot quantify non-clinical FTD. FTD has been assessed historically as a categorical variable (Hart & Lewine, 2017), which fails to adequately describe the breadth of dysfunction captured under the umbrella of FTD. By

ignoring the multidimensional nature of FTD, these scales impede the ability of researchers to identify causal link between thought disorder and its neural and cognitive basis.

A number of validated manually scored instruments have been developed to assess multiple domains of disordered thought such as the Thought Language Index (Liddle et al., 2002), Thought Disorder Index(TDI) (Solovay et al., 1986) and Thought and Language Disorder (TALD) scale (Kircher et al., 2014). By capturing FTD as a trans-diagnostic set of features, these scales make it possible to begin the work of linking specific elements of FTD with specific neural processes. While progress has been made in this regard (Wensing et al., 2017), several barriers remain in place. First, while potentially mitigated through their manualized approach, these scales rely heavily on clinical judgement and require highly trained staff. Second, because FTD items in these scales are tailored to identify illness specific symptomology (often centered on schizophrenia), these measures may not be adequately sensitive to subtler forms of thought disorder. Finally, many of these scales can be time consuming to complete in busy clinical settings, presenting substantial pragmatic barriers to their wide application.

To extend the study of the linguistic underpinnings of FTD, recent research has leveraged automated speech analysis software. These programs allow clinicians to assess larger speech samples, across multiple linguistic domains with consistent, fast, and inexpensive results. The use of these software programs allows researchers to assess speech samples at multiple levels of analysis (e.g., at the level of the word, phrase, sentence, or full speech sample) at multiple time points without the burdens presented by hand scored instruments. This allows researchers to assess poor conceptual integration not only based on traditional clinical measures of thought disorder, but also at structural language levels. Because it remains unclear whether thought disorder presents at basic linguistic or higher order levels, the ability to efficiently and simultaneously assess factors at multiple levels of analysis is advantageous.

Over the past several years these technologies have been applied in the assessment of

language dysfunction in both schizophrenia (Elvevåg, Foltz, Rosenstein, & DeLisi, 2010;

Minor, Willits, Marggraf, Jones, & Lysaker, 2019; Willits, Rubin, Jones, Minor, & Lysaker,

2018) and clinical high risk samples(Bedi et al., 2015; Gupta, Hespos, Horton, & Mittal,

2018). These studies have shown speech disturbances at multiple levels of linguistic analysis

ranging from basic speech descriptors (e.g., word counts) (Willits et al., 2018) up to higher

order linguistic variables, such as text cohesion (Bedi et al., 2015; Gupta et al., 2018; Willits

et al., 2018). In a 2018 study, Willits et al (Willits et al., 2018) showed that patients,

compared to healthy controls, used fewer causal, logical and contrastive connectives.

Connectives are words used to link concepts/thoughts when speaking, and when used

incorrectly the listener must exert more cognitive resources to interpret speech(Cain & Nash,

2011). However, it remains unclear whether the output from automated speech analysis

software effectively captures the clinical construct of FTD or if they represent a different

construct that is not clinically apparent, albeit critical for the disease process.

Because connectives are the linguistic basis upon which cohesion is built through a text, we

hypothesized that the aberrant use of connectives in speech samples may underlie the clinical

impression of thought disorder. We studied the use of linguistic connectives in relation to

clinically quantified dimensions of thought disorder using Coh-Metrix, an automated speech

analysis software, in untreated first episode psychosis patients (FEPs) and healthy controls

(HCs).

2. Method

2.1 Participants:

Data were collected from 39 treatment-naïve, psychotic first episode patients (FEPs)

recruited from the Prevention and Early Intervention Program for Psychoses in London,

Ontario, Canada. All participants were in the acute phase of the illness, and often were

referred for services when disorganization was noted by the attending clinician. We also

recruited an additional 23 healthy controls from the community group matched for age, sex,

and parental SES. Exclusion criteria included greater than 2 weeks of lifetime antipsychotic

exposure, a diagnosis of affective psychosis, active substance dependence or the presence of

intellectual/developmental disorder or the inability to provide informed consent. All

participants included in the analysis were native English speakers.

