The visible and the latent linguistic clues of mental health in Brazilian Portuguese textual posts

Rodrigo Wilkens[1], Helena Caseli[2], Vânia Neris[2], Aline Villavicencio[1]
[1]University of Exeter,
[2]Universidade Federal de São Carlos (UFSCar),
Correspondence: r.wilkens@exeter.ac.uk

Abstract

Depressive symptomatology may be reflected in the language used by people with possible depressive profiles (PDP). This paper investigates to what extent symptoms of depression are manifested in Brazilian Portuguese narrative texts, and whether these can be used to identify relevant linguistic clues related to PDP. Moreover, the relation between these symptoms and PDP is explored, characterising the lexical, syntactic, and psycholinguistic aspects of texts produced by PDP. We found that texts associated with PDPs differed in some of these characteristics from non-PDP texts. The interactions between symptoms and PDP can also shed light on patterns of communication differentiation and the relationship between them. The results of this paper can help to characterise and understand the linguistic indicators that can be used to train more bespoke and accurate language models.

1 Introduction

Depression is a condition that affects 4% of the global population (Global Burden of Disease Collaborative Network, 2025). In Brazil, the prevalence of depressive disorder is even higher, affecting almost 17.4% of the adult population (Errazuriza et al., 2023). Depression is a clinically significant psychological disorder that causes a range of impairments as insomnia, memory loss, loss of pleasure, and a resulting decline in quality of life. In its most severe forms, depression may result in death by suicide. Despite its widespread prevalence and serious consequences, only about a quarter of individuals with mental disorders receive appropriate care (World Health Organization, 2021).

These challenges motivate the creation of complementary depression detection techniques, with broader reach, to enable the allocation of mental health resources to individuals and populations that would otherwise have no or poor access to them. These techniques may be used to analyse user-generated content published on social media platforms and other text-based reporting tools, such as personal diaries. Written expressions — whether shared on social media or recorded in personal diaries — can serve as valuable early indicators of depression. Recognizing these signs enables for timely prevention and prompt initiation of mental health care. Indeed, Natural Language Processing (NLP) can be used to support depression detection, as texts produced by individuals suffering from depression may contain clear signs of the condition — such as explicit mentions of symptoms (Yazdavar et al., 2017; Yadav et al., 2020; Mendes and Caseli, 2024) and self-declarations of diagnosis/ongoing treatment (Santos et al., 2023) — as well as more subtle clues like language style.

Although prior research demonstrates the utility of NLP for identifying depression (Yazdavar et al., 2017; Yadav et al., 2020), cultural factors can significantly shape how individuals express themselves. Therefore, it is critical to investigate these signs in different languages before starting NLP tools development. Investigating linguistic clues across diverse corpora is essential for advancing accurate computational technologies that can automatically process and understand the language linked to Possible Depressive Profile (PDP) in various cultural contexts. To do so for Brazilian Portuguese texts, we investigated some specific phenomena in this language (e.g. null subjects/person morphology) and also the combination of lexical, syntactic, and psycholinguistic features. This work is carried out in the scope of the AIM-Health[1] research project.

This paper investigates how symptoms of depression are manifested in Brazilian Portuguese narrative texts to identify relevant linguistic markers. The data consists of anonymous Facebook posts from the Amive corpus (Mendes and Caseli,

[1]https://www.aim-health.ufscar.br/

Proceedings of the 17th International Conference on Computational Processing of Portuguese (PROPOR 2026)
Vol. 2, pages 135–147
April 13-16, 2026 ©2026 Association for Computational Linguistics

2024) annotated at symptom-level (18 symptoms of depression) and user-level (indicating PDP). We address the following three research questions:

RQ1 Can the overt linguistic clues in narratives be used to identify cases of PDP?

RQ2 What are the linguistic characteristics of the symptoms presented in the Amive corpus?

RQ3 What are the symptoms linked to PDP?

These research questions lay the groundwork for using this corpus in Portuguese-language automatic processing of depression and mental health contexts. They also provide insights that can guide the analysis of other health-related corpora.

