Felletár *et al.*
Borderline Personality Disorder and Emotion Dysregulation (2025) 12:49
https://doi.org/10.1186/s40479-025-00324-0

Borderline Personality Disorder and Emotion Dysregulation

RESEARCH Open Access

Increased self-focus and diminished informativity: referential and structural properties of narrative speech production in borderline personality disorder

Fanni Felletár[1,2], Veronika Vincze[3], Gábor Gosztolya[3,4], Ildikó Hoffmann[1,5], Anna Babarczy[1,2] and Zsolt Szabolcs Unoka[2,6*]

Abstract

Background Narrative speech production (NSP), i.e., the conceptualization, linguistic formulation, and articulation of a story, is a multifaceted process underpinned by cognitive functions and mentalization ability, often impaired in individuals with borderline personality disorder (BPD). This study examines differences in linguistic formulation between individuals with BPD and healthy controls (HCs), and explores how task type influences linguistic formulation, as well as how linguistic formulation relates to temporal parameters of speech uniquely in BPD.

Methods Speech of 33 BPD and 31 HC individuals was recorded in three task types: telling their previous day, retelling a story, and picture sequences. Features of linguistic formulation were extracted with natural language processing methods, while temporal parameters were extracted using automatic speech recognition. Hypothesis-driven generalized linear mixed-effects models (GLMMs) were applied to test predefined group differences in four linguistic features (content words, first- and third-person singular verbs, and syntactic complexity). Additional exploratory GLMMs examined other linguistic features and task effects. Within-group Spearman correlations assessed associations between linguistic and temporal measures, controlling for task.

Results Hypothesis testing showed that the NSP in BPD is characterized by fewer content words, more first-person singular verbs, and lower syntactic complexity than that of HCs. Exploratory analyses revealed that individuals with BPD used pronouns more frequently than HCs, particularly demonstrative pronouns (e.g., *this*) and first-person singular pronouns (e.g., *I*). In BPD, higher first-person singular reference (pronouns and verbs) correlated with fewer silent pauses, while greater syntactic complexity correlated with more filled pauses. Task modulated verbosity and the use of other pronoun types.

Conclusions Findings suggest that NSP in BPD is characterized by dominant self-referential thought content, reflected in elevated first-person singular reference, and by qualitatively impoverished language use, marked by

*Correspondence:
Zsolt Szabolcs Unoka
unoka.zsolt@semmelweis.hu

Full list of author information is available at the end of the article

reduced content word production, increased pronoun use, and lower syntactic complexity. Heightened self-focus may hinder the efficient allocation of cognitive resources required for cohesive, listener-oriented NSP.

Keywords Borderline personality disorder, Narrative speech production, Content words, Pronouns, First-person singular, Syntactic complexity, Pauses

Introduction

Narrative speech production as a complex behavior

Narrative speech production (NSP) involves the conceptualization, linguistic formulation, and articulation of a monologue that presents a temporal-causal sequence of events from one's viewpoint [1–3]. After constructing the intended message during conceptualization, the speaker assigns lexical items and grammatical structure to the message in the linguistic formulation phase.

First, conceptual elements of the message – such as events (e.g., actions, states), entities (e.g., people, objects, ideas), attributes, and circumstances (e.g., manner, time) – activate lemmas (i.e., the dictionary form of words) in the mental lexicon that contain conceptual (i.e., contextual usage), semantic (e.g., meaning, synonyms), and syntactic (i.e., syntactic category, e.g., noun; and grammatical function, e.g., subject) properties [2]. This rapid, association-based process is driven by the speaker's declarative memory [4]. Lexical access requires an intact semantic memory to successfully retrieve appropriate lemmas [2, 5] and intact executive control to inhibit additionally activated inappropriate lemmas in the semantic memory [5, 6]. Lexical selection is mediated by the speaker's perspective-taking [7], mutual semantic knowledge with the listener [8, 9], and the discourse context [8].

Conceptual elements can be denoted either by content words (i.e., nouns, verbs, adjectives, adverbs), which have an inherent meaning, or by pronouns, a type of function word [10]. Definite pronouns (e.g., *she*) replace content words (e.g., *girl*) to maintain communication efficiency. They either refer to entities in the physical context by the speaker's gesture (e.g., pointing, gaze) or to previously introduced content words in the discourse context by grammatical dependencies [11, 12]. Definite pronouns are less informative and so require less cognitive effort from the speaker to produce them [13]. Their processing, however, is much more cognitively demanding for listeners, as they have to infer the intended referent [14]. The speaker thus has to make assumptions about the listener's attentional focus and inferential capacity, which requires mentalization ability [15]. In contrast to definite pronouns, indefinite pronouns quantify a set of entities or circumstances (e.g., *nothing*, *somehow*, *everybody*) irrespective of the context [16].

Then, sentence structure is formed by combining lemmas based on their syntactic properties [2]. This is a slower, rule-based process driven by the speaker's procedural memory [4, 17]. Most function words (e.g., articles, conjunctions, adpositions) are accessed at this stage to create grammatical relationships between content words as well as between larger syntactic units such as phrases and clauses [18]. The more clauses a sentence contains, the more it is syntactically complex [19, 20]. The construction of complex sentences requires a higher processing speed and a greater working-memory capacity [21–24]. These enable speakers to build complicated syntactic trees and maintain their constituents in their working memory. When deciding about the complexity of sentences, the speaker must also estimate the listener's information need [8] as the more complex a sentence is, the more information it conveys. Words receive their final forms (i.e., lexemes) during morphophonological encoding, where their inflected forms are transferred into a sequence of phonemes [2]. During the whole process of NSP, the speaker must simultaneously monitor their own speech and the listener's reactions [25–27].

NSP is, therefore, a complex, goal-directed task requiring domain-general cognitive functions and mentalization ability [3, 5, 25, 26, 28]. In its articulation phase, speech pauses – divided into silent (i.e., the absence of speech) and filled (e.g., *uh*, *um*, *er*) pauses – indicate the cognitive load of conceptualization, linguistic formulation [29, 30], and the monitoring of speech output [27]. While silent pauses occur relatively more frequently before content words signaling lexical search, filled pauses occur relatively more frequently before function words and at syntactic boundaries, implying grammatical planning [31]. The articulation rate, i.e., the speed of producing consecutive syllables, is another temporal aspect of speech that has been found to be affected by the speaker's arousal [32], processing speed [33], and working memory capacity [34]. Speech pauses and articulation rate together constitute the speech rate. These three features are collectively referred to as temporal parameters.

Borderline personality disorder, cognitive functioning, and mentalization

Individuals with borderline personality disorder (BPD) experience rapid shifts between affective states due to a marked reactivity of mood; their self and other representations alternate between extremes of idealization and devaluation; their self-image and sense of self are unstable [35–38]. Self-referential thinking is frequently observed in BPD. These individuals tend to overattribute mental states to others, often interpreting ambiguous interpersonal cues as rejection [39, 40]. Negative

affectivity (i.e., frequent, intense feelings of sadness, anxiety, and anger) and emotion dysregulation (i.e., deficient top-down control of affective states) are core factors of BPD [38, 41], underpinned by the hyperreactivity of the amygdala and the hypoactivity of the prefrontal cortex [39, 42]. In response to overwhelming emotional states, individuals with BPD frequently experience dissociative symptoms, such as depersonalization (i.e., the sense of detachment from the self) and derealization (i.e., the sense of detachment from the environment) [35, 43].

Besides the symptoms of BPD, many individuals also exhibit subtle cognitive impairments and difficulties in mentalization, possibly due to prefrontal hypoactivation. Studies investigating cognitive functioning in BPD [44–49] consistently found a slower processing speed as well as poor attention, short- and long-term verbal memory, inhibition, and planning in individuals with BPD compared to healthy controls (HCs). Attention, working memory, and long-term memory deficits are associated with dissociative states in BPD [43, 49, 50]. Individuals with BPD are also known to have difficulties in understanding the mental states of others, especially when they are exposed to complex stimuli (e.g., pictures about interpersonal interactions) requiring cognitive as opposed to affective mentalization [40, 51, 52].

Given the symptoms and neuropsychological impairment mentioned above, it seems probable that NSP is compromised in people living with BPD.

Narrative speech production and psychopathology

Having a complex nature, NSP can be viewed as a gateway to one's mind and personality [1, 30, 53–55], making it a valuable biomarker of psychopathology [56–61]. During psychiatric interviews, patients share their personal history (anamnesis) in a narrative, which serves as a platform for clinicians to assess the content, structural organization, and fluency of speech [62, 63]. In self-referential thinking, thought content predominantly centers on the self, either as a result of the hyperactivation of internally generated, often ruminative thoughts about the self's past or anticipated future combined with the hypoactivation of externally derived sensory information [63, 64], or due to a tendency to associate neutral stimuli with the self [56, 63]. Negative formal thought disorder (NTD) is a transdiagnostic construct, characterized by the reduction or impoverishment of thought, language, and communication in their quality or quantity [65]. The NSP of individuals with NTD often displays "poverty of content" (i.e., speech that is adequate in amount but conveys little information) [56], lower syntactic complexity, and a slower speech rate [66].

In the BPD population, several studies have focused on word usage. According to the review of Močnik et al. [67], BPD narratives are characterized by an impersonal tone lacking personal attribution and presented from an external viewpoint. They contain an elevated number of function words, such as conjunctions, negations, and pronouns [68, 69]. Supporting this, others found that autobiographical narratives of individuals with BPD are less specific compared to HCs [70–75]. Furthermore, individuals with BPD use more first-person singular [74, 76] and third-person singular pronouns [68, 77] relative to HCs. The elevated use of first-person singular pronouns is also a characteristic of depression [78–80] and has been associated with self-referential thinking [55, 78].

Only Carter and Grenyer [77, 81] studied the structural properties of speech in BPD and found that individuals with BPD produce less complex sentences compared to HCs. However, lower syntactic complexity is well documented in schizophrenia [82–85] and is also detectable in depression [80, 86]. In a few studies, researchers also examined temporal parameters of speech in BPD during psychotherapy sessions [87] and interviews [81, 88, 89], and it transpired that both the number and duration of silent pauses are distinguishing aspects of speech production in BPD. In our previous study [90], we found that the speech of individuals with BPD is characterized by significantly lower articulation and speech rate, a significantly higher number of silent and filled pauses, and a significantly higher frequency and duration of filled pauses relative to HCs. Lower syntactic complexity and frequent pausing have been associated with NTD [56, 66].

