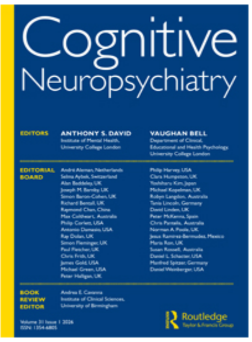

Cognitive Neuropsychiatry

ISSN: 1354-6805 (Print) 1464-0619 (Online) Journal homepage: www.tandfonline.com/journals/pcnp20

Characterisation of alogia in people with schizophrenia using natural language processing. a comparative study with a group of people with bipolar disorder

Alvaro Cavieres, Francisca Alonso, Amanda Cifuentes, Vanessa Acuña, Jorge Portal, Juan Satt, Monica Otero & Pablo Lopez-Silva

To cite this article: Alvaro Cavieres, Francisca Alonso, Amanda Cifuentes, Vanessa Acuña, Jorge Portal, Juan Satt, Monica Otero & Pablo Lopez-Silva (27 Jul 2026): Characterisation of alogia in people with schizophrenia using natural language processing. a comparative study with a group of people with bipolar disorder, Cognitive Neuropsychiatry, DOI: 10.1080/13546805.2026.2703583

To link to this article: https://doi.org/10.1080/13546805.2026.2703583

Published online: 27 Jul 2026.

COGNITIVE NEUROPSYCHIATRY
https://doi.org/10.1080/13546805.2026.2703583

Characterisation of alogia in people with schizophrenia using natural language processing. a comparative study with a group of people with bipolar disorder

Alvaro Cavieres[a], Francisca Alonso[b], Amanda Cifuentes[a], Vanessa Acuña[a], Jorge Portal[c], Juan Satt[b], Monica Otero[d] and Pablo Lopez-Silva[e,f,g,h]

[a]Departamento de Psiquiatría, Universidad de Valparaíso, Valparaíso, Chile; [b]CIDCL, Escuela de Fonoaudiología, Facultad de Medicina, Universidad de Valparaíso, Valparaíso, Chile; [c]Advanced Center for Electrical and Electronic Engineering (AC3E), Universidad Tecnica Federico Santa Maria, Valparaıso, Chile; [d]Facultad de Ingeniería, Universidad San Sebastián, Santiago, Chile; [e]School of Psychology, Universidad de Valparaíso, Valparaíso, Chile; [f]Research Centre in Cognition and Language Development, Universidad de Valparaíso, Valparaíso, Chile; [g]Institute for Complex Systems, Valparaíso, Chile; [h]Collaborating Centre for Value-Based Research, St Catherine's College, University of Oxford, Oxford, UK

ABSTRACT

Using Natural Language Processing (NLP)-derived metrics, we examined poverty of speech in a group of people diagnosed with schizophrenia compared to a group of people with bipolar disorder and a control group, intending to describe a pattern of alterations specific to schizophrenia and to help clarify the degree of association with deficit symptoms and cognitive deficits.

Methods: Ninety participants were recruited for the study, distributed equally in three groups: schizophrenia (SZ), bipolar disorder (BD), and healthy controls (HC). All participants described scenes depicting people interacting with each other in everyday social situations, and their cognitive functioning was assessed using the screen for cognitive impairment in Psychiatry (SCIP). Clinical status was assessed with the Brief Psychiatric Rating Scale (BPRS). Speech samples were explored with metrics that quantify speech production and complexity.

Results: People with schizophrenia exhibited lower verbal production and reduced sentence complexity compared to healthy controls, while the bipolar disorder group showed values largely overlapping with controls. Across the pooled patient sample (SZ + BD), negative symptom severity was associated with multiple linguistic measures, including mean length of utterance ($r = -0.384$, $BF_{10} = 14.03$, 95% CI [−0.571, −0.138]), syntagms, discourse complexity, and lexical–morphological indices, whereas positive symptoms showed little or no association with language metrics**.** Better cognitive functioning was associated with greater linguistic richness across most measures.

Conclusions: The findings from this study suggest that poverty of speech, as captured by NLP-derived measures, is more prominent in schizophrenia than in bipolar disorder and healthy controls within the present sample. The observed associations with cognitive performance highlight the close relationship between language

ARTICLE HISTORY
Received 8 April 2025
Accepted 1 July 2026

KEYWORDS
Natural language processing; Alogia; Schizophrenia

CONTACT Alvaro Cavieres cavieres.alvaro@gmail.com Departamento de Psiquiatría, Universidad de Valparaíso, Carvallo 200 Playa Ancha, Valparaíso, Chile

production and cognition, but do not establish a causal dependence. Rather than prompting a re-examination of alogia's classification among negative symptoms, these results underscore the need for integrative models that account for the well-known covariation between negative symptoms and cognitive functioning.

Introduction

Abnormalities in language and speech are common in schizophrenia (Bhati, 2005; Kuperberg, 2010). They are frequently related to impairments in several dimensions of patients' everyday functioning (Bowie & Harvey, 2008). These alterations are often categorised as (a) positive thought disorder, characterised by disorganised, circumstantial, or tangential speech, and (b) negative thought disorder or scarce speech production (Kircher et al., 2014). In this context, the term "poverty of speech," used to refer to a decrease in the amount of speech and/or information conveyed by the subject, has been recognised as one of the symptoms of schizophrenia since the earliest descriptions of the disease. While Kraepelin (Diefendorf & Kraepelin, 1907) noted that patients "seem to have a cessation of the need to express themselves," Bleuler (1950) observed that "the blocking, the poverty of ideas, the incoherence, the obfuscation, the delusions and emotional delusions and emotional abnormalities are expressed in the patients' language." Over the years, the precise definition of the very term, the phenomena it captures, and the aspects considered most relevant for its assessment have undergone significant modifications.

Poverty of speech has also been called "alogia." Possibly the first author to use the term in a description of the symptomatology of persons with schizophrenia was Kleist (1914). The author introduced the term to refer to a failure of thought, which he also described as present in cases of dementia. Kleist would characterise alogia as poverty in the generation and expression of thought. In 1961, Wing (1961) incorporated the section "Poverty of content, or restriction of quantity of speech" in the first version of the Present State Examination with items assessing both the quantity and the content of the speech. The final version (1974) only included an item about the information provided by the subject. Later, in 1979, in the Thought Language and Communication Scale, Andreasen (1979) distinguished between Poverty of Speech – as the restriction of the quantity of spontaneous speech – and Poverty of Speech Content, defined as the delivery of a reduced amount of information without a decrease in verbal production.

