ARTICLE   OPEN

Check for updates

Progressive changes in descriptive discourse in First Episode Schizophrenia: a longitudinal computational semantics study

Maria Francisca Alonso-Sánchez [1,2], Sabrina D. Ford[2,3], Michael MacKinley[2,3], Angélica Silva[2], Roberto Limongi[2] and Lena Palaniyappan [2,4,5 ✉]

Computational semantics, a branch of computational linguistics, involves automated meaning analysis that relies on how words occur together in natural language. This offers a promising tool to study schizophrenia. At present, we do not know if these word-level choices in speech are sensitive to the illness stage (i.e., acute untreated vs. stable established state), track cognitive deficits in major domains (e.g., cognitive control, processing speed) or relate to established dimensions of formal thought disorder. In this study, we collected samples of descriptive discourse in patients experiencing an untreated first episode of schizophrenia and healthy control subjects (246 samples of 1-minute speech; $n = 82$, FES = 46, HC = 36) and used a co-occurrence based vector embedding of words to quantify semantic similarity in speech. We obtained six-month follow-up data in a subsample (99 speech samples, $n = 33$, FES = 20, HC = 13). At baseline, semantic similarity was evidently higher in patients compared to healthy individuals, especially when social functioning was impaired; but this was not related to the severity of clinically ascertained thought disorder in patients. Across the study sample, higher semantic similarity at baseline was related to poorer Stroop performance and processing speed. Over time, while semantic similarity was stable in healthy subjects, it increased in patients, especially when they had an increasing burden of negative symptoms. Disruptions in word-level choices made by patients with schizophrenia during short 1-min descriptions are sensitive to interindividual differences in cognitive and social functioning at first presentation and persist over the early course of the illness.

INTRODUCTION

Language disorganization is a prominent feature in psychosis, and it is commonly observed initially as a disorder in generating interpersonal discourse. This produces a significant functional impairment for the patient as it interferes with one's ability to describe or explain attributes and thus socialize in everyday life[1]. When engaged in a descriptive discourse of a concrete referent, such as a picture, to a second person, patients with schizophrenia make unusual word choices[2], exhibit repetitiveness and convey less information (referred to as 'weakening of goal'[3] or 'poverty of content'[4]) than healthy controls[3,5]. In particular, the restricted repertoire of word selection, characterized by smaller loops of word-to-word connectivity that occurs with more proximal repeats in selected words, becomes apparent even before overt psychosis[6], predicts later onset of psychosis[6,7], becomes more pronounced during the first episode[7], and relates to reduced social and occupational functioning[8].

Descriptive discourse involves multiple levels of cognitive processing[9] to integrate parts and attributes of the whole to produce a descriptive schema[10]. We often employ descriptions in the service of rhetorical functions (i.e., ways to inform, argue, persuade someone) through our choice of words. In psycholinguistic terms, descriptive discourse requires semantic competence[1] and appropriate lexical access to a connectionist system of words organized by their conceptual relationships with one another[10]. In this context, lexical units (words) with a higher likelihood of occurring together have a stronger connection or a smaller distance between them (distributional semantics)[11]. This idea follows the original spreading-activation hypothesis of lexical

representations in the brain[12]. Competitive theories of lexical selection assume that lexical representations must overcome interference from the neighbour's activation through lateral inhibition[13]. Applying this to the picture description task, a failure of appropriate selection via inhibition at the lexical level may give rise to a description that is replete with words that are highly associated with each other, without capturing the different attributes of the picture at hand.

A proactive 'top-down' contextual guidance during discourse can reduce the overreliance on the bottom-up activation of the lexico-semantic network for word selection[14]. A breakdown in this contextual guidance, implemented as top-down inhibition from inferior frontal to semantic storage systems[15], has been variously described in schizophrenia[16]. A large body of literature demonstrates frontal cognitive control deficits in schizophrenia, exemplified by reduced performance in the colour-word Stroop Task that tests one's ability to inhibit competing semantic categorical representations when choosing a word[17]. In particular, the increased Stroop interference effect, in both response time and accuracy measures, has been interpreted as a marker of impaired inhibitory feature of cognitive control[17]. Abnormalities in this aspect of cognitive control have been previously related to conceptual disorganization[18], a symptom related to linguistic aberrations in schizophrenia[19,20]. In addition, inter-individual variations in processing speed also influences lexical access[21]. In fact, reduced processing speed is the neurocognitive domain with the strongest correlation with disorganisation[22,23]. On this basis, we can expect deficits in cognitive control and processing speed

[1]CIDCL, Fonoaudiología, Facultad de Medicina, Universidad de Valparaíso, Valparaíso, Chile. [2]Robarts Research Institute, Western University, London, ON, Canada. [3]Lawson Health Research Institute, London, ON, Canada. [4]Douglas Mental Health University Institute , McGill University, Montreal, QC, Canada. [5]Department of Psychiatry, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada. ✉email: lpalaniy@uwo.ca

to influence word selection during a descriptive discourse in patients with schizophrenia.

When examining similarity among the words used during discourse, there are broadly 2 approaches. One approach is to count the instances of repetition of a word. This phenomenon is described as perseveration in clinical rating scales[3,4]. A measure of lexical diversity called Type-Token Ratio (TTR; the ratio of unique to total words in a text) is computed based on such repetitions. As exact repetitions are relatively rare, perseveration is often not detectable in cross-sectional interviews[24,25], and results from TTR studies are inconclusive[22–25] with more recent studies showing both increased[26] and reduced[27,28] TTR in schizophrenia. Graph theoretical approaches that rely on the proximity between two repetitions, rather than counting the instances of repetitions, appear to carry more diagnostic and prognostic information in schizophrenia[8,29,30]. However, this approach cannot distinguish meaningful repetitions of informational value (e.g., "He liked the idea of travel, and the memory of travel, but not travel itself" [—— Julian Barnes, Flaubert's *Parrot*]) from the problematic repetitions that affect communication. The second approach is to employ distributional semantics to estimate the similarity, rather than exact repetition, among a set of words. This taps on a network-based distributional model of words. If lexical units are interconnected based on their co-occurrence in everyday language, then similarity among a set of words used during a discourse can be quantified based on this distributional co-occurrence.

Approaches from distributional semantics have been applied to study the relationship among words produced during various speech elicitation tasks in schizophrenia. The most popular approach, introduced by Elvevåg[31], involves the use of latent semantic analysis (LSA) that taps on the document-level statistical co-occurrence of words in a large corpus of written texts; this determines their position in the semantic space based on the "company they keep". The cosine similarity of this spatial index can then be computed among the words spoken by a patient. Several studies have demonstrated the potential utility of distributional semantics in predicting the onset of psychosis[2,32,33], examining thought disorder[34–36] and its neuroanatomical basis of linguistic disruptions in psychosis[37]. Other similar methods improved on LSA, by weighting the statistics of co-occurrence based on the actual proximity of words in the sentences occurring in the reference corpora[38–44]. We employ one such improved method (CoVec), that has been used previously in the study of semantic fluency tasks in schizophrenia[45,46].

Cosine similarity can be computed between words that are adjacent to each other within a frame, indicating if words proximal to each other are sampled from a narrow semantic space[43–46]. Cosine similarity among the full frame of words in a descriptive text (termed Mean Similarity in CoVec) indicates the semantic diversity of all words employed to provide the complete description of a referent. As spoken text rarely assumes the form of sentences, a finite moving window (e.g., 5, 10 or 20 words size[45–48]) is also used to define frames of measurement. In our case, the full 1-minute description of a picture constitutes the frame of interest (ASW-F or Average Similarity of Words in Full Frame) to define semantic similarity, with the average similarity estimated from a 10-word moving window (ASW-10) as a secondary measure.

