

Western University

Scholarship@Western

Brain and Mind Institute Researchers'
Publications

Brain and Mind Institute

4-1-2021

Linguistic determinants of formal thought disorder in first episode psychosis

Michael Mackinley

Robarts Research Institute

Jenny Chan

Robarts Research Institute

Hanna Ke

Lawson Health Research Institute

Kara Dempster

Schulich School of Medicine & Dentistry

Lena Palaniyappan

Robarts Research Institute, lpalaniy@uwo.ca

Follow this and additional works at: <https://ir.lib.uwo.ca/brainpub>

Citation of this paper:

Mackinley, Michael; Chan, Jenny; Ke, Hanna; Dempster, Kara; and Palaniyappan, Lena, "Linguistic determinants of formal thought disorder in first episode psychosis" (2021). *Brain and Mind Institute Researchers' Publications*. 1027.

<https://ir.lib.uwo.ca/brainpub/1027>

Running head: CONNECTIVES AND THOUGHT DISORDER

Linguistic Determinants of Formal Thought Disorder in First Episode Psychosis.

Michael Mackinley^{1,2}, Jenny Chan¹, Hannah Ke², Kara Dempster³, Lena Palaniyappan^{1,2,3}

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

² Lawon Health Research Institute, London, Ontario, Canada

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

Corresponding Author: Michael MacKinley, mmacki4@uwo.ca

Word Count: 2732

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

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

1
2
3 ignoring the multidimensional nature of FTD, these scales impede the ability of researchers
4
5 to identify causal link between thought disorder and its neural and cognitive basis.
6
7

8
9 A number of validated manually scored instruments have been developed to assess multiple
10
11 domains of disordered thought such as the Thought Language Index (Liddle et al., 2002),
12
13 Thought Disorder Index(TDI) (Solovay et al., 1986) and Thought and Language Disorder
14
15 (TALD) scale (Kircher et al., 2014). By capturing FTD as a trans-diagnostic set of features,
16
17 these scales make it possible to begin the work of linking specific elements of FTD with
18
19 specific neural processes. While progress has been made in this regard (Wensing et al.,
20
21 2017), several barriers remain in place. First, while potentially mitigated through their
22
23 manualized approach, these scales rely heavily on clinical judgement and require highly
24
25 trained staff. Second, because FTD items in these scales are tailored to identify illness
26
27 specific symptomology (often centered on schizophrenia), these measures may not be
28
29 adequately sensitive to subtler forms of thought disorder. Finally, many of these scales can be
30
31 time consuming to complete in busy clinical settings, presenting substantial pragmatic
32
33 barriers to their wide application.
34
35
36
37
38

39 To extend the study of the linguistic underpinnings of FTD, recent research has leveraged
40
41 automated speech analysis software. These programs allow clinicians to assess larger speech
42
43 samples, across multiple linguistic domains with consistent, fast, and inexpensive results. The
44
45 use of these software programs allows researchers to assess speech samples at multiple levels
46
47 of analysis (e.g., at the level of the word, phrase, sentence, or full speech sample) at multiple
48
49 time points without the burdens presented by hand scored instruments. This allows
50
51 researchers to assess poor conceptual integration not only based on traditional clinical
52
53 measures of thought disorder, but also at structural language levels. Because it remains
54
55 unclear whether thought disorder presents at basic linguistic or higher order levels, the ability
56
57 to efficiently and simultaneously assess factors at multiple levels of analysis is advantageous.
58
59
60

1
2
3 Over the past several years these technologies have been applied in the assessment of
4
5 language dysfunction in both schizophrenia (Elvevåg, Foltz, Rosenstein, & DeLisi, 2010;
6
7 Minor, Willits, Marggraf, Jones, & Lysaker, 2019; Willits, Rubin, Jones, Minor, & Lysaker,
8
9 2018) and clinical high risk samples (Bedi et al., 2015; Gupta, Hespos, Horton, & Mittal,
10
11 2018). These studies have shown speech disturbances at multiple levels of linguistic analysis
12
13 ranging from basic speech descriptors (e.g., word counts) (Willits et al., 2018) up to higher
14
15 order linguistic variables, such as text cohesion (Bedi et al., 2015; Gupta et al., 2018; Willits
16
17 et al., 2018). In a 2018 study, Willits et al (Willits et al., 2018) showed that patients,
18
19 compared to healthy controls, used fewer causal, logical and contrastive connectives.
20
21 Connectives are words used to link concepts/thoughts when speaking, and when used
22
23 incorrectly the listener must exert more cognitive resources to interpret speech (Cain & Nash,
24
25 2011). However, it remains unclear whether the output from automated speech analysis
26
27 software effectively captures the clinical construct of FTD or if they represent a different
28
29 construct that is not clinically apparent, albeit critical for the disease process.
30
31
32
33
34