The negative symptoms subscale of the PANSS (Kay, 1991) is less explicit about the importance of the quantity of speech. Instead, it includes two items: (a) "Lack of Spontaneousness and Fluidity in Conversation" defined as a reduction in normal conversational fluency associated with apathy, unwillingness, helplessness, or cognitive deficit, and (b) "Stereothypical Thinking" defined as decreased fluency, spontaneity, and flexibility of thought, evidenced by rigid, repetitive, or sterile content. However, this last item and "difficulty in abstract thinking" were not considered within the negative symptom domain in later factor analyses (Daniel, 2013). Moreover, to avoid the overlap between negative symptoms and cognitive deficit, these items were excluded from assessment instruments developed after the NIMH-MATRICS Consensus on Negative Symptoms (Kirkpatrick et al., 2006).

Nevertheless, the relationship between verbal production and cognitive functioning is still not fully understood. Early studies used verbal fluency as a quantitative measure of speech and found a negative correlation with scores in a cognitive battery (Joyce et al., 1996). Using the PANSS alogia items, another piece of research found that disorganisation of semantic memory underlies alogia in schizophrenia (Sumiyoshi et al., 2005). Another study using the SANSS found an association between poverty of speech and difficulties in goal maintenance and verbal working memory tasks (Becker et al., 2012). A different study examined whether the two negative symptom dimensions of diminished expression and apathy are differentially related to neurocognitive impairment in schizophrenia (Hartmann-Riemer et al., 2015), finding that only diminished expression (alogia, blunted affect) meaningfully predicts neurocognitive deficits. Here, it is worth mentioning a study that found a specific overlap in variance between alogia ratings and verbal fluency scores, leading the authors to conclude that these variables tap into a common underlying construct – reduced verbal output – that may be classified as a negative symptom or as a cognitive deficit depending on the assessment context rather than on distinctions in underlying mechanisms (Fervaha et al., 2016).

Common to all of the studies described is that the assessment of poor speech was made by assigning a score on a clinical scale according to the examiner's subjective opinion of the degree of deviation of the subject from what should be considered normal. In contrast, Natural Language Processing (NLP) studies employ computational algorithms to study the semantic and emotional content, semantic coherence, syntactic structure, and complexity of spontaneous speech samples (Le Glaz et al., 2021). For the case of alogia, these studies represent a clear advantage over the application of verbal fluency tests that place conditions that limit the verbal productivity of the subjects or the evaluation based solely on clinical observation. At the same time, these new resources introduce new terminology. Alterations previously described as tangentiality, derailment, and circumstantiality are now captured by the notion of semantic coherence, while "language poverty" can be described as a decrease in discourse complexity (Corcoran & Cecchi, 2020).

From a contemporary perspective, these methodological advances bring into focus a mechanistic question that remains insufficiently addressed within the negative symptom framework: whether alogia primarily reflects limitations in cognitive resources, deficits in behavioural initiation, or disturbances in discourse planning and linguistic organisation. Importantly, these mechanisms are not specific to diagnosis and may be differentially influenced by sociodemographic variables, cognitive load, and task demands.

In a meta-analysis, Cohen et al. (2014) examined 13 studies collectively including 480 patients with schizophrenia and 326 nonpsychiatric controls employing various objective analyses of natural speech. Overall, speech production (reflecting alogia) was impaired at a large effect size level ($d = -.80$; $k = 13$), whereas speech variability (reflecting blunted affect) was much more modest ($d = -.36$; $k = 2$). Regarding the former, this was largely driven by measures of pause behaviour, as opposed to other aspects of speech (e.g. number of words or utterances).

Against this background, the present study does not aim to establish diagnostic specificity per se, but rather to examine how NLP-derived speech metrics vary across clinical groups while explicitly considering alternative explanatory factors such as educational background and cognitive performance.

The objective of this research is to expand on these results and characterise poverty of speech in a group of people diagnosed with schizophrenia (SZ) compared to a group of people with Bipolar Disorder (BD) and a control group from the general population using NLP-derived metrics. Rather than attributing observed differences exclusively to diagnosis, we examine the extent to which these differences may be accounted for by cognitive performance and other covariates**.** In contrast to the limited amount of information that can be obtained through a clinical assessment, we expect that this method will allow the description of patterns of speech alteration across groups and that correlating the results with scores on the BPRS (Overall & Gorham, 1962) and the SCIP neuropsychological battery (Gómez-Benito et al., 2018) will help clarify the relative contributions of negative symptoms and cognitive deficits to poverty of speech.

Methods

Participants

Ninety participants were recruited for the study, with 30 individuals diagnosed with schizophrenia (SZ), 30 individuals diagnosed with bipolar disorder (BD), and 30 healthy controls (HC). The three groups did not differ on age and sex All participants were between 18 and 55 years of age. Subjects with severe medical or neurological comorbidities, susbstance related disorder (DSM5), or those who were unable to provide informed consent or follow study instructions were not enrolled. In the case of controls, the presence of a psychiatric disorder was ruled out by a psychiatrist through a structured interview using the MINI and having first-degree relatives with a diagnosis of severe psychiatric pathology was considered an exclusion criterion. Schizophrenia and Bipolar disorder were diagnosed according to DSM 5 criteria (American Psychiatric Association, 2013). The Ethics Committee of the Valparaiso San Antonio Health Service approved the study protocol, and all participants gave their written informed consent before participating. All patients were clinically stable and receiving outpatient care at the Hospital del Salvador de Valparaíso. Their treating physicians invited them to participate. One of the research psychiatrists confirmed the subjects' eligibility and administered the structured assessment using the BPRS.

Procedures

All participants diagnosed with bipolar disorder and schizophrenia were evaluated by a psychiatrist using the BPRS. For the neurocognitive assessment, the SCIP neuropsychological battery was used, which comprises five distinct subscales: immediate verbal learning (VL-I), working memory (WM), verbal fluency (VF), delayed verbal learning (VL-D), and processing speed (PS). A total score can also be calculated as the sum of all items.

