Schizophrenia Research 259 (2023) 48–58

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Assessing coherence through linguistic connectives: Analysis of speech in patients with schizophrenia-spectrum disorders

H. Corona-Hernández [a,*], J.N. de Boer [a,b], S.G. Brederoo [a,c], A.E. Voppel [a], I.E.C. Sommer [a,c]

[a] Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, University of Groningen, the Netherlands
[b] Department of Psychiatry, University Medical Center Utrecht, Utrecht University & Brain Center Rudolf Magnus, Utrecht, the Netherlands
[c] Department of Psychiatry, University Medical Center Groningen, University of Groningen, the Netherlands

ARTICLE INFO

Keywords:
Classification
Disorganized speech
Grammatical connectivity
Schizophrenia-spectrum disorders
Semantic cosine similarity
Word2vec

ABSTRACT

Background: Incoherent speech is a core diagnostic symptom of schizophrenia-spectrum disorders (SSD) that can be studied using semantic space models. Since linguistic connectives signal relations between words, they and their surrounding words might represent linguistic loci to detect unusual coherence in speech. Therefore, we investigated whether connectives' measures are useful to assess incoherent speech in SSD.
Methods: Connectives and their surrounding words were extracted from transcripts of spontaneous speech of 50 SSD-patients and 50 control participants. Using word2vec, two different cosine similarities were calculated: those of connectives and their surrounding words (connectives-related similarity), and those of free-of-connectives words-chunks (non-connectives similarity). Differences between groups in proportion of five types of connectives were assessed using generalized logistic models, and connectives-related similarity was analyzed through non-parametric multivariate analysis of variance. These features were evaluated in classification tasks to differentiate between groups.
Results: SSD-patients used less contingency (e.g., because) ($p = .008$) and multiclass connectives (e.g., as) ($p < .001$) than control participants. SSD-patients had higher minimum similarity of multiclass (adj-$p = .04$) and temporality connectives (e.g., after) (adj-$p < .001$), narrower similarity-range of expansion (e.g., and) (adj-$p = .002$) and multiclass connectives (adj-$p = .04$), and lower maximum similarity of expansion connectives (adj-$p = .005$). Using connectives' features alone, SSD-patients and controls could be distinguished with 85 % accuracy.
Discussion: Our results show that SSD-speech can be distinguished from speech of control participants with high accuracy, based solely on connectives' features. We conclude that including connectives could strengthen computational models to categorize SSD.

1. Introduction

Disorganized speech is a core feature of schizophrenia-spectrum disorders (SSD) (American Psychiatric Association, 2013) that has been increasingly assessed using semantic space models (Corcoran et al., 2020; Corcoran and Cecchi, 2020; Hitczenko et al., 2021). Such computational models create *n* dimensions, each standing for an abstract feature of word meaning. The represented meaning (i.e., vector) of a given word can thus be located within the semantic space of *n* dimensions, and it is posited that words with similar meaning are found close to each other within a given semantic space (Landauer et al., 1998; Mikolov et al., 2013b). Using these models, it has been shown that patients with SSD can be distinguished from healthy controls with accuracies between 70 % and 93 % (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2020; Tang et al., 2021; Voppel et al., 2021), while predicting psychosis onset in at-high-risk individuals has accuracies ranging from 72 % to 100 % (Bedi et al., 2015; Corcoran et al., 2018; Rezaii et al., 2019).

Disorganization in speech is considered to signal a reduction in the underlying semantic coherence of a given message (Corcoran and Cecchi, 2020; Hitczenko et al., 2021). To be attained, coherence requires *thematic continuity* and *grammatical connectivity* (Givón, 2020). While *thematic continuity* is reflected in the maintenance of semantic content, *grammatical connectivity* refers to the use of explicit markers to

* Corresponding author at: Cognitive Neuroscience Center, Department of Biomedical Sciences of Cells and Systems, University Medical Center Groningen, University of Groningen, Internal zip code FA32, Antonius Deusinglaan 2, 9713 AW Groningen, the Netherlands.
E-mail address: h.corona.hernandez@umcg.nl (H. Corona-Hernández).

https://doi.org/10.1016/j.schres.2022.06.013
Received 28 February 2022; Received in revised form 14 June 2022; Accepted 15 June 2022
Available online 29 June 2022
0920-9964/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

hierarchically organize the sequence of the content. Syntactically, *grammatical connectivity* is most clearly instantiated by linguistic connectives, which relate two or more words, clauses, or sentences to each other (Maat and Sanders, 2006; Sanders and Maat, 2006; van der Vliet and Redeker, 2014). Importantly, connectives establish different types of explicit coherence relations in discourse, such as comparative (e.g., this flower is red, *while* that other is white), contingent (e.g., I have to replace this piece *because* it is damaged), expansive (e.g., *besides* being mammals, gorillas are primates), and temporal (e.g., we will go outside *after* the rain stops) (Bourgonje et al., 2018; Stede et al., 2019). Thus, in fine-tuning semantic space models to better quantify disorganized speech, it could be valuable to separately assess *thematic continuity* and *grammatical connectivity*.

General semantic content has been the main focus of interest in previous studies using semantic space models to quantify coherence. Specifically, in speech of patients with SSD, unusual general semantic content has often been assessed across entire interviews or conversations. Typically, the measures for analysis are obtained by averaging series of semantic distances (i.e., cosine similarities) across all words or sentences uttered by patients, one after each other (Hitczenko et al., 2021). This form of assessment is reasonable considering that semantic space models perform better if they are built upon sentences rather than upon speech samples from word-association or verbal-fluency tasks (de Boer et al., 2018). However, this procedure hampers the possibility to distinctively quantify the coherence that relates to syntactic markers of connectivity (i.e., connectives).

While the use of connectives has not been separately assessed in semantic space models yet, previous studies have examined their occurrence in speech. Patients with SSD have been found to use less connectives of differentiation in comparison to control participants, (Just et al., 2020), and less causal, contrastive, and logical connectives when compared to adults with a diagnosis of HIV+ (Willits et al., 2018). In contrast, another study showed that untreated first episode psychosis patients with high scores in conceptual disorganization (PANSS Item P2) overall used more connectives than control participants (Mackinley et al., 2021).

These inconsistencies in results currently limit our knowledge about the frequency and coherence with which different types of connectives are used by patients with SSD compared to control participants. Moreover, even though semantic space models have been shown to be reliable tools to assess disorganized speech in patients with SSD, no previous research has specifically focused on grammatical connectivity. Considering this, in the present study we first evaluated whether patients with SSD and control participants use different types of connectives in similar proportions. Second, by calculating cosine similarity between connectives and their surrounding words (i.e., connectives-related similarity), we assessed whether connectives and their surrounding words can be used as linguistic loci to detect unusual coherence in speech of patients with SSD. Third, we tested how automatic classification driven by connectives-related similarity compares to another driven by non-connectives similarity, and how accurately connectives-related similarity and proportions per type of connective together could distinguish patients with SSD from control participants.

2. Methods

2.1. Participants

Fifty individuals with a schizophrenia-spectrum disorder and fifty healthy control participants, all native Dutch speakers, took part in this study. These participants had been previously investigated in Voppel et al. (2021). Their inclusion took place at the University Medical Center Utrecht. Patients' diagnoses were established by a trained physician, and confirmed using either the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992) or the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Patients' severity of symptoms was assessed with the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). Control participants were included if they had neither current nor a history of psychiatric disorders. All participants gave written informed consent before obtaining the measurements.