35
36 Because connectives are the linguistic basis upon which cohesion is built through a text, we
37
38 hypothesized that the aberrant use of connectives in speech samples may underlie the clinical
39
40 impression of thought disorder. We studied the use of linguistic connectives in relation to
41
42 clinically quantified dimensions of thought disorder using Coh-Metrix, an automated speech
43
44 analysis software, in untreated first episode psychosis patients (FEPs) and healthy controls
45
46 (HCs).
47
48
49

50 2. Method

51 2.1 Participants:

52
53
54 Data were collected from 39 treatment-naïve, psychotic first episode patients (FEPs)
55
56 recruited from the Prevention and Early Intervention Program for Psychoses in London,
57
58
59
60

1
2
3 Ontario, Canada. All participants were in the acute phase of the illness, and often were
4 referred for services when disorganization was noted by the attending clinician. We also
5 recruited an additional 23 healthy controls from the community group matched for age, sex,
6 and parental SES. Exclusion criteria included greater than 2 weeks of lifetime antipsychotic
7 exposure, a diagnosis of affective psychosis, active substance dependence or the presence of
8 intellectual/developmental disorder or the inability to provide informed consent. All
9 participants included in the analysis were native English speakers.

20 2.2 Instruments:

23 2.2.1 The Thought Language Index

24
25 The Thought and Language Index (TLI) (Liddle et al., 2002) is an instrument for assessing
26 formal thought disorder under standardised conditions. Participants are asked to produce
27 three one-minute speech samples in response to 3 photographs from the Thematic
28 Apperception Test (Murray, 1943). Participant responses were scored on 8 domains: poverty
29 of speech, weakening of goal, looseness, peculiar word usage, peculiar sentence structure,
30 peculiar logic, distractability, and perseveration. These domains were also used to compute 3
31 aggregate scores: impoverishment of thinking (the summed scores of Poverty of speech,
32 weakening of goal, and perseveration), disorganization (the summed scores of looseness,
33 peculiar word usage, peculiar sentence construction, peculiar logic, and distractability), and
34 overall thought disorder (the summed scores of all 8 domains).

49 2.2.2 Positive and Negative Syndrome Scale – 8 Item Scale

50
51 The Positive and Negative Syndrome Scale – 8 Item scale (PANSS-8) is an abbreviated
52 version of the 30 Item PANSS clinical assessment of symptomology in schizophrenia and
53 psychosis with acceptable internal consistency and highly correlated with the full PANSS(Lin
54 et al., 2018). The PANSS Item P2 (conceptual disorganization) is a single item measure
55
56
57
58
59
60

1
2
3 intended to assess FTD characterized by “disruption of goal-directed sequencing, e.g.,
4 circumstantiality, tangentiality, loose associations non sequiturs, gross illogicality, or thought
5 block” (Kay, Fiszbein, & Opler, 1987) . Scored on a 1 (absent) to 7 (extreme) Likert type
6 scale, we classified those with scores of 1-3 (absent to mild) as “Low Conceptual
7 disorganization” and those 4-7 (moderate to extreme) as high conceptual disorganization.
8
9

15 2.2.3 Coh-Metrix 3.0

16
17
18 The Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014) system is a web-based
19 automated speech analysis software that computes several word, sentence and passage level
20 linguistic variables from written and spoken speech samples. The software automatically
21 computes a number of lower order (e.g., word counts, frequency of pronoun use, use of
22 connectives) and higher order (e.g., readability, cohesion, text formality) linguistic variables
23 using several common computational linguistics methods including analyzing lexicon, part-
24 of-speech classifiers, syntactic parsers, and through the use of latent semantic analysis
25 (Graesser, McNamara, Louwerse, & Cai, 2004). We based our project on the work in Willits
26 et al (2018), focusing primarily on the use of 5 connectives variables (See Table 1 for
27 details). Connectives were scored based on their frequency per 1000 words spoken.
28
29
30
31
32
33
34
35
36
37
38
39
40