Studies employing distributional semantics have often used the term coherence to describe the degree of similarity (e.g. local coherence[4], semantic coherence[31], or cohesion[49]) or incoherence when describing its pathological reduction[34,44] (see[38,50] for a review). While several NLP studies have employed the term coherence in this sense, we use the term 'similarity' rather than coherence when employing cosine similarity. Hoffman pointed out that coherence is a psychological experience of a listener and not a property of a text[51]. To experience a text as coherent, the listener must employ a subjective interpretive synthesis that depends on their experience of the referent (i.e., drawing the linkage between the described object and the presented text) and directionality (i.e. which word or idea came first), in addition to the dependency among the lexical/semantic units. Furthermore, words with a low probability of co-occurrence can be coherently juxtaposed in certain contexts, that may not be apparent from the text itself. Also, metadiscursive (frameshifting[51]) elements can improve coherence for a listener (e.g., changing topics by saying "to go on a tangent for a bit"). For these reasons, we do not infer semantic *coherence* but only *similarity* from the indices of distributional semantics employed here.

We hypothesize that when faced with the task of describing an unfamiliar concrete referent[52] (a picture), patients with schizophrenia will employ words with a higher probability of semantic co-occurrence. We expect abnormal semantic similarity to be evident in the untreated, first episode phase of illness and relate to formal thought disorder, reduced cognitive control and processing speed in patients. To test if the abnormality in semantic similarity was specific to the picture description task, wherein the word choices we make depend on the descriptive nature of discourse, we studied similarity of word choices in a conventional category fluency task. We will also address several confounds such as years of education[53], migrant status, parental socioeconomic status, bilingualism[54] and antipsychotic use (especially those with high occupancy of dopamine D2 receptors)[55] that are critical for the current study as they typically influence schizophrenia prognosis[56].

Several previous cross-sectional studies have related language and communication difficulties to social functioning among patients[57,58]. Interestingly, studies investigating longitudinal changes of language remains scarce in psychosis[59], even though worsening of formal thought disorder over time has been shown to relate to progressive worsening of social and occupational outcome[60]. Furthermore, exposure to antipsychotics, that occurs when treatment is initiated in FES, is also associated with worsening of speech measures, especially word selection measures[55]. We anticipate that, unlike healthy controls who will show no changes in their word-level choices over the time, a persistent or worsening deficit in semantic similarity over time will be seen among FES patients.

To this end, we recruited a sample of acutely unwell, first-episode patients with < 14 days of lifetime exposure to antipsychotics at baseline. These patients were then treated in an early intervention clinic and followed up after 6 months to examine their discourse stability. This allowed us to relate treatment variables (antipsychotic exposure) as well as outcome variables (SOFAS scores) to word similarity measures over time.

RESULTS

Demographic and clinical characteristics

Healthy controls and the FES group (First Episode Schizophrenia) did not significantly differ in age, gender distribution or educational level. In the FES group, 20% of the participants were first-generation immigrants (determined from self-report) while 30% of the matched HC group were first-generation immigrants. There was no group difference in the use of English as the first language (82% FES and 88% HC had English as the first language). All the participants had English as their only transactional language. As expected, the HC group performed better on a modified digit-symbol substitution task (DSST) measuring processing speed and the Colour-Word Stroop task. Clinical and demographic characteristics are provided in Table 1. In the FES group, 50% of the sample were fully antipsychotic naïve while the other 50% were exposed to a mean of 2.8 days of a lifetime daily dose to antipsychotics. Of those in the FES sample exposed to

Table 1. Clinical and demographic characteristics of the sample at baseline.

	HC Mean ± SD	FES Mean ± SD	BF$_{10}$	Effect size δ 95% CI
Age	21.4 ± 3.2	22.0 ± 3.6	0.308	−0.56, 0.24
Gender	67% male	77% male	0.509	−1.48, 0.46
Educational level (<12/>12 years)	27%/73%	37%/63%	0.474	−1.41, 0.46
PANSS-8 Positive	–	12.1 ± 3.0	–	–
PANSS-8 Negative	–	7.4 ± 4.3	–	–
PANSS-8 total	–	25.6 ± 6.8	–	–
SOFAS	80.2 ± 10	39.3 ± 13.3	>10000	
Parental SES (< 3/ >3)	42% / 58%	33% / 67%	0.387	−0.55, 1.34
CDS	–	3.5 ± 3.3	–	–
CGI	–	5.2 ± 0.9	–	–
TLI total	0.28 ± 0.3	1.60 ± 1.3	>10000	−1.65, −0.69
TLI Disorganization of Thinking	0.153 ± 0.2	1.01 ± 1.1	674	−1.38, −0.45
TLI Impoverishment of Thinking	0.13 ± 0.2	0.58 ± 0.7	41.4	−1.17, −0.27
TLI Dysregulation	0.06 ± 0.16	0.17 ± 0.29	1.69	−0.85, −0.00
DSST	68.6 ± 11.3	52.8 ± 13.9	>10000	0.66, 1.63
Semantic Verbal Fluency	26.6 ± 6.9	19.8 ± 6.2	646	0.47, 1.45
Stroop total correct	78.2 ± 3.1	70.8 ± 13.1	19.93	0.22, 1.33
Stroop total time	74.6 ± 11.3	84.8 ± 17.0	11.12	−1.07, −0.17
Stroop IG	8.89 ± 1.5	7.09 ± 3.5	12.2	0.14, 1.02
Daily dose	–	0.81 ± 0.49		
Total dose		160.7 ± 110		

Mean and Standard deviations are shown for continuous variables, with percentages for categorical variables. BF$_{10}$: Bayes Factor. *SOFAS* Social and Occupational Functioning Assessment Scale, SES: Parental socioeconomic status score. *CDS* Calgary Depression Scale, CGI-S Clinical Global Impressions Scale Severity of Illness, *TLI* Thought and Language Index, Impoverishment: Poverty of Speech + Weakening of Goal; Disorganized Thinking: Peculiar words + sentences + illogicality; Dysregulation: Perseveration + Distractibility. *DSST* Modified Digit Symbol Substitution Test. Stroop IG: Stroop interference score - Golden method. Daily dose: average Daily Defined Dose, Total Dose: total exposure calculated based on Daily Dose and number of days of exposure.

Table 2. Summary group differences at baseline.

	HC Mean ±SD	FES Mean ±SD	BF$_{10}$	Effect size δ 95% CI
Number of words	70.6 ± 14.9	68.4 ± 30.3	0.249	−0.32, 0.48
ASW-F	0.334 ± 0.025	0.352 ± 0.034	6.53	−1.05, −0.17
ASW-10	0.400 ± 0.023	0.421 ± 0.031	32.76	−1.14, −0.25

ASW-F Average similarity of words – full picture description, *ASW-10* Average similarity of words – 10 words moving window. Note that the variables reported here are individually averaged across 3 speech samples per subject. BF$_{10}$: Bayes Factor (alternate vs. null hypothesis).

antipsychotics, 50% were on antipsychotics with low dopamine occupancy and the other 50% were on antipsychotics with high dopamine occupancy (as defined by de Boer and colleagues[55]).

Baseline differences in word similarity

In the description task, the groups did not differ in the number of words spoken but FES had higher similarity (ASW-F, BF$_{10}$ = 6.53; ASW-10, BF$_{10}$ = 32.76) compared to the HC group. These results are shown in Table 2 and Fig. 1. The increase in semantic similarity was specific to the picture description task; when we studied similarity of word choices in a category fluency task in a subsample of subjects (HC $n = 33$, FES $n = 39$), there was no difference among groups (ASW-F, HC: 0.497 ± 0.04; FES: 0.477 ± 0.05, BF$_{10}$ = 0.696), indicating discourse-related specificity of increased semantic similarity in schizophrenia.