2.2. Speech sampling

Speech was elicited using a semi-structured interview, comprising 60 open-ended questions, from which a subset was presented in a semi-randomized order across all participants. The questions were designed to elicit spontaneous speech but prevent excessive emotional arousal, with topics such as life experiences, current daily habits, hobbies, and hypothetical situations, avoiding health-related and psychopathology-related topics. To prevent participants from adjusting their own speech, they were informed about the research aim only after concluding the interview. All interviews were conducted by trained researchers. The elicited verbal samples were audio-recorded and later transcribed following the CHILDES protocol (MacWhinney, 2000). Transcribers were blind to participants' group. Procedures were approved by the ethical committee of the University Medical Center Utrecht.

2.3. Linguistic features

2.3.1. Data preprocessing

Using the CHILDES transcripts, individual plain text files were derived for each participant. Speech of the interviewers, punctuation marks, metadata, headers, special characters, and markers of events (e.g., & = laughs) were all excluded from these files. In the resulting files, all words were set to lowercase, and grammatical contractions (e.g., *t'is*, standing for *het is*, "it is") were retained. Fillers (e.g., *ehm*) and repetitions have been shown to add noise to semantic similarity calculations (Iter et al., 2018). However, they were also retained in the transcripts used for analysis for three reasons. First, it is currently unknown whether fillers and repetitions might influence connectives-related similarity. Second, despite recent attempts to control for "inadequate repetitions" of speech (Just et al., 2020), there is no standard procedure for avoiding this bias. Third, fillers (Tang et al., 2021) and repetitions (Andreasen, 1979; Hong et al., 2015; Maher, 1972) have been shown to be important to distinguish patients with SSD from control participants.

2.3.2. Selection and occurrence of connectives

Based on Bourgonje et al. (2018), 188 different Dutch connectives were selected for this study. These connectives were originally divided in four broad categories, and we created a fifth category (i.e., multiclass) to be filled in with all connectives that were listed in more than one category, excluding such connectives from their initial lists and being relocated to this new one. Connectives ultimately belonged to only one of the following categories: comparison ($n = 37$), contingency ($n = 48$), expansion ($n = 44$), multiclass ($n = 23$), and temporality ($n = 36$) (see supplementary materials: 1. List of connectives in Dutch).

For each preprocessed transcript, all occurrences of the 188 different Dutch connectives were automatically extracted using the R "quanteda" package (Benoit et al., 2018). Subsequently, all the extracted connectives were automatically given the label of the category they belonged to. For each occurrence, along with the connective, the previous and the following three words were retained, considering them as part of the surrounding context of the connective, resulting in a seven-words fixed window size. This length was chosen for two main reasons: shorter word-windows are less suited to reveal differences in cosine similarity between groups (Elvevåg et al., 2007), and larger word-windows were found to be poorer informative features for classification tasks (Voppel et al., 2021). All cases in which there were fewer words in the surrounding context of the connective were also preserved (i.e., connectives occurring at the end of an interview), leading a subset of instances to

have less than seven words.

2.3.3. Semantic space model

A semantic space with 300 dimensions was modeled using the skip-gram method of the word2vec learning algorithm (Mikolov et al., 2013b). It was trained on more than five-million words from The Netherlands' transcripts collection of the *Corpus Gesproken Nederlands* (CGN) (van Eerten, 2007), using the R "word2vec" package (Wijffels, 2020). In this model, each dimension might be taken to represent an abstract feature of word meaning, and the meaning of each word (i.e., its word embedding) is the vector indicating the position of the word relative to the 300 semantic dimensions of the model. Using the skip-gram method, the word2vec algorithm computes each word embedding in a few steps. For each word, first a random embedding is created. Next, using all instances of the word and its surrounding words as constraint, the random embedding is iteratively changed to resemble the embeddings of its surrounding words more, and the embeddings of words which do not appear nearby less (Mikolov et al., 2013a, 2013b). Finally, each word is assigned a fixed and unique embedding, which can then be used to measure the semantic (i.e., cosine) similarity between embeddings (Mikolov et al., 2013b).

2.3.4. Computation of connectives-related and non-connectives cosine similarities

For each seven-words chunk having a connective, the connectives-related cosine similarity was operationalized as the cosine similarity between the embedding of the connective and the as-a-whole averaged embedding of the three previous and three following words. All segments of the transcripts no longer containing connectives (henceforth, free-of-connectives segments) were also split into chunks of seven words. For each free-of-connectives segment, the non-connectives similarity was operationalized as the cosine similarity between the embedding of the fourth word and the as-a-whole averaged embedding of the three previous and three following words. Thus, the same procedure was used to obtain the two different types of cosine similarities.

Connectives-related and non-connectives similarities were first

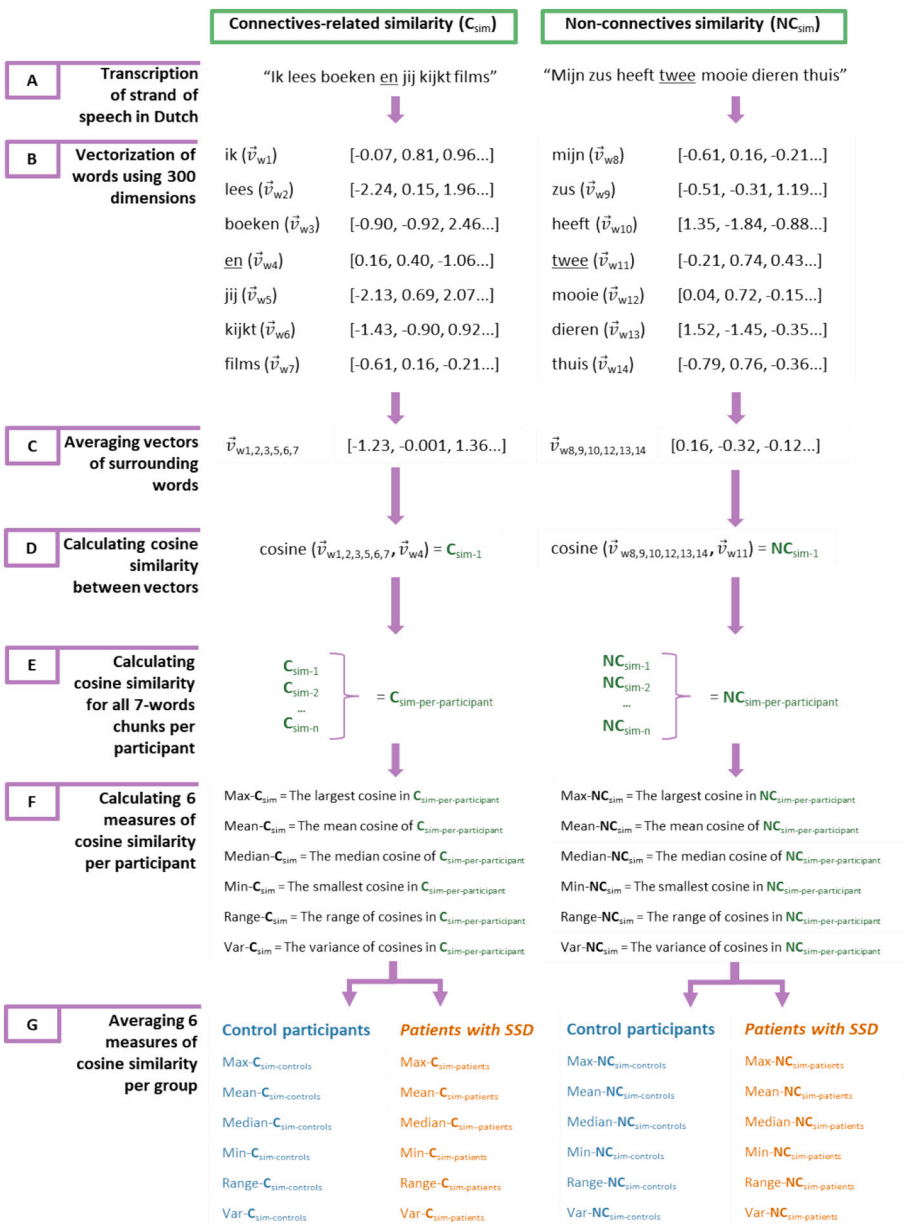

Fig. 1. Steps (A-G) taken to calculate the connectives-related and the non-connectives cosine similarity. Note that the word "en" is a connective, while the word "twee" is not.

calculated separately for all the chunks of words per participant. Then, six different measures of the cosine similarities were independently obtained per participant: maximum, mean, median, minimum, range, and variance. Finally, these six different measures of cosine similarity were averaged per group for each type of cosine similarity (see Fig. 1). In all cases, cosine similarities could range from −1 to 1, with −1 standing for the lowest possible similarity, and 1 for the highest possible similarity.