41 2.3 Procedure

42
43
44 This study was approved by the local REB and informed consent was received from
45 participants prior to participation. Clinical interviews, diagnoses and symptom rating scales,
46 including a patient history and demographic assessment were completed by a board certified
47 psychiatrist. Cognitive assessments and the TLI interview and rating were completed by
48 graduate level research assistants. Speech samples were limited to 3 1-minute speech samples
49 from the TLI interview, with researcher speech removed prior to analysis, and were blinded
50
51
52
53
54
55
56
57
58
59
60

1
2
3 to participant status consistent with the procedure described by Sommer et al. (Sommer et al.,
4
5 2010)

6 7 8 2.4 Statistical Analysis 9

10
11 Our analysis was completed in four steps. First, recorded and transcribed speech samples
12
13 were entered in Coh-Metrix 3.0 web tool to calculate the incidence scores (use per 1000
14
15 words) for the each of the 5 connectives variables. We then conducted a principal component
16
17 analysis on the five connective variables to extract factors based on a scree plot that
18
19 cumulatively explained 72.14% of variance in raw scores and compared the selected factor
20
21 scores between FEP and control groups using independent t tests. We then investigated the
22
23 relationship between the factor scores and clinically observed using the TLI and PANSS P2.
24
25 To establish specificity of these relationships, SOFAS and cognitive scores were correlated
26
27 with the factor scores.
28
29
30
31

32 33 3.0 Results 34

35 36 3.1 Demographics 37

38
39 When comparing healthy controls to FEPs, no statistically significant differences existed for
40
41 gender, age, or socioeconomic status (Table 2). As expected, statistically significant
42
43 differences were found between patients and controls on the number of words spoken,
44
45 however no statistically significant differences were found between high and low
46
47 disorganization FEPs on this measure. Additionally, there were expected differences in
48
49 cognitive scores between patient groups and healthy controls on cognitive and clinical
50
51 variables.
52
53
54
55
56
57
58
59
60

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

1
2
3 where no coherent link exists (inappropriate temporal connective use). The resulting of both
4 mechanisms is an increased cognitive load placed on the listener to decipher the intended
5 output of speech (Cain & Nash, 2011). If these efforts at interpreting speech fail, the result is
6 a perception of tangentiality or incoherence, features of positive thought disorder.
7
8
9
10
11
12

13 We have shown that the traditional measures of disorganization (TLI & PANSS-P2)
14 map reliably on to the specific outputs of automated speech analysis in psychosis. Automated
15 speech analysis avoids the cost, time, and reliability issues that impact hand scoring
16 instruments for the assessment of thought disorder, while simultaneously allowing for
17 considerably more comprehensive analyses than classic clinician rated scales. This allows for
18 the examination of multiple linguistic domains of thought disorder with relatively few
19 disruptions to the clinical setting. Further, if Coh-Metrix is able to identify elements of
20 disordered thought reliably during the first episode of psychosis, this approach could be
21 leveraged to identify subtle longitudinal changes in speech that may be imperceptible to
22 clinicians, allowing insights into how thought disorder evolves throughout the course of
23 illness.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 A number of limitations with the current work should be noted. Firstly, speech
40 analyses were gathered from a provided visual stimuli. These stimulus-evoked responses
41 likely involve unique cognitive processes that may not precisely reflect spontaneous language
42 production. Secondly, aberrant connectives use in speech does not necessarily imply
43 internally generated thoughts are not meaningfully connected and this unobservable thought-
44 language gap may differ between patients and healthy controls. Finally, as we collected only
45 a single cross-sectional speech sample during the first episode of psychosis sample, it is
46 unclear if our findings would be stable over the course of illness following exposure to
47 antipsychotic medications, as positive thought disorder is less observable later stages of
48 illness (Roche, Creed, et al., 2015).
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Persistent positive FTD in chronic schizophrenia is linked to aberrant glutamate transmission
4 as well as structural changes in language areas of the brain ¹³. To date, it is not clear if the
5
6 disorganization seen in early stages of schizophrenia is also linked to the same neural
7
8 pathways. One important challenge in elucidating such neural underpinnings is the difficulty
9
10 in separating positive and negative FTD during early stages of psychosis (see Palaniyappan et
11
12 al., 2015³²). Factor analyses of symptoms in early psychosis often indicate the presence of a
13
14 single dimension of Bleulerian symptoms embracing both impoverishment and
15
16 disorganisation (McGorry, Bell, Dudgeon, & Jackson, 1998; Tonna et al., n.d.). The approach
17
18 of using linguistic connectives identifies a distinct positive FTD factor in unmedicated early
19
20 stage of psychosis, thus promising further enquiries to understand the neural underpinnings of
21
22 early positive FTD.
23
24
25
26
27