Our research approach involved exploring the lexical, syntactic, and psycholinguistic characteristics of the annotated texts. The results indicate clear differences between texts expressing signs of depression and those that do not. Accordingly, we present and discuss these characteristics that may serve as valuable features for training NLP models.

2 Related Work

De Choudhury et al. 2013 were one of the first to investigate the use of social media text for the automatic prediction of depression. They investigated the task of user-level depression classification based on features like emotional content from LIWC (Tausczik and Pennebaker, 2010) categories and activation/arousal from ANEW (Bradley and Lang, 1999), and depression related terms, language style (syntactic features) among others.

Ji et al. 2018 investigated Reddit and Twitter posts focusing on suicidal ideation. They also applied LIWC (Pennebaker et al., 2015) to analyse suicide-related texts and suicide-free posts. They point out differences in language style with a more direct and aggressive language for Twitter users, who tend to make direct mention of their suicidal ideation and somatic symptoms, while Reddit users employ a more indirect speech, contextualizing their feelings and talking about other subjects like finance, social life, and family. The authors also observed similarities between the two datasets, such as the increased frequency of first-person usage in suicidal users. The authors also used LDA (Blei et al., 2003) to topic analysis from posts containing suicidal ideation and summarized the results in three categories: internal factors (words expressing people's feelings, intentions, and desires like "know", "want", "feel" and "like", and "hope"), external social factors (topics containing words such as "money" and "working", "friend", "school", "surgery", "crisis", and "accident"), and the mixed internal/external factors.

Trifu et al. 2024 investigated the linguistic markers for major depressive disorder using a Romanian version of LIWC[2]. They analysed word categories related to: part-of-speech tags, affect words (e.g. positive, negative, anxiety), social words (e.g. family, friends), cognitive process (e.g. insight, causations, certainty), perceptual process (e.g. see, hear, feel), biological process (e.g. body, health, sexual), drives (e.g. affiliation, achievement, power), time-oriented (past, present, future), personal concern (e.g. work, leisure, money), and informal language. They concluded that the language used by depressive patients is significantly different from people without mood disorders in both form – such as the prevalence of short sentences, impersonal pronouns, first-person pronouns in plural form, conjunctions, auxiliary verbs and negations – and content – words indicating negative affects, anxiety in contrast to words indicating positive affects.

For the Portuguese language, a few studies investigate language technologies in mental health domain. Santos et al. 2023 delivered SetembroBR, a large corpus for user-level depression and anxiety classification composed of posts from self-declared depressed Brazilian Twitter users. Mann et al. 2020 developed multimodal classifiers for user-level depression detection based on textual and visual information from Brazilian college students' Instagram posts and their answers to a BDI (Beck Depression Inventory) questionnaire. Text features were TF-IDF, BoW, pretrained FastText and ELMo embeddings, and visual features were ResNet/ResNext embeddings, fused in a fully connected multimodal layer. LIWC categories were also used in traditional machine learning models. Mann et al. 2025 focused on topic-detection in Instagram, Twitter and Reddit datasets in Brazilian Portuguese in a user-level classification approach. They performed experiments with LDA-based models, BERTopic (Grootendorst, 2022) and large language models (LLMs) and found that the LLM-based method yielded a broader and more varied range of topics than conventional techniques. They found an overlap in the topics discussed by users with and

[2] Available at: https://www.liwc.app/.

without depression, which highlights the difficulty of classifying a user based solely on their topics.

Casani et al. 2021 and Mendes and Caseli 2024 investigated multiclass depression symptom classification in Portuguese Twitter and Facebook posts, respectively. In both cases, the corpora were manually labelled by mental health experts. In (Casani et al., 2021), only 3 symptoms categories were considered – psychological, physiological, and behavioural – and a dataset of 2,008 annotated posts was generated. In (Mendes and Caseli, 2024), 21 signals of depression (18 symptoms and 3 other signals) were defined and a dataset of 2,304 annotated symptoms was generated. However, none of these works for Portuguese language brings a linguistic analysis of the built datasets.