2.4. Analysis

2.4.1. Statistics

Groups were compared with regard to demographic continuous variables through independent one-way analysis of variance (ANOVA), and nominal variables through Chi-square tests without continuity correction.

To assess differences between groups in the proportion of types of connectives relative to all words used, we carried out generalized linear mixed-effects logistic regression models using the *glmer* function from the R "lme4" package (Bates et al., 2015), with proportion as dependent variable. Following Baayen (2008) and Winter (2020), group (patients with SSD vs control participants) and type of connective (comparison vs contingency vs expansion vs multiclass vs temporality) were considered as fixed-effects factors, and participants as random-effects factors (allowing by-participant varying intercepts and varying slopes). Implementing a forward-testing approach, we carried out stepwise model comparison between a series of independent regression models in order to arrive at the model that best fitted the data. Likelihood ratio tests were performed to assess whether there were significant differences between each pair of models being compared (Baayen, 2008).

To assess group differences for the connectives-related similarity, we used non-parametric multivariate analysis of variance (MANOVA), followed by post-hoc Wilcoxon rank sum tests with Holm correction.

For all correlations, we used the Spearman's rank non-parametric test, correcting for multiple comparisons when necessary. Statistical results with (adjusted) p-values $< .05$ were considered to be significant. All analyses were done in *RStudio* version 1.4.1103 (RStudio Team, 2019) running *R* version 4.1.0 (R Core Team, 2020).

2.4.2. Classification tasks

For reliable results, in all classification tasks the models were trained using 10-fold cross-validation, repeated ten times. This means that, for each iteration, the learning algorithm used nine-tenths partitions of the data for training, and one-tenth for testing. Accuracy, sensitivity, specificity, and area under the curve (AUC) for receiver operating characteristic (ROC) were obtained in order to assess the performance of each classifier. All tasks were conducted using the R "caret" package (Kuhn, 2021).

2.4.2.1. Connectives' vs non-connectives' features. To assess whether connectives-related and non-connectives similarities might yield similar classification results, we first obtained and evaluated the performance of a control-classifier. Afterwards, we assessed how much improvement in performance this control-classifier could gain by independently adding either connectives' features or non-connectives similarity to it.

The control-classifier was built by training a random forest algorithm, using the mean, minimum and variance of general semantic similarity from sliding-windows between 5 and 10 words. This algorithm has been proven to be one of the best in performing binary classification (Fernández-Delgado et al., 2014). Mean, minimum, variance, and the 5–10 range of windows were chosen as parameters based on previous results showing that they were highly informative for classification (Voppel et al., 2021).

For the comparison between the connectives' features and the non-connectives similarity, number of features was controlled for (see

Table 1
Details of the different features used for the classification tasks.

Name of features used for classification	Use in classification task		Number of features	Description
	Connectives' features vs non-connectives similarity	Performance of connectives' features alone		
Proportions per type of connective	✓	✓	5	Proportions of use of the comparison, contingency, expansion, multiclass, and temporality connectives.
Non-connectives similarity	✓		6	Maximum, mean, median, minimum, range and variance of all the free-of-connectives chunks.
General connectives-related similarity (random sub-sampling)	✓		6	Maximum, mean, median, minimum, range and variance of all connectives taken together (i.e., regardless of type of connective). These measures were calculated based on a random subsampling of connectives-chunks, which were matched to the number of free-of-connectives chunks per participant[a].
General connectives-related similarity (all chunks)	✓		6	Maximum, mean, median, minimum, range and variance of all connectives taken together (i.e., regardless of type of connective) calculated based on all connectives-chunks.
Per-type-connectives-related similarity with smallest adj. p-values	✓		6	Connectives-related similarity measures that, across the types of connectives, have the smallest adj. p-values as found through post-hoc analyses (see Table 7).
Proportions and per-type-connectives-	✓		7	Connectives-related similarity measures and

(*continued on next page*)

Table 1 (*continued*)

Name of features used for classification	Use in classification task		Number of features	Description
	Connectives' features vs non-connectives similarity	Performance of connectives' features alone		
related similarity significantly different between groups				proportions per type of connective that were found to be significantly different between groups (see Table 5 and Table 7).
Sliding-window similarity	✓		18	Mean, minimum and variance of similarity from sliding-windows between 5 and 10 words. Partial set of the cosine similarities used in Voppel et al. (2021).
Connectives-related similarities per type of connective		✓	30	Maximum, mean, median, minimum, range and variance calculated for each of the 5 different types of connectives.

[a] Across all participants, there were more connectives-chunks than free-of-connectives chunks. This was due to partial overlap of some connectives-chunks when two connectives were used too close to each other. Accordingly, for each participant, a random subsampling of the connectives-chunks was carried out.

Table 1). To rule out that amount of data influenced the results, a set of connectives-related similarity controlling for this was also calculated (see Table 1). These procedures to control for amount of data and number of features were exclusively done for this series of classification tasks.

2.4.2.2. Performance of connectives' features alone. Using connectives-related similarity per type of connective either alone or along with proportions of use per type of connective, random forest and support vector machine with polynomial kernel algorithms were trained to perform a binary classification between patients with SSD and control participants. For these classification tasks, the six measures of connectives-related similarity were calculated independently for each of the five types of connectives (see Table 1).

3. Results

3.1. Demographics and speech sample

The majority of patients in our sample had a diagnosis of psychosis not otherwise specified (46 %), followed by schizophrenia (38 %), schizoaffective (14 %) and schizophreniform disorder (2 %). There were no significant differences between groups in age ($p = .98$) and sex ($p = .23$). Years of education were significantly less for patients with SSD than for control participants ($p = .001$). In patients with SSD, the mean dose of antipsychotic medication as measured in chlorpromazine equivalence was 226.1 mg. Thirty-two patients used tight-binding antipsychotics, sixteen patients used loose-binding medication, and two patients were not receiving antipsychotic medication (see Table 2).

General characteristics of the participants' speech sample and use of connectives are presented in Table 3. Since patients with SSD had significantly less years of education than control participants, possible correlations between years of education and basic features of the speech sample were assessed (i.e., tokens and types). Neither number of running words (tokens) nor number of different word forms (types) correlated to years of education, (both *rho*(ρ) < 0.10, both $p > .05$).