28 In conclusion, FTD remains a poorly understood aspect of psychotic illness. The practical
29
30 difficulty in capturing robust measures of thought disorder through clinician rated or hand
31
32 scored instruments is an ongoing barrier to research in this area. Our work adds to the body
33
34 of literature promoting automated speech analysis, and argues that the use of connectives in
35
36 evoked speech could objectively identify the specific phenomenon of disorganization
37
38 (positive FTD), independent of the functional and cognitive status of patients experiencing
39
40 first episode psychosis. This raises the question of whether the elusive Bleulerian concept of
41
42 loosening of association can now be reliably tracked across the different stages of psychosis.
43
44
45
46
47 If this becomes feasible, we may inch closer to understanding the neural basis of a core
48
49 deficit in schizophrenia.
50

51 5.0 Funding

52
53
54 This study was funded by CIHR Foundation Grant (375104/2017) to LP; Schulich School of
55
56 Medicine Clinical Investigator Fellowship to KD; AMOSO Opportunities fund to LP; Bucke
57
58 Family Fund to LP; Canada Graduate Scholarship to KD. Data acquisition was supported by
59
60

CONNECTIVES AND THOUGHT DISORDER

13

1
2
3 the Bucke Family Fund, The Chrysalis Foundation and The Arcangelo Rea Family
4
5 Foundation (London, Ontario).
6
7

8
9 6.0 Disclosures

10
11 LP reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian
12
13 Psychiatric Association; book royalties from Oxford University Press; investigator-initiated
14
15 educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted
16
17 work. All other authors report no relevant conflicts.
18
19

20
21 Data is available upon request.
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

For Peer Review

7.0 References

- 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
- 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/npjischz.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](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

1
2
3 experiences and subsequent clinical psychosis: A 10-year study. *The American*
4 *Journal of Psychiatry*, 167(9), 1075–1082.

5
6
7 <https://doi.org/10.1176/appi.ajp.2010.09060883>

8
9
10 Elvevåg, B., Foltz, P. W., Rosenstein, M., & DeLisi, L. E. (2010). An automated method to
11 analyze language use in patients with schizophrenia and their first-degree relatives.
12 *Journal of Neurolinguistics*, 23(3), 270–284.

13
14
15 <https://doi.org/10.1016/j.jneuroling.2009.05.002>

16
17
18 Eussen, M. L. J. M., de Bruin, E. I., Van Gool, A. R., Louwse, A., van der Ende, J.,

19
20
21 Verheij, F., ... Greaves-Lord, K. (2015). Formal thought disorder in autism spectrum
22 disorder predicts future symptom severity, but not psychosis prodrome. *European*
23 *Child & Adolescent Psychiatry*, 24(2), 163–172. [https://doi.org/10.1007/s00787-014-](https://doi.org/10.1007/s00787-014-0552-9)
24
25
26
27
28
29 0552-9

30
31 Graesser, A. C., McNamara, D. S., Louwse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis
32 of text on cohesion and language. *Behavior Research Methods, Instruments, &*
33 *Computers*, 36(2), 193–202. <https://doi.org/10.3758/BF03195564>

34
35
36
37 Gupta, T., Hespos, S. J., Horton, W. S., & Mittal, V. A. (2018). Automated analysis of
38 written narratives reveals abnormalities in referential cohesion in youth at ultra high
39 risk for psychosis. *Schizophrenia Research*, 192, 82–88.

40
41
42
43
44 <https://doi.org/10.1016/j.schres.2017.04.025>

45
46
47 Hart, M., & Lewine, R. R. J. (2017). Rethinking Thought Disorder. *Schizophrenia Bulletin*,
48 43(3), 514–522. <https://doi.org/10.1093/schbul/sbx003>

49
50
51 Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale
52 (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13(2), 261–276.