2.2 Instruments:

2.2.1 The Thought Language Index

The Thought and Language Index (TLI) (Liddle et al., 2002) is an instrument for assessing

formal thought disorder under standardised conditions. Participants are asked to produce

three one-minute speech samples in response to 3 photographs from the Thematic

Apperception Test (Murray, 1943). Participant responses were scored on 8 domains: poverty

of speech, weakening of goal, looseness, peculiar word usage, peculiar sentence structure,

peculiar logic, distractability, and perseveration. These domains were also used to compute 3

aggregate scores: impoverishment of thinking (the summed scores of Poverty of speech,

weakening of goal, and perseveration), disorganization (the summed scores of looseness,

peculiar word usage, peculiar sentence construction, peculiar logic, and distractibility), and

overall thought disorder (the summed scores of all 8 domains).

2.2.2 Positive and Negative Syndrome Scale – 8 Item Scale

The Positive and Negative Syndrome Scale – 8 Item scale (PANSS-8) is an abbreviated

version of the 30 Item PANSS clinical assessment of symptomology in schizophrenia and

psychosis with acceptable internal consistency and highly correlated with the full PANSS(Lin

et al., 2018). The PANSS Item P2 (conceptual disorganization) is a single item measure

intended to assess FTD characterized by "disruption of goal-directed sequencing, e.g., circumstantiality, tangentiality, loose associations non sequiturs, gross illogicality, or thought block" (Kay, Fiszbein, & Opler, 1987) . Scored on a 1 (absent) to 7 (extreme) Likert type scale, we classified those with scores of 1-3 (absent to mild) as "Low Conceptual disorganization" and those 4-7 (moderate to extreme) as high conceptual disorganization.

2.2.3 Coh-Metrix 3.0

The Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014) system is a web-based automated speech analysis software that computes several word, sentence and passage level linguistic variables from written and spoken speech samples. The software automatically computes a number of lower order (e.g., word counts, frequency of pronoun use, use of connectives) and higher order (e.g., readability, cohesion, text formality) linguistic variables using several common computational linguistics methods including analyzing lexicon, part-of-speech classifiers, syntactic parsers, and through the use of latent semantic analysis (Graesser, McNamara, Louwerse, & Cai, 2004). We based our project on the work in Willits et al (2018), focusing primarily on the use of 5 connectives variables (See Table 1 for details). Connectives were scored based on their frequency per 1000 words spoken.

2.3 Procedure

This study was approved by the local REB and informed consent was received from participants prior to participation. Clinical interviews, diagnoses and symptom rating scales, including a patient history and demographic assessment were completed by a board certified psychiatrist. Cognitive assessments and the TLI interview and rating were completed by graduate level research assistants. Speech samples were limited to 3 1-minute speech samples from the TLI interview, with researcher speech removed prior to analysis, and were blinded

CONNECTIVES AND THOUGHT DISORDER   8

to participant status consistent with the procedure described by Sommer et al. (Sommer et al.,

2010)

2.4 Statistical Analysis

Our analysis was completed in four steps. First, recorded and transcribed speech samples

were entered in Coh-Metrix 3.0 web tool to calculate the incidence scores (use per 1000

words) for the each of the 5 connectives variables. We then conducted a principal component

analysis on the five connective variables to extract factors based on a scree plot that

cumulatively explained 72.14% of variance in raw scores and compared the selected factor

scores between FEP and control groups using independent t tests. We then investigated the

relationship between the factor scores and clinically observed using the TLI and PANSS P2.

To establish specificity of these relationships, SOFAS and cognitive scores were correlated

with the factor scores.

3.0 Results

3.1 Demographics

 When comparing healthy controls to FEPs, no statistically significant differences existed for

gender, age, or socioeconomic status (Table 2). As expected, statistically significant

differences were found between patients and controls on the number of words spoken,

however no statistically significant differences were found between high and low

disorganization FEPs on this measure. Additionally, there were expected differences in

cognitive scores between patient groups and healthy controls on cognitive and clinical

variables.

3.2 PCA analysis

 Principal component analysis of the five Coh-Metrix based connective scores resulted in two factors (Table 1), each explaining 51.6% and 20.6% of variance respectively. Factor 1 positively loaded all 5 connectives variables (overall connective use factor), and Factor 2 positively loaded temporal and additive connectives while negatively loading for causal connectives (acausal temporal linkage factor). Both Factor 1 loading and Factor 2 loaded higher among FEPs [Factor 1: M(sd)=0.10 (1.15); Factor 2: M (sd)= 0.13 (0.99)] than healthy controls[Factor 1: M(sd)= -0.09 (0.81); Factor 2: M (sd)=-0.04 (1.2), however the difference was not significant. The distribution of connectives factor 1 and factor 2 scores can be found in Figure 1 (supplemental Materials).