3.2. Proportion of connectives

Following stepwise model comparison with a forward-testing approach, the model that best fitted the data included an interaction between group and type of connective, as well as by-participant varying intercepts and varying slopes for type of connective per participant (see Table 4). Relative to connectives of comparison, patients with SSD had a lower probability of using connectives of contingency ($p < .001$), multiclass ($p < .001$), and temporality ($p < .001$), and a higher probability of using connectives of expansion ($p < .001$). Similarly, control participants had a lower probability of using connectives of contingency ($p < .001$) and temporality ($p < .001$), and a higher probability of using connectives of expansion ($p < .001$). When comparing the groups, patients with SSD had a lower probability of using connectives of contingency ($p = .008$) and multiclass ($p < .001$) than control participants (see Table 5 and Fig. 2). The structure of the random-effects factors is shown in Table 6.

Considering patients with SSD had significantly less years of education, it was assessed whether this could have confounded the above-mentioned results. Independent two-tailed bivariate correlations were conducted, and false positive results were controlled using Holm correction. Results showed no significant correlations (all *rho*(ρ) < 0.21,

Table 2
Demographic characteristics of participants.

Characteristic	Patients (n = 50)	Controls (n = 50)	Statistic value	*P*-value
Age in years, mean (SD)	31.6 (12.46)	31.5 (12.69)	$F(1,98) = 0.001$	.98
Females, n (%)	14 (28 %)	9 (18 %)	$\chi^2(1) = 1.41$	.23
Years of education, mean (SD)[a]	12.8 (2.22)	14.3 (2.39)	$F(1,97) = 10.49$	.001
Antipsychotic medication				
Chlorpromazine equivalence (mg), mean (SD)[b]	226.1 (142.81)			
Loose binding, n (%)	16 (32 %)			
None	2 (4 %)			
Tight binding	32 (64 %)			
Diagnosis, n (%)				
Psychosis not otherwise specified	23 (46 %)			
Schizophrenia	19 (38 %)			
Schizoaffective	7 (14 %)			
Schizophreniform	1 (2 %)			
PANSS, mean (SD)				
Negative	13.4 (5.19)			
Positive	11.1 (4.41)			
General	26.5 (7.74)			
Total	51.2 (13.85)			

n = sample size, SD = standard deviation.
[a] Data available for all control participants, but only for 49 patients.
[b] Data available for 44 patients alone: two patients were not taken antipsychotic medication, and, for four patients, there was no information about dose equivalence.

Table 3
Characteristics of the speech produced by the participants per group.

Feature	Patients (n = 50)	Controls (n = 50)
Basic description of words		
Number of running words (*tokens*)		
Mean (SD)	1161.8 (753.27)	1819.4 (458.07)
Total	58,090	90,971
Number of different word forms (*types*)		
Mean (SD)	342.5 (152.46)	498.6 (89.02)
Total	17,125	24,930
Connectives		
Number of connectives of comparison		
Mean (SD)	39.6 (29.43)	62.4 (25.53)
Total	1,982	3,122
Number of connectives of contingency		
Mean (SD)	21.6 (19.40)	39.9 (18)
Total	1,080	1,997
Number of connectives of expansion		
Mean (SD)	46.9 (37.39)	79.9 (28.28)
Total	2,348	3,996
Number of connectives of multiclass		
Mean (SD)	28.6 (29.05)	58.4 (24.78)
Total	1,430	2,924
Number of connectives of temporality		
Mean (SD)	14.1 (11.51)	23.6 (10.40)
Total	709	1,181

n = sample size, SD = standard deviation.

Table 4
Stepwise procedure followed to obtain the logistic model that best fitted the data.

Factors in the model	Likelihood ratio test		
	Log-lik. increase	Statistic value	*P*-value
Participants as varying intercepts			
+Group	7	χ^2 (1) = 13.53	< .001
+Type of connective	3,556,062	χ^2 (4) = 7,112,125	< .001
+Varying slopes for type of connective per participant	1,010,919	χ^2 (14) = 2,021,838	< .001
+Interaction (Group*Type of connective)	9	χ^2 (4) = 17.19	.0017

Note: the first row stands for the initial random model. Each subsequent row shows how the goodness of fit increased when the factor in the row was added to the model that included all preceding factors.

all adj-$p > .05$).

3.3. Connectives-related similarity

Multivariate analysis of variance showed that there were significant differences between groups in connectives-related similarity, F (10.3, 1018.3) = 5.2, $p < .001$. Post-hoc analyses showed that patients with SSD had higher minimum similarity of temporality connectives (adj-$p <$ .001), as well as narrower range (adj-$p = .002$) and lower maximum similarity of expansion connectives (adj-$p = .005$) than control participants. Additionally, compared to controls, patients had narrower range (adj-$p = .04$) and higher minimum similarity of multiclass connectives (adj-$p = .04$) (see Table 7). Maximum similarity of expansion connectives positively correlated to years of education (*rho*(ρ) = 0.30, adj-$p =$ .01), while the other four connectives-related similarities did not (all *rho* (ρ) ≤ 0.16, all adj-$p > .05$). In performing these analyses, a missing value in the variance of contingency connectives in one patient, and a missing value in the variance of temporality connectives in another patient, were substituted with zeros. Running the same analyses with the exclusion of these two patients did not change the results of these variables. For a full overview of the results with the exclusion of these two patients, see supplementary materials: 2. Additional analyses on connectives-related similarity.

3.4. Classification tasks

3.4.1. Connectives' vs non-connectives' features

The control classifier (RF-c) yielded 83.5 % accuracy. By adding non-connectives similarity to the classification (RF-non-conn), accuracy resulted in 84.9 %. Matched in number of features and amount of data, the classifier that rather included connectives-related similarity (RF-conn-I) yielded 85 % accuracy. The combination of the sliding-window measures and the connectives-related similarity matched in features and data (RF-conn-I) yielded the highest sensitivity (81.2 %). The combination of the sliding-window measures and the connectives' features that were significantly different between groups (RF-conn-V) yielded the highest specificity (93.6 %) (see Table 8).

3.4.2. Performance of connectives' features alone

Using connectives-related similarity per type of connective alone, the best classifier (RF-I) yielded 79.4 % accuracy, 75 % sensitivity and 83.8 % specificity. Combining these features with the proportions of use per type of connective, the best classifier (SVM-II) yielded 85 % accuracy, 83.8 % sensitivity and 86.2 % specificity (see Table 9).

4. Discussion

In this study, we analyzed linguistic coherence by comparing the relative use of different types of connectives and connectives-related similarity between patients with SSD and control participants. In parallel, we assessed how much connectives' features might improve a control-classifier, followed by an evaluation of the usefulness of connectives' features to achieve accurate results in automatically distinguishing patients with SSD from control participants.

Patients with SSD used significantly less contingency and multiclass connectives, while their use of the other types of connectives was not different from that of control participants. Although years of education differed between groups, it did not seem to affect these results. Regarding connectives-related similarity, patients with SSD had higher minimum similarity in both multiclass and temporality connectives, narrower range in both expansion and multiclass connectives, and lower maximum similarity in expansion connectives.

In the classification tasks comparing connectives' features and non-connectives similarity, both types of measures yielded similar overall performance in distinguishing patients with SSD from control participants. In the second series of classification tasks, combining connectives-related similarity per connective type with proportions of use per type of connective, the best classifier yielded 85 % accuracy.

4.1. Proportion of connectives

We found significant differences between groups in their use of contingency connectives (subsuming cause, condition, and purpose connectives). This is partially in line with previous studies showing that connectives of cause are used relatively less by patients with SSD (Willits et al., 2018), but opposite results have also been found (Just et al., 2020). Aligning with previous research, we found no significant differences between patients with SSD and control participants in their use of expansion (also referred as additive connectives) and temporality connectives (Willits et al., 2018). Our results also showed no significant differences between groups in their use of comparison connectives (including concession, contrast, and similarity connectives). This is partially inconsistent with previous studies reporting that patients with SSD use less contrastive and differentiation connectives than control participants (Just et al., 2020; Willits et al., 2018).