Speech task: All the participants were asked to describe a set of images displayed on a computer. These scenes depict people interacting with each other in various everyday social situations. The participants were instructed to describe the scene freely, with no time limit. After each scene was shown, the interviewer asked: "What do you think is happening in the scene?" All the responses were recorded and transcribed.

Natural language processing analysis

Two types of language metrics were used to explore discourse structure: metrics that quantify the amount of speech production and metrics that indicate linguistic structure and expressive complexity. Speech production metrics were defined as measures indexing the amount of verbal output, independently of its internal organisation. The amount of speech production considered were mean length utterance (MLU), which is the average length of the sentences in the text, calculated by dividing the total number of words by the total number of sentences; the number of words in the text, counted by separating it into individual words; and the number of sentences in the text, calculated by separating the text into sentences using periods as delimiters.

Linguistic complexity metrics were defined as measures indexing the internal organisation, hierarchical structure, richness of discourse, rather than its quantity**.** The complexity metrics considered were the number of relative clauses or clauses that connect ideas by pronouns referring to something previously mentioned; the number of completive clauses, which complete the meaning of a noun or verb; the number of proper adverbial clauses, which play the same role as adverbs (time, place, manner, and quantity); and the number of improper adverbial clauses, which play the role of circumstantial complements but cannot be replaced by adverbs. Four clause-level metrics were derived from the dependency parse assigned by the es_core_news_lg spaCy model. Relative clauses were identified by the presence of at least one token tagged acl:relcl (adjectival relative clause) per sentence. Completive clauses were counted when at least one token carried the ccomp (clausal complement) tag, marking subordinate clauses that complete the meaning of the main verb. Proper adverbial clauses were identified via the advcl tag (adverbial clause modifier), corresponding to full subordinate clauses functioning as adverbial modifiers, whereas improper adverbial clauses were identified via advmod (adverbial modifier), capturing single words or phrases – rather than full clauses – that function adverbially. For all four metrics, the unit of analysis is the sentence, and values reflect raw counts across the transcript. Overall syntactic complexity was calculated by a normalised composite index computed per sentence, incorporating: number of tokens, number of clauses, maximum parse tree depth, mean word length, and lexical diversity. The raw composite is divided by the total number of tokens in the sentence to allow comparison across utterances of different lengths.

The number of syntagmas was defined as the number of words that functioned together in a sentence structure. Syntactic dependency was calculated as an index of hierarchical relations between words within a sentence, capturing the degree to which sentence elements depend on one another in a structured manner; this measure reflects syntactic organisation and planning demands rather than speech production per se.

The morphology index was calculated by considering the number of unique morphological forms in relation to the total number of words. Coordinating conjunctions were calculated by counting the number of words that function as coordinating connectors, such as “and,” “or,” and “nor,” among others. Subordinating conjunctions were calculated by counting the number of words that function as subordinating connectors, such as “because,” “if,” and “when,” among others. Causal connectors were calculated by counting words that link two sentences together describing cause and effect, such as “so,” “since,” or “as.”

Statistical analysis

We analyzed these metrics using Bayesian ANOVA to compare the two hypotheses, with the group as a factor against the null model. We used an r-scale fixed effect of 0.5 as a prior and, for the Post Hoc test, a Cauchy prior (0, r = 1/sqrt(2)). This design is sensitive to effect sizes that are meaningful within our prior specification, which reflects an expectation of moderate effect sizes and is consistent with recommendations in the literature (Rouder et al., 2012). Bayesian apprach was selected over frequentist alternatives because it quantifies graded evidence for both the alternative and the null hypothesis via the Bayes Factor (BF_{10}), avoiding the binary limitations of null hypothesis significance testing (Rouder et al., 2012). BF_{10} values are interpreted as follows: 1–3 anecdotal, 3–10 moderate, 10–30 strong, and >30 extreme evidence in favour of the alternative; $BF_{10} < 1/3$ constitutes evidence for the null. Post hoc Posterior Odds represent the fully updated belief ratio combining prior odds with the likelihood of the data and quantify the relative probability that each pair of group means differs after observing the data. We used an independent multinomial test with a prior concentration of 1 to analyze the nominal variables. Then, we used a general linear model to analyze the covariance within factors between the BD and the SZ groups. We incorporated the cognitive performance and the positive and negative symptoms as covariates and compared the models. Finally, Bayesian correlation analysis was performed to examine the relationship between discourse structure metrics and SCIP and BPRS scores to account for interactions with cognitive and symptom variables. Within this framework, the sample size was adequate to detect medium-to-large effects, as reflected in the strength of evidence observed for several key group comparisons. All statistical analyses were performed using JASP version 0.14.0.1.

Results

Demographic and clinical characteristics

The groups did not differ in age or sex; however, their educational levels did not match. All the participants had Spanish as their only transactional language. As expected, the control group outperformed the other groups regarding all cognitive tasks. The demographic and clinical features are shown in Table 1.

Discourse results across diagnostic categories

The groups consistently maintained the same performance pattern, in which healthy controls scored higher than both groups of patients except for the morphology index. The scores of the BD group were more similar to those of the HC group than to those of the SZ group in most metrics. When considering all groups, it was found that the SZ group had fewer words but a similar number of sentences in the quantity-related variables. Additionally, the average sentence length was higher in the control group. Significant differences are evident between the SZ and HC groups, but no differences are observed between the BD and HC groups. On the other hand, in language complexity-related variables, differences were only observed in the use of relative clauses between the HC and BD groups, with no other differences found between the BD and HC groups in the remaining variables. The results are shown in Table 2 and Figure 1.

Table 1. Demographic and clinical characteristics of the sample.