53
54
55
56 <https://doi.org/10.1093/schbul/13.2.261>

- 1
2
3 Kircher, T., Bröhl, H., Meier, F., & Engelen, J. (2018). Formal thought disorders: From
4
5 phenomenology to neurobiology. *The Lancet Psychiatry*, *5*(6), 515–526.
6
7 [https://doi.org/10.1016/S2215-0366\(18\)30059-2](https://doi.org/10.1016/S2215-0366(18)30059-2)
8
9
10 Kircher, T., Krug, A., Stratmann, M., Ghazi, S., Schales, C., Frauenheim, M., ... Nagels, A.
11
12 (2014). A rating scale for the assessment of objective and subjective formal Thought
13
14 and Language Disorder (TALD). *Schizophrenia Research*, *160*(1), 216–221.
15
16 <https://doi.org/10.1016/j.schres.2014.10.024>
17
18
19 Kuperberg, G. R. (2010). Language in schizophrenia Part 1: An Introduction. *Language and*
20
21 *Linguistics Compass*, *4*(8), 576–589. <https://doi.org/10.1111/j.1749->
22
23 [818X.2010.00216.x](https://doi.org/10.1111/j.1749-818X.2010.00216.x)
24
25
26 Liddle, P. F., Ngan, E. T. C., Caissie, S. L., Anderson, C. M., Bates, A. T., Quedsted, D. J., ...
27
28 Weg, R. (2002). Thought and Language Index: An instrument for assessing thought
29
30 and language in schizophrenia. *The British Journal of Psychiatry*, *181*(4), 326–330.
31
32 <https://doi.org/10.1192/bjp.181.4.326>
33
34
35 Lin, C.-H., Lin, H.-S., Lin, S.-C., Kuo, C.-C., Wang, F.-C., & Huang, Y.-H. (2018). Early
36
37 improvement in PANSS-30, PANSS-8, and PANSS-6 scores predicts ultimate
38
39 response and remission during acute treatment of schizophrenia. *Acta Psychiatrica*
40
41 *Scandinavica*, *137*(2), 98–108. <https://doi.org/10.1111/acps.12849>
42
43
44 McGorry, P. D., Bell, R. C., Dudgeon, P. L., & Jackson, H. J. (1998). The dimensional
45
46 structure of first episode psychosis: An exploratory factor analysis. *Psychological*
47
48 *Medicine*, *28*(4), 935–947.
49
50
51 McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated Evaluation*
52
53 *of Text and Discourse with Coh-Metrix*. Cambridge University Press.
54
55
56 Minor, K. S., Willits, J. A., Marggraf, M. P., Jones, M. N., & Lysaker, P. H. (2019).
57
58 Measuring disorganized speech in schizophrenia: Automated analysis explains
59
60

1
2
3 variance in cognitive deficits beyond clinician-rated scales. *Psychological Medicine*,
4 49(3), 440–448. <https://doi.org/10.1017/S0033291718001046>

7 Morgan, C. J., Coleman, M. J., Ulgen, A., Boling, L., Cole, J. O., Johnson, F. V., ... Levy, D.
8 L. (2017). Thought Disorder in Schizophrenia and Bipolar Disorder Probands, Their
9 Relatives, and Nonpsychiatric Controls. *Schizophrenia Bulletin*, 43(3), 523–535.
10 <https://doi.org/10.1093/schbul/sbx016>

16 Roche, E., Creed, L., MacMahon, D., Brennan, D., & Clarke, M. (2015). The Epidemiology
17 and Associated Phenomenology of Formal Thought Disorder: A Systematic Review.
18 *Schizophrenia Bulletin*, 41(4), 951–962. <https://doi.org/10.1093/schbul/sbu129>

23 Roche, E., Lyne, J., O'Donoghue, B., Segurado, R., Behan, C., Renwick, L., ... Clarke, M.
24 (2016). The prognostic value of formal thought disorder following first episode
25 psychosis. *Schizophrenia Research*, 178(1–3), 29–34.
26 <https://doi.org/10.1016/j.schres.2016.09.017>

32 Roche, E., Lyne, J. P., O'Donoghue, B., Segurado, R., Kinsella, A., Hannigan, A., ... Clarke,
33 M. (2015). The factor structure and clinical utility of formal thought disorder in first
34 episode psychosis. *Schizophrenia Research*, 168(1), 92–98.
35 <https://doi.org/10.1016/j.schres.2015.07.049>

41 Solovay, M. R., Shenton, M. E., Gasperetti, C., Coleman, M., Kestnbaum, E., Carpenter, J.
42 T., & Holzman, P. S. (1986). Scoring manual for the Thought Disorder Index.
43 *Schizophrenia Bulletin*, 12(3), 483–496. <https://doi.org/10.1093/schbul/12.3.483>