No previous studies have paid specific attention to polysemic

Table 5
Fixed-effects factors in the model that best fitted the data on proportions of connectives.

Model		Estimate	Standard error	*z*-Value	*P*-value
Control participants and connectives of comparison as baseline levels	(Intercept)	−3.24	0.04	−72.16	< .001
	Group				
	Patients	0.009	0.06	0.14	.88
	Type of connective				
	Contingency	−0.43	0.07	−6.11	< .001
	Expansion	0.34	0.05	6.03	< .001
	Multiclass	−0.07	0.06	−1.17	.23
	Temporality	−0.97	0.07	−12.57	< .001
	Interactions				
	Patients*Contingency	−0.26	0.10	−2.61	.008
	Patients*Expansion	−0.08	0.08	−0.98	.32
	Patients*Multiclass	−0.35	0.08	−4.16	< .001
	Patients*Temporality	−0.10	0.10	−0.99	.31
Patients and connectives of comparison as baseline levels	(Intercept)	−3.23	0.04	−71.94	< .001
	Group				
	Controls	−0.009	0.06	−0.14	.88
	Type of connective				
	Contingency	−0.70	0.07	−9.80	< .001
	Expansion	0.26	0.05	4.64	< .001
	Multiclass	−0.42	0.06	−7.06	< .001
	Temporality	−1.08	0.07	−13.98	< .001
	Interactions				
	Controls*Contingency	0.26	0.10	2.61	.008
	Controls*Expansion	0.08	0.08	0.98	.32
	Controls*Multiclass	0.35	0.08	4.16	< .001
	Controls*Temporality	0.10	0.10	0.99	.31

Note: estimates are "log odds" (i.e., logits). Positive estimates reflect an increase in probability and negative ones reflect a decrease. For a general guide to mixed-effects models in linguistics and their interpretation, see Baayen (2008) and Winter (2020).

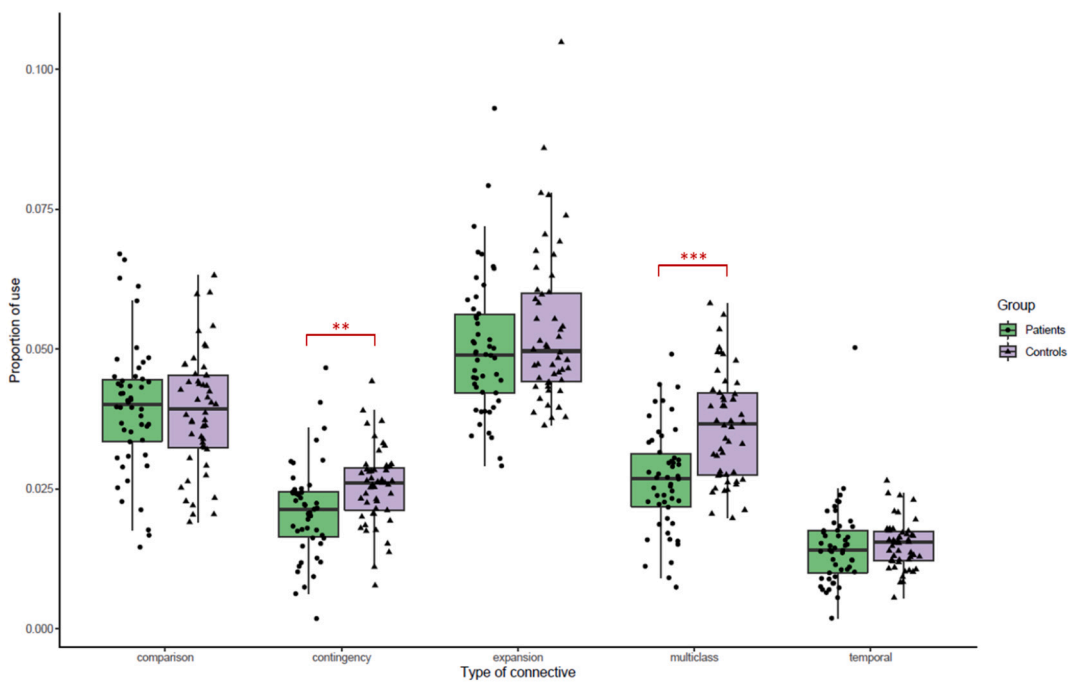

Fig. 2. Mean proportion of use for each type of connective per group.

Table 6
Random-effects parameters in the best model fitted to proportion of connectives.

Groups	Name	Variance	Standard deviation
Participant	(Intercept)	0.1012	0.3181
	Contingency	0.2587	0.5086
	Expansion	0.1672	0.4089
	Multiclass	0.1844	0.4294
	Temporality	0.3004	0.5481

connectives, as the ones included in our multiclass category. In our study, the proportion of use of multiclass connectives was significantly smaller for patients with SSD than for control participants. Polysemic words are processed faster than unambiguous words (Eddington and Tokowicz, 2015; Klepousniotou and Baum, 2007). It is possible that this fast cognitive processing of polysemic connectives was different between patients with SSD and control participants in our sample, reflecting itself in the smaller proportion of polysemic connectives used by the patients. Yet, some variables that were not controlled for might

Table 7
Differences in connectives-related similarity measures between groups.

Type of connective	Coherence measure	Patients (n = 50), average (SD)	Controls (n = 50), average (SD)	Statistic		*r* effect size
				W	Adj. *P*-Value	
Comparison	Mean	0.51 (0.01)	0.50 (0.01)	881	.23	
	Median	0.51 (0.01)	0.51 (0.01)	920	.41	
	Maximum	0.62 (0.04)	0.63 (0.04)	1328	≈.99	
	Minimum	0.38 (0.04)	0.37 (0.02)	987	.91	
	Range	0.23 (0.05)	0.25 (0.05)	1421	≈.99	
	Variance	0.003 (0.001)	0.003 (0.001)	964	.68	
Contingency	Mean	0.48 (0.03)	0.47 (0.01)	827	.08	
	Median	0.49 (0.03)	0.48 (0.01)	814	.06	
	Maximum	0.59 (0.06)	0.59 (0.03)	1307	≈.99	
	Minimum	0.36 (0.06)	0.33 (0.03)	832	.09	
	Range	0.23 (0.09)	0.26 (0.06)	1565	.48	
	Variance	0.004 (0.003)	0.004 (0.001)	1129	≈.99	
Expansion	Mean	0.46 (0.01)	0.46 (0.01)	1190	≈.99	
	Median	0.45 (0.01)	0.45 (0.01)	1205	≈.99	
	Maximum	0.61 (0.05)	0.66 (0.06)	1792	.005	0.37
	Minimum	0.33 (0.03)	0.31 (0.03)	937	.48	
	Range	0.28 (0.07)	0.34 (0.06)	1814	.002	0.38
	Variance	0.005 (0.002)	0.005 (0.002)	1307	≈.99	
Multiclass	Mean	0.50 (0.02)	0.50 (0.01)	1235	≈.99	
	Median	0.49 (0.02)	0.49 (0.01)	1133	≈.99	
	Maximum	0.62 (0.06)	0.64 (0.03)	1592	.35	
	Minimum	0.38 (0.04)	0.35 (0.04)	797	.04	−0.31
	Range	0.23 (0.08)	0.28 (0.05)	1704	.04	0.31
	Variance	0.004 (0.002)	0.004 (0.001)	1239	≈.99	
Temporality	Mean	0.43 (0.02)	0.42 (0.01)	882	.23	
	Median	0.43 (0.03)	0.42 (0.01)	929	.46	
	Maximum	0.52 (0.06)	0.52 (0.05)	1342	≈.99	
	Minimum	0.35 (0.03)	0.31 (0.03)	603	.0002	−0.44
	Range	0.16 (0.07)	0.20 (0.06)	1638	.16	
	Variance	0.003 (0.002)	0.003 (0.001)	1328	≈.99	

n = sample size, SD = standard deviation.

have influenced our results. For instance, semantic activation of polysemic words is influenced by word frequency and context (Rice et al., 2019), and pragmatic inferences play a role in this as well (Carston, 2021). In parallel, it has been argued that connections between the word-forms and meaning of words in the mental lexicon are weaker in patients with SSD than in control participants (Kuperberg et al., 2019). Whether any of these factors relate to our results of the use of these

Table 8
Comparison of classification performance between connectives' features and non-connectives similarity.