	HC	BD	SZ	BF_{10}
Age	35.7 (12.4)	37.2 (8.5)	39.2 (12.9)	0.176
Sex (Male/Female)	40% / 60%	33% / 66%	43% / 56%	0.549
Educational level (≤12 / > 12)	4% / 96%	36% / 64%	86% / 14%	>1000***
SCIP - VL-I	21.9 (3.4)	18.5 (4.5)	15.9 (4)	>1000***
SCIP - WM	19.9 (3.4)	17.0 (3.7)	15.9 (3.2)	350***
SCIP – VF	21.3 (3.7)	17.7 (5.3)	10.5 (3.7)	>1000***
SCIP - VL-D	6.73 (2.3)	5.56 (2.0)	3.96 (2.6)	310***
SCIP - PS	12.4 (2.9)	10.3 (3.2)	7.13 (2.3)	>1000***
SCIP - Total	82.3 (10)	69.2 (14)	53.5 (8.5)	>1000***
BPRS – Positive subscale	-	4.0 (0.1)	6.1 (3.4)	>1000***
BPRS – Negative subscale	-	4.0 (0.3)	9.7 (2.7)	>1000***
BPRS - Total	-	19.2 (2.8)	26 (5.7)	>1000***

HC: healthy controls, BD: bipolar disorder, SZ: schizophrenia, SCIP: Screen for Cognitive Impairment in Psychiatry, VL-I: verbal learning immediate, WM: working memory,VF: verbal fluency, VL-D: verbal learning delayed, PS: processing speed

*$BF_{10} > 3$, **$BF_{10} > 30$, ***$BF_{10} > 100$

The covariance analysis conducted to determine the effects of covariates and covariate interactions with factors considering cognitive performance and clinical symptoms showed that the group model outperformed the other models in most linguistic measures. The results are shown in Table 3.

Symptoms, cognition, and the structure of language

Among the schizophrenia (SZ) and bipolar disorder (BD) groups considered jointly, mean length of utterance (MLU) was inversely correlated with negative symptoms as assessed with the BPRS ($r = -0.384$, $BF_{10} = 14.03$, 95% CI [−0.571, −0.138]), but not with positive symptoms ($r = -0.214$, $BF_{10} = 0.60$, 95% CI [−0.435, 0.042]). When patient groups were pooled, several additional linguistic metrics showed substantial to decisive evidence of association with negative symptom severity, including syntagms, discourse complexity, syntactic dependency, and morphological indices (Table 4). In contrast, positive symptoms showed little or no evidence of association with any of the linguistic measures analyzed.

Importantly, not all symptom–language associations were in the same direction. While most production- and structure-related metrics showed inverse relationships with negative symptom severity, the morphology index exhibited a positive association with negative symptoms ($r = 0.373$, $BF_{10} = 77.96$). This counterintuitive finding suggests that different components of linguistic organisation may be differentially affected and warrants cautious interpretation. In particular, this association may reflect the influence of education level, overall text length, or compensatory strategies rather than a direct effect of negative symptoms per se.

When all participants were considered, better performance on the SCIP neuropsychological battery was associated with greater linguistic richness across several measures, particularly those related to verbal domains such as working memory and verbal fluency. The morphology index was the only measure that showed an inverse relationship with cognitive performance. These results are shown in Table 5.

Regarding clause production, no conclusive evidence for overall group differences was observed at the global level ($BF_{10} = 0.926$). However, descriptive analyses indicated a

Table 2. Natural language processing analysis results for all groups.

	HC Mean (SD)	BD Mean (SD)	SZ Mean (SD)	BF_{10}	Post Hoc: Posterior Odds			Model posterior 95% CrI	Model R^2 Mean 95% CrI
					HC-BD	HC-SZ	BD-SZ		
MLU	25.7 (6.2)	22.7 (12.9)	15.2 (7.6)	213.4***	0.271	33179.4	3.251	HC: 1.29, 6.88 BD: −1.44, 4.02 SZ: −8.24, −254	0.159 0.05, 0.27
Number of Words	750 (322)	600 (311)	358 (383)	229***	0.617	183	2.941	HC: 58.8, 261 BD: −68.3, 125 SZ: −293, −84.7	0.162 0.05, 0.26
Sentences	33.4 (12)	33.0 (14.9)	25.5 (15.7)	0.942	0.155	1.08	0.661	HC: −1.54, 6.27 BD: −2.00, 5.86 SZ: −8.30, −0.27	0.025 0.00, 0.11
Relative Clauses	17.1 (6.6)	11.1 (5.2)	7.2 (10.4)	1189.4***	55.93	248.6	0.638	HC: 2.43, 7.14 BD: −2.87, 1.58 SZ: −6.49, −1.87	0.194 0.08, 0.30
Completive Clauses	8.6 (5.0)	7.3 (4.8)	3.9 (5.7)	14.4*	0.237	13.87	1.864	HC: 0.22, 3.25 BD: −0.89, 2.09 SZ: −3.92, −0.79	0.100 0.00, 0.20
Proper Adverbial Clauses	17.7 (7.2)	14.8 (6.3)	9.5 (9.4)	76.1**	0.465	42.16	2.382	HC: 1.00, 5.54 BD: −1.46, 2.92 SZ: −6.41, −1.73	0.140 0.03, 0.24
Improper Adverbial Clauses	20.0 (9.7)	18.7 (11.2)	9.6 (11.0)	65.6**	0.169	54.42	8.726	HC: 0.35, 6.54 BD: −0.76, 5.35 SZ: −9.01, −2.53	0.136 0.03, 0.24
Complexity	858.8 (365)	693.9 (342)	419.4 (422)	260.6***	0.596	191.1	3.429	HC: 65.5, 293 BD: −75.3, 142 SZ: −329.3, −99	0.165 0.06, 0.27
Syntagmas	188.7 (71)	144.4 (65)	89.06 (89.5)	1662.8***	2.032	792	3.251	HC: 20.3, 66.1 BD: −18.4, 24.8 SZ: −70.0, −23.4	0.200 0.09, 0.31
Morphology index	0.428 (0.04)	0.464 (0.05)	0.501 (0.08)	145.6***	3.46	110.7	0.741	HC: −0.05,−0.01 BD: −0.02, 0.02 SZ: 0.01, 0.05	0.155 0.05, 0.26
Coordinating conjunction	942.8 (409)	745 (394)	428 (498)	310***	0.693	224.6	3.144	HC: 82.8, 341 BD: −87, 157 SZ: −379, −117	0.167 0.06, 0.27
Subordinating conjunction	30.06 (22)	28.1 (22)	12.06 (24)	6.724*	0.162	4.918	2.71	HC: −0.91, 12.3 BD: −2.55, 10.5 SZ: −16.8, −2.89	0.080 0.00, 0.18
Causal Connectors	4.5 (5.2)	4.4 (5.0)	1.5 (3.0)	2.532	0.154	2.905	2.808	HC: −0.39, 2.14 BD: −0.48, 2.04 SZ: −2.99, −0.34	0.052 0.00, 0.14

MLU: mean length utterance *$BF_{10} > 3$, **$BF_{10} > 30$, ***$BF_{10} > 100$. CrI: Credible interval. The post hoc test are reported with posterior odds that have been corrected for multiple testing.