Longitudinal changes in word similarity

In the 6-month follow-up sample ($n = 33$, FES = 20, HC = 13), the 2 groups were matched for age (FES: 22.5 ± 5.0; HC: 21.5 ± 3.1, BF$_{10}$ = 0.390) and gender (FES: 80% male; HC: 70% male, BF$_{10}$ = 0.611). The follow-up sample of patients had no more than anecdotal evidence of differences at baseline with the dropped-out sample (PANSS BF$_{10}$ = 0.302; TLI BF$_{10}$ = 0.327; DSST BF$_{10}$ = 1.699; ASW-F BF$_{10}$ = 1.718). Patients with FES had strong evidence for functional improvement based on SOFAS scores (Baseline: 41.5 ± 13.5; follow-up: 61.0 ± 12.9; mean change = 19.5 ± 14.3; paired t test BF$_{10}$ = 4868), and clinical improvement based on a reduction in PANSS-8 total score (Baseline: 25.2 ± 5.7; Follow-up: 15.1 ± 5.0,

mean change = -10.25 ± 4.9; paired t test BF$_{10}$ > 10000) from baseline to follow-up assessment, as expected following clinical intervention (medication doses detailed below). While average positive symptom scores improved (Baseline: 12.5 ± 2.6; Follow-up: 5.2 ± 1.7, BF$_{10}$: > 10000), the average negative symptom scores of the PANSS did not show a notable change between baseline and the follow-up (Baseline: 6.8 ± 3.7; Follow-up: 7.1 ± 4.1, BF$_{10}$: 0.255), indicating the persistent nature of this core feature of schizophrenia.

To study the longitudinal trajectory of word usage during descriptive discourse, we performed a Bayesian paired t-test from baseline to 6-month follow up in both groups. As shown in Table 3, the null model was more likely than the difference-between-measures model for the HC group across all linguistic variables, indicating relative stability of semantic similarity and the number of produced words among healthy subjects, when the same pictures were described twice in a period of ~6 months. In the FES group, the most notable difference between measures was seen in semantic similarity which was estimated from the full 1-min picture description (ASW-F; BF$_{10}$ = 6.3; see Fig. 1). We did not see the same level of evidence for linear change in ASW-10 or the number of words. For further correlational analysis with cognitive and symptom factors, we selected ASW-F as the linguistic measure of interest.

Symptoms, functioning, and word similarity

Among FES subjects, ASW-F at the time of illness onset was higher in the presence of more severe positive symptoms (PANSS-8 positive r: 0.39, BF$_{10}$: 9.24) and reduced functioning (SOFAS scores r: −0.41, BF$_{10}$: 128) but this relationship was not seen with PANSS-8 negative (r: 0.08, BF$_{10}$: 0.18) scores, TLI impoverishment (r: 0.21, BF$_{10}$: 0.49), disorganization (r: 0.14, BF$_{10}$: 0.28) or dysregulation (r: −0.06 BF$_{10}$: 0.20) scores (Fig. 2). Among FES subjects that were followed-up, there was moderate evidence for increasing ASW-F in patients with increasing PANSS-8 negative (r: 0.592, BF$_{10}$: 18.7) but not with change in PANSS-8 positive (r: −0.125 BF$_{10}$: 0.435), or SOFAS scores (r: −0.04 BF$_{10}$: 0.322).

Cognition and word similarity

When all subjects (patients and controls) at the baseline were considered together, ASW-F was higher in subjects with reduced Stroop accuracy (r: −0.31, BF$_{10}$: 13.3). The within-group effects were weaker, but in the same direction (FES only: r: −0.22, BF$_{10}$: 1.01; HC only: r: −0.29, BF$_{10}$: 1.61). Higher ASW-F scores also related to a lower Stroop Interference score (of Golden: IG) (r: −0.29, BF$_{10}$ of 8.24,; FES only: r: −0.20, BF$_{10}$: 0.81; HC only: r: −0.25, BF$_{10}$: 1.13) and prolonged reaction time for the Stroop incongruent condition (r: 0.29, BF$_{10}$: 8.6; FES only: r: 0.28, BF$_{10}$: 1.97; HC only: r: 0.06, BF$_{10}$: 0.29). This indicates that semantic co-occurrence in discourse production was higher in the presence of a cognitive control deficit indexed by reduced inhibitory control (poor accuracy) and information processing speed. A more specific
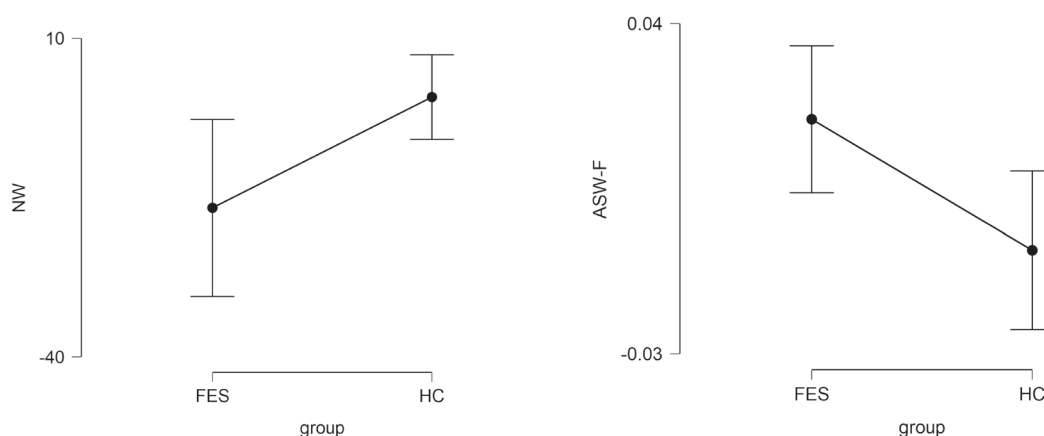
Baseline
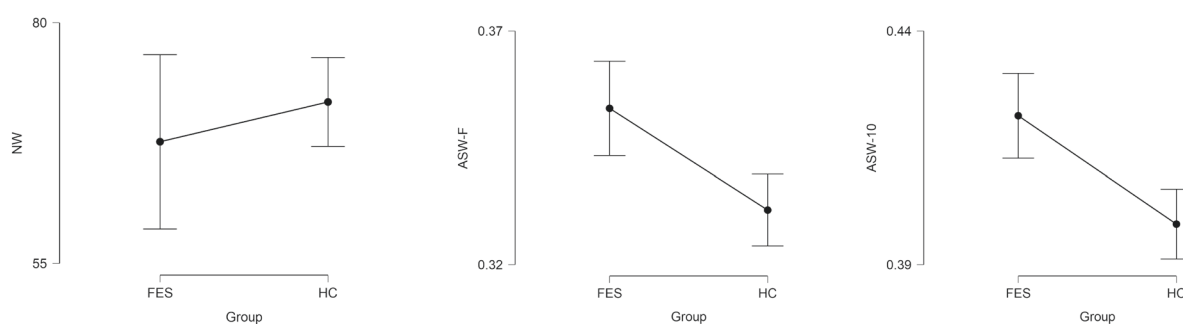
Linear change

Fig. 1 Group differences in linguistic variables at baseline and the change over time of linguistic variables. Descriptive plots of 95% credible interval between groups. NW Number of words, ASW-F Average Similarity of Words in Full picture description, ASW-10 Average Similarity of Words over moving window of 10 words, FES First Episode Schizophrenia, HC Healthy control.

index of serial processing speed derived from a modified Digit Symbol Substitution Test was also lower in the presence of increased ASW-F across the entire sample (r: −0.48, BF_{10}: 304). This association was largely driven by the FES group (r: −0.41, BF_{10}: 7.99), and not the HC (r: −0.03, BF_{10}: 0.21) (see more details in the supplementary materials).