Purpose and novelty of the present study

This study aims to investigate differences between individuals with BPD and HCs in terms of linguistic formulation (i.e., lexical, syntactic, and morphological features of NSP) and to explore how task type influences linguistic formulation, as well as how linguistic formulation relates to temporal parameters of speech uniquely in BPD. Based on previous research, we hypothesize that the narratives of individuals with BPD (1) contain fewer content words, (2) more first-person singular verbs, (3) more third-person singular verbs, and (4) are characterized by lower syntactic complexity relative to those of HCs. In Hungarian, the first-person singular pronoun is often omitted, as verbal suffixes mark first-person singular reference. Thus, first-person singular verbs offer a more accurate measure of self-reference than pronouns. Also, we explore (1) the difference between BPD and HC groups in terms of other features of linguistic formulation, (2) the effect of speech elicitation tasks (recalling the previous day, recalling a read story, picture sequencing tasks) on between-group differences, and (3) unique associations between linguistic formulation and temporal parameters within the BPD group.

To address these questions, we recorded the NSP of 33 BPD and 31 HC individuals across three speech elicitation tasks (recounting the previous day, retelling a story, and describing picture sequences). Linguistic formulation was measured in terms of lexical, syntactic, and morphological features using *magyarlanc* linguistic processing toolkit [91], while temporal parameters were measured via automatic speech recognition (ASR) using the Speech-GAP Test® [92]. The novelty of this study lies in its cognitive approach to NSP in BPD, using various speech elicitation tasks to assess different cognitive functions, and its multi-level approach, which examines the interaction between linguistic formulation and articulation phases of NSP.

Materials and methods

Participants

Participants were recruited by convenience sampling, on the one hand, i.e., from social media groups for individuals with BPD and other groups keen on advertising experiments, and by snowball sampling, on the other hand, i.e., by advertising the experiment for potential participants with the help of existing participants. For both groups, the inclusion criteria were to be a native Hungarian speaker and to have intact hearing and speech abilities. For individuals with BPD, a further inclusion criterion was to present clinical documentation confirming the BPD diagnosis (F60.3 Emotionally unstable personality disorder) [93]. Comorbid disorders were not considered exclusion criteria, as they are also common in the BPD population. For HC individuals, the exclusion criteria were to have any psychiatric or neurological disorders. They self-reported any diagnoses of language or speech disorders (e.g., aphasia, stuttering), endocrinological disorders (e.g., PCOS, insulin resistance), neurological disorders or injuries (e.g., epilepsy, brain injury), or mental disorders (e.g., panic disorder, ADHD). As a result, 33 individuals with BPD and 31 HC individuals took part in the experiment. The two groups were matched based on the participants' gender, age, and years of education they had completed. Table 1 lists the demographic data of the BPD and HC groups.

Table 1 Descriptive statistics of demographic data of the BPD and HC groups

Variable	BPD	HC	Test	*p*
N	33	31	–	–
Sex, f:m	28:5	26:5	X^2 (1) = 0.012	0.914
Age, M (SD)	27.18 (6.92)	26.68 (8.08)	W = 571.000	0.436
Years of education, M (SD)	14.64 (2.56)	15.29 (2.74)	W = 453.500	0.435
Individuals with comorbid disorders, N (%)	23 (69.70)	–	–	–
- bipolar disorders	8 (37.78)	–	–	–
- depressive disorders	7 (30.43)	–	–	–
- anxiety disorders	6 (26.09)	–	–	–
- substance use disorders	4 (17.39)	–	–	–
- other personality disorders	3 (13.04)	–	–	–
- eating disorders	2 (8.70)	–	–	–
- attention deficit hyperactivity disorder	2 (8.70)	–	–	–
- sleep disorders	2 (8.70)	–	–	–
- schizophrenia spectrum disorders	1 (4.35)	–	–	–

Note. BPD – borderline personality disorder, HC – healthy controls, Test – type and value of statistical tests applied, p – significance value of statistical tests, f:m – female: male ratio, M – mean, SD – standard deviation, N – number of subjects

Materials and procedure

Data collection was conducted in the following steps: (1) participants were asked to read a story about the origin of dishwashing without explicit information on the purpose of reading; (2) as a distractor task, they evaluated the story on three 5-point Likert scales assessing comprehensibility, interest, and modernity [94]; (3) they were asked to recall their previous day; (4) participants were asked to recall the story about the origin of dishwashing [94]; (5) participants had to arrange three interrelated pictures showing a *family* ("family pictures") in chronological order, then construct a story based on the arranged sequence; (6) the previous task was repeated with three interrelated pictures showing *peer relations* ("peers pictures"); (7) the previous task was repeated with three interrelated pictures showing a *romantic relationship* ("romance pictures"). Figure 1 shows the picture sequences used.

Picture orders of the participants were documented, and their speech was recorded with their written consent. The recording was done in a noise-free environment, using a Sony ICD-PX470 dictaphone and a RØDE Lavalier Go clip microphone. The procedure was conducted in accordance with the Declaration of Helsinki, and resulted in five speech recordings for each subject: "previous day," "read story," "family pictures," "peers pictures," and "romance pictures."

Feature extraction

Extracting temporal parameters of speech

Audio recordings were first manually split into different recordings by tasks. This step resulted in five speech samples for each participant, in accordance with the recording procedure.

To automatically estimate temporal parameters, a standard ASR system was utilized [92]. We used the HTK tool [95], modified to allow the use of a Hidden Markov Model/Deep Neural Network (DNN) hybrid set-up [96]. As acoustic features, we used 40 raw Mel-frequency filter bank energy values along with log-energy and the

Fig. 1 Picture sequences used for the story construction. Note. Each sequence contains three pictures. A unique identifier is assigned to each picture with a letter representing the sequence (F – family, A – peers/age group, R – romance) and a number between 1 and 3 in a randomized order, written on the back of the pictures

first and second order derivatives ('FBANK + Δ + ΔΔ'). The DNN acoustic model was trained on a subset of 60 h of recordings from the BEA corpus [97]; to better suit noisy acoustic conditions, it was extended to 240 h by adding noise, background speech, and reverberation to the recordings of the BEA corpus. Recognition was performed at the level of phones, consisting of Hungarian phonemes, silent and filled pauses, breath intakes, and sighs [98]. The output of the ASR system for a speech recording is a list of phones along with the starting and ending time points of each phone.

The acoustic temporal parameters investigated were divided into three categories:

- *Utterance length*: the duration between the beginning and end of the response of the subject (the initial and final silent pauses excluded).
- *Speech rate and Articulation rate*: the number of phones uttered over the whole duration of the utterance or over the duration excluding pauses.
- *Duration of pauses*, *Number of pauses*, *Average length of pauses*, *and Frequency of pauses*: describing the amount of pauses in some way. These were calculated in two different ways: for silent pauses only, and for filled pauses only.

Table 2 Temporal parameters

Parameter	Metric	Description
Utterance length	s	the duration of the whole narrative
Articulation rate	phone/s	the number of phones per second excluding pauses
Speech rate	phone/s	the number of phones per second including pauses
Duration of pauses	ratio	the aggregated duration of pauses relative to the duration of the whole narrative
Number of pauses	ratio	the number of pauses relative to the number of speech sounds
Average length of pauses	s	the aggregated duration of pauses relative to the number of pauses
Frequency of pauses	phone/s	the number of pauses per second

Note. s – second

These temporal parameters can all be derived from the output of the ASR system (i.e., from the time-aligned phone sequence) by simple calculations. This process led to 11 Speech-GAP temporal parameters overall (see Table 2).

Extracting features of linguistic formulation

Automatic transcriptions were generated from the speech samples using Alrite (Alrite©; https://alrite.io/ai/hu/). The accuracy of the transcripts was verified by the first author, who manually reviewed each transcript while

listening to the corresponding audio recordings. This process involved correcting misspellings and other recognition errors produced by the ASR system. Additionally, disfluencies such as repetitions and self-corrections were removed to ensure consistency in linguistic analysis. The transcripts were then automatically analysed with *magyarlanc*, a linguistic preprocessing toolkit for Hungarian [91]. The pipeline included sentence segmentation, from which the number of sentences was extracted, followed by tokenization, allowing for the extraction of word counts. Lemmatization permitted the calculation of the type/token ratio. Part-of-speech tagging was then applied, from which variables such as the ratio of nouns were derived. Next, dependency- and constituency-based syntactic and morphological analyses were conducted, allowing for the extraction of features like the ratio of first-person singular verbs and the proportion of complex sentences. The full list of variables is given in Table 3.

Statistical analyses

Statistical analyses were conducted using JASP statistical software [99].

For hypothesis testing, a series of generalized linear mixed-effects models were applied in a 2 × 5 design, with each model testing a single hypothesis. Each model included only one dependent variable – the ratio of content words, the ratio of first-person singular verbs, the ratio of third-person singular verbs, or the ratio of complex sentences – depending on the linguistic feature targeted by the given hypothesis. The fixed effects were Group (BPD, HC) and Task (previous day, story recall, family pictures, peers pictures, romance pictures), and

Table 3 Descriptive statistics of the BPD and HC groups, the main effects of group and task factors, and group × task interactions