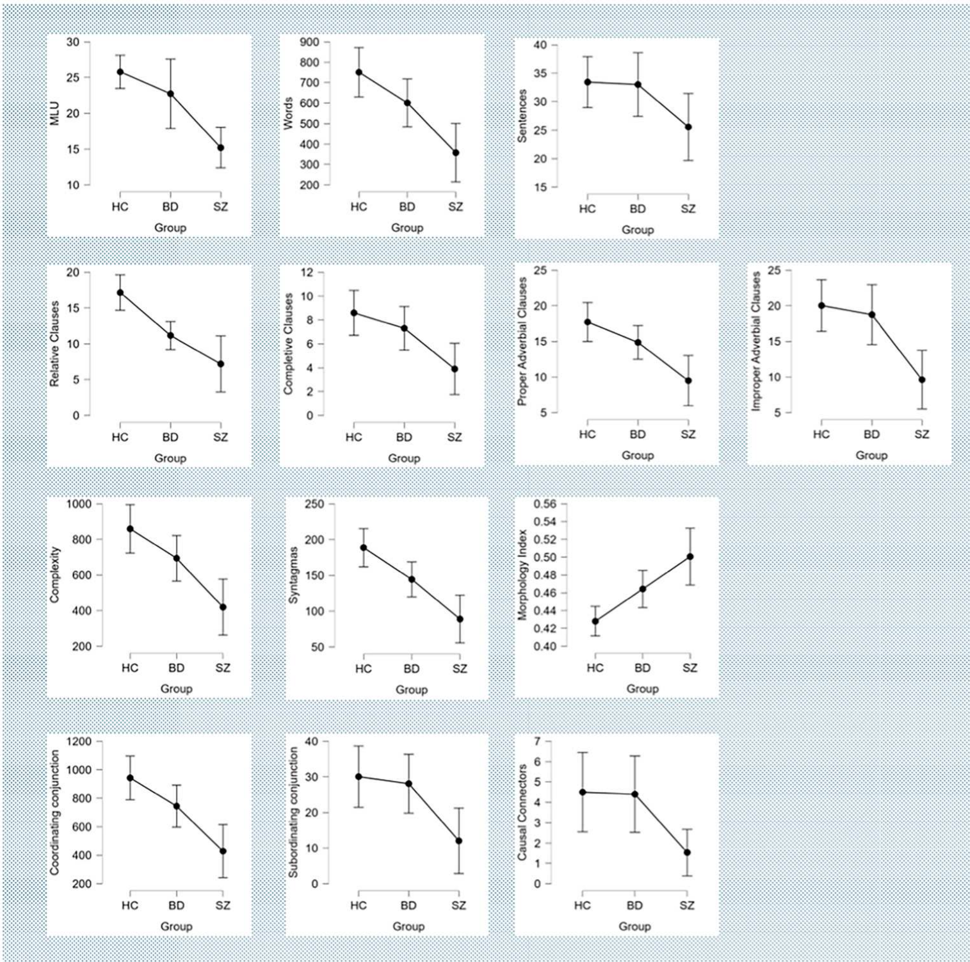

Figure 1. Natural language processing analysis results for all groups.

tendency for the SZ group to produce sentences with fewer clauses on average compared to the BD and healthy control groups, suggesting subtle differences that did not reach evidential thresholds in the present sample.

All symptom–language correlations were estimated using pooled patient data (SZ + BD) to examine dimensional relationships across diagnostic categories and to maximise statistical power. The potential influence of education level and cognitive performance was examined separately through model-comparison analyses.

Discussion

The principal aim of the study was to assess poverty of speech across different diagnoses, including SZ and BD, using natural language processing techniques. In contrast to the typical clinical evaluation, where poverty of speech is rated on scales based on clinicians' subjective assessment of patients' utterances, this method provides objective

Table 3. Covariance analysis including cognitive performance and clinical symptoms.

	Models	BF_M	BF_{10}
MLU	Negative symptoms	5.567	1
	Group + Negative symptoms	2.229	0.475
	Group	1.725	0.381
Words	Group	3.332	1
	Negative symptoms	1.712	0.564
	Group + Cognitive	1.604	0.531
Sentences	Group	2.744	1
	Null model	2.388	0.888
	Group + Positive symptoms	1.839	0.706
Relative clauses	Group	2.845	1
	Null model	2.582	0.921
	Cognitive	1.584	0.599
Completive Clauses	Group	3.471	1
	Cognitive	1.566	0.503
	Group + Cognitive	1.476	0.477
Proper Adverbial Clauses	Group	3.801	1
	Negative symptoms	1.873	0.549
	Group + Negative symptoms	1.484	0.445
Improper Adverbial Clauses	Group	6.69	1
	Negative symptoms	2.077	0.394
	Group + Negative symptoms	1.803	0.348
Complexity	Group	3.596	1
	Negative symptoms	1.856	0.569
	Group + Negative symptoms	1.616	0.503
Syntagmas	Group	3.231	1
	Negative symptoms	2.191	0.719
	Positive + Negative symptoms	1.833	0.614
Morphology index	Group	3.837	1
	Null model	2.891	0.793
	Negative symptoms	1.739	0.51
Coordinating conjunctions	Group	3.493	1
	Group + Cognitive	1.608	0.513
	Negative symptoms	1.515	0.486
Subordinating conjunctions	Group	4.057	1
	Group + Cognitive	1.973	0.546
	Cognitive	1.754	0.492
Causal Connectors	Cognitive	2.621	1
	Group	2.509	0.963
	Group + Cognitive	2.259	0.88