Name of classifier	Name of features included in classification task							Number of features included in the model	Accuracy	Sensitivity	Specificity	AUC-ROC
	Sliding-window similarity	Non-connectives similarity	General connectives-related similarity (random sub-sampling)	General connectives-related similarity (all chunks)	Per-type-connectives-related similarity with smallest adj. *p*-values	Proportions per type of connective	Proportions and per-type-connectives-related similarity significantly different between groups					
RF-c	✓							18	83.5 %	79.4 %	87.6 %	0.89
RF-non-conn	✓	✓						24	84.9 %	79.4 %	90.4 %	0.89
RF-conn-I	✓		✓					24	**85 %**	**81.2 %**	88.8 %	0.88
RF-conn-II	✓			✓				24	84.7 %	79.2 %	90.2 %	0.89
RF-conn-III	✓				✓			24	84.9 %	80.4 %	89.4 %	0.90
RF-conn-IV	✓					✓		23	83.8 %	80 %	87.6 %	0.88
RF-conn-V	✓						✓	25	84.4 %	75.2 %	**93.6 %**	**0.90**

RF = Random Forest algorithm.

Table 9
Results of using connectives-related cosine similarities either alone or along with proportion of connectives in distinguishing patients from controls.

Algorithm	Name of classifier	Name of features included in classification task		Number of features included in the model	Accuracy	Sensitivity	Specificity	AUC-ROC
		Connectives-related similarities per type of connective	Proportions per type of connective					
RF	RF-I	✓		30	79.4 %	75 %	83.8 %	0.84
	RF-II	✓	✓	35	80.1 %	77.4 %	82.8 %	0.89
SVM-pk	SVM-I	✓		30	78.4 %	77.6 %	79.2 %	0.85
	SVM-II	✓	✓	35	**85 %**	**83.8 %**	**86.2 %**	**0.91**

RF = Random Forest; *SVM-pk* = Support Vector Machine with polynomial kernel.

polysemic connectives remains an open question.

Overall, our results suggest that speech of patients with SSD is characterized by a relative reduction in the use of contingency connectives (i.e., markers of cause, condition, and purpose) and multiclass connectives (i.e., markers that can establish more than one explicit type of semantic relation between clauses and/or sentences).

4.2. Connectives-related similarity

Previous research has shown that, compared to control participants, patients with SSD reach lower scores in semantic similarity (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2019). There is consistency between those findings and our results showing that patients with SSD had lower scores in three out of the five connectives-related similarity measures that significantly differed between groups.

In detail, our results showed that patients with SSD had narrower range in similarity of expansion connectives. Expansion connectives establish additive relations with either positive (e.g., books *and* notebooks) or negative polarity (e.g., books *or* notebooks) (Evers-Vermeul and Sanders, 2009; Sanders et al., 1992). The narrower range in similarity of expansion connectives might indicate that there is less semantic variation in the words, clauses and sentences added together by patients with SSD. Patients also had a lower maximum similarity of expansion connectives. This might mean that, compared to control participants, patients with SSD added together words, clauses and/or sentences that shared less semantic features. Of notice, maximum similarity of expansion connectives positively correlated to years of education. The results of maximum similarity of expansion connectives could therefore be a reflection of education level, rather than a difference between patients with SSD and control participants.

Intriguingly, in our study, patients with SSD showed higher minimum similarity of temporality connectives. With few exceptions (Panicheva and Litvinova, 2019), this is opposite to the majority of previous reports showing that patients with SSD often have lower semantic similarity scores than control participants (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2019). Temporality connectives establish ordered relations between series of events (Evers-Vermeul and Sanders, 2009; Sanders et al., 1992). Patients with SSD use temporality connectives as core linguistic devices to achieve coherence in narrative discourse (Saavedra, 2010). Interestingly, cognitively well-functioning patients with SSD can achieve temporal coherence similar to that of control participants (Holm et al., 2016). In our sample of patients with SSD, their total PANSS score (see Table 2) indicates that, cognitively, they were well-functioning (Leucht et al., 2005). Thus, our results show that patients with SSD can use temporality connectives as coherently as control participants during semi-structured interviews, suggesting that their use of temporality connectives might be related to cognitive well-functioning.

As well, patients with SSD had narrower range and higher minimum similarity of multiclass connectives. The narrower range might mean that patients had less semantic variation in the words conforming the previous and following context of multiclass connectives. Accordingly, the higher minimum similarity of multiclass connectives would reflect the low-end of such narrower range of semantic variation.

4.3. Classification tasks

4.3.1. Connectives' vs non-connectives' features

Among all classifiers, connectives' features and non-connectives cosine similarity yielded similar accuracies in distinguishing patients with SSD from control participants. When controlling for amount of data and number of predictors, general connectives-related similarity (RF-conn-I) seemed to increase sensitivity to classify patients with SSD. However, it is likely that this was due to the random sub-sampling procedure, because general connectives-related similarity (RF-conn-II) no longer increased sensitivity when the similarity was calculated based on all connectives-chunks. In contrast, the classifier using the connectives' features that were significantly different between groups (RF-conn-V) yielded 6 % more specificity than the control classifier (RF-c), suggesting that connectives' features that were significantly different between groups are useful to correctly classify true negatives. Overall, our results suggest that connectives' features and non-connectives similarity can reach similar results in distinguishing patients with SSD from control participants.

4.3.2. Performance of connectives' features alone

SVM-II yielded 85 % accuracy, 83.8 % sensitivity and 86.2 % specificity. These percentages are within the accuracy range reported in previous studies (for reviews, see Corcoran et al., 2020; Corcoran and Cecchi, 2020; Hitczenko et al., 2021). Accuracy of 85 % had been previously obtained by our group using a full sliding-window general-semantic-similarity classifier (Voppel et al., 2021). SVM-II included connectives' features alone, and it was trained based on less speech input (i.e., only connectives and their surrounding words). This suggests that connectives and their surrounding words are linguistic loci that might concentrate important patterns to detect atypical coherence related to SSD.

4.4. Limitations and future directions

We acknowledge our study has limitations. Our proportion-of-connectives' results could not be straightforwardly compared to previous findings due to differences in the control group (healthy participants in our study and HIV+ in Willits et al., 2018) and in the analysis technique that was employed (mixed-effects logistic regression models in our study and principal component analysis in Mackinley et al., 2021). Also, the number of categories of connectives varied across studies (ranging from two to five), as well as the number of connectives per category and the number of types of connectives inside each main category. Similarly, there were inconsistencies in annotation schemes for connectives. For instance, in contrast to previous studies (Just et al., 2020; Mackinley et al., 2021; Willits et al., 2018), the annotation scheme that we used (Bourgonje et al., 2018; Stede et al., 2019) did not include a separate category for logical connectives, following the line of reasoning that there are no *logical connectives*, but rather abstract *logical operators* that then can have linguistic correlates in different types of connectives (Sanders et al., 1992). Furthermore, fillers and repetitions were not removed from the transcripts used for our analyses, even though they are known to influence cosine similarity calculations (Iter et al., 2018).