In this paper we investigate the linguistic clues and the symptoms relationship in the same corpus of (Mendes and Caseli, 2024): the Amive corpus.

3 Materials and Methods

3.1 The Amive Corpus

A key methodological aspect of this study concerns corpus selection. Identifying naturally occurring expressions of depression in large-scale online data is challenging due to the relatively low prevalence of explicit references to depressive symptoms in general discourse. Random sampling of social media content would therefore yield an extremely low signal-to-noise ratio. To address this issue, we rely on the Amive corpus (Mendes and Caseli, 2024), which was collected from public Facebook pages dedicated to sharing posts written by Brazilian college students. Although this targeted setting introduces potential sampling limitations (e.g., demographic concentration and platform-specific discourse norms), it substantially increases the likelihood of retrieving content relevant to depressive experiences. Given the volume of available material, further filtering was necessary to make human annotation feasible. According to the authors, posts were collected based on keywords related to depression ("suicide", "depression", "cut myself", "will to live", "kill myself" and "want to die") and publication date from January 1st, 2012 to December 31st, 2021.

After corpus construction, it was annotated[3] in two levels: user-level and sentence-level. In the user-level, each post was annotated as a Possible Depressive Profile (PDP) or not (non-PDP). In the sentence-level, texts spans were annotated regarding 18 symptoms of depression: (1) *Agitation/Restlessness*, (2) *Attention/memory deficit*, (3) *Alteration in weight/eating habits*, (4) *Sleep disorder*, (5) *Physical symptom*, (6) *Difficulty in decision-making*, (7) *Feeling of emptiness*, (8) *Loss/Diminishment of pleasure/libido*, (9) *Feeling of guilt*, (10) *Irritation/Aggressiveness*, (11) *Alteration in efficiency/functionality*, (12) *Tiredness/Discouragement/Fatigue*, (13) *Despair*, (14) *Worry/Fear/Anxiety*, (15) *Feeling of worthlessness/Low self-esteem*, (16) *Suicide/Self-extermination*, (17) *Helplessness/Social harm/Loneliness*, and (18) *Sadness/Depressed mood*.

For this study, the original annotation was revised to standardize the boundaries of the text spans and to certify that all annotated segments were present in the final corpus. Thus, the corpus under analysis in this paper has 604 textual posts, 336 of them annotated as PDP and 268 as non-PDP. It is worth mention that it is possible to have annotated symptoms in non-PDP posts. Table 1 shows the amount of instances annotated for each symptom. Although notes may extend beyond the sentence boundary, this rarely occurs.[4] Table 2 shows the total and average number of sentences, tokens, types, symptoms, and unique symptoms per text.

Table 3 brings some instances from the original corpus translated to English and slightly modified to anonymization purpose.

3.2 Linguistic characterization of symptoms and PDP

To address our first two research questions, we examined the lexical, syntactic, and psycholinguistic characteristics of the annotated texts summarized in Table 4. For both symptom-annotated spans and texts labelled as PDP or non-PDP, we computed the number of tokens, types, Type–Token Ratio (TTR), span length, mean syllables per word, and the Flesch readability score adapted for Brazilian Portuguese (Martins et al., 1996). These measures were selected to capture lexical richness, verbosity,

[3]According to Mendes and Caseli 2024, annotation was performed by four students from psychology, psychiatry or occupational therapy.

[4]Notes on symptoms that exceed the sentence boundary occur in 3% of cases (4) *Sleep disorder* and (16) *Suicide/Self-extermination*, and in 1% of cases (11) *Alteration in efficiency/functionality*, (12) *Tiredness/Discouragement/Fatigue*, (13) *Despair*, (14) *Worry/Fear/Anxiety*, (18) *Sadness/Depressed mood* and (17) *Helplessness/Social harm/Loneliness*.