Dependent variable	BPD	HC	Group		Task		Group × Task	
	M (SD)	M (SD)	X^2	p	X^2	p	X^2	p
Lexical features			df = 1		df = 4		df = 4	
Number of words	235.006 (351.527)	249.703 (226.822)	2.555	0.110	22329.031	< 0.001	1027.813	< 0.001
Type/token ratio (lexical diversity)	0.490 (0.100)	0.486 (0.093)	0.098	0.755	116.694	< 0.001	5.216	0.266
Ratio of content words (lexical density)	0.479 (0.039)	0.501 (0.037)	15.970	< 0.001	59.349	< 0.001	3.305	0.508
Syntactic features								
Ratio of nouns	0.116 (0.033)	0.130 (0.040)	6.331	0.012	156.289	< 0.001	5.535	0.237
Ratio of verbs	0.168 (0.030)	0.172 (0.026)	0.403	0.526	59.937	< 0.001	2.842	0.585
Ratio of adjectives	0.041 (0.023)	0.047 (0.025)	2.106	0.147	76.419	< 0.001	1.336	0.855
Ratio of numerals	0.010 (0.010)	0.013 (0.010)	3.398	0.065	63.197	< 0.001	8.408	0.078
Ratio of adverbs	0.140 (0.041)	0.136 (0.039)	0.453	0.501	56.660	< 0.001	9.617	0.047
Ratio of pronouns	0.096 (0.029)	0.083 (0.027)	10.710	0.001	64.555	< 0.001	3.435	0.488
Ratio of conjunctions	0.132 (0.033)	0.124 (0.029)	1.164	0.281	77.629	< 0.001	5.702	0.223
Ratio of adpositions	0.005 (0.006)	0.006 (0.006)	0.374	0.541	18.457	0.001	2.980	0.561
Ratio of negations	0.021 (0.014)	0.018 (0.012)	2.336	0.126	26.194	< 0.001	5.488	0.241
Number of sentences	5.285 (8.290)	5.103 (7.487)	0.000	0.976	403.567	< 0.001	41.806	< 0.001
Mean length of sentences	70.612 (60.033)	82.469 (61.579)	1.071	0.301	4.873	0.301	3.228	0.520
Ratio of complex sentences	0.823 (0.220)	0.902 (0.146)	4.183	0.041	17.555	0.002	3.218	0.522
Ratio of coordinations	0.096 (0.027)	0.095 (0.025)	0.090	0.764	61.752	< 0.001	1.920	0.750
Ratio of subordinations	0.051 (0.018)	0.054 (0.018)	0.473	0.492	35.399	< 0.001	4.526	0.340
Morphological features								
Ratio of plural nouns	0.068 (0.072)	0.067 (0.063)	0.027	0.870	15.510	0.004	1.736	0.784
Ratio of sg1 verbs	0.172 (0.201)	0.129 (0.179)	9.062	0.003	129.965	< 0.001	2.834	0.586
Ratio of sg2 verbs	0.006 (0.015)	0.011 (0.026)	1.222	0.269	113.083	< 0.001	9.073	0.059
Ratio of sg3 verbs	0.572 (0.210)	0.575 (0.183)	0.030	0.863	231.428	< 0.001	3.676	0.452
Ratio of pl1 verbs	0.030 (0.059)	0.042 (0.076)	2.342	0.126	83.860	< 0.001	5.920	0.205
Ratio of pl2 verbs	0.001 (0.007)	0.000 (0.001)	0.862	0.353	18.212	0.001	3.666	0.453
Ratio of pl3 verbs	0.105 (0.121)	0.132 (0.110)	0.114	0.735	89.815	< 0.001	3.782	0.436
Ratio of past tense verbs	0.344 (0.276)	0.398 (0.272)	2.397	0.122	193.353	< 0.001	8.852	0.065
Ratio of present tense verbs	0.533 (0.270)	0.483 (0.276)	2.030	0.154	180.867	< 0.001	6.760	0.149
Ratio of comparative adjectives	0.057 (0.121)	0.059 (0.141)	1.788	0.181	146.681	< 0.001	7.200	0.126
Ratio of superlative adjectives	0.002 (0.017)	0.006 (0.026)	1.670	0.196	31.257	< 0.001	4.151	0.386

Note. BPD – borderline personality disorder, HC – healthy controls, sg1 – first-person singular, sg2 – second-person singular, sg3 – third-person singular, pl1 – first-person plural, pl2 – second-person plural, pl3 – third-person plural, M – mean, SD – standard deviation, X^2 – statistical score of the likelihood-ratio test, df – degrees of freedom, p - significance value of the X^2-score

the random effect was the participant ID. To identify the differences between the BPD and HC groups, the main effect of the Group was calculated. To explore the effect of tasks, the interaction of Group and Task was calculated. Post-hoc pairwise group comparisons were conducted within each task in variables showing a significant Group × Task interaction. The same model structure was used to explore between-group differences in additional features of linguistic formulation.

To explore the patterns of NSP in BPD, a Spearman correlation analysis was carried out between the variables of linguistic formulation and temporal parameters of speech in the BPD and HC groups separately, controlled for tasks. Only those variables were examined that previously showed significant between-group differences.

Results

Hypothesis-based between-group differences

The main effect of the Group was significant in models with the following dependent variables: the ratio of content words, with the narratives of the BPD group containing fewer content words ($X^2(1) = 15.970$, $p < .001$), the ratio of first-person singular verbs, with the narratives of the BPD group containing more first-person singular verbs ($X^2(1) = 9.062$, $p = .001$), and the ratio of complex sentences, with the narratives of the BPD group containing less complex sentences ($X^2(1) = 4.183$, $p = .041$), relative to HCs' (see Table 3). We have not found a significant between-group difference regarding third-person singular verbs.

Exploratory analyses

Between-group differences

The main effect of the Group was significant in models with the following dependent variables: the ratio of pronouns, with the narratives of the BPD group containing more pronouns ($X^2(1) = 10.710$, $p = .001$), and the ratio of nouns, with the narratives of the BPD group containing fewer nouns ($X^2(1) = 6.331$, $p = .012$) relative to HCs' (see Table 3).

We conducted a post-hoc analysis to explore between-group differences in terms of distinct pronoun types. The main effect of the Group was significant in models with the following dependent variables: the ratio of demonstrative pronouns (e.g., *this*, *those*), with the narratives of the BPD group containing more demonstrative pronouns ($X^2(1) = 5.820$, $p = .016$), and the ratio of first-person singular pronouns (e.g., *I*, *me*), the narratives of the BPD group containing more first-person singular pronouns ($X^2(1) = 3.886$, $p = .049$) relative to HCs' (see Table 4).

The effect of speech elicitation task

In the exploratory analyses, Group × Task interaction was significant in models with the following dependent variables: the number of words ($X^2(4) = 1027.813$, $p < .001$), the number of sentences ($X^2(4) = 41.806$, $p < .001$), and the ratio of adverbs ($X^2(4) = 9.617$, $p = .047$) (see Table 3). Post-hoc pairwise group comparisons revealed a pattern of BPD individuals producing the highest number of words ($z = 1.252$, $p = .242$) and sentences ($z = 2.216$, $p = .134$), and the highest ratio of adverbs ($z = 2.084$, $p = .186$) in the "previous day" task, and the lowest number of words ($z = -2.644$, $p = .041$) and sentences ($z =$

Table 4 Descriptive statistics of pronoun types in the BPD and HC groups, the main effects of group and task factors, and group × task interactions

Dependent variable	BPD	HC	Group		Task		Group × Task	
	M (SD)	M (SD)	X^2	p	X^2	p	X^2	p
Definite pronouns			df = 1		df = 4		df = 4	
Ratio of sg1 pronouns	0.080 (0.110)	0.045 (0.063)	3.886	0.049	3.404	0.493	5.392	0.249
Ratio of sg2 pronouns	0.013 (0.052)	0.010 (0.029)	0.000	0.976	271.145	< 0.001	6.340	0.175
Ratio of sg3 pronouns	0.225 (0.182)	0.220 (0.134)	0.013	0.908	2.967	0.563	1.715	0.788
Ratio of pl1 pronouns	0.005 (0.019)	0.006 (0.030)	0.319	0.572	38.083	< 0.001	13.386	0.010
Ratio of pl2 pronouns	0.001 (0.010)	0.000 (0.002)	0.767	0.381	12.266	0.015	10.250	0.036
Ratio of pl3 pronouns	0.033 (0.067)	0.033 (0.077)	0.114	0.735	89.815	< 0.001	3.782	0.436
Ratio of reciprocal pronouns	0.001 (0.004)	0.002 (0.006)	2.623	0.105	98.638	< 0.001	5.117	0.275
Ratio of demonstrative pronouns	0.038 (0.020)	0.032 (0.015)	5.820	0.016	17.139	0.002	3.274	0.513
Ratio of relative pronouns	0.008 (0.009)	0.008 (0.009)	0.027	0.868	18.254	0.001	1.667	0.797
Ratio of interrogative pronouns	0.003 (0.005)	0.004 (0.005)	1.650	0.199	11.166	0.025	3.613	0.461
Indefinite pronouns								
Ratio of universal pronouns	0.003 (0.005)	0.003 (0.004)	0.421	0.517	31.228	< 0.001	7.261	0.123
Ratio of existential pronouns	0.006 (0.009)	0.006 (0.009)	0.067	0.796	66.383	< 0.001	11.720	0.020
Ratio of negative pronouns	0.001 (0.003)	0.001 (0.002)	0.003	0.955	28.011	< 0.001	5.622	0.229

Note. BPD – borderline personality disorder, HC – healthy controls, sg1 – first-person singular, sg2 – second-person singular, sg3 – third-person singular, pl1 – first-person plural, pl2 – second-person plural, pl3 – third-person plural, M – mean, SD – standard deviation, X^2 – statistical score of the likelihood-ratio test, df – degrees of freedom, p - significance value of the X^2-score

– 1.773, $p = .305$), and the lowest ratio of adverbs (z = – 1.060, $p = .867$) in the "peers pictures" task compared to HCs (see Table 5).

In the post-hoc analysis of pronouns, Group × Task interaction was significant in models with the following dependent variables: the ratio of first-person plural pronouns (e.g., *we*, *us*; $X^2(4) = 13.386$, $p = .010$), the ratio of second-person plural pronouns (e.g., *you*; $X^2(4) = 10.250$, $p = .036$), and the ratio of existential pronouns (e.g., *somewhere*, *anything*; $X^2(4) = 11.720$, $p = .020$) (see Table 4). Regarding plural personal pronouns, BPD individuals produced the highest ratio of first-person plural pronouns in the "family pictures" task ($z = 1.843$, $p = .326$), and the lowest in the "peers pictures" task (z = – 1.489, $p = .546$) relative to HCs. In contrast, BPD individuals produced the highest ratio of second-person plural pronouns in the "peers pictures" task compared to HCs ($z = 1.942$, $p = .261$). Task had a significant effect on the use of existential pronouns in the two groups. Individuals with BPD produced a significantly higher ratio of existential pronouns in the "read story" ($z = 23.993$, $p < .001$), "family pictures" ($z = 39.537$, $p < .001$), and "peers pictures" task ($z = 6.072$, $p < .001$), and a significantly lower ratio of existential pronouns in the "previous day" (z = – 28.825, $p < .001$) and "romance pictures" tasks (z = – 18.242, $p < .001$) relative to HCs (see Table 5).