In replicating or expanding our results, future studies on the use of connectives in speech of patients with SSD should take into account a series of methodological challenges. The first is a consistent use of connectives categories across studies, which would aid knowledge accumulation. Secondly, our theory-based decision of using a seven-words window alone for our analyses decreased the Type I error rate of our findings. However, it remains to be determined whether this is the most appropriate window for the assessment of cosine similarity between connectives and their surrounding words. This was beyond the scope of the current study, but future research may specifically address the role of window size by directly comparing a range of different window sizes. In relation to this, it would be necessary to analyze what procedure to calculate the connectives-related similarity is the most reliable, valid and theoretically sounded. More recent computational semantic representations could be used to obtain the word embeddings for analyses, such as Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2019), possibly in combination with time-series analyses of semantic coherence (Xu et al., 2022). Also, mixed designs (e.g., Holm et al., 2016; Saavedra, 2010) might provide valuable details overlooked by quantitative approaches alone, increasing our comprehension of how thematic continuity and syntactic connectivity (in)dependently build up (in)coherence in patients with SSD.

Additionally, we need to understand how the use of connectives might be influenced by speech elicitation techniques and cognitive factors. For instance, speech elicitation techniques (e.g., re-telling a story or reading a text out loud) have been shown to influence linguistic outcomes (Kapantzoglou et al., 2017; Niebuhr and Michaud, 2015), and some types of connectives are more cognitively demanding to use than others (Evers-Vermeul and Sanders, 2009; Zufferey and Gygax, 2020). For these reasons, in future studies it would be informative to assess whether the use of connectives differs among speech elicitation techniques, and whether this might also depend on the varying cognitive demands of different connectives. This emphasizes the importance of exploring possible relations between patterns of connectives' use and cognitive outcomes in patients with SSD.

Also, future research should examine the generalizability of our findings on the use of connectives to other languages and across generations of speakers. There is evidence that patterns of word use are consistent across different linguistic families (Calude and Pagel, 2011). However, word order related to grammatical connectivity varies across languages (Lehmann, 2011), and word order changes throughout the history of languages (Gell-Mann and Ruhlen, 2011; Maurits and Griffiths, 2014). For instance, these days Subject-Object-Verb (SOV) is the canonical word order of a subordinated declarative sentence introduced by a connective in Dutch (Jordens, 1988; Koster, 1975), while English has a fixed SVO structure (Comrie, 1981) and Spanish can have either of these (López Meirama, 2006). These syntactic structures might be different for future generations of Dutch, English or Spanish speakers. Thus, both cross-linguistic and historical-grammar factors await an exploration to further our understanding of the use of connectives as signifying patterns of speech (dis)organization in patients with SSD.

5. Conclusions

Connectives' features are informative and explainable variables that can be used to reliably assess disorganized speech in patients with SSD. The combination of this method with other linguistic components is a promising venue to further improve accuracy in categorizing individuals with SDD and control participants. Such fine-tuned automatic analyses of speech samples will help to reach the ultimate aim of advancing clinical practice.

CRediT authorship contribution statement

H. Corona-Hernández: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft. **J.N. de Boer:** Conceptualization, Methodology, Writing – review & editing. **S.G. Brederoo:** Conceptualization, Methodology, Writing – review & editing. **A.E. Voppel:** Formal analysis, Writing – review & editing. **I.E.C. Sommer:** Supervision, Writing – review & editing.

Declaration of competing interest

All authors declare that they have no conflicts of interest.

Acknowledgement

HCH was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604).

Role of the funding source

The founding source(s) had no role neither in conducting this research nor in deciding to publish this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2022.06.013.

References

American Psychiatric Association, 2013. Diagnostic And Statistical Manual of Mental Disorders: DSM-5, 5th ed. American Psychiatric Association, Washington, D.C.
Andreasen, N.C., 1979. Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability. Arch. Gen. Psychiatry 36, 1315–1321.
Andreasen, N.C., Flaum, M., Arndt, S., 1992. The Comprehensive Assessment of Symptoms and History (CASH): an instrument for assessing diagnosis and psychopathology. Arch. Gen. Psychiatry 49, 615–623.
Baayen, R.H., 2008. Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Cambridge University Press, New York, NY.
Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67, 1–48.
Bedi, G., Carrillo, F., Cecchi, G.A., Slezak, D.F., Sigman, M., Mota, N.B., Ribeiro, S., Javitt, D.C., Copelli, M., Corcoran, C.M., 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophr. 1, 1–7.
Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., Matsuo, A., 2018. quanteda: an R package for the quantitative analysis of textual data. J. Open Source Softw. 3, 774.
Bourgonje, P., Hoek, J., Evers-Vermeul, J., Redeker, G., Sanders, T., Stede, M., 2018. Constructing a lexicon of Dutch discourse connectives. Comput. Linguist. Neth. J. 8, 163–175.
Calude, A.S., Pagel, M., 2011. How do we use language? Shared patterns in the frequency of word use across 17 world languages. Philos. Trans. Biol. Sci. 366, 1101–1107.
Carston, R., 2021. Polysemy: pragmatics and sense conventions. Mind Lang. 36, 108–133.
Comrie, B., 1981. Language Universals And Linguistic Typology: Syntax And Morphology, 2nd ed. The University of Chicago Press, Chicago, Ill.
Corcoran, C.M., Cecchi, G.A., 2020. Using language processing and speech analysis for the identification of psychosis and other disorders. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 5, 770–779.
Corcoran, C.M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D.C., Bearden, C.E., Cecchi, G.A., 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry 17, 67–75.
Corcoran, C.M., Mittal, V.A., Bearden, C.E., Gur, R.E., Hitczenko, K., Bilgrami, Z., Savic, A., Cecchi, G.A., Wolff, P., 2020. Language as a biomarker for psychosis: a natural language processing approach. Schizophr. Res. 226, 158–166.
de Boer, J.N., Voppel, A.E., Begemann, M.J.H., Schnack, H.G., Wijnen, F., Sommer, I.E.C., 2018. Clinical use of semantic space models in psychiatry and neurology: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 93, 85–92.
Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In: Presented at the NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf, pp. 4171–4186.
Eddington, C.M., Tokowicz, N., 2015. How meaning similarity influences ambiguous word processing: the current state of the literature. Psychon. Bull. Rev. 22, 13–37.
Elvevåg, B., Foltz, P.W., Weinberger, D.R., Goldberg, T.E., 2007. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. Schizophr. Res. 93, 304–316.
Evers-Vermeul, J., Sanders, T., 2009. The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition. J. Child Lang. 36, 829–854.