Symptom ID	Count	Tokens	Tokens annotation
1	15	235	15.7 (11.3)
2	16	138	8.6 (3.8)
3	17	158	9.3 (4.7)
4	31	301	9.7 (4.9)
5	31	329	10.6 (6.5)
6	35	288	8.2 (5.0)
7	40	397	9.9 (6.8)
8	43	436	10.1 (4.8)
9	73	1006	13.8 (8.2)
10	149	2104	14.1 (10.2)
11	155	1745	11.3 (6.5)
12	153	1576	10.3 (6.8)
13	164	2074	12.6 (8.2)
14	210	2523	12.0 (7.8)
15	234	2636	11.3 (8.4)
16	278	3103	11.2 (7.3)
17	377	5094	13.5 (9.1)
18	469	4580	9.8 (6.9)

Table 1: Amount of annotated instances for each symptom and its total number of tokens and average (and standard deviation) length number of tokens.

	Total	Avg (std)
Sentences	4711	7.80 (8.51)
Tokens	100013	165.58 (159.4)
Types	59106	97.86 (73.05)
Symptom	25112	41.58 (46.7)
Symptom (unique)	1568	2.60 (2.26)

Table 2: Corpus size in sentences, tokens, types, symptoms, and unique symptoms total and mean (and standard deviation) per text.

and processing difficulty, which are frequently associated with affective and cognitive states in mental health research.

We further analyzed the distribution of part-of-speech (POS) tags and dependency relations to detect potential shifts in syntactic organization and discourse structure[5]. Syntactic distributions provide a structural perspective on language production that may reflect changes in cognitive load, self-focus, or narrative framing.

In addition, we explored person usage, with particular attention to first-person singular forms. We measured the distribution of grammatical person (first, second, and third) in both singular and plural forms. The focus on person marking is motivated by evidence linking self-referential language to depressive symptomatology (De Choudhury et al., 2013; Nambisan et al., 2015; Ji et al., 2018).

Because Portuguese allows null subjects, we distinguished between explicit personal pronouns and subjects inferred from verbal morphology. This distinction enables a more accurate estimation of self-referential tendencies in a pro-drop language. We therefore computed counts both including and excluding explicit pronoun realizations.

Finally, we employed the Brazilian Portuguese adaptation of LIWC[6] (Filho et al., 2013) to quantify the use of words associated with affective and cognitive dimensions, including positive emotion, negative emotion, anxiety, anger, sadness, negation, certainty, death, cognitive processes, and general affect. LIWC categories were included to provide a psychologically grounded lexical profile aligned with established mental health research. We additionally examined the use of absolutist expressions[7], as such language has been associated with cognitive rigidity in depressive discourse.

To examine whether linguistic features differed as a function of depressive profile and symptom presence, statistical analyses were conducted at the subject level. Because most linguistic measures exhibited non-normal distributions, non-parametric tests were employed. Differences between PDP and non-PDP subjects were assessed using the Mann–Whitney U test. When comparing linguistic patterns across symptom categories, the Kruskal–Wallis test was applied, followed by Dunn's post-hoc pairwise comparisons when global significance was observed. To control for multiple comparisons across features and symptom groups, p-values were adjusted using the Benjamini–Hochberg False Discovery Rate (FDR) procedure. All analyses were conducted using subject-level normalized measures in order to control for variation in text length.

3.3 Symptoms and PDP linking

To investigate which symptoms are linked to possible depressive profiles (RQ3), we fitted a Bayesian Generalised Linear Mixed Model (GLMM) with a Bernoulli likelihood and logit link function. The dependent variable was the binary subject-level label (`is_pdp`). Predictors consisted of subject-level symptom counts derived from manual annotations.

Symptom counts were standardised (z-scores) to ensure comparability of regression coefficients. A varying intercept was included for each subject in order to account for subject-level heterogeneity and

[5] Syntactic analysis was performed using the *pt_core_news_lg* model in SpaCy.

[6] `http://nilc.icmc.usp.br/portlex/index.php/pt/projetos/liwc`

[7] List of absolutist expressions explored: tudo, todos, todas, sempre, nunca, ninguém, jamais, inteiramente, and completamente.