Associations between linguistic formulation and temporal parameters

Table 6 summarizes the Spearman correlation coefficients between features of linguistic formulation and temporal parameters within the BPD and HC groups. In both groups, first-person singular pronouns and first-person singular verbs were significantly positively correlated with articulation rate.

Within the BPD group, the ratio of first-person singular pronouns (r = –.22, $p = .005$), first-person singular verbs (r = –.20, $p = .011$), and demonstrative pronouns (r = –.19, $p = .014$) were significantly negatively correlated with the number of silent pauses. Also, there was a significant positive association between the ratio of demonstrative pronouns and articulation rate ($r = .22$, $p = .004$), between the ratio of first-person singular verbs and speech rate ($r = .18$, $p = .023$), as well as a significant negative association between the ratio of first-person singular pronouns and the number of filled pauses (r = –.17, $p = .028$). The ratio of complex sentences was significantly positively correlated with the number ($r = .19$, $p = .015$), frequency ($r = .22$, $p = .005$), and duration of filled pauses ($r = .25$, $p = .001$).

Within the HC group, no associations were found regarding the aforementioned feature pairs. However, significant associations were found between the ratio of content words and the number (r = –.27, $p < .001$), frequency (r = –.27, $p < .001$), and duration of filled pauses (r = –.24, $p = .003$), and between first-person singular pronouns and the duration of filled pauses (r = –.20, $p = .015$).

Discussion

This study aimed to examine differences between individuals with BPD and HCs in terms of linguistic formulation. It also explored how task type influences linguistic formulation and how linguistic formulation relates to the temporal parameters of speech uniquely in BPD. To our best knowledge, this is the first study to examine NSP in BPD across tasks differing systematically in emotional and attentional demands, and the first to explore associations between linguistic formulation and temporal parameters of NSP.

Our first hypothesis posited that narratives produced by individuals with BPD should contain fewer content words compared to those of HCs. Consistent with findings reported by Močnik et al. [67], narratives by participants with BPD contained a significantly lower proportion of content words. These results indicate that BPD narratives are less informative as they contain fewer words that carry concrete meaning. More specifically, individuals with BPD used significantly fewer nouns, suggesting that these individuals refer to fewer specific entities or entity groups. The reduced number of content words might be a symptom of "poverty of content of

Table 5 Main effects of group, for those variables where group × task interaction was significant

Dependent variable	Previous day		Read story		Family pictures		Peers pictures		Romance pictures	
	z	p	z	p	z	p	z	p	z	p
Number of words	1.252	0.242	–1.550	0.242	–2.429	0.060	–2.644	0.041	–2.390	0.060
Number of sentences	2.216	0.134	–0.707	1.000	0.710	1.000	–1.773	0.305	–0.326	1.000
Ratio of adverbs	2.084	0.186	0.256	1.000	–0.464	1.000	–1.060	0.867	1.378	0.673
Ratio of pl1 pronouns	0.520	1.000	0.602	1.000	1.843	0.326	–1.489	0.546	0.543	1.000
Ratio of pl2 pronouns	0.107	1.000	0.233	1.000	0.187	1.000	1.942	0.261	0.234	1.000
Ratio of existential pronouns	–28.825	< 0.001	23.993	< 0.001	39.537	< 0.001	6.072	< 0.001	–18.242	< 0.001

z – score of the z-test, p – significance value of the z-score

Holm correction was applied to p values

Table 6 Correlation matrix of spearman correlation coefficients between features of linguistic formulation and temporal parameters within the BPD and HC groups, controlled for tasks

	AR		SR		NSP		NFP		FFP		DFP	
	BPD	HC	BPD	HC	BPD	HC	BPD	HC	BPD	HC	BPD	HC
Content words	0.06	0.12	0.08	0.09	−0.09	−0.14	−0.08	−0.27***	−0.07	−0.27***	−0.06	−0.24**
Nouns	−0.01	−0.14	−0.03	−0.12	0.03	0.05	0.07	0.03	0.06	−0.00	0.04	0.04
Demonstrative pronouns	0.22**	0.05	0.14	−0.00	−0.19*	−0.00	0.00	0.09	0.06	0.12	0.02	0.00
Sg1 pronouns	0.27***	0.25***	0.15	0.06	−0.22**	−0.08	−0.17*	−0.12	−0.12	−0.10	−0.08	−0.20*
Sg1 verbs	0.29***	0.25**	0.18*	0.04	−0.20*	0.01	−0.06	0.05	0.00	0.09	0.03	0.06
Syntactic complexity	−0.12	−0.11	0.02	0.11	−0.00	−0.08	0.19*	−0.05	0.22**	−0.02	0.25**	−0.05

Note. AR – articulation rate, SR – speech rate, NSP – number of silent pauses, NFP – number of filled pauses, FFP – frequency of filled pauses, DFP – duration of filled pauses, BPD – borderline personality disorder, HC – healthy control, sg1 – first-person singular

* $p < .05$

** $p < .01$

*** $p < .001$

speech", a form of NTD, in BPD. Poverty of content has also been reported in schizophrenia, bipolar disorder, and depression [65, 100, 101], disorders genetically linked to BPD [102]. NTD was found to be related to compromised attention, verbal memory, and planning [103]. In BPD, poverty of content may be triggered by dissociative states. Feeling detached from themselves or their surroundings, they may fail to recall specific episodic memories or focus on details in the presented pictures, resulting in less informative narratives. This mechanism is a hypothesis and should be directly tested in future work.

In addition, the present study found that individuals with BPD produced a significantly higher ratio of pronouns relative to HCs. Our post-hoc analysis revealed that BPD narratives contain a significantly higher ratio of first-person singular (e.g., *I*, *me*) and demonstrative pronouns (e.g., *this*, *those*) compared to HCs. The usage of these types of pronouns anchors the narratives to the speaker's deictic center – the subjective point of reference from which spatial, temporal, and interpersonal relations are organized [104]. First-person singular pronouns directly locate the speaker at the center of the narrative, while demonstrative pronouns position other entities and circumstances (e.g., time and place) in relation to that center. We also found that both pronoun types are significantly negatively correlated with the number of silent pauses. The increased and fluent use of both pronoun types suggests a strong egocentric anchoring of narratives, where the speaker orients the narrative heavily around their own subjective perspective. This narrative style may reflect reduced consideration of the listener's perspective, potentially stemming from impaired mentalization abilities commonly associated with BPD [40, 51, 52]. Because we did not directly measure mentalization, this link should be treated as a hypothesis for future research.

Besides their deictic use, demonstrative pronouns can also serve anaphoric functions, referring back to previously mentioned entities (e.g., *I put that on.*) or events (e.g., *I did that.*) in the discourse. Another explanation for the frequent use of demonstrative pronouns is that individuals with BPD prefer using demonstrative pronouns to content words or larger language units when referring back to specific entities or entire events. Pronouns – similar to other function words – are much more predictable based on preceding words and therefore are accessed faster than content words [105–107]. The increased reliance on demonstrative pronouns may reflect a compensatory lexical selection or grammatical encoding strategy in BPD, demanding less lexical retrieval and planning. The observed negative association between demonstrative pronoun use and the number of silent pauses supports this interpretation, suggesting that such expressions

are accessed more fluently. We frame this compensatory-strategy account as a hypothesis, not a confirmed mechanism. This strategy is potentially driven by verbal memory deficits and impaired planning, both of which have been documented in individuals with BPD [44, 46, 48, 49]. However, the excessive use of demonstrative pronouns may lead to referential ambiguity at the expense of the listener's elevated cognitive load. If the speaker uses the same referential expression (e.g., *that*) to refer back to multiple distinct entities or events, the burden of disambiguation shifts to the listener. In these cases, the lexical choice is not informative enough for them to determine the intended meaning. Ambiguous pronoun use – found in autism spectrum disorder [108] and schizophrenia [83, 109] – may reflect deficient mentalization in BPD. Again, we did not assess mentalization directly; this pathway should be tested in future studies.

Our second hypothesis stated that individuals with BPD use more first-person singular verbs than HCs. This was confirmed by the data. In addition, BPD individuals produced a significantly higher ratio of first-person singular pronouns relative to HCs. Results aligned with prior findings [74, 76]. The frequent use of first-person singular pronouns is associated with self-referential thinking [55, 78], which in turn is linked to activation of the default mode network [110, 111]. In BPD, self-referential thinking can be manifested in rejection sensitivity [39] and rumination [112]. We also found that first-person singular reference (verbs and pronouns) significantly negatively correlated with the number of silent pauses in BPD but not in HCs. One interpretation is that self-referential information is salient in BPD individuals' memory due to frequent retrieval, so they are easily accessible without the need for pausing during NSP. Supporting this, experimental studies have demonstrated that BPD individuals exhibit a stronger interference effect [50] and enhanced recall [113] for self-relevant words relative to HCs. On the one hand, the salience of self-referential thoughts may hinder shifting from an internal, self-focused mode to an external, task-related focus, likely reflecting attentional deficits in BPD [44–49]. On the other hand, the same bias may increase the tendency to associate external stimuli (e.g., read story, picture sequences) with the self, possibly due to poor inhibition in BPD [45, 48, 49]. This interpretation is a hypothesis and should be tested in future work through direct assessment of cognitive functions.

Our third hypothesis anticipated the greater use of third-person singular verbs in BPD narratives compared to those for HCs, but this was not supported by the data.

Our fourth hypothesis said that narratives of individuals with BPD would be characterized by lower syntactic complexity relative to those of HCs. In line with Carter and Grenyer [77, 81], the narratives of individuals with BPD consisted of fewer complex sentences compared to those of HCs. Lower syntactic complexity – also found in schizophrenia [83] and depression [101] – is associated with NTD, processing speed, attention, and executive functioning [82, 84]. Furthermore, our correlation analysis revealed a significant positive association between syntactic complexity and various measures of filled pauses among individuals with BPD but not in HCs. Similarly, Goldman-Eisler [30] found no association between syntactic complexity and hesitation pauses in the general population. Due to their lower processing speed and reduced verbal working memory capacity [44–46], individuals with BPD may struggle to construct syntactically complex sentences. This can yield simpler, isolated sentences, making narratives fragmented and less cohesive. Increased filled pauses indicate elevated cognitive load during the production of more complex sentences. Alternatively, speakers may compensate for processing demands by simplifying syntax to reduce the likelihood of breakdowns in fluency and cohesion, or by pausing more frequently to allow for greater planning time. From the listener's perspective, syntactically less complex sentences may convey less information. This may reflect underlying mentalization deficits in individuals with BPD, limiting their ability to model the listener's background knowledge. These proposed links to cognitive functions and mentalization are hypotheses and should be directly tested in future research.