3.3 Relationship with clinical measures of disorganization

Correlations between linguistic connectives factors and TLI scores in FEPs showed that both linguistic connective factors were associated with the overall TLI scores among patients (Factor 1: *R=0.419 p=0.010, Factor 2:R=0.408, p=0.012),* but not among controls (Table 3). When broken down into type of thought disorder, the association between connectives factors and TLI scores were driven by the association between both overall and acausal connectives use and disorganization (Factor 1: R=0.416, p=0,01; Factor 2: R=0.379, p=0.02); no statistically significant associations existed between connectives use and TLI impoverishment scores, suggesting that the aberrant use of connectives is associated with positive, rather than negative, elements of thought disorder.

An independent t-test comparing low and high disorganization FEPs showed higher load of acausal temporal connectives in high disorganization FEPs compared to low disorganization FEPs (mean [SD] in high vs. low disorganization FEPs = 0.64 (1.1) vs. -0.37 (1.02) ; t=2.91, p=0.006).

3.4 Specificity of Connective Factors

None of the correlations between acausal temporal connective use (factor 2) and measures of disease severity (SOFAS, CGI-Severity) and cognition (Digit symbol substitution, Trails-B, and Category Fluency) were significant, even at an uncorrected threshold, suggesting that aberrant connective use is a specific index of disorganized thinking rather than overall illness severity or cognitive dysfunction (Table 4).

4.0 Discussion

In this study we compared the use of linguistic connectives to traditional clinically quantified dimensions of formal thought disorder among drug-naïve first episode psychosis patients and matched healthy controls. We report 3 major findings. 1. Disorganization in first episode psychosis is characterised by excessive use of connectives 2. Connectives use during stimulus-evoked speech is correlated with disorganization, but not impoverishment, indicating the specificity of linguistic connectives to positive rather than negative thought disorder in FEP 3. Aberrant connectives use is specifically related to conceptual disorganization but not to processing speed, verbal fluency or social functioning. Thus, automated speech analysis for connectives captures a distinct facet of loosened associations that is not influenced by the cognitive and functional status of an individual. These findings suggest that aberrant use of connectives (specifically increased temporal but decreased use of causal connectives) contribute to the clinical impression of 'disorganization' and 'thought disorder' captured using instruments such as PANSS and TLI.

The importance of connectives to the clinical impression of thought disorder likely operates through one of two potential mechanisms: 1. Patients fail to elucidate the connections between meaningfully linked concepts or ideas (reduced causal, logical, or contrastive connectives) or 2. Patients successfully, but erroneously, link ideas over time

where no coherent link exists (inappropriate temporal connective use). The resulting of both mechanisms is an increased cognitive load placed on the listener to decipher the intended output of speech (Cain & Nash, 2011). If these efforts at interpreting speech fail, the result is a perception of tangentiality or incoherence, features of positive thought disorder.

We have shown that the traditional measures of disorganization (TLI & PANSS-P2) map reliably on to the specific outputs of automated speech analysis in psychosis. Automated speech analysis avoids the cost, time, and reliability issues that impact hand scoring instruments for the assessment of thought disorder, while simultaneously allowing for considerably more comprehensive analyses than classic clinician rated scales. This allows for the examination of multiple linguistic domains of thought disorder with relatively few disruptions to the clinical setting. Further, if Coh-Metrix is able to identify elements of disordered thought reliably during the first episode of psychosis, this approach could be leveraged to identify subtle longitudinal changes in speech that may be imperceptible to clinicians, allowing insights into how thought disorder evolves throughout the course of illness.

A number of limitations with the current work should be noted. Firstly, speech analyses were gathered from a provided visual stimuli. These stimulus-evoked responses likely involve unique cognitive processes that may not precisely reflect spontaneous language production. Secondly, aberrant connectives use in speech does not necessarily imply internally generated thoughts are not meaningfully connected and this unobservable thought-language gap may differ between patients and healthy controls. Finally, as we collected only a single cross-sectional speech sample during the first episode of psychosis sample, it is unclear if our findings would be stable over the course of illness following exposure to antipsychotic medications, as positive thought disorder is less observable later stages of illness (Roche, Creed, et al., 2015).