Sentence	Symptom	From a PDP post?
When I'm alone, I feel a deep emptiness inside myself.	Feeling of emptiness	PDP
I try to stay alive, just waiting for another day to pass.	Suicide/Self-extermination	PDP
My heart is beating so hard it hurts, and I just feel like trash and a total failure.	Physical symptom & Feeling of worthlessness/Low self-esteem	PDP
What do I do???	Difficulty in decision-making	PDP
People simply don't like me.	Helplessness/Social harm/Loneliness	PDP
If all this crap is well-being, then I don't know what BAD-BEING is.	Irritation/Aggressiveness	non-PDP
In these cases, I'm afraid of relapsing and doing things I promised myself I would never do again.	Worry/Fear/Anxiety	non-PDP
It hurts to see my friend depressed.	Sadness/Depressed mood	non-PDP
Parties and bars aren't as fun anymore; it's a completely empty world for me.	Loss/Diminishment of pleasure/libido & Feeling of emptiness	non-PDP

Table 3: Instances from the original corpus translated to English and slightly modified to anonymization purpose.

Linguistic characteristic	Level
Number of tokens	Lexical
Number of types	Lexical
Type-Token Ratio (TTR)	Lexical
Span length	Lexical
Mean syllables per word	Lexical
Flesch readability score	Lexical
Part-of-speech tags	Syntactic
Dependency relations	Syntactic
Grammatical person	Syntactic
LIWC categories	Psycholinguistic
Absolutist expressions	Psycholinguistic

Table 4: Linguistic characteristics investigated in this paper

Linguistic characteristic	PDP
Mean syllables per word	↓
Flesch readability score	↑
TTR	↓
POS tags	↑ VERB, ADV ↓ NOUN, ADJ, ADP, PROPN
Dependency relations	↑ advmod, obj, mark, xcomp ↓ amod, case, nmod, expl
Grammatical person	↑ 1_Sing VERB ↓ 3_Sing, 3_Plur VERB ↓ explicit subjects
LIWC categories	↑ negative emotion, anger, negation, certainty, death
Absolutist expressions	↑

Table 5: Linguistic characteristics that showed significant differences between PDP vs non-PDP posts. ↑ (↓) indicates a higher (lower) proportion of the linguistic phenomenon in PDP posts.

repeated symptom annotations within individuals. Formally, the model can be expressed as:

$$\text{logit}(P(\text{PDP}_i = 1)) = \alpha + \sum_{k=1}^{K} \beta_k X_{ik} + u_i, \quad (1)$$

where α is the global intercept, β_k are fixed effects for each symptom k, X_{ik} is the standardised count of symptom k for subject i, and $u_i \sim \mathcal{N}(0, \sigma_{\text{subject}})$ is the subject-level random intercept.

4 Results

4.1 PDP (RQ1)

The results for PDP versus non-PDP posts are described in this section. Table 5 summarizes the linguistic characteristics with significant differences in PDP posts regarding non-PDP ones.

Lexical. To investigate whether texts associated with possible depressive profiles (PDP) differ lexically from non-PDP texts, we compared several measures of lexical complexity.

No significant differences were observed in overall text length (tokens), vocabulary size (types), or mean sentence length (all p > .05). This suggests that both groups produce texts of comparable size and global lexical variety.

However, significant differences emerged in three readability-related measures. Texts in the PDP group exhibited a lower average number of syllables per word (p = .0002), higher Flesch readability scores (p = .0475) and lexical diversity (TTR; p = .0268), indicating simpler lexical choices and greater overall readability.

These results suggest that while lexical diversity remains stable across groups, texts associated with PDP tend to use shorter words and display lower lexical complexity. This pattern may reflect a more direct, less elaborated linguistic style.