Effects of antipsychotics exposure
We did not observe differences in ASW-F between the anti-psychotic naïve, or low and high D2 receptor occupancy medication sub-samples at baseline (ANOVA, $BF_{10} = 0.239$), or between patients taking low and high occupancy drugs by the time of follow-up (t-test, $BF_{10} = 0.607$). To investigate possible dose effects of antipsychotics, we related both the Daily dose (average Daily Defined Dose) and Total Dose (total exposure calculated based on Daily Dose and number of days of exposure) to number of words and ASW-F at both time points. As shown in Table 4, the difference between the baseline and follow-up measures on number of words and ASW-F were not correlated with Daily Dose or Total Dose.

Effect of social factors on word similarity
To investigate possible effects of immigrant status and the use of a language other than English at home[56], we removed 20% of subjects that satisfied this criterion and analyzed the difference in ASW-F at baseline. We continued to see evidence in favour of increased semantic similarity among patients with FES (ASW-F

$BF_{10} = 6.46$). Similarly, when patients were stratified according to education status (<12/>12 years) and by parental socioeconomic status (higher than median vs. lower than median) and were compared with each other, there was no difference in ASW-F or number of words (Educational background: ASW-F $BF_{10} = 0.594$; number of words $BF_{10} = 0.173$; Socio-economic status: ASW-F $BF_{10} = 0.194$; number of words $BF_{10} = 0.148$). These results indicate that word similarity is affected by the diagnosis of schizophrenia, rather than social factors that are often associated with the diagnosis.

DISCUSSION
Using a computational semantics approach, we examined word similarity during a controlled descriptive discourse task in untreated first-episode schizophrenia at baseline and after 6 months of treatment. We report four major findings. First, when faced with the task of describing an unfamiliar concrete referent (a picture), patients with schizophrenia choose words with a higher probability of semantic co-occurrence. The likelihood of this phenomenon is more pronounced when psychotic symptoms are severe and functional deficits are profound. Interestingly, this objectively verifiable linguistic feature of higher similarity is seen irrespective of the degree of clinically detectable thought disorder. Second, higher word similarity during the discourse was related to lower cognitive control (in the whole sample), as indexed by the Stroop task, and reduced processing speed (especially in patients), indicating a role for domain-general processes in aberrant word

Table 3. Summary of baseline and follow-up 6 months comparison.

	HC					FES					Comparison
	Baseline	6 months	Paired BF$_{10}$	Linear change	δ 95% CI	Baseline	6 months	Paired BF$_{10}$	Linear change	δ 95% CI	BF$_{10}$ linear change *groups
	Mean ± SD	Mean ± SD		Mean ± SD		Mean ± SD	Mean ± SD		Mean ± SD		
Number of words	69.2 ± 13.9	70.0 ± 12.4	0.28	0.76 ± 11.0	−5.88, 7.41	66.9 ± 30.0	52.1 ± 19.8	1.74	−16.62 ± 29.6	−30.50, −2.750	0.13
ASW-F	0.332 ± 0.02	0.324 ± 0.01	0.44	−0.008 ± 0.028	−0.025, 0.009	0.337 ± 0.02	0.353 ± 0.03	2.07	0.020 ± 0.033	0.004, 0.035	6.32
ASW-10	0.398 ± 0.02	0.391 ± 0.02	0.38	−0.007 ± 0.028	−0.024, 0.010	0.407 ± 0.02	0.414 ± 0.02	0.61	0.011 ± 0.026	−0.001, 0.023	2.37

NW Number of words, *ASW-F* Average Similarity of Words in Full picture description, *ASW-10* Average Similarity of Words over moving window of 10 words. BF$_{10}$: Bayes Factor. δ 95% CI Effect Size 95% credible interval.

choices in schizophrenia. Third, the higher semantic similarity in patients was only present in the discursive task and not in the verbal fluency task. Four, despite symptomatic improvement with treatment (i.e., reduction of positive symptoms), the aberrant semantic similarity persisted with time, worsening especially in those with increasing burden of negative symptoms, but this was not explained by exposure to antipsychotics. Taken together, restricted sampling from the putative semantic space during a discursive discourse is likely to be a specific, persistent deficit in early stages of schizophrenia that follows the trajectory of negative symptoms.

Semantic impairment in people with schizophrenia is widely reported[61], however, this evidence relies mostly on comprehension based experimental paradigms[62–64] or experiments where the semantic retrieval demand, or route in the semantic space, is set by the researchers (stimulus with prime and target) and not chosen by the participants. Studies of the latter type generally involve category fluency tests, wherein patients have either no reduction in overall word similarity or choose adjacent words that are less similar[4,65]. In contrast to verbal fluency tasks, in a discursive task, there is a necessity to 'forage' widely to accomplish the goal of description. Such wide foraging appears to be diminished in schizophrenia[66]. We also note that such a narrowing of semantic sampling space relates to a higher Stroop interference effect; thus, a failure of the prefrontal executive control, either in a general- or domain-specific manner[67], may influence word choice. The lack of control in the selection of the lexical items may lead to a restricted repertoire wherein a word and its activated associates[68,69] dominate the unfolding discourse.

Contrary to our expectation, we did not find a relationship between semantic similarity and severity of formal thought disorder in this sample of FES. In general, the degree of shared variance between computational linguistic measures and observer-rated formal thought disorder scores have not been consistent[29,42,52,70–72]. In particular, while some sentence level structural measures (e.g., connective use[70], narrativity and referential cohesion[73]) relate to thought disorder, the overall shared variance is small for word-level measures[73]. This also supports the view that semantic similarity (i.e., the distance among words inferred from distributional semantics) is a latent variable; pathological changes in semantic similarity are not immediately discernible in a clinical interview, even when qualitative word peculiarities are sought from transcribed speech. Nevertheless, greater variance in clinical ratings may be required to conclusively study this issue[44].