Persistent positive FTD in chronic schizophrenia is linked to aberrant glutamate transmission as well as structural changes in language areas of the brain [13]. To date, it is not clear if the disorganization seen in early stages of schizophrenia is also linked to the same neural pathways. One important challenge in elucidating such neural underpinnings is the difficulty in separating positive and negative FTD during early stages of psychosis (see Palaniyappan et al., 2015[32]). Factor analyses of symptoms in early psychosis often indicate the presence of a single dimension of Bleulerian symptoms embracing both impoverishment and disorganisation (McGorry, Bell, Dudgeon, & Jackson, 1998; Tonna et al., n.d.). The approach of using linguistic connectives identifies a distinct positive FTD factor in unmedicated early stage of psychosis, thus promising further enquiries to understand the neural underpinnings of early positive FTD.

In conclusion, FTD remains a poorly understood aspect of psychotic illness. The practical difficulty in capturing robust measures of thought disorder through clinician rated or hand scored instruments is an ongoing barrier to research in this area. Our work adds to the body of literature promoting automated speech analysis, and argues that the use of connectives in evoked speech could objectively identify the specific phenomenon of disorganization (positive FTD), independent of the functional and cognitive status of patients experiencing first episode psychosis. This raises the question of whether the elusive Bleulerian concept of loosening of association can now be reliably tracked across the different stages of psychosis. If this becomes feasible, we may inch closer to understanding the neural basis of a core deficit in schizophrenia.

5.0 Funding

CONNECTIVES AND THOUGHT DISORDER   13

7.0 References

Andreasen, N. C. (1986). Scale for the assessment of thought, language, and communication (TLC). *Schizophrenia Bulletin*, *12*(3), 473–482. https://doi.org/10.1093/schbul/12.3.473

Bedi, G., Carrillo, F., Cecchi, G. A., Slezak, D. F., Sigman, M., Mota, N. B., … Corcoran, C. M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. *Npj Schizophrenia*, *1*, 15030. https://doi.org/10.1038/npjschz.2015.30

Bleuler, E. (1950). *Dementia praecox or the group of schizophrenias*. Oxford, England: International Universities Press.

Bowie, C. R., Gupta, M., & Holshausen, K. (2011). Disconnected and underproductive speech in schizophrenia: Unique relationships across multiple indicators of social functioning. *Schizophrenia Research*, *131*(1), 152–156. https://doi.org/10.1016/j.schres.2011.04.014

Bowie, C. R., & Harvey, P. D. (2008). Communication abnormalities predict functional outcomes in chronic schizophrenia: Differential associations with social and adaptive functions. *Schizophrenia Research*, *103*(1), 240–247. https://doi.org/10.1016/j.schres.2008.05.006

Cain, K., & Nash, H. M. (2011). The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, *103*(2), 429–441. https://doi.org/10.1037/a0022824

Cuesta, M. J., Peralta, V., & De Leon, J. (1994). Schizophrenic syndromes associated with treatment response. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *18*(1), 87–99. https://doi.org/10.1016/0278-5846(94)90026-4

Dominguez, M.-G., Saka, M. C., can Saka, M., Lieb, R., Wittchen, H.-U., & van Os, J. (2010). Early expression of negative/disorganized symptoms predicting psychotic

experiences and subsequent clinical psychosis: A 10-year study. *The American Journal of Psychiatry*, *167*(9), 1075–1082.

https://doi.org/10.1176/appi.ajp.2010.09060883

Elvevåg, B., Foltz, P. W., Rosenstein, M., & DeLisi, L. E. (2010). An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. *Journal of Neurolinguistics*, *23*(3), 270–284.

https://doi.org/10.1016/j.jneuroling.2009.05.002

Eussen, M. L. J. M., de Bruin, E. I., Van Gool, A. R., Louwerse, A., van der Ende, J., Verheij, F., … Greaves-Lord, K. (2015). Formal thought disorder in autism spectrum disorder predicts future symptom severity, but not psychosis prodrome. *European Child & Adolescent Psychiatry*, *24*(2), 163–172. https://doi.org/10.1007/s00787-014-0552-9

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193–202. https://doi.org/10.3758/BF03195564

Gupta, T., Hespos, S. J., Horton, W. S., & Mittal, V. A. (2018). Automated analysis of written narratives reveals abnormalities in referential cohesion in youth at ultra high risk for psychosis. *Schizophrenia Research*, *192*, 82–88.