BF_{10}: Bayes factor for the alternative hyphotesis; BF_M: Bayes Factor of the the intercept-only (null) model.

quantification of various parameters, thereby increasing the accuracy and complexity of the evaluation. Our results showed differences between people with schizophrenia and the control group in most of the parameters assessed, with patients showing lower production and less sentence complexity. More interestingly, for most metrics, the scores of the BD group were more similar to those of the HC group than those of the SZ group. Importantly, all participants in the bipolar disorder group were clinically stable and in euthymic states at the time of assessment. While this reduces heterogeneity related to acute mood symptoms, it also limits the interpretation of the BD ≈ HC pattern observed in this study. Specifically, the absence of marked alogia-like features in the BD group cannot be taken as evidence against state-related language alterations in mood disorders, as such effects may be more evident during depressive or mixed episodes. Future studies incorporating longitudinal designs and explicit mood-state stratification will be necessary to clarify state versus trait contributions to speech poverty in bipolar disorder.

Table 4. Correlations between linguistic metrics and BPRS scores for both patient groups.

		Pearson's r	BF_{10}	Lower 95% CI	Upper 95% CI
PS	MLU	−0.36	51.6**	−0.521	−0.162
PS	Relative Clauses	−0.302	7.97*	−0.473	−0.098
PS	Completive Clauses	−0.198	0.73	−0.384	0.01
PS	Proper Adverbial Clauses	−0.245	1.87	−0.424	−0.038
PS	Improper Adverbial Clauses	−0.252	2.22	−0.43	−0.046
PS	Complexity	−0.282	4.60*	−0.456	−0.077
PS	Syntagmas	−0.313	11.2*	−0.482	−0.111
PS	Morphology index	0.31	10.2*	0.107	0.48
PS	Coordinating conjunction	−0.279	4.32*	−0.454	−0.074
PS	Subordinating conjunction	−0.237	1.58	−0.417	−0.03
NS	MLU	−0.449	1945***	−0.595	−0.261
NS	Relative Clauses	−0.407	296***	−0.561	−0.213
NS	Completive Clauses	−0.32	12.9*	−0.489	−0.116
NS	Proper Adverbial Clauses	−0.378	93.8**	−0.537	−0.18
NS	Improper Adverbial Clauses	−0.345	29.0*	−0.51	−0.144
NS	Complexity	−0.412	356***	−0.565	−0.218
NS	Syntagmas	−0.457	2858***	−0.601	−0.27
NS	Morphology index	0.373	77.9**	0.175	0.533
NS	Coordinating conjunction	−0.409	326***	−0.563	−0.215
NS	Subordinating conjunction	−0.259	2.54	−0.437	−0.052
TS	MLU	−0.403	245***	−0.557	−0.208
TS	Relative Clauses	−0.421	534***	−0.572	−0.228
TS	Completive Clauses	−0.289	5.40*	−0.463	−0.084
TS	Proper Adverbial Clauses	−0.337	22.0*	−0.503	−0.135
TS	Improper Adverbial Clauses	−0.298	7.01*	−0.471	−0.094
TS	Complexity	−0.382	107***	−0.54	−0.184
TS	Syntagmas	−0.429	748***	−0.578	−0.237
TS	Morphology index	0.388	135***	0.191	0.545
TS	Coordinating conjunction	−0.382	107***	−0.54	−0.184
TS	Subordinating conjunction	−0.245	1.86	−0.426	−0.038

PS: Positive score NS: Negative score TS: Total score MLU: mean length utterance
*$BF_{10} > 3$, **$BF_{10} > 30$, ***$BF_{10} > 100$

Importantly, this pattern remained when negative symptom severity was taken into account, suggesting that the observed linguistic alterations cannot be fully explained by negative symptoms as clinically rated. This finding could imply a specific relationship between speech poverty and schizophrenia.

Taken together, our results suggest that alogia, as captured by NLP-derived measures, reflects not only a quantitative reduction in speech output but also a disruption in the organisation of discourse, pointing to a qualitative alteration in verbal behaviour characteristic of schizophrenia and not in bipolar patients. While these features were more evident in the schizophrenia group, the present cross-sectional design does not allow attribution of these alterations to diagnosis per se, as they may also reflect differences in cognitive resources, educational background, and task demands.

Regarding the relationship between alogia and the severity of patients' symptoms, we expected these to be associated, especially with negative symptoms. However, when considering both groups of patients, only the MLU was more related symptom severity as determined with the BPRS then the group. The lack of association with the negative symptoms assessed (emotional withdrawal, blunted affect) is striking since alogia is considered part of the deficit syndrome of schizophrenia (Kirkpatrick et al., 1989). Moreover, negative symptoms are commonly grouped into two separable dimensions, i.e. diminished expression and avolition, with diminished expression comprising symptoms of

Table 5. Correlations between linguistic metrics and cognitive performance for all participants.

		Pearson's r	BF_{10}	Lower 95% CI	Upper 95% CI
SCIP	MLU	0.392	173.5***	0.197	0.548
SCIP	Sentences	0.19	0.641	−0.018	0.377
SCIP	Relative Clauses	0.379	104.6***	0.183	0.537
SCIP	Completive Clauses	0.351	38.0**	0.152	0.514
SCIP	Proper Adverbial Clauses	0.388	150.0***	0.193	0.545
SCIP	Improper Adverbial Clauses	0.366	63.3**	0.168	0.526
SCIP	Complexity	0.412	390.9***	0.219	0.564
SCIP	Syntagmas	0.417	497.8***	0.225	0.568
SCIP	Morphology index	−0.336	22.99**	−0.502	−0.136
SCIP	Coordinating conjunction	0.423	627.1***	0.231	0.573
SCIP	Conj Subordinan	0.339	25.0*	0.138	0.504
SCIP	Causal Connectors	0.318	12.89*	0.116	0.486

SCIP: Screen for Cognitive Impairment in Psychiatry Total score MLU: mean length utterance
*$BF_{10} > 3$, **$BF_{10} > 30$, ***$BF_{10} > 100$

blunted affect and alogia (Marder & Galderisi, 2017), although other analyses (Ahmed et al., 2022) show better fit with five-factor models in which each negative symptom would represent a separate domain. Taken together, these findings suggest that the linguistic features identified in this study are not reducible to negative symptoms as a broad clinical dimension, but rather reflect a speech-specific disturbance that overlaps only partially with negative symptomatology.