48 Sommer, I. E., Derwort, A. M. C., Daalman, K., de Weijer, A. D., Liddle, P. F., & Boks, M.
49 P. M. (2010). Formal thought disorder in non-clinical individuals with auditory verbal
50 hallucinations. *Schizophrenia Research*, 118(1), 140–145.
51 <https://doi.org/10.1016/j.schres.2010.01.024>

- 1
2
3 Tonna, M., Ossola, P., Marchesi, C., Bettini, E., Lasalvia, A., Bonetto, C., ... Ruggeri, M.
4
5 (n.d.). Dimensional structure of first episode psychosis. *Early Intervention in*
6
7 *Psychiatry*, 0(0). <https://doi.org/10.1111/eip.12789>
8
9
10 Wensing, T., Cieslik, E. C., Müller, V. I., Hoffstaedter, F., Eickhoff, S. B., & Nickl-
11
12 Jockschat, T. (2017). Neural Correlates of Formal Thought Disorder: An Activation
13
14 Likelihood Estimation Meta-Analysis. *Human Brain Mapping*, 38(10), 4946–4965.
15
16 <https://doi.org/10.1002/hbm.23706>
17
18
19 Willits, J. A., Rubin, T., Jones, M. N., Minor, K. S., & Lysaker, P. H. (2018). Evidence of
20
21 disturbances of deep levels of semantic cohesion within personal narratives in
22
23 schizophrenia. *Schizophrenia Research*, 197, 365–369.
24
25 <https://doi.org/10.1016/j.schres.2017.11.014>
26
27
28 Ziermans, T., de Wit, S., Schothorst, P., Sprong, M., van Engeland, H., Kahn, R., & Durston,
29
30 S. (2014). Neurocognitive and clinical predictors of long-term outcome in adolescents
31
32 at ultra-high risk for psychosis: A 6-year follow-up. *PloS One*, 9(4), e93994.
33
34 <https://doi.org/10.1371/journal.pone.0093994>
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

8.0 Supplemental Materials

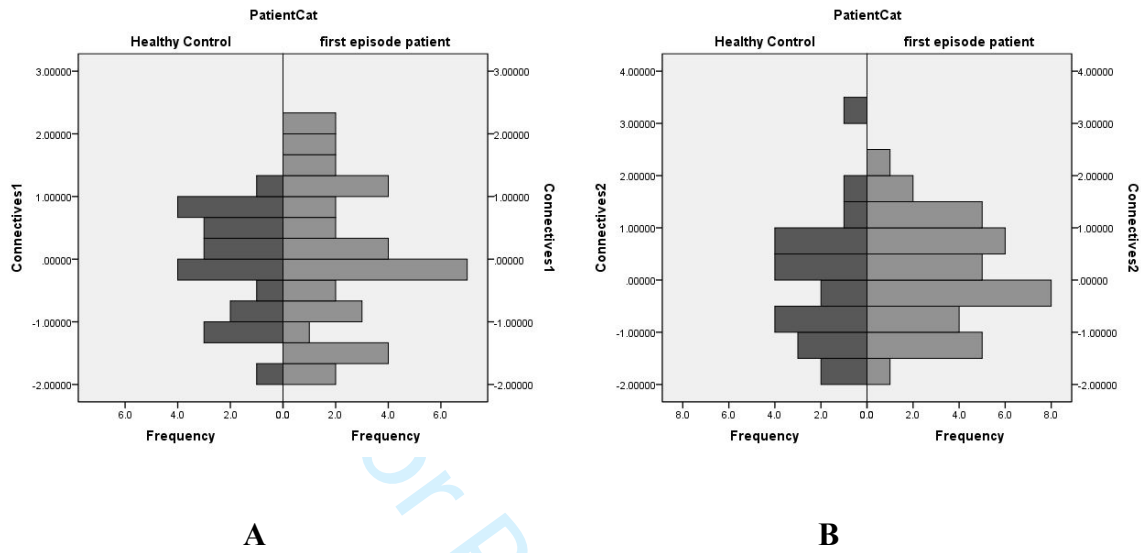


Figure 1. A. Distribution of Connectives Factor 1 scores among healthy controls and first episode patients. **B.** Distribution of Connectives Factor 2 scores among healthy controls and first episode patients.

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)*

For Peer Review

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

	SOFAS <i>R</i> (p-value)	CGI-Severity <i>R</i> (p-value)	DSST <i>R</i> (p-value)	Category Fluency <i>R</i> (p-value)	Trails-B <i>R</i> (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)

For Peer Review