We also examined whether the type of speech elicitation task moderated the between-group differences. In the "previous day" task, the BPD group produced the highest number of words and sentences, the highest ratio of adverbs, and the lowest ratio of existential pronouns compared to HCs relative to other tasks. The elevated word and sentence counts reflect greater verbosity. This suggests that for individuals with BPD, accessing and elaborating on self-referential content is both more rewarding and cognitively less demanding than recalling or constructing stories based on external stimuli. The higher proportion of adverbs and lower use of existential pronouns indicate that individuals with BPD produced not only longer but also more detailed narratives in the "previous day" task. Adverbs (e.g., *yesterday*, *outside*) and existential pronouns (e.g., *anytime*, *somewhere*) can function as conceptual opposites in language use. Adverbs define the circumstances of narrated events – such as time, place, and manner – thereby enhancing narrative specificity. In contrast, existential pronouns introduce vagueness or uncertainty by referring to non-specific circumstances. Our finding differs from prior research on narrative specificity in BPD, which has reported overgeneral memory during autobiographical recall [70–75] and reduced adverb use when recounting early attachment experiences [77]. Speaking about the previous day may have been less triggering than recalling more significant

autobiographical memories. Integrating our findings with prior work, we hypothesize that individuals with BPD produce longer, more specific narratives when recalling endogenous material (vs. exogenous attentional focus), potentially reflecting self-referential thinking, and when describing less triggering (vs. triggering) events, possibly due to emotion dysregulation [39, 42]. Future studies should test this hypothesis using concurrent affective measures across tasks varying in emotional and attentional demands.

The BPD group produced the highest ratio of first-person plural pronouns and existential pronouns in the "family pictures" task, and the highest ratio of second-person plural pronouns, along with the lowest use of first-person plural pronouns, adverbs, and overall narrative output (i.e., words and sentences) in the "peers pictures" task, compared to HCs relative to other tasks. The increased use of first- and second-person plural pronouns suggests that individuals with BPD may be more inclined to speak from a character's point of view by adopting a "field perspective" [114]. In the "family pictures" task, the elevated use of first-person plural pronouns (e.g., *we*, *us*) may indicate a stronger identification with familial roles (e.g., the child) and a tendency to reconstruct the family unit as a cohesive group. However, this apparent sense of unity is accompanied by a simultaneously high use of existential pronouns (e.g., *somewhere*, *anything*), which reflect uncertainty [115]. This duality is consistent with features of insecure attachment in BPD, wherein proximity seeking coexists with low epistemic trust in caregivers' intentions [116]. In the "peers pictures" task, the increased use of second-person plural pronouns (e.g., *you all*) alongside a decrease in first-person plural pronouns may indicate feelings of social exclusion, possibly arising from the identification with the bullied character. Perceived social exclusion has been associated with the hyperactivation of the anterior cingulate cortex [117, 118], which reflects rejection sensitivity in BPD [39]. The lower verbosity and adverb use – along with the elevated number of silent pauses reported in our previous study [90] – suggest that NSP about social exclusion is emotionally overwhelming, potentially disrupting linguistic formulation in individuals with BPD. Lower adverb use aligns with prior studies reporting reduced specificity in rejection-related recall among BPD individuals [74] and fewer adverbs when recounting early attachment experiences [77]. Our findings suggest a dual pattern in BPD in the aforementioned picture sequences: individuals embody characters through a field perspective yet simultaneously disconnect from the narrative setting, as indicated by less specific narratives. This pattern is coupled with reduced global coherence found in our previous study [90]. Picture sequences depicting family and peer relations may be emotionally salient for individuals with BPD, prompting identification with vulnerable characters. This identification, in turn, may elicit a stress response, which may subsequently disrupt cognitive functions required for constructing original narratives, resulting in shorter, less specific, and less coherent speech output. This interpretation is a hypothesis and should be tested in future work using concurrent measures of affect.

Limitations

Despite the insights gained from this study, several limitations should be mentioned.

Most importantly, the study was not preregistered, which limits transparency and increases the number of possible analyses. The cross-sectional design of the study prevents us from capturing stable characteristics of NSP in BPD. Additionally, the sample size was relatively small, making the findings more susceptible to the influence of outliers and reducing statistical power. The gender imbalance in the sample may have influenced NSP and limited the generalizability of the results across genders. Another important limitation is the absence of neuropsychological testing. Without measures of executive functioning and attention, we cannot rule out that general cognitive factors affected NSP.

Moreover, HCs did not undergo psychological screening, which raises the possibility that undiagnosed psychiatric conditions may have confounded the group comparisons. The diagnostic procedure for BPD was based on clinical documentation rather than standardized interviews conducted within the study, which introduces a risk of diagnostic variability. As expected in clinical populations, the majority of participants in the BPD group had comorbid psychiatric conditions. While this reflects real-world complexity, it complicates interpretation: observed group differences may reflect general psychopathology rather than mechanisms specific to BPD. Another limitation is the lack of detailed data on pharmacotherapy. Given that medications can influence cognitive and linguistic functioning, their unmeasured effects may have contributed to variation in speech patterns.

Finally, the study was conducted exclusively in a Hungarian-speaking sample. Owing to its Uralic origin and distinct grammatical structure, Hungarian may limit cross-linguistic generalizability. Replication in other languages and cultural contexts is needed to validate these findings more broadly. Nevertheless, the NSP patterns observed in BPD (e.g., heightened first-person singular reference) closely resemble those reported in Indo-European languages, suggesting that these findings may reflect universal features of BPD.

Conclusions

Overall, the NSP of individuals with BPD – marked by the reduced use of content words, increased reliance on pronouns, heightened frequency of first-person singular verbs, and lower syntactic complexity – reveals a pattern of reduced informativity in both referential and structural domains, alongside an intensified self-focus. The observed negative correlation between first-person singular reference and the number of silent pauses further supports increased self-focus, which may arise from the overactivation of self-referential thoughts and a heightened tendency to attribute external events to the self. This tendency can direct attention away from the listener's needs and linguistic formulation in general, resulting in reduced informativity of NSP. Lower syntactic complexity – reflecting reduced structural informativity – was associated with an elevated number of filled pauses, possibly indicating an increased cognitive effort when constructing complex sentences. While lower structural informativity appears to be a stable feature of BPD, lower referential informativity (e.g., pronoun use) seems more context-dependent, varying with the emotional (i.e., triggering or non-triggering) and attentional (i.e., endogenous or exogenous) demands of the task. Lower informativity suggests that individuals with BPD may have difficulty mobilizing cognitive resources necessary for cohesive, listener-oriented linguistic formulation, possibly due to heightened self-focus.

Clinically, the observed linguistic patterns provide practical targets for evaluation and intervention in BPD. Our findings indicate that NSP in BPD is marked by heightened self-referential focus (elevated first-person singular reference) and reduced informativity (fewer content words, more pronouns, lower syntactic complexity), with task-dependent fluctuations under socially triggering, exogenous-attention demands. These linguistic signatures can help clinicians (i) recognize moments of self-focus and potential dissociation when narratives become pronoun-heavy and less specific, (ii) tailor interview prompts toward listener-oriented reformulation (e.g., asking for concrete nouns, agents, places) to scaffold coherence, and (iii) monitor therapy progress using brief narrative probes across tasks that differentially tax social cognition (e.g., peer-exclusion images vs. neutral daily recall). Because increased first-person reference was produced with fewer silent pauses, readily accessible self-related content may compete with mentalization for the listener, suggesting utility for treatments that train perspective-taking and structural elaboration (e.g., MBT, skills for syntactic expansion). These speech features are feasibly extractable with automated tools and could complement standard assessments as digital markers for emotion dysregulation and mentalization difficulties in BPD, pending replication across languages and larger, clinically diverse samples.

We are conducting a follow-up study with a larger sample that includes HC prescreening, standardized BPD diagnostics, detailed pharmacotherapy documentation, and a battery of neuropsychological, affective, and mentalization measures alongside narrative tasks systematically varying in attentional and emotional demands. This design will directly test the proposed links between self-referential bias, affectivity, cognitive control, and NSP.

Abbreviations

NSP Narrative speech production
BPD Borderline personality disorder
HC Healthy control
NTD Negative formal thought disorder
ASR Automatic speech recognition
DNN Deep Neural Network

Glossary

Anaphore	A pronoun that refers back to a previously mentioned element in the discourse, relying on grammatical and/or semantic cues for interpretation.
Articulation	The physical production of speech sounds by the movement of speech organs.
Clause	A syntactic unit that contains a subject and a predicate.
Cohesion	The use of linguistic devices (e.g., anaphores, conjunctions) that link sentences and clauses, creating formal connectedness in a text.
Complex Sentence	A sentence that includes more than one clause.
Conceptualization	The first stage of speech production, where the speaker constructs the intended message based on events, entities, attributes, and circumstances.
Constituency-based Syntax	A type of syntactic analysis that breaks sentences into hierarchically nested constituents or phrases.
Content Words	Words with inherent meaning (nouns, verbs, adjectives, adverbs) that denote conceptual elements of a message.
Definite Pronoun	A pronoun (e.g., *she*, *it*) that refers to a specific entity.