Our study has several strengths as well as certain limitations. To our knowledge, this is the first longitudinal report on the nature of word choices made during a controlled discourse in patients with psychosis. Although the evolution of lexical and semantic deficits in schizophrenia is still not fully understood, meta-analytical evidence indicates no temporal change when category fluency is tested -indicating its fixed, endophenotype-like stability over time[74]. In contrast, we report that discourse-specific word choice deteriorates over time in the early stages of schizophrenia. Secondly, we estimated antipsychotic exposure meticulously over the follow-up period. The discourse-related word similarity did not change in proportion to antipsychotic dose exposure, in contrast with the reported influence of antipsychotic dose on other NLP measures such as syntactic complexity and percentage of time speaking[55]. We were limited in terms of the number of healthy controls for whom we had follow-up assessment of word similarity; nevertheless, this did not diminish our ability to demonstrate group differences in the longitudinal change scores based on within-subject variance. Secondly, our descriptive discourse was constrained by time; we do not know if the choice of words would have been less similar if the discourse was unconstrained and spontaneous. This needs to be examined in future studies with speech elicited in different contexts. Lastly, our
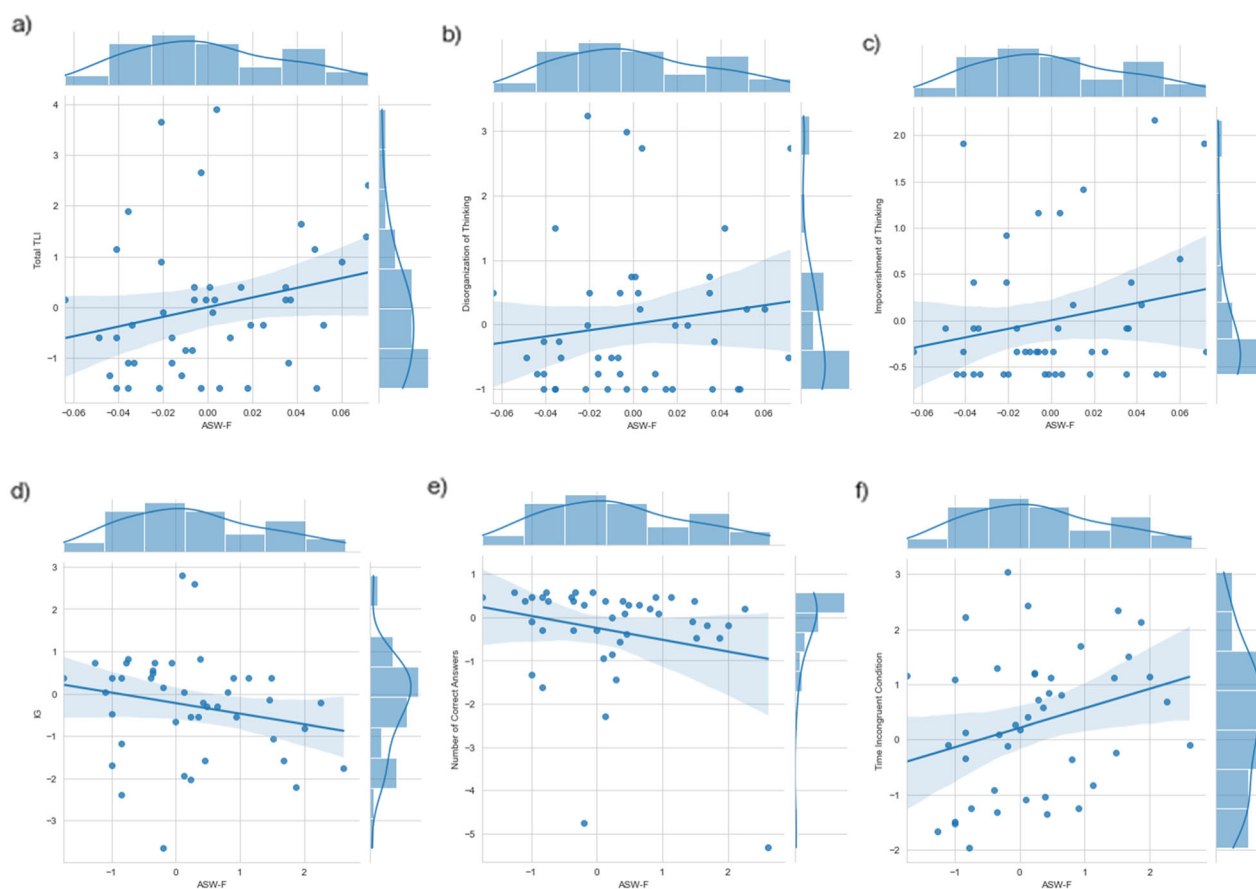
Fig. 2 Correlation between ASW-F, TLI symptoms and Stroop scores in the patient group at baseline. ASW-F Average Similarity of Words in Full picture description with TLI (Thought Language Index) scores **a**) Total, **b**) Disorganization of thinking subscore and **c**) Impoverishment of thinking subscore; and with Stroop **d**) IG: Interference score, **e**) Number of correct answers and **f**) Response time incongruent condition.

Table 4. Relationship between 6-months change in linguistic variables and medication dose.

	Pearson's r	BF_{10}	Lower 95% CI	Upper 95% CI
Daily Dose - NW	0.105	0.303	−0.330	0.491
Daily Dose- ASW-F	−0.161	0.343	−0.530	0.283
Total Dose - NW	0.083	0.293	−0.348	0.475
Total Dose- ASW-F	−0.225	0.424	−0.574	0.227

Daily dose = average Daily Defined Dose, Total Dose: total exposure calculated based on Daily Dose and number of days of exposure. *NW* Number of words, *ASW-F* Average Similarity of Words in Full picture description, *CI* credible intervals.

sample of first-episode schizophrenia did not include the most unwell patients (not referred by clinicians) and those who were involuntarily hospitalised (deemed to lack capacity to consent) and drop-outs were substantial. While the patients who were unavailable for follow-up had a similar profile to those who were retained, we cannot rule out the possibility that they had better outcomes; we urge caution in generalising our results to this group.

In conclusion, we demonstrate that descriptive discourse in first episode of schizophrenia is characterized by an aberrantly high semantic co-occurrence that relates to functional deficits at initial presentation and persists despite treatment in the early stages. Given its relevance to social functioning, and our ability to

measure it objectively in a non-invasive, repeated manner, we propose this measure to be a suitable computational linguistic measure that indexes one aspect of the hitherto unclear but persistent pathophysiology of schizophrenia.

METHODS

Participants

Eighty-two English-speaking participants were recruited, including 46 experiencing their First Episode of Schizophrenia (FES) and 36 healthy controls (HC). FES participants were enrolled through the Prevention and Early Intervention for Psychosis Program of London Health Sciences Centre (London, Ontario, Canada) and were diagnosed with Schizophrenia according to the DSM-5 criteria, using the consensus procedure described by Leckman and the Structured Clinical Interview for DSM-5 to confirm diagnosis 6 months after the first presentation[75]. The severity of symptoms was confirmed with the Positive and Negative Syndrome Scale-8 items version (PANSS)[76]. The FES participants were in the acute phase of the illness and drug-naïve for antipsychotics at the time of the first assessment with a maximum equal to or less than 14 days of total lifetime antipsychotic use. We used a consecutive referral strategy for patient recruitment whereby all patients referred to the only first episode clinic in the catchment area between April 2017 and June 2019 were approached, if deemed to have the capacity to consent for the study by the clinicians.

We also recruited a HC group from the same geographical catchment as patients, through pamphlets and word-of-mouth advertisement. Healthy subjects were group-matched with FES for age, sex, level of completed formal education and parental socio-economic status. The inclusion criteria for HC group included no personal or family mental illness or neurological diseases, prior or current antipsychotic exposure, active substance dependence or the inability to provide informed consent.

All participants provided written informed consent before assessment and ethics approval was granted by the Human Research Ethics Board at Western University, London, Ontario.

Thirty-three participants, 20 with schizophrenia (SZ) and 13 HC, were followed up approximately 6 months from the first assessment ($\bar{x}$ = 214.9 ± 44.9 days). The medication exposure of the FES group was calculated according to the Daily Defined Dose (DDD) methodology[77], and D2-occupancy based classification followed the description of de Boer and colleagues[55]. To calculate total exposure, we considered the type of medication, the dose prescribed, the number of days of effective exposure based on treatment compliance over the follow-up time measured using an established instrument[78] for adherence that correlates well with pill counts[79]. As reported in our prior study[80], nearly 50% of patients went on long-acting injection by the 1st month of treatment, further ensuring treatment compliance.