https://doi.org/10.1016/j.schres.2017.04.025

Hart, M., & Lewine, R. R. J. (2017). Rethinking Thought Disorder. *Schizophrenia Bulletin*, *43*(3), 514–522. https://doi.org/10.1093/schbul/sbx003

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276.

https://doi.org/10.1093/schbul/13.2.261

Kircher, T., Bröhl, H., Meier, F., & Engelen, J. (2018). Formal thought disorders: From phenomenology to neurobiology. *The Lancet Psychiatry*, *5*(6), 515–526. https://doi.org/10.1016/S2215-0366(18)30059-2

Kircher, T., Krug, A., Stratmann, M., Ghazi, S., Schales, C., Frauenheim, M., … Nagels, A. (2014). A rating scale for the assessment of objective and subjective formal Thought and Language Disorder (TALD). *Schizophrenia Research*, *160*(1), 216–221. https://doi.org/10.1016/j.schres.2014.10.024

Kuperberg, G. R. (2010). Language in schizophrenia Part 1: An Introduction. *Language and Linguistics Compass*, *4*(8), 576–589. https://doi.org/10.1111/j.1749-818X.2010.00216.x

Liddle, P. F., Ngan, E. T. C., Caissie, S. L., Anderson, C. M., Bates, A. T., Quested, D. J., … Weg, R. (2002). Thought and Language Index: An instrument for assessing thought and language in schizophrenia. *The British Journal of Psychiatry*, *181*(4), 326–330. https://doi.org/10.1192/bjp.181.4.326

Lin, C.-H., Lin, H.-S., Lin, S.-C., Kuo, C.-C., Wang, F.-C., & Huang, Y.-H. (2018). Early improvement in PANSS-30, PANSS-8, and PANSS-6 scores predicts ultimate response and remission during acute treatment of schizophrenia. *Acta Psychiatrica Scandinavica*, *137*(2), 98–108. https://doi.org/10.1111/acps.12849

McGorry, P. D., Bell, R. C., Dudgeon, P. L., & Jackson, H. J. (1998). The dimensional structure of first episode psychosis: An exploratory factor analysis. *Psychological Medicine*, *28*(4), 935–947.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press.

Minor, K. S., Willits, J. A., Marggraf, M. P., Jones, M. N., & Lysaker, P. H. (2019). Measuring disorganized speech in schizophrenia: Automated analysis explains

variance in cognitive deficits beyond clinician-rated scales. *Psychological Medicine*,

49(3), 440–448. https://doi.org/10.1017/S0033291718001046

Morgan, C. J., Coleman, M. J., Ulgen, A., Boling, L., Cole, J. O., Johnson, F. V., … Levy, D.

L. (2017). Thought Disorder in Schizophrenia and Bipolar Disorder Probands, Their

Relatives, and Nonpsychiatric Controls. *Schizophrenia Bulletin*, *43*(3), 523–535.

https://doi.org/10.1093/schbul/sbx016

Roche, E., Creed, L., MacMahon, D., Brennan, D., & Clarke, M. (2015). The Epidemiology

and Associated Phenomenology of Formal Thought Disorder: A Systematic Review.

Schizophrenia Bulletin, 41(4), 951–962. https://doi.org/10.1093/schbul/sbu129

Roche, E., Lyne, J., O'Donoghue, B., Segurado, R., Behan, C., Renwick, L., … Clarke, M.

(2016). The prognostic value of formal thought disorder following first episode

psychosis. *Schizophrenia Research, 178*(1–3), 29–34.

https://doi.org/10.1016/j.schres.2016.09.017

Roche, E., Lyne, J. P., O'Donoghue, B., Segurado, R., Kinsella, A., Hannigan, A., … Clarke,

M. (2015). The factor structure and clinical utility of formal thought disorder in first

episode psychosis. *Schizophrenia Research*, *168*(1), 92–98.

https://doi.org/10.1016/j.schres.2015.07.049

Solovay, M. R., Shenton, M. E., Gasperetti, C., Coleman, M., Kestnbaum, E., Carpenter, J.

T., & Holzman, P. S. (1986). Scoring manual for the Thought Disorder Index.