Deixis	The use of expressions (e.g., *this*, *that*, *here*, *now*) whose meaning depends on the speaker's spatial, temporal, or interpersonal context. Deictic expressions are anchored in the deictic center, typically the speaker's point of view.
Dependency-based Syntax	A syntactic analysis method that represents grammatical relationships as directed links (dependencies) between words.
Discourse Context	The linguistic context formed by prior discourse, which helps determine the referent of an anaphore.
Disfluency	Interruptions in the normal flow of speech, including repetitions, false starts, and self-corrections.
Disfluency	Interruptions in the normal flow of speech, including repetitions, false starts, and self-corrections.
Function Words	Words that primarily serve grammatical purposes (e.g., articles, conjunctions, adpositions, pronouns) and carry little semantic content on their own.
Global coherence	The degree to which all parts of a narrative relate to the main topic.
Grammatical Encoding	The stage in speech production where retrieved lexical items (lemmas) are organized into syntactic structures and assigned grammatical functions (e.g., subject, tense, agreement) to form well-structured sentences.
Indefinite Pronoun	A pronoun that quantifies an unspecified or unknown set of entities (e.g., *nothing*, *somehow*, *everybody*), independent of discourse context.
Lemma	The base or dictionary form of a word.
Lemmatization	The process of reducing a word to its lemma, enabling consistent linguistic analysis across inflected forms.
Lexeme	The fully inflected, phonologically encoded form of a word ready for articulation.
Lexical Access	The process of retrieving the appropriate lemmas from the mental lexicon.
Lexical Selection	Choosing the most contextually appropriate lemma among competing activated candidates.
Linguistic Formulation	A stage in speech production that involves the grammatical and morphological encoding of the intended message.
Mental Lexicon	The mental repository of words, including their meanings, syntactic properties, and usage contexts.
Morphophonological Encoding	The stage in speech production where lemmas are transformed into lexemes.
Part-of-Speech Tagging	A computational method that assigns word categories (e.g., noun, verb, adjective) to each word in a text based on context and grammar.
Phone	A unit of speech sound, including all perceptible vocal events such as phonemes, pauses, breaths, and sighs.
Phoneme	The smallest unit of language that can differentiate words (e.g., /p/ vs. /b/).
Phrase	A syntactic unit consisting of a word (e.g., *walk*) or group of words (e.g., *the dog*) that functions as a constituent within a sentence.
Pronoun	A type of function word (e.g., *this*) used to replace a noun or noun phrase (e.g., *the pen*), either to refer to something in the physical context or previously mentioned in discourse.
Syntax	The set of rules governing how words and phrases are combined to form grammatically correct sentences.
Syntactic Category	The grammatical classification of a word (e.g., noun, verb, adjective), which determines its function in a sentence.
Syntactic Complexity	The degree to which a sentence contains multiple or nested clauses.

Syntactic Tree	A hierarchical diagram representing the grammatical relationships between different parts of a sentence.
Token	An individual occurrence of a word in a given text. Repeated uses of the same word count as multiple tokens.
Tokenization	The process of segmenting text into its tokens.
Type	A unique word form in a given text. Repeated uses of the same word count as one type.
Type/Token Ratio	A measure of lexical diversity calculated by dividing the number of types by the total number of tokens.

Acknowledgements
We would like to extend our sincere gratitude to all those individuals who participated in this study.

Author contributions
**Fanni Felletár: ** Writing – original draft, Conceptualization, Methodology, Investigation, Formal Analysis. **Veronika Vincze: ** Writing – original draft, Writing – review & editing, Software. **Gábor Gosztolya: ** Writing – original draft, Software. **Ildikó Hoffmann: ** Writing – review & editing, Supervision, Methodology. **Anna Babarczy: ** Writing – review & editing, Supervision, Methodology. **Zsolt Szabolcs Unoka: ** Writing – review & editing, Supervision.

Funding
The research was implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation (NRDI) Fund, financed under the TKP2021-NVA funding scheme. Fanni Felletár's research was funded by the Hungarian Academy of Sciences. Gábor Gosztolya was supported by the NRDI Office of the Hungarian Ministry of Innovation and Technology (grant no. TKP2021-NVA-09) and within the framework of the Artificial Intelligence National Laboratory Program (RRF-2.3.1-21-2022-00004).

Data availability
Data used in this study have been deposited in the Open Science Framework (i.e., www.osf.io), currently accessible through a private weblink. The latter will be made available upon reasonable request to the corresponding author.

Declarations

Ethics approval and consent to participate
Written informed consent was obtained from all the participants. The study was approved by the ethics committee of HUN-REN Hungarian Research Centre for Linguistics (No. 1/2021).

Consent for publication
The participants who took part in the study provided written informed consent for the publication of the data collection results in anonymized, aggregated form.

Competing interests
The authors declare no competing interests.

Author details
[1]Psycholinguistics and Neurolinguistics Research Group, Institute for General and Hungarian Linguistics, ELTE Research Centre for Linguistics, Budapest, Hungary
[2]Department of Cognitive Science, Faculty of Natural Sciences, Budapest University of Technology and Economics, Budapest, Hungary
[3]HUN-REN-SZTE Research Group on Artificial Intelligence, Szeged, Hungary
[4]Institute of Informatics, University of Szeged, Szeged, Hungary
[5]Department of Psychiatry, Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary
[6]Department of Psychiatry and Psychotherapy, Semmelweis University, 6 Balassa Street, Budapest 1083, Hungary

Received: 20 May 2025 / Accepted: 8 October 2025
Published online: 17 November 2025

References

1. Bruner JS. Acts of meaning. Cambridge, MA: Harvard University Press; 1990.
2. Levelt WJM. Speaking: from intention to articulation. Cambridge, MA: MIT Press; 1989.
3. Mar RA. The neuropsychology of narrative: story comprehension, story production and their interrelation. Neuropsychologia. 2004;42:1414–34. https://doi.org/10.1016/j.neuropsychologia.2003.12.016.
4. Pinker S. Rules of Language. Science. 1991;253(5019):530–5. https://doi.org/10.1126/science.1857983.
5. Ye Z, Zhou X. Executive control in Language processing. Neurosci Biobehav Rev. 2009;33(8):1168–77. https://doi.org/10.1016/j.neubiorev.2009.03.003.
6. Dell GS. A spreading-activation theory of retrieval in sentence production. Psychol Rev. 1986;93(3):283–321. https://doi.org/10.1037/0033-295X.93.3.283.
7. Clark EV. Conceptual perspective and lexical choice in acquisition. Cognition. 1997;64(1):1–37. https://doi.org/10.1016/s0010-0277(97)00010-3.
8. Grice HP. Logic and conversation. In: Cole P, Morgan JL, editors. Syntax and semantics. Volume 3. New York: Academic; 1975. pp. 41–58.
9. Searle JR. Intentionality: an essay in the philosophy of Mind. Cambridge: Cambridge University Press; 1983.
10. Schmitt BM, Meyer AS, Levelt WJ. Lexical access in the production of pronouns. Cognition. 1999;69(3):313–35. https://doi.org/10.1016/s0010-0277(98)00073-0.
11. Diessel H. Demonstratives, joint attention, and the emergence of grammar. Cogn Linguist. 2006;17(4):463–89. https://doi.org/10.1515/COG.2006.015.
12. Levinson SC. Deixis. In: Horn LR, Ward G, editors. The handbook of pragmatics. Oxford: Blackwell Publishing; 2004. pp. 97–120.
13. Arnold JE. Reference production: Production-internal and addressee-oriented processes. Lang Cogn Process. 2008;23(4):495–527. https://doi.org/10.1080/01690960801920099.
14. Gordon PC, Grosz BJ, Gilliom LA. Pronouns, names, and the centering of attention in discourse. Cogn Sci. 1993;17(3):311–47. https://doi.org/10.1207/s15516709cog1703_1.
15. Sperber D, Wilson D, Relevance. Communication and cognition. 2nd ed. Oxford: Blackwell Publishing; 1995.
16. Huddleston R, Pullum GK. The Cambridge grammar of the english Language. Cambridge: Cambridge University Press; 2002.
17. Kempen G, Hoenkamp E. An incremental procedural grammar for sentence formulation. Cogn Sci. 1987;11(2):201–58. https://doi.org/10.1207/s15516709cog1102_5.
18. Garrett MF. The analysis of sentence production. In: Bower GH, editor. Psychology of learning and motivation. Volume 9. New York: Academic; 1975. pp. 133–77. https://doi.org/10.1016/S0079-7421(08)60270-4.
19. Ferreira F. Effects of length and syntactic complexity on initiation times for prepared utterances. J Mem Lang. 1991;30(2):210–33. https://doi.org/10.1016/0749-596X(91)90004-4.
20. Givón T. Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. Stud Lang. 1991;15(2):335–70. https://doi.org/10.1075/sl.15.2.05giv.
21. Gibson E. Linguistic complexity: locality of syntactic dependencies. Cognition. 1998;68(1):1–76. https://doi.org/10.1016/s0010-0277(98)00034-1.
22. Hawkins J. Efficiency and complexity in grammars. Oxford: Oxford University Press; 2004. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001.