Assessments

All participants were assessed with the Social and Occupational Functioning Assessment Scale (SOFAS) to quantify the level of functioning in social and occupational domains, without overlapping with symptom measurements[81] and with the Socioeconomic Status (SES) to measure the parental level of occupation and employment from 1 (Managerial and professional occupations) to 5 (routine occupations)[82]. The FES group was assessed with the Calgary Depression Scale (CDS)[83] covering depressive symptoms over the past 2 weeks and with the Clinical Global Impression Scale Severity of Illness (CGI-S) to assess the overall severity from 1 (normal) to 7 (among the most extremely ill patients)[84].

Participants were assessed using a modified version of the digit symbol substitution task (oral and written version) used in our previous studies[22,85,86], semantic verbal fluency task in its original version and the Colour-Word Stroop test in an adapted version used in other studies[87,88]. The DSST oral and written versions were scored by counting the number of correct symbols within the allowed time, with the total DSST score being calculated by averaging the oral and written version scores. For the fluency task, participants were instructed to generate as many words as possible within one minute from the semantic category of animals, and the metric of average similarity across the full set of response was measured using CoVec (see next section). In the Stroop test, the performance was measured by the number of correct answers, the response time in incongruent conditions and the Interference score (IG). The IG was computed with Golden method[89], in which we calculated the number of correctly named items in each condition: Word score = number of words read correctly, Colour score = number of colour hues named correctly, and Colour-Word score = number of colour hues named correctly. Then we estimated the Predicted Colour-Word score with the product of the Word and Colour scores with the following formula: Predicted Colour-Word score = (Word score x Colour score) / (Word score + Colour score). Finally, the interference score (IG) was computed subtracting the Predicted Colour-Word score from the actual number of correctly named items in the Colour-Word incongruous condition[90].

The discourse task was the description of 3 images and the scoring was done using the Thought Language Index (TLI). The TLI is a reliable instrument to assess formal thought disorders under standardized conditions[3]. The participants were asked to describe Thematic Apperception Test[91] images and were given one minute for each image. The interviewer prompted the participants to continue if they stopped speaking before the stipulated time. The interview was recorded and later transcribed by research assistants. The transcriptions were then analyzed with the Covington Vector semantic tool[92].

Semantic Analysis

The Covington Vector semantic tool (CoVec) is a natural language processing tool based on data from the Global Vectors for Word Representation (GloVe) Project, with 840 billion words in English on a 300-element vectors[93]. GloVe measures the likelihood of co-occurrence of words through vector cosine similarity based on overall statistics of how often the word appears given the context (P(w | c)). The GloVe project is a count-based model with a large matrix of (words*context) co-occurrence information that is normalized by log-smoothing the matrix. Covec reports the average of similarity, that is, whether successive words are commonly used in the same context (or together), with an n-word frame segment, using all the positions of the frame. Before processing the text, CoVec removes punctuation, marks 'stop words' (eg. "a", "the", "is", "at", among

others), and finally, ignores words that are not found in the GloVe dataset (displays a warning of all the missing words). The metrics used include the Number of words (NW), average similarity of words in the full-frame of the text (ASW-F) or in 10 words moving window (ASW-10). Note that ASW is described as Coherence in CoVec's output.

Data analysis

Clinical and demographic data were analyzed using descriptive and Bayesian statistics (Bayesian t-test for continuous variables and Bayesian Chi-square between categorical variables). We first compared group performance with a Bayesian t-test on the number of words and semantic similarity variables. To compare the progression of language features, we conducted a Bayesian paired t-test between baseline and 6-month follow-up measures, then, we estimated the linear change between measures and compared it between groups. We used a Bayesian ANOVA to explore the differences between the types of medication in the FES sample. We conducted a Bayesian Pearson correlation to explore the effect of antipsychotics on our language variables. To address the interaction with cognitive and symptom variables, a Bayesian correlation was conducted between semantic co-occurrence and Stroop, DSST, TLI and PANSS scores. The variables were correlated considering the linear change between baseline and follow-up and were standardized by dividing the linear change with the baseline. Finally, we tested the effect of the use of a language other than English, educational background and socio-economic status of the parents with Bayesian t-test for two groups stratification and Bayesian ANOVA for three groups stratification. The prior distribution for the parameter was set by default and all reported statistical tests were two-sided; no transformations were undertaken on any data. Effect sizes are presented as correlation coefficients [r] or Cohen's delta [δ], with 95% credible intervals reported for both measures. All the statistical analyses used JASP version 0.14.0.1[94] and the figures were made on Python in Jupyter Notebook 6.1.5[95].

REFERENCES

1. Merlo, S. & Mansur, L. L. Descriptive discourse: Topic familiarity and disfluencies. *J. Commun. Disorders* **37**, 489–503 (2004).
2. Rosenstein, M., Foltz, P. W., DeLisi, L. E. & Elvevåg, B. Language as a biomarker in those at high-risk for psychosis. *Schizophr. Res.* **165**, 249–250 (2015).
3. Liddle, P. F. et al. Thought and language index: An instrument for assessing thought and language in schizophrenia. *British J. Psychiatry* **181**, 326–330 (2002).
4. Andreasen, N. C. & Grove, W. M. Thought, language, and communication in schizophrenia: diagnosis and prognosis Schizophrenia bulletin. *Schizophr. Bull* **12**, 348–359 (1986).
5. Ayer, A. et al. Formal thought disorder in first-episode psychosis. *Compr. Psychiatry* **70**, 209–215 (2016).
6. Mota, N. B., Copelli, M. & Ribeiro, S. Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *npj Schizophrenia* **3**, 1–10 (2017).
7. Spencer, T. J. et al. Lower speech connectedness linked to incidence of psychosis in people at clinical high risk. *Schizophr. Res.* **228**, 493–501 (2021).
8. Palaniyappan, L. et al. Speech structure links the neural and socio-behavioural correlates of psychotic disorders. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **88**, 112–120 (2019).
9. Sherratt, S. Multi-level discourse analysis: A feasible approach. *Aphasiology* **21**, 375–393 (2007).
10. Dell, G. Connectionist models of language production: lexical access and grammatical encoding. *Cogn. Sci.* **23**, 517–542 (1999).

M.F. Alonso-Sánchez et al.

11. Turney, P. D. & Pantel, P. From frequency to meaning: Vector space models of semantics. *J. Artif. Intell. Res.* **37**, 141–188 (2010).

12. Collins, A. M. & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychol. Rev.* **82**, 407–428 (1975).

13. Roelofs, A. A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition* **172**, 59–72 (2018).

14. Rabagliati, H., Delaney-busch, N., Snedeker, J. & Kuperberg, G. Spared bottom-up but impaired top-down interactive effects during naturalistic language processing in schizophrenia: evidence from the visual-world paradigm. *Psychol. Med.* **49**, 1335–1345 (2018).

15. Chiou, R., Humphreys, G. F., Jung, J. Y. & Lambon Ralph, M. A. Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex* **103**, 100–116 (2018).

16. Kuperberg, G. R. et al. Multimodal neuroimaging evidence for looser lexico-semantic networks in schizophrenia:Evidence from masked indirect semantic priming. *Neuropsychologia* **124**, 337–349 (2019).

17. Westerhausen, R., Kompus, K. & Hugdahl, K. Impaired cognitive inhibition in schizophrenia: A meta-analysis of the Stroop interference effect. *Schizophr. Res.* **133**, 172–181 (2011).

18. Palaniyappan, L. Dissecting the neurobiology of linguistic disorganisation and impoverishment in schizophrenia. *Sem. Cell Dev. Biol.* (2021) https://doi.org/10.1016/j.semcdb.2021.08.015.

19. Lesh, T. A. et al. Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia. *NeuroImage: Clin.* **2**, 590–599 (2013).