Schizophrenia Bulletin, *12*(3), 483–496. https://doi.org/10.1093/schbul/12.3.483

Sommer, I. E., Derwort, A. M. C., Daalman, K., de Weijer, A. D., Liddle, P. F., & Boks, M.

P. M. (2010). Formal thought disorder in non-clinical individuals with auditory verbal

hallucinations. *Schizophrenia Research*, *118*(1), 140–145.

https://doi.org/10.1016/j.schres.2010.01.024

Tonna, M., Ossola, P., Marchesi, C., Bettini, E., Lasalvia, A., Bonetto, C., … Ruggeri, M.

(n.d.). Dimensional structure of first episode psychosis. *Early Intervention in

Psychiatry*, *0*(0). https://doi.org/10.1111/eip.12789

Wensing, T., Cieslik, E. C., Müller, V. I., Hoffstaedter, F., Eickhoff, S. B., & Nickl-

Jockschat, T. (2017). Neural Correlates of Formal Thought Disorder: An Activation

Likelihood Estimation Meta-Analysis. *Human Brain Mapping*, *38*(10), 4946–4965.

https://doi.org/10.1002/hbm.23706

Willits, J. A., Rubin, T., Jones, M. N., Minor, K. S., & Lysaker, P. H. (2018). Evidence of

disturbances of deep levels of semantic cohesion within personal narratives in

schizophrenia. *Schizophrenia Research*, *197*, 365–369.

https://doi.org/10.1016/j.schres.2017.11.014

Ziermans, T., de Wit, S., Schothorst, P., Sprong, M., van Engeland, H., Kahn, R., & Durston,

S. (2014). Neurocognitive and clinical predictors of long-term outcome in adolescents

at ultra-high risk for psychosis: A 6-year follow-up. *PloS One*, *9*(4), e93994.

https://doi.org/10.1371/journal.pone.0093994

CONNECTIVES AND THOUGHT DISORDER
19
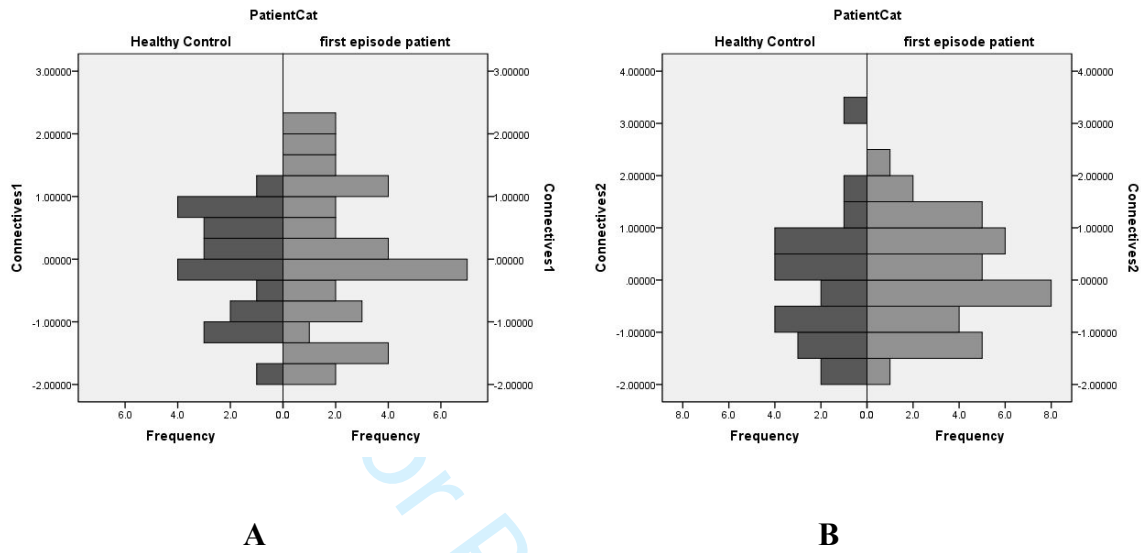
8.0 Supplemental Materials

Figure 1. A. Distribution of Connectives Factor 1 scores among healthy controls and first episode patients. **B.** Distribution of Connectives Factor 2 scores among healthy controls and first episode patients.