23. Kolk H. A time-based approach to agrammatic production. Brain Lang. 1995;50(3):282–303. https://doi.org/10.1006/brln.1995.1049.
24. Kolk HHJ, Chwilla DJ, van Herten M, Oor PJW. Structure and limited capacity in verbal working memory: A study with event-related potentials. Brain Lang. 2003;85(1):1–36. https://doi.org/10.1016/S0093-934X(02)00548-5.
25. Alexander MP. Impairments of procedures for implementing complex Language are due to disruption of frontal attention processes. J Int Neuropsychol Soc. 2006;12(2):236–47. https://doi.org/10.1017/s1355617706060309.
26. Barker MS, Nelson NL, Robinson GA. Idea formulation for spoken Language production: the interface of cognition and Language. J Int Neuropsychol Soc. 2020;26(2):226–40. https://doi.org/10.1017/s1355617719001097.
27. Levelt WJ. Monitoring and self-repair in speech. Cognition. 1983;14(1):41–104. https://doi.org/10.1016/0010-0277(83)90026-4.
28. Tomasello M. Constructing a Language: A usage-based theory of Language acquisition. Cambridge, MA: Harvard University Press; 2003.
29. Butterworth B. Evidence from pauses in speech. In: Butterworth B, editor. Language production: speech and talk. New York: Academic; 1980. pp. 155–76.
30. Goldman-Eisler F, Psycholinguistics. Experiments in spontaneous speech. London/New York: Academic; 1968.
31. Maclay H, Osgood CE. Hesitation phenomena in spontaneous english speech. WORD. 1959;15(1):19–44. https://doi.org/10.1080/00437956.1959.11659682.
32. Scherer KR. Vocal communication of emotion: A review of research paradigms. Speech Commun. 2003;40(1–2):227–56. https://doi.org/10.1016/S0167-6393(02)00084-5.
33. Kail R. Processing speed, speech rate, and memory. Dev Psychol. 1992;28(5):899–904. https://doi.org/10.1037/0012-1649.28.5.899.
34. Hulme C, Thomson N, Muir C, Lawrence A. Speech rate and the development of short-term memory span. J Exp Child Psychol. 1984;38(2):241–53. https://doi.org/10.1016/0022-0965(84)90124-3.
35. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC: American Psychiatric Association; 2013.
36. Houben M, Vansteelandt K, Claes L, Sienaert P, Berens A, Sleuwaegen E, et al. Emotional switching in borderline personality disorder: A daily life study. Personal Disord. 2016;7(1):50–60. https://doi.org/10.1037/per0000126.
37. Kernberg OF. Borderline conditions and pathological narcissism. New York: Jason Aronson; 1975.
38. Linehan MM. Cognitive-Behavioral treatment of borderline personality disorder. New York: Guilford; 1993.
39. Gunderson J, Herpertz S, Skodol A, Torgersen S, Zanarini MC. Borderline personality disorder. Nat Rev Dis Primers. 2018;4:18029. https://doi.org/10.1038/nrdp.2018.29.
40. McLaren V, Gallagher M, Hopwood CJ, Sharp C. Hypermentalizing and borderline personality disorder: A Meta-Analytic review. Am J Psychother. 2022;75(1):21–31. https://doi.org/10.1176/appi.psychotherapy.20210018.
41. Bach B, Sellbom M. Continuity between DSM-5 categorical criteria and traits criteria for borderline personality disorder. Can J Psychiatry. 2016;61(8):489–94. https://doi.org/10.1177/0706743716640756.
42. Sicorello M, Schmahl C. Emotion dysregulation in borderline personality disorder: A fronto-limbic imbalance? Curr Opin Psychol. 2021;37:114–20. https://doi.org/10.1016/j.copsyc.2020.12.002.
43. Krause-Utz A, Frost R, Chatzaki E, Winter D, Schmahl C, Elzinga BM. Dissociation in borderline personality disorder: recent Experimental, Neurobiological Studies, and implications for future research and treatment. Curr Psychiatry Rep. 2021;23(6):37. https://doi.org/10.1007/s11920-021-01246-8.
44. Abramovitch A, Short T, Schweiger A, The C, Factor. Cognitive dysfunction as a transdiagnostic dimension in psychopathology. Clin Psychol Rev. 2021;86:102007. https://doi.org/10.1016/j.cpr.2021.102007.
45. D'Iorio A, Di Benedetto GL, Santangelo G. A meta-analysis on the neuropsychological correlates of borderline personality disorder: an update. Neurosci Biobehav Rev. 2024;165:105860. https://doi.org/10.1016/j.neubiorev.2024.105860.
46. Ruocco AC. The neuropsychology of borderline personality disorder: a meta-analysis and review. Psychiatry Res. 2005;137(3):191–202. https://doi.org/10.1016/j.psychres.2005.07.004.
47. Unoka Z, Richman MJ. Neuropsychological deficits in BPD patients and the moderator effects of co-occurring mental disorders: A meta-analysis. Clin Psychol Rev. 2016;44:1–12. https://doi.org/10.1016/j.cpr.2015.11.009.
48. Fertuck EA, Lenzenweger MF, Clarkin JF, Hoermann S, Stanley B. Executive neurocognition, memory systems, and borderline personality disorder. Clin Psychol Rev. 2006;26(3):346–75. https://doi.org/10.1016/j.cpr.2005.05.008.
49. LeGris J, van Reekum R. The neuropsychological correlates of borderline personality disorder and suicidal behaviour. Can J Psychiatry. 2006;51(3):131–42. https://doi.org/10.1177/070674370605100303.
50. Winter D, Elzinga B, Schmahl C. Emotions and memory in borderline personality disorder. Psychopathology. 2014;47(2):71–85. https://doi.org/10.1159/000356360.
51. Lazarus SA, Cheavens JS, Festa F, Rosenthal MZ. Interpersonal functioning in borderline personality disorder: A systematic review of behavioral and laboratory-based assessments. Clin Psychol Rev. 2014;34(3):193–205. https://doi.org/10.1016/j.cpr.2014.01.007.
52. Németh N, Mátrai P, Hegyi P, Czéh B, Czopf L, Hussain A, et al. Theory of Mind disturbances in borderline personality disorder: A meta-analysis. Psychiatry Res. 2018;270:143–53. https://doi.org/10.1016/j.psychres.2018.08.049.
53. Allport GW. The use of personal documents in psychological science. New York: Social Science Research Council; 1942. https://doi.org/10.1037/11389-000.
54. Chafe WL. The Pear stories: Cognitive, cultural and linguistic aspects of narrative production. Norwood (NJ): Ablex; 1980.
55. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural Language use: our words, our selves. Annu Rev Psychol. 2003;54:547–77. https://doi.org/10.1146/annurev.psych.54.101601.145041.
56. Andreasen NC. Thought, language, and communication disorders: I. Clinical assessment, definition of terms, and evaluation of their reliability. Arch Gen Psychiatry. 1979;36:1315–21. https://doi.org/10.1001/archpsyc.1979.01780120045006.
57. Corcoran CM, Cecchi GA. Using Language processing and speech analysis for the identification of psychosis and other disorders. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020;5(8):770–9. https://doi.org/10.1016/j.bpsc.2020.06.004.
58. Dikaios K, Rempel S, Dumpala SH, Oore S, Kiefte M, Uher R. Applications of speech analysis in psychiatry. Harv Rev Psychiatry. 2023;31(1):1–13. https://doi.org/10.1097/hrp.0000000000000356.
59. Elvevåg B, Cohen AS, Wolters MK, Whalley HC, Gountouna VE, Kuznetsova KA, et al. An examination of the Language construct in nimh's research domain criteria: time for reconceptualization! Am J Med Genet B Neuropsychiatr Genet. 2016;171(6):904–19. https://doi.org/10.1002/ajmg.b.32438.
60. Foltz PW, Chandler C, Diaz-Asper C, Cohen AS, Rodriguez Z, Holmlund TB, et al. Reflections on the nature of measurement in language-based automated assessments of patients' mental state and cognitive function. Schizophr Res. 2022;255:1–8. https://doi.org/10.1016/j.schres.2022.07.011.
61. Hinzen W, Palaniyappan L. The 'L-factor': Language as a transdiagnostic dimension in psychopathology. Prog Neuropsychopharmacol Biol Psychiatry. 2024;131:110952. https://doi.org/10.1016/j.pnpbp.2024.110952.
62. MacKinnon RA, Michels R, Buckley PJ. The psychiatric interview in clinical practice. 3rd ed. Washington: American Psychiatric Publishing; 2016.
63. Trzepacz PT, Baker RW. The psychiatric mental status examination. Oxford, U.K.: Oxford University Press; 1993.
64. Ingram RE. Self-focused attention in clinical disorders: review and a conceptual model. Psychol Bull. 1990;107(2):156–76. https://doi.org/10.1037/0033-2909.107.2.156.
65. Andreasen NC. Thought, language, and communication disorders. II. Diagnostic significance. Arch Gen Psychiatry. 1979;36(12):1325–30. https://doi.org/10.1001/archpsyc.1979.01780120055007.
66. Olah J, Spencer T, Cummins N, Diederen K. Automated analysis of speech as a marker of sub-clinical psychotic experiences. Front Psychiatry. 2024;14:1265880. https://doi.org/10.3389/fpsyt.2023.1265880.
67. Močnik S, Smrke U, Mlakar I, Močnik G, Gregorič Kumperščak H, Plohl N. Beyond clinical observations: a scoping review of AI-detectable observable cues in borderline personality disorder. Front Psychiatry. 2024;15:1345916. https://doi.org/10.3389/fpsyt.2024.1345916.
68. Jeanneau M, Armelius BA. Linguistic characteristics of neurotic, borderline and psychotic personality organization. Scand J Psychol. 1993;34(1):64–75. https://doi.org/10.1111/j.1467-9450.1993.tb01101.x.
69. Sundbom E, Jeanneau M. Multivariate modelling and personality organization: a comparative study of the defense mechanism test and linguistic expressions. Scand J Psychol. 1996;37(1):74–83. https://doi.org/10.1111/j.1467-9450.1996.tb00640.x.
70. Beran E, Richman MJ, Unoka Z. Autobiographical memory impairment in borderline personality disorder: A quantitative meta-analysis interpreted in terms of the CaR-FA-X model. J Pers Disord. 2019;33(6):818–31. https://doi.org/10.1521/pedi_2018_32_368.