20. Barch, D. M. et al. Increased Stroop facilitation effects in schizophrenia are not due to increased automatic spreading activation. *Schizophr.a Res.* **39**, 51–64 (1999).

21. Minzenberg, M. J., Poole, J. H., Vinogradov, S., Shenaut, G. K. & Ober, B. A. Slowed lexical access is uniquely associated with positive and disorganised symptoms in schizophrenia. *Cogn. Neuropsychiatry* **8**, 107–127 (2003).

22. Rathnaiah, M. et al. Quantifying the Core Deficit in Classical Schizophrenia. *Schizophr. Bull. Open* **1**, 1–11 (2020).

23. Ventura, J., Thames, A. D., Wood, R. C., Guzik, L. H. & Hellemann, G. S. Disorganization and Reality Distortion in Schizophrenia. *Schizophr. Res.* **121**, 1–14 (2010).

24. Kircher, T. et al. A rating scale for the assessment of objective and subjective formal thought and language disorder (TALD). *Schizophr. Res.* **160**, 216–221 (2014).

25. Sommer, I. E. et al. Formal thought disorder in non-clinical individuals with auditory verbal hallucinations. *Schizophrenia Research* **118**, 140–145 (2010).

26. Ziv, I. et al. Morphological characteristics of spoken language in schizophrenia patients – an exploratory study. *Scandinavian J. Psychol.* (2021) https://doi.org/10.1111/sjop.12790.

27. de Boer, J. N. et al. Language in schizophrenia: relation with diagnosis, symptomatology and white matter tracts. *npj Schizophrenia* **6**, 1–10 (2020).

28. Tan, E. J., Meyer, D., Neill, E. & Rossell, S. L. Investigating the diagnostic utility of speech patterns in schizophrenia and their symptom associations. *Schizophr. Res.* **238**, 91–98 (2021).

29. Morgan, S. E. et al. Natural Language Processing markers in first episode psychosis and people at clinical high-risk. *Transl.lPpsychiatry* (2021) https://doi.org/10.1038/s41398-021-01722-y.

30. Mota, N. B. et al. Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS ONE* **7**, 1–9 (2012).

31. Elvevåg, B., Foltz, P. W., Weinberger, D. R. & Goldberg, T. E. Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophr. Res.* **93**, 304–316 (2007).

32. Bedi, G. et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia* **1**, 15030 (2015).

33. Corcoran, C. M. et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* **17**, 67–75 (2018).

34. Iter, D., Yoon, J. & Jurafsky, D. Automatic Detection of Incoherent Speech for Diagnosing Schizophrenia. 136–146 (2018) https://doi.org/10.18653/v1/w18-0615.

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

36. Holshausen, K., Harvey, P. D., Elvevåg, B., Foltz, P. W. & Bowie, C. R. Latent semantic variables are associated with formal thought disorder and adaptive behavior in older inpatients with schizophrenia. *Cortex* **55**, 88–96 (2014).

37. Tagamets, M. A., Cortes, C. R., Griego, J. A. & Elvevåg, B. Neural correlates of the relationship between discourse coherence and sensory monitoring in schizophrenia. *Cortex* **55**, 77–87 (2014).

38. Voleti, R., Member, S., Liss, J. M. & Berisha, V. A. Review of Automated Speech and Language Features for Assessment of Cognitive and Thought Disorders. *IEEE J. Sel. Top Signal. Process.* **14**, 282–298 (2019).

39. Voppel, A., de Boer, J., Brederoo, S., Schnack, H. & Sommer, I. Quantified language connectedness in schizophrenia-spectrum disorders. *Psychiatry Res.* **304,** 114130 (2021).

40. Voleti, R. et al. Objective assessment of social skills using automated language analysis for identification of schizophrenia and bipolar disorder. *Proc. Ann. Conf. Int.Speech Commun. Assoc., INTERSPEECH* **2019-Septe**, 1433–1437 (2019).

41. Rezaii, N., Walker, E. & Wolff, P. OPEN A machine learning approach to predicting psychosis using semantic density and latent content analysis. *npj Schizophrenia* (2019) https://doi.org/10.1038/s41537-019-0077-9.

42. Just, S. A. et al. Modeling Incoherent Discourse in Non-Affective Psychosis. *Front. Psychiatry* **11**, 1–11 (2020).

43. Sarzynska-Wawer, J. et al. Detecting formal thought disorder by deep contextualized word representations. *Psychiatry Res.* **304**, 114135 (2021).

44. Tang, S. X. et al. Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. *npj Schizophrenia* **7**, 1–8 (2021).

45. Pauselli, L. et al. Computational linguistic analysis applied to a semantic fluency task to measure derailment and tangentiality in schizophrenia. *Psychiatry Res.* **263**, 74–79 (2018).

46. Ku, B. S., Pauselli, L., Covington, M. A. & Compton, M. T. Computational linguistic analysis applied to a semantic fluency task: A replication among first-episode psychosis patients with and without derailment and tangentiality. *Psychiatry Res.* **304**, 114105 (2021).

47. Hoffman, P., Cogdell-Brooke, L. & Thompson, H. E. Going off the rails: Impaired coherence in the speech of patients with semantic control deficits. *Neuropsychologia* **146**, 107516 (2020).

48. Elvevåg, B., Foltz, P. W., Weinberger, D. R. & Goldberg, T. E. *Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia.* http://lsa.colorado.edu/.

49. Bar, K. et al. Semantic Characteristics of Schizophrenic Speech. 84–93 (2019) https://doi.org/10.18653/v1/w19-3010.

50. Hitczenko, K., Mittal, V. A. & Goldrick, M. Understanding Language Abnormalities and Associated Clinical Markers in Psychosis: The Promise of Computational Methods. 1–19 (2020).

51. Hoffman, R. E., Kirstein, L., Stopek, S. & Cicchetti, D. V. Apprehending schizophrenic discourse: A structural analysis of the Listener's task. *Brain Lang.* **15**, 207–233 (1982).

52. Silva, A., Limongi, R., MacKinley, M. & Palaniyappan, L. Small Words That Matter: Linguistic Style and Conceptual Disorganization in Untreated First-Episode Schizophrenia. *Schizophr. Bull. Open* **2**, 1–10 (2021).

53. Mota, N. B., Sigman, M., Cecchi, G., Copelli, M. & Ribeiro, S. The maturation of speech structure in psychosis is resistant to formal education. *npj Schizophrenia* **4**, 25 (2018).

54. Ratana, R., Sharifzadeh, H., Krishnan, J. & Pang, S. A Comprehensive Review of Computational Methods for Automatic Prediction of Schizophrenia With Insight Into Indigenous Populations. *Front. Psychiatry* **10**, 1–15 (2019).

55. de Boer, J. N., Voppel, A. E., Brederoo, S. G., Wijnen, F. N. K. & Sommer, I. E. C. Language disturbances in schizophrenia: the relation with antipsychotic medication. *npj Schizophrenia* **6**, 1–9 (2020).

56. Palaniyappan, L. More than a biomarker: could language be a biosocial marker of psychosis? *npj Schizophrenia* **7**, 13–15 (2021).

57. Bowie, C. R. & Harvey, P. D. Communication Abnormalities Predict Functional Outcomes in Chronic Schizophrenia: Differential Associations with Social and Adaptive Functions. *Schizophr Res* **103**, 240–247 (2008).

58. Tan, E. J., Thomas, N. & Rossell, S. L. Speech disturbances and quality of life in schizophrenia: Differential impacts on functioning and life satisfaction. *Compr. Psychiatry* **55**, 693–698 (2014).