On the contrary, when analyzing the correlation with performance on the SCIP neuropsychological battery, better cognitive functioning was associated with greater language richness. It should be clarified that the richness of the thought content was not evaluated but rather the complexity of language structuring. Our results are consistent with previous findings of an association between the diminished expression domain of negative symptoms and decreased overall cognitive functioning (Hartmann-Riemer et al., 2015) and with the cognitive resource limitation theory that states that fewer cognitive resources are available for complex, expressive action and interaction, including speech, in schizophrenia and other psychiatric conditions (Cohen et al., 2012). Recently, Compton et al. (2023) found a strong difference in number of words produced and other language variables between a group of people with schizophrenia and controls, with most metrics correlated with alogia and other negative symptoms, as well as with cognitive functioning. Alternatively, alogia could be a preliminary and simple indicator to estimate the cognitive functioning of patients.

Our findings should also be interpreted in light of prior NLP studies comparing schizophrenia with affective psychoses and mood disorders, which have generally reported overlapping but quantitatively attenuated language alterations outside schizophrenia, particularly when controlling for education and cognitive load. Moreover, extensive work in Spanish-language corpora has shown robust associations between educational attainment and lexical diversity, syntactic complexity, and morphological richness (Ochoa, 2020). The marked educational differences across groups in the present sample therefore represent a critical contextual factor and reinforce the need for cautious interpretation of group contrasts as diagnostic rather than sociodemographic in nature.

From this perspective, our results do not challenge the inclusion of poverty of speech among negative symptoms, but rather highlight the difficulty of disentangling negative symptomatology from cognitive functioning using cross-sectional data. The observed associations between linguistic measures and cognition are consistent with extensive prior literature showing substantial covariation between negative symptoms and cognitive deficits in schizophrenia, and should therefore be interpreted as compatible with, rather than contradictory to, existing negative symptom frameworks**.** More importantly, when comparing models that included cognitive performance as a covariate, only MLU and causal connectors yielded BFM values indicating that the data provided stronger evidence for an effect of cognition than of group membership. For the remaining measures, the inclusion of cognitive performance as a covariate did not result in greater belief updating relative to the group factor, suggesting that group differences in these variables are not primarily mediated by cognitive ability.

These findings may be interpreted as being compatible with accounts that relate alogia to expressive behaviour and cognitive resource demands (Cohen et al., 2012; Hartmann-Riemer et al., 2015), rather than treating it exclusively as a unitary negative symptom driven by motivational or affective deficits.

One of the limitations of our study is that we did not directly analyze the effect of medication. However, this could have been reflected in the negative symptom score of the BPRS, and we could not find a relationship with the metrics used to estimate alogia. This absence of association further supports the interpretation that the linguistic measures captured in this study are not merely epiphenomena of overall clinical severity**.** Additionally, the group of patients with a diagnosis of bipolar disorder was exposed to similar medications, although with different dosages. Nevertheless, future studies explicitly designed to disentangle medication effects from cognitive and linguistic variables will be necessary to further validate these findings.

Previous research with chronically institutionalised patients with schizophrenia found that poverty of speech was specifically associated with deficits in social outcomes, but not in other functional domains like activities of daily living and recreation (Bowie & Harvey, 2008). Significantly, alogia worsens with age, especially among the oldest patients. In a study with community-dwelling middle-aged patients, alogia predicted difficulties in communicating skilfully in semi-structured social situations, as well as community functioning across interpersonal, occupational, and everyday living domains (Muralidharan et al., 2018). Additional studies using NLP methodology could expand on these results, adding more complex information that could potentially be useful in designing rehabilitation interventions. In this context, objective linguistic markers of alogia may represent promising targets for monitoring functional outcomes and response to psychosocial interventions.

Natural language processing is a fast, innocuous, non-invasive, and ecologically valid analysis method. While the study design requires a relatively unstructured task, i.e. describing a scene, it would be interesting to know what results would be obtained by analyzing speech in an actual clinical situation where the subject was describing their symptoms or reaction to medications. Another limitation of our study is that we cannot isolate the effect that fewer years of education may have on the outcomes exhibited by the group of people with schizophrenia. Future studies combining NLP-derived speech measures with dedicated negative symptom instruments (e.g. BNSS, CAINS) will be necessary to further disentangle alogia from other components of diminished expression.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

The data that support the findings of this study are available from the corresponding author, AC, upon reasonable request.

References

Ahmed, A. O., Kirkpatrick, B., Granholm, E., Rowland, L. M., Barker, P. B., Gold, J. M., Buchanan, R. W., Outram, T., Bernardo, M., Paz García-Portilla, M., Mane, A., Fernandez-Egea, E., & Strauss, G. P. (2022). Two factors, five factors, or both? External validation studies of negative symptom dimensions in schizophrenia. *Schizophrenia Bulletin*, *48*(3), 620–630. https://doi.org/10.1093/SCHBUL/SBAB148

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders. diagnostic and statistical manual of mental disorders. https://doi.org/10.1176/APPI.BOOKS.9780890425596.

Andreasen, N. C. (1979). Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. *Archives of General Psychiatry*, *36*(12), 1315–1321. https://doi.org/10.1001/ARCHPSYC.1979.01780120045006

Becker, T. M., Cicero, D. C., Cowan, N., & Kerns, J. G. (2012). Cognitive control components and speech symptoms in people with schizophrenia. *Psychiatry Research*, *196*(1), 20–26. https://doi.org/10.1016/J.PSYCHRES.2011.10.003

Bhati, M. T. (2005). The brain, language, and schizophrenia. *Current Psychiatry Reports*, *7*(4), 297–303. https://doi.org/10.1007/S11920-005-0084-6

Bleuler, E. (1950). Dementia praecox or the group of schizophrenias.