Table 1: Linguistic Connectives Definitions and PCA Factor Loading

Connective Variable	Variable Description	Connective factor 1	Connective Factor 2
Causal Connectives	Words that link a cause with an effect. (e.g., Because, due to)	0.703	-0.455
Logical Connectives	Words conjoining two logically linked statements (e.g., and, as well as)	0.951	-0.084
Temporal Connectives	Words that help put events in order of time (E.g., First, then, After)	0.342	0.791
Contrastive connectives	Words used to compare and contrast ideas (e.g., although, but, despite)	0.843	-0.135
Additive Connectives	Words used to add information or connect ideas (e.g., in addition, Moreover, another)	0.598	0.406

Table 2: Descriptive Characteristics by Patient Category

Variable	Healthy Control n=23	All FEP n=39	Low Disorganization FEPs n=21	High Disorganization FEPs n=18
Gender (M/F)	17/6	32/7	16/5	16/2
Age [m (sd)]	22.08 (3.47)	23.63 (4.48)	24.37 (5.37)	22.65 (2.79)
NS-SEC[M (sd)]	3.41 (1.34)	3.76 (1.20)	3.75 (1.12)	3.78 (1.31)
Words Spoken [M (sd)]	145.83 (28.54)	127.58 (35.31)*	125.55 (31.34)*	129.96 (40.35)
DUP Months [M (sd)]	N/A	8.83 (11.86)	8.83 (13.79)	8.82 (9.11)
Lifetime Antipsychotic Exposure (DDD)	N/A	2.31 (3.68)	2.73 (3.33)	1.81 (4.10)
SOFAS [M (sd)]	81.39 (4.41)	38.60 (12.42)**	41.39 (13.35)**	34.82 (10.23)**
TLI Total [M (sd)]	0.26 (0.39)	1.39 (1.31) **	0.93 (0.91)**	1.96 (1.52)**
TLI disorganization [M (sd)]	0.17 (0.24)	0.93 (1.12) **	0.57 (0.61)*	1.35 (1.41)
TLI Impoverishment [M (sd)]	0.09 (0.21)	0.47 (0.61) **	0.35 (0.48)*	0.60 (0.72) **
PANSS Total [M (sd)]	8.0 (0.00)	26.46 (7.2)**	23.66 (5.53) **	29.72 (7.69)**
PANSS Negative [M (sd)]	3.0 (0.00)	6.94 (4.46)**	5.81 (3.69)**	8.27 (4.99)**
DSST	68.39 (8.1)	51.83 (14.65)**	55.19 (13.70)**	47.91 (8.14)**
Trails-B	56.21 (15.35)	87.21 (42.6) **	81.64 (42.95)*	93.83 (42.61)**
Category Fluency [M (sd)]	24.45 (6.44)	18.67 (5.28)**	18.95 (4.63)**	18.31 (6.14) **

SD, standard deviation; NS-SEC, national statistics socio-economic classification; SOFAS, Social and Occupational Functioning Assessment Score; *Significantly different compared to healthy controls (p < 0.05); **Significantly different compared to healthy controls (p < 0.01).

Table 3: Linguistic Features Healthy Controls Vs. First Episode Patients

Variable		TLI: Overall R (p-value)	TLI: Impoverishment R (p-value)	TLI: Disorganization R (p-value)
Healthy Control	Factor 1	-0.026 (0.91)	-0.136 (0.55)	0.079 (0.73)
	Factor 2	-.339 (0.13)	-0.371 (0.10)	-0.327 (0.14)
FEP	Factor 1	**0.419 (0.010)****	0.149 (0.38)	**0.416 (0.01)****
	Factor 2	**0.408(0.012)***	0.193 (0.25)	**0.379 (0.02)***

Table 4: Correlations Between Connective Use, Cognition and Disease Severity

	SOFAS R (p-value)	**CGI-Severity** R (p-value)	**DSST** R (p-value)	**Category Fluency** R (p-value)	**Trails-B** R (p-value)
Factor 1 (overall connective use)					
Healthy Control	-0.175 (0.47)	N/A	-0.231 (0.29)	0.105 (0.66)	0.05 (0.98)
FEP:	-0.054 (0.76)	0.163 (0.336)	-0.221 (0.19)	0.022 (0.90)	0.105 (0.55)
Factor 2 (acausal temporal connectives)					
Healthy Control	0.091 (0.14)	N/A	0.098 (0.66)	0.104 (0.67)	-0.201 (0.41)
FEP	0.018 (0.92)	-0.246 (0.15)	0.171 (0.31)	-0.066 (0.97)	-0.170 (0.33)