Syntactic. The distribution of part-of-speech (PoS) categories revealed robust differences between PDP and non-PDP texts. Texts produced by individuals classified as PDP showed a significantly higher proportion of verbs (VERB) ($p < .001$) and adverbs (ADV) ($p < .001$), suggesting a more action- and modification-oriented discourse profile. In contrast, PDP texts contained significantly fewer nouns (NOUN) ($p < .001$), adjectives (ADJ; $p = .005$), prepositions (ADP; $p = .0002$), proper nouns (PROPN; $p = .0001$), and subordinating conjunction (**SCONJ**; $p < .001$). These results indicate a structural shift from nominal and descriptive constructions toward verbal and adverbial constructions in PDP discourse. No significant differences were observed for determiners, auxiliaries, coordinating conjunctions, punctuation, numerals, or particles. Taken together, these findings suggest a reduction in nominal density and lexical specification in PDP texts, accompanied by increased predicate and modifier use.

PDP texts exhibited significantly higher proportions of adverbial modifiers (advmod) ($p < .001$), direct objects (obj) ($p = .0044$), subordinating markers (mark) ($p = .0007$), and open clausal complements (xcomp) ($p = .0422$). These increases suggest greater syntactic embedding and modification. Conversely, PDP texts showed significantly lower proportions of adjectival modifiers (amod) ($p < .001$), prepositional markers (case) ($p = .0001$), nominal modifiers (nmod) ($p < .001$), and expletives (expl) ($p = .0277$). The reduction in nominal modification and compounding further supports the interpretation of diminished nominal complexity in PDP discourse. Overall, the syntactic profile of PDP texts appears characterized by increased verbal projection and clausal embedding alongside reduced nominal modification.

PDP texts showed a significantly higher proportion of first-person singular verbs (1_Sing VERB) ($p < .001$), accompanied by significantly lower proportions of third-person singular (3_Sing VERB) ($p < .001$) and plural (3_Plur VERB) verbs ($p = .0002$). This pattern indicates a marked shift toward self-referential verbal constructions. Additionally, PDP texts contained significantly fewer explicit subjects ($p = .0292$). This suggests that increased self-focus may be structurally encoded through null subjects rather than overt pronouns.

Psycholinguistc. At the diagnostic level, several psycholinguistic dimensions significantly distinguished PDP from non-PDP posts. *Anger*-related lexicon was also elevated in PDP ($p = 0.033$), with means of 0.0172 (PDP) vs. 0.0133 (non-PDP). This suggests that irritability or anger-related expressions contribute to the linguistic differentiation of PDP. *Death*-related words strongly distinguished groups ($p < 0.001$). The PDP group (0.0146) used nearly twice as many death-related terms as the non-PDP group (0.0065), indicating marked lexical salience. *Affect*-related words were significantly elevated in PDP ($p = 0.0035$; 0.1036 vs 0.0959), reflecting a higher overall emotional lexical load. This pattern indicates that PDP narratives are characterized a generalized increase in affective language. PDP individuals produced significantly more *negative emotion* words ($p < 0.001$). Mean values were higher in the PDP group (0.0392) than in the non-PDP group (0.0327), indicating increased negative affective expression. *Certainty* words were more frequent in PDP ($p = 0.0043$; 0.0359 vs 0.0306). This may reflect more categorical or definitive cognitive positioning in PDP narratives. Indicating a small impact, *Absolutist* language was significantly higher in PDP ($p < 0.001$; 0.0084 vs. 0.0051), supporting previous literature linking cognitive rigidity and depressive symptomatology. *Negation* showed a robust difference ($p = 0.0001$). PDP individuals used more negation markers (0.0077) than non-PDP ones (0.0046), consistent with increased negative cognitive framing. *Sadness*-related words were also significantly more frequent in PDP ($p = 0.0055$; 0.0129 vs 0.0101). However, this category exhibited the smallest absolute difference between groups, suggesting that while sadness is statistically associated with PDP, it may represent a more pervasive baseline emotional marker rather than a strongly discriminative linguistic feature. Notably, *positive emotion* vocabulary did not reach statistical significance at the group level.

4.2 Symptom (RQ2)

In this section, we present the analysis to answer RQ2 on the symptoms of depression. Table 6 provides an overview of how the various linguistic features differ across each symptom. A systematic comparison of the differences