Bowie, C. R., & Harvey, P. D. (2008). Communication abnormalities predict functional outcomes in chronic schizophrenia: Differential associations with social and adaptive functions. *Schizophrenia Research*, *103*(1-3), 240–247. https://doi.org/10.1016/J.SCHRES.2008.05.006

Clinical psychiatry: A text-book for students and physicians / abstracted and adapted from the 7th German edition of Kraepelin's "Lehrbuch der Psychiatrie" by A. Ross Diefendorf., Wellcome Collection. n.d. Retrieved April 2, 2022, from https://wellcomecollection.org/works/md763v8u.

Cohen, A. S., McGovern, J. E., Dinzeo, T. J., & Covington, M. A. (2014). Speech deficits in serious mental illness: A cognitive resource issue? *Schizophrenia Research*, *160*(1-3), 173–179. https://doi.org/10.1016/J.SCHRES.2014.10.032

Cohen, A. S., Morrison, S. C., Brown, L. A., & Minor, K. S. (2012). Towards a cognitive resource limitations model of diminished expression in schizotypy. *Journal of Abnormal Psychology*, *121*(1), 109–118. https://doi.org/10.1037/A0023599

Compton, M. T., Ku, B. S., Covington, M. A., Metzger, C., & Hogoboom, A. (2023). Lexical diversity and other linguistic measures in schizophrenia: Associations With negative symptoms and neurocognitive performance. *Journal of Nervous & Mental Disease*, *211*(8), 613–620. https://doi.org/10.1097/NMD.0000000000001672

Corcoran, C. M., & Cecchi, G. A. (2020). Using language processing and speech analysis for the identification of psychosis and other disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 770–779. https://doi.org/10.1016/J.BPSC.2020.06.004

Daniel, D. G. (2013). Issues in selection of instruments to measure negative symptoms. *Schizophrenia Research*, *150*(2-3), 343–345. https://doi.org/10.1016/J.SCHRES.2013.07.005

Fervaha, G., Takeuchi, H., Foussias, G., Agid, O., & Remington, G. (2016). Using poverty of speech as a case study to explore the overlap between negative symptoms and cognitive dysfunction. *Schizophrenia Research*, *176*(2-3), 411–416. https://doi.org/10.1016/J.SCHRES.2016.05.019

Gómez-Benito, J., Berrío, ÁI, Guilera, G., Rojo, E., Purdon, S., & Pino, O. (2018). The screen for cognitive impairment in psychiatry: Proposal for a polytomous scoring system. *International Journal of Methods in Psychiatric Research*, 27. https://doi.org/10.1002/MPR.1598

Hartmann-Riemer, M. N., Hager, O. M., Kirschner, M., Bischof, M., Kluge, A., Seifritz, E., & Kaiser, S. (2015). The association of neurocognitive impairment with diminished expression and apathy in schizophrenia. *Schizophrenia Research*, *169*(1-3), 427–432. https://doi.org/10.1016/J.SCHRES.2015.10.032

Joyce, E. M., Collinson, S. L., & Crichton, P. (1996). Verbal fluency in schizophrenia: Relationship with executive function, semantic memory and clinical alogia. *Psychological Medicine*, *26*(1), 39–49. https://doi.org/10.1017/S0033291700033705

Kay, S. (1991). *Positive and negative syndromes in schizophrenia: Assessment and research*. Psychology Press.

Kircher, T., Krug, A., Stratmann, M., Ghazi, S., Schales, C., Frauenheim, M., Turner, L., Fährmann, P., Hornig, T., Katzev, M., Grosvald, M., Müller-Isberner, R., & Nagels, A. (2014). A rating scale for the assessment of objective and subjective formal Thought and Language Disorder (TALD). *Schizophrenia Research*, *160*(1-3), 216–221. https://doi.org/10.1016/J.SCHRES.2014.10.024

Kirkpatrick, B., Buchanan, R. W., McKenny, P. D., Alphs, L. D., & Carpenter, W. T. (1989). The schedule for the deficit syndrome: An instrument for research in schizophrenia. *Psychiatry Research*, *30*(2), 119–123. https://doi.org/10.1016/0165-1781(89)90153-4

Kirkpatrick, B., Fenton, W. S., Carpenter, W. T., & Marder, S. R. (2006). The NIMH-MATRICS consensus statement on negative symptoms. *Schizophrenia Bulletin*, *32*(2), 214–219. https://doi.org/10.1093/SCHBUL/SBJ053

Kleist, K. (1914). *Aphasie und Geisteskrankheit*. Münchener Medizinische Wochenschrift.

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

Le Glaz, A., Haralambous, Y., Kim-Dufor, D. H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., Devylder, J., Walter, M., Berrouiguet, S., & Lemey, C. (2021). Machine learning and natural language processing in mental health: Systematic review. *Journal of Medical Internet Research*, e15708. https://doi.org/10.2196/15708

Marder, S. R., & Galderisi, S. (2017). The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry*, *16*(1), 14. https://doi.org/10.1002/WPS.20385

Muralidharan, A., Finch, A., Bowie, C. R., & Harvey, P. D. (2018). Thought, language, and communication deficits and association with everyday functional outcomes among community-dwelling middle-aged and older adults with schizophrenia. *Schizophrenia Research*, *196*, 29–34. https://doi.org/10.1016/J.SCHRES.2017.07.017

Ochoa, L. (2020). Complejidad lingüística: Revisión bibliográfica. *Signo y Pensamiento*, *39*(77), https://doi.org/10.11144/Javeriana.syp39-77.clrb

Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological Reports*, *10*(3), 799–812. https://doi.org/10.2466/pr0.1962.10.3.799

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Sumiyoshi, C., Sumiyoshi, T., Nohara, S., Yamashita, I., Matsui, M., Kurachi, M., & Niwa, S. (2005). Disorganization of semantic memory underlies alogia in schizophrenia: An analysis of verbal fluency performance in Japanese subjects. *Schizophrenia Research*, *74*(1), 91–100. https://doi.org/10.1016/J.SCHRES.2004.05.011

Wing, J. K. (1961). A simple and reliable subclassification of chronic schizophrenia. *Journal of Mental Science*, *107*(450), 862–875. https://doi.org/10.1192/BJP.107.450.862