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we need hundreds of classifiers to solve real world classification Problems? J. Mach. Learn. Res. 15, 3133–3181.
Gell-Mann, M., Ruhlen, M., 2011. The origin and evolution of word order. Proc. Natl. Acad. Sci. 108, 17290–17295.
Givón, T., 2020. Coherence. John Benjamins Publishing Company, Amsterdam, The Netherlands.
Hitczenko, K., Mittal, V.A., Goldrick, M., 2021. Understanding language abnormalities and associated clinical markers in psychosis: the promise of computational methods. Schizophr. Bull. 47 (2), 344–362.
Holm, T., Thomsen, D.K., Bliksted, V., 2016. Life story chapters and narrative self-continuity in patients with schizophrenia. Conscious. Cogn. 45, 60–74.
Hong, K., Nenkova, A., March, M.E., Parker, A.P., Verma, R., Kohler, C.G., 2015. Lexical use in emotional autobiographical narratives of persons with schizophrenia and healthy controls. Psychiatry Res. 225, 40–49.
Iter, D., Yoon, J., Jurafsky, D., 2018. Automatic detection of incoherent speech for diagnosing schizophrenia. In: Proceedings of the Fifth Workshop on Computational Linguistics And Clinical Psychology: From Keyboard to Clinic. Association for Computational Linguistics, New Orleans, LA, pp. 136–146.
Jordens, P., 1988. The acquisition of word order in Dutch and German as L1 and L2. Second. Lang. Res. 4, 41–65.
Just, S., Haegert, E., Kořánová, N., Bröcker, A.-L., Nenchev, I., Funcke, J., Montag, C., Stede, M., 2019. Coherence models in schizophrenia. In: Proceedings of the Sixth Workshop on Computational Linguistics And Clinical Psychology. Presented at the Proceedings of the Sixth Workshop on Computational Linguistics And Clinical Psychology. Association for Computational Linguistics, Minneapolis, Minnesota, pp. 126–136.
Just, S.A., Haegert, E., Kořánová, N., Bröcker, A.-L., Nenchev, I., Funcke, J., Heinz, A., Bermpohl, F., Stede, M., Montag, C., 2020. Modeling incoherent discourse in non-affective psychosis. Front. Psychiatry 11. https://doi.org/10.3389/fpsyt.2020.00846.
Kapantzoglou, M., Fergadiotis, G., Restrepo, M.A., 2017. Language sample analysis and elicitation technique effects in bilingual children with and without language impairment. J. Speech Lang. Hear. Res. 60, 2852–2864.
Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276.
Klepousniotou, E., Baum, S.R., 2007. Disambiguating the ambiguity advantage effect in word recognition: an advantage for polysemous but not homonymous words. J. Neurolinguistics 20, 1–24.
Koster, J., 1975. Dutch as an SOV language. Linguist. Anal. 1, 111–136.
Kuhn, M., 2021. caret: Classification And Regression Training.
Kuperberg, G.R., Weber, K., Delaney-Busch, N., Ustine, C., Stillerman, B., Hämäläinen, M., Lau, E., 2019. Multimodal neuroimaging evidence for looser lexico-semantic networks in schizophrenia: evidence from masked indirect semantic priming. Neuropsychologia 124, 337–349.
Landauer, T.K., Foltz, P.W., Laham, D., 1998. An introduction to latent semantic analysis. Discourse Process. 25, 259–284.
Lehmann, C., 2011. Towards a typology of clause linkage. In: Haiman, J., Thompson, S.A. (Eds.), Clause Combining in Grammar And Discourse. John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 181–226.
Leucht, S., Kane, J.M., Kissling, W., Hamann, J., Etschel, E., Engel, R.R., 2005. What does the PANSS mean? Schizophr. Res. 79, 231–238.
López Meirama, B., 2006. Semantic and discourse-pragmatic factors in Spanish word order. In: Clements, J.C., Yoon, J. (Eds.), Functional Approaches to Spanish Syntax: Lexical Semantics, Discourse And Transitivity. Palgrave Macmillan, London, pp. 7–52.
Maat, H.P., Sanders, T., 2006. Connectives in text. In: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 33–41.
Mackinley, M., Chan, J., Ke, H., Dempster, K., Palaniyappan, L., 2021. Linguistic determinants of formal thought disorder in first episode psychosis. Early Interv. Psychiatry 15, 344–351.
MacWhinney, B., 2000. The CHILDES Project: Tools for Analyzing Talk, 3rdEdition. Lawrence Erlbaum Associates, Mahwah, NJ.
Maher, B., 1972. The language oi schizophrenia: a review and interpretation. Br. J. Psychiatry 120, 3–17.
Maurits, L., Griffiths, T.L., 2014. Tracing the roots of syntax with Bayesian phylogenetics. Proc. Natl. Acad. Sci. 111, 13576–13581.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. In: ICLR Workshop.
Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013b. Distributed representations of words and phrases and their compositionality. In: Procof the 26th International Conference on Neural Information Processing Systems, pp. 3111–3119.
Niebuhr, O., Michaud, A., 2015. Speech data acquisition: the underestimated challenge. Kiel. Arb. Linguist. Phon. KALIPHO 3, 1–42.
Panicheva, P., Litvinova, T., 2019. Semantic coherence in schizophrenia in Russian written texts. In: Presented at the 25th Conference of Open Innovations Association (FRUCT), pp. 241–249.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rezaii, N., Walker, E., Wolff, P., 2019. A machine learning approach to predicting psychosis using semantic density and latent content analysis. NPJ Schizophr. 5, 9. https://doi.org/10.1038/s41537-019-0077-9.
Rice, C.A., Tokowicz, N., Fraundorf, S.H., Liburd, T.L., 2019. The complex interactions of context availability, polysemy, word frequency, and orthographic variables during lexical processing. Mem. Cogn. 47, 1297–1313.
RStudio Team, 2019. RStudio: Integrated Development Environment for R. RStudio Inc, Boston, MA.
Saavedra, J., 2010. Quantitative criteria of narrative coherence and complexity in persons with paranoid schizophrenia. J. Nerv. Ment. Dis. 198, 349–355.
Sanders, T., Maat, H.P., 2006. Cohesion and coherence: linguistic approaches. In: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 591–595.
Sanders, T.J.M., Spooren, W.P.M., Noordman, L.G.M., 1992. Toward a taxonomy of coherence relations. Discourse Process. 15, 1–35.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 quiz 34–57.
Stede, M., Scheffler, T., Mendes, A., 2019. Connective-Lex: a web-based multilingual lexical resource for connectives. Discours 24.
Tang, S.X., Kriz, R., Cho, S., Park, S.J., Harowitz, J., Gur, R.E., Bhati, M.T., Wolf, D.H., Sedoc, J., Liberman, M.Y., 2021. Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. NPJ Schizophr. 7.
van der Vliet, N., Redeker, G., 2014. Explicit and implicit coherence relations in Dutch texts. In: Gruber, H., Redeker, G. (Eds.), The Pragmatics of Discourse Coherence: Theories And Applications. John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 23–52.
van Eerten, L., 2007. Over het Corpus Gesproken Nederlands. Ned. Taalkunde 12, 194–2015.
Voppel, A.E., de Boer, J.N., Brederoo, S.G., Schnack, H.G., Sommer, I.E.C., 2021. Quantified language connectedness in schizophrenia-spectrum disorders. Psychiatry Res. 304, 114130 https://doi.org/10.1016/j.psychres.2021.114130.
Wijffels, J., 2020. word2vec: Distributed Representations of Words.
Willits, J.A., Rubin, T., Jones, M.N., Minor, K.S., Lysaker, P.H., 2018. Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. Schizophr. Res. 197, 365–369.
Winter, B., 2020. Statistics for Linguists: An Introduction Using R. Routledge, New York, NY.
Xu, W., Wang, W., Portanova, J., Chander, A., Campbell, A., Pakhomov, S., Ben-Zeev, D., Cohen, T., 2022. Fully automated detection of formal thought disorder with Time-series Augmented Representations for Detection of Incoherent Speech (TARDIS). J. Biomed. Inform. 126, 103998.
Zufferey, S., Gygax, P., 2020. Do teenagers know how to use connectives from the written mode? Lingua 234, 102779.