59. Oeztuerk, O. F., Pigoni, A., Antonucci, L. A. & Koutsouleris, N. Association between formal thought disorders, neurocognition and functioning in the early stages of psychosis: a systematic review of the last half - century studies. *Eur. Arch. Psychiatry Clin. Neurosci.* 272, 381–393 (2022)

60. Roche, E. et al. Language disturbance and functioning in first episode psychosis. *Psychiatry Res.* **235**, 29–37 (2015).

61. Minzenberg, M. J., Ober, B. A. & Vinogradov, S. Semantic priming in schizophrenia: A review and synthesis. *J. Int. Neuropsychol. Soc.* **8**, 699–720 (2002).

62. Kuperberg, G. R. Building meaning in schizophrenia. *Clin. EEG Neurosci.* **39**, 99–102 (2008).

63. Kuperberg, G. R. Separate streams or probabilistic inference? What the N400 can tell us about the comprehension of events. *Lang. Cogn. Neurosci.* **3798**, 602–616 (2016).

64. Kuperberg, G. R. Language in Schizophrenia Part 1: An Introduction. *Linguistics and Language Compass* (2010) https://doi.org/10.1111/j.1749-818X.2010.00216.x.

65. Crider, A. Perseveration in schizophrenia. *Schizophr. Bull.* **23**, 63–74 (1997).

66. Lundin, N. B. et al. Semantic Search in Psychosis: Modeling Local Exploitation and Global Exploration. *Schizophr. Bull. Open* **1**, 1–11 (2020).

67. Waford, R. N. & Lewine, R. Is perseveration uniquely characteristic of schizophrenia? *Schizophr. Res.* **118**, 128–133 (2010).
68. Maher, B. A., Manschreck, T. C. & Molino, M. A. C. Redundancy, pause distributions and thought disorder in schizophrenia. *Language and Speech* **26**, 191–199 (1983).
69. Manschreck, T. C., Ames, D., Maher, B. A. & Hoover, T. M. Repetition in schizophrenic speech. *Language and Speech* **28**, 255–268 (1985).
70. Mackinley, M., Chan, J., Ke, H., Dempster, K. & Palaniyappan, L. Linguistic determinants of formal thought disorder in first episode psychosis. *Early Interv. Psychiatry* **15**, 344–351 (2021).
71. Morgan, S. E. et al. Assessing psychosis risk using quantitative markers of disorganised speech. *npj Schizophrenia* (2021) https://doi.org/10.1101/2021.01.04.20248717.
72. Lott, P. R., Guggenbühl, S., Schneeberger, A., Pulver, A. E. & Stassen, H. H. Linguistic analysis of the speech output of schizophrenic, bipolar, and depressive patients. *Psychopathology* **35**, 220–227 (2002).
73. Minor, K. S., Willits, J. A., Marggraf, M. P., Jones, M. N. & Lysaker, P. H. Measuring disorganized speech in schizophrenia: Automated analysis explains variance in cognitive deficits beyond clinician-rated scales. *Psychol. Med.* **49**, 440–448 (2019).
74. Szöke, A. et al. Longitudinal studies of cognition in schizophrenia: Meta-analysis. *British J. Psychiatry* **192**, 248–257 (2008).
75. Leckman, J. F., Sholomskas, D., Thompson, D., Belanger, A. & Weissman, M. M. Best Estimate of Lifetime Psychiatric Diagnosis: A Methodological Study. *Arch. Gen. Psychiatry* **39**, 879–883 (1982).
76. Kay, S. R. & Qpjer, L. A. The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1982).
77. WHO Collaborating Centre for Drug Statistics Methodology. *Guidelines for ATC classification and DDD assignment 2021*. vol. 148 (Norwegian Institute of Public Health, 2021).
78. Malla, A. et al. Predictors of rate and time to remission in first-episode psychosis: A two-year outcome study. *Psychol. Med.* **36**, 649–658 (2006).
79. Cassidy, C. M., Rabinovitch, M., Schmitz, N., Joober, R. & Malla, A. A comparison study of multiple measures of adherence to antipsychotic medication in first-episode psychosis. *J. Clin. Psychopharmacol.* **30**, 64–67 (2010).
80. Dempster, K. et al. Early treatment response in first episode psychosis: a 7-T magnetic resonance spectroscopic study of glutathione and glutamate. *Mol. Psychiatry* **25**, 1640–1650 (2020).
81. Rybarczyk, B. *Encyclopedia of Clinical Neuropsychology. Encyclopedia of Clinical Neuropsychology* (Springer, 2011).
82. Boyd, M. A socioeconomic scale for Canada: Measuring occupational status from the census. *Can. Rev. Sociol.* **45**, 51–91 (2008).
83. Addington, D., Addington, J. & Maticka-Tyndale, E. Assessing depression in schizophrenia: The Calgary Depression Scale. The British Journal of Psychiatry. *British J. Psychiatry,* **163**, 39–44 (1993).
84. Guy W. E. *ECDEU: Assessment Manual for Psychopharmacology (revised)*. Nimh vol. 1 (DHEW, 1976).
85. Palaniyappan, L., Al-Radaideh, A., Mougin, O., Gowland, P. & Liddle, P. F. Combined white matter imaging suggests myelination defects in visual processing regions in schizophrenia. *Neuropsychopharmacology* **38**, 1808–1815 (2013).
86. Palaniyappan, L., Simmonite, M., White, T. P., Liddle, E. B. & Liddle, P. F. Neural primacy of the salience processing system in schizophrenia. *Neuron* **79**, 814–828 (2013).
87. Taylor, R. et al. Functional magnetic resonance spectroscopy of glutamate in schizophrenia and major depressive disorder: Anterior cingulate activity during a color-word Stroop task. *npj Schizophrenia* **1**, 15028 (2015).
88. Limongi, R. et al. Glutamate and Dysconnection in the Salience Network: Neurochemical, Effective Connectivity, and Computational Evidence in Schizophrenia. *Biol. Psychiatry* **88**, 273–281 (2020).
89. Golden, C. *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses*. (Stoelting C, 1978). https://doi.org/10.1007/978-3-319-57111-9_660.
90. Scarpina, F. & Tagini, S. The stroop color and word test. *Fronti.Psychol.* **8**, 1–8 (2017).
91. Murray, H. *Thematic Apperception Test*. (Harvard University Press, 1943).
92. Covington, M. A. Covington Vector Semantics Tools. (2016).
93. Pennington, J., Socher, R. & Manning, C. D. GloVe: Global Vectors for Word Representation. *Conference: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* 1532–1543 (2014). https://doi.org/10.3115/v1/D14-1162.
94. Love, J. et al. JASP: Graphical Statistical Software for Common Statistical Designs. *J. Stat. Softw.* **88**, 1–17, https://doi.org/10.18637/jss.v088.i02 (2019).
95. Kluyver, T. et al. Jupyter Notebooks – a publishing format for reproducible computational workflows. in *In Positioning and Power in Academic Publishing: Players, Agents and Agendas*. (eds. Loizides, F. & Scmidt, B.) 87–90 (IOS Press., 2016). https://doi.org/10.3233/978-1-61499-649-1-87.

AUTHOR CONTRIBUTIONS

MFA and LP conceptualized the project, SF and MM collected the data while MF analyzed the data, MF and LP wrote the manuscript; RL, AS and all authors critically reviewed and approved the final version of the manuscript.

COMPETING INTERESTS

LP reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. LP is the convenor of the DISCOURSE in psychosis research consortium (www.discourseinpsychosis.org). All other authors report no relevant conflicts.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41537-022-00246-8.

Correspondence and requests for materials should be addressed to Lena Palaniyappan.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.