71. Jones B, Heard H, Startup M, Swales M, Williams JMG, Jones RSP. Autobiographical memory and dissociation in borderline personality disorder. Psychol Med. 1999;29(6):1397–404. https://doi.org/10.1017/S0033291799001208.
72. Maurex L, Lekander M, Nilsonne A, Andersson EE, Asberg M, Ohman A. Social problem solving, autobiographical memory, trauma, and depression in women with borderline personality disorder and a history of suicide attempts. Br J Clin Psychol. 2010;49(Pt 3):327–42. https://doi.org/10.1348/014466509X454831.
73. Rasmussen AS, Jørgensen CR, O'Connor M, Bennedsen BE, Godt KD, Bøye R, et al. The structure of past and future events in borderline personality disorder, eating disorder, and obsessive–compulsive disorder. Psychol Conscious. 2017;4(2):190–210. https://doi.org/10.1037/cns0000109.
74. Rosenbach C, Renneberg B. Remembering rejection: specificity and linguistic styles of autobiographical memories in borderline personality disorder and depression. J Behav Ther Exp Psychiatry. 2015;46:85–92. https://doi.org/10.1016/j.jbtep.2014.09.002.
75. Reid T, Startup M. Autobiographical memory specificity in borderline personality disorder: associations with co-morbid depression and intellectual ability. Br J Clin Psychol. 2010;49(3):413–20. https://doi.org/10.1348/014466510X487059.
76. Lyons M, Aksayli ND, Brewer G. Mental distress and Language use: linguistic analysis of discussion forum posts. Comput Hum Behav. 2018;87:207–11. https://doi.org/10.1016/j.chb.2018.05.035.
77. Carter PE, Grenyer BFS. Expressive Language disturbance in borderline personality disorder in response to emotional autobiographical stimuli. J Pers Disord. 2012;26(3):305–21. https://doi.org/10.1521/pedi.2012.26.3.305.
78. Brockmeyer T, Zimmermann J, Kulessa D, Hautzinger M, Bents H, Friederich HC, et al. Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. Front Psychol. 2015;6:1564. https://doi.org/10.3389/fpsyg.2015.01564.
79. Newell EE, McCoy SK, Newman ML, Wellman JD, Gardner SK. You sound so down: capturing depressed affect through depressed Language. J Lang Soc Psychol. 2018;37(4):451–74. https://doi.org/10.1177/0261927X17731123.
80. Smirnova D, Cumming P, Sloeva E, Kuvshinova N, Romanov D, Nosachev G. Language patterns discriminate mild depression from normal sadness and euthymic state. Front Psychiatry. 2018;9:105. https://doi.org/10.3389/fpsyt.2018.00105.
81. Carter PE, Grenyer BFS. The effect of trauma on expressive Language impairment in borderline personality disorder. Pers Ment Health. 2012;6:183–95. https://doi.org/10.1002/pmh.1177.
82. Ciampelli S, de Boer JN, Voppel AE, Corona Hernandez H, Brederoo SG, van Dellen E, et al. Syntactic network analysis in schizophrenia-spectrum disorders. Schizophr Bull. 2023;49(Suppl2):S172–82. https://doi.org/10.1093/schbul/sbac194.
83. Covington MA, He C, Brown C, Naçi L, McClain JT, Fjordbak BS, et al. Schizophrenia and the structure of language: the linguist's view. Schizophr Res. 2005;77(1):85–98. https://doi.org/10.1016/j.schres.2005.01.016.
84. Schneider K, Leinweber K, Jamalabadi H, Teutenberg L, Brosch K, Pfarr JK, et al. Syntactic complexity and diversity of spontaneous speech production in schizophrenia spectrum and major depressive disorders. Schizophrenia. 2023;9(1):35. https://doi.org/10.1038/s41537-023-00359-8.
85. Silva AM, Limongi R, MacKinley M, Ford SD, Alonso-Sánchez MF, Palaniyappan L. Syntactic complexity of spoken Language in the diagnosis of schizophrenia: A probabilistic Bayes network model. Schizophr Res. 2023;259:88–96. https://doi.org/10.1016/j.schres.2022.06.011.
86. Trifu RN, Nemeș B, Herta DC, Bodea-Hategan C, Talaș DA, Coman H. Linguistic markers for major depressive disorder: a cross-sectional study using an automated procedure. Front Psychol. 2024;15:1355734. https://doi.org/10.3389/fpsyg.2024.1355734.
87. Zimmermann R, Fürer L, Schenk N, Koenig J, Roth V, Schlüter-Müller S, et al. Silence in the psychotherapy of adolescents with borderline personality pathology. Personal Disord. 2021;12(2):160–70. https://doi.org/10.1037/per0000402.
88. Wang B, Wu Y, Taylor N, Lyons T, Liakata M, Nevado-Holgado AJ, et al. Learning to detect bipolar disorder and borderline personality disorder with Language and speech in non-clinical interviews. INTERSPEECH 2020. 2020;437–41. https://doi.org/10.21437/interspeech.2020-3040.
89. Wang B, Wu Y, Vaci N, Liakata M, Lyons T, Saunders KEA. Modelling paralinguistic properties in conversational speech to detect bipolar disorder and borderline personality disorder. ICASSP. 2021;7243–47. https://doi.org/10.1109/icassp39728.2021.9413891.
90. Felletár F, Gosztolya G, Yang ZG, Hoffmann I, Babarczy A, Unoka ZS. Associations between coherence and Temporal parameters of narrative speech production in borderline personality disorder. J Psychiatr Res. 2025;188:218–28. https://doi.org/10.1016/j.jpsychires.2025.05.063.
91. Zsibrita J, Vincze V, Farkas R. Magyarlanc: a tool for morphological and dependency parsing of Hungarian. In: Proceedings of Recent Advances in Natural Language Processing (RANLP). INCOMA Ltd; 2013. p. 763–71. https://aclanthology.org/R13-1099/.
92. Kálmán J, Devanand DP, Gosztolya G, Balogh R, Imre N, Tóth L, et al. Temporal speech parameters detect mild cognitive impairment in different languages: validation and comparison of the Speech-GAP Test® in english and Hungarian. Curr Alzheimer Res. 2022;19(5):373–86. https://doi.org/10.2174/1567205019666220418155130.
93. World Health Organization. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneva: WHO; 1992. https://iris.who.int/handle/10665/37958.
94. Pléh C, Oláh A, Zétényi T. Individual differences in story recall. Text. 1988;8(4):395–410. https://doi.org/10.1515/text.1.1988.8.4.395.
95. Young S, Evermann G, Gales MJF, Hain T, Kershaw D, Moore G, et al. The HTK book. Cambridge, UK: Cambridge University Engineering Department; 2006.
96. Hinton G, Deng L, Yu D, Dahl G, Mohamed AR, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag. 2012;29(6):82–97. https://doi.org/10.1109/msp.2012.2205597.
97. Neuberger T, Gyarmathy D, Gráczi TE, Horváth V, Gósy M, Beke A. In: Proc TSD, Brno, editors. Development of a large spontaneous speech database of agglutinative Hungarian Language. Czech Republic; 2014. pp. 424–31. https://doi.org/10.1007/978-3-319-10816-2_51.
98. Moore RK, Skidmore L. On the use/misuse of the term 'phoneme'. In: Proc Interspeech. 2019. pp. 2340–4. https://doi.org/10.21437/Interspeech.2019-2711
99. JASP Team. JASP (Version 0.18.3) [Computer software]. 2024. https://jasp-stats.org/
100. Cohen AS, McGovern JE, Dinzeo TJ, Covington MA. Speech deficits in serious mental illness: a cognitive resource issue? Schizophr Res. 2014;160(1–3):173–9. https://doi.org/10.1016/j.schres.2014.10.032.
101. Palaniyappan L, Wang YL, Meister F. Disorganisation and depression: a re-examination of how we think and speak when depressed. Eur Arch Psychiatry Clin Neurosci. 2025. https://doi.org/10.1007/s00406-025-01994-1.
102. Witt SH, Streit F, Jungkunz M, Frank J, Awasthi S, Reinbold CS, et al. Genome-wide association study of borderline personality disorder reveals genetic overlap with bipolar disorder, major depression and schizophrenia. Transl Psychiatry. 2017;7(6):e1155. https://doi.org/10.1038/tp.2017.115.
103. Bora E, Yalınçetin B, Akdede BB, Alptekin K. Neurocognitive and linguistic correlates of positive and negative formal thought disorder: a meta-analysis. Schizophr Res. 2019;209:2–11. https://doi.org/10.1016/j.schres.2019.05.025.
104. Fillmore CJ. Santa Cruz lectures on deixis. Bloomington (IN): Indiana University Linguistics Club; 1975.
105. Bell A, Brenier JM, Gregory M, Girand C, Jurafsky D. Predictability effects on durations of content and function words in conversational english. J Mem Lang. 2009;60(1):92–111. https://doi.org/10.1016/j.jml.2008.06.003.
106. Herron D, Bates EA. Sentential and acoustic factors in the recognition of open- and closed-class words. J Mem Lang. 1997;37:217–39. https://doi.org/10.1006/JMLA.1997.2514.
107. Segalowitz SJ, Lane KC. Lexical access of function versus content words. Brain Lang. 2000;75(3):376–89. https://doi.org/10.1006/brln.2000.2361.
108. Finnegan EG, Asaro-Saddler K, Zajic MC. Production and comprehension of pronouns in individuals with autism: A meta-analysis and systematic review. Autism. 2020;25(1):3–17. https://doi.org/10.1177/1362361320949103.
109. Hinzen W, Rosselló J. The linguistics of schizophrenia: thought disturbance as Language pathology across positive symptoms. Front Psychol. 2015;6:971. https://doi.org/10.3389/fpsyg.2015.00971.
110. Menon V. 20 years of the default mode network: A review and synthesis. Neuron. 2023;111(16):2469–87. https://doi.org/10.1016/j.neuron.2023.04.023. Epub 2023 May 10. PMID: 37167968; PMCID: PMC10524518.
111. Ruocco AC, Marceau EM. Update on the neurobiology of borderline personality disorder: a review of structural, resting-state and task-based brain imaging studies. Curr Psychiatry Rep. 2024;26:807–15. https://doi.org/10.1007/s11920-024-01553-w.
112. Richman Czégel MJ, Unoka Z, Dudas RB, Demetrovics Z. Rumination in borderline personality disorder: a meta-analytic review. J Pers Disord. 2022;36(4):399–412. https://doi.org/10.1521/pedi.2022.36.4.399.

113. Maffly-Kipp J, McCredie MN, Morey LC. The self-reference effect as a behavioral indicator of identity disturbances associated with borderline personality features in a non-clinical sample. Borderline Personal Disord Emot Dysregul. 2022;9(1):20. https://doi.org/10.1186/s40479-022-00189-7.
114. Nigro G, Neisser U. Point of view in personal memories. Cogn Psychol. 1983;15(4):467–82. https://doi.org/10.1016/0010-0285(83)90016-6.
115. Vincze V. Uncertainty detection in Hungarian texts. In: Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers; Dublin. Ireland. Dublin City University and Association for Computational Linguistics. pp. 1844–53.
116. Luyten P, Campbell C, Fonagy P. Borderline personality disorder, complex trauma, and problems with self and identity: a social-communicative approach. J Pers. 2020;88(1):88–105. https://doi.org/10.1111/jopy.12483.
117. Rotge JY, Lemogne C, Hinfray S, Huguet P, Grynszpan O, Tartour E, George N, Fossati P. A meta-analysis of the anterior cingulate contribution to social pain. Soc Cogn Affect Neurosci. 2015;10(1):19–27. https://doi.org/10.1093/scan/nsu110.
118. Domsalla M, Koppe G, Niedtfeld I, Vollstädt-Klein S, Schmahl C, Bohus M, Lis S. Cerebral processing of social rejection in patients with borderline personality disorder. Soc Cogn Affect Neurosci. 2014;9(11):1789–97. https://doi.org/10.1093/scan/nst176.

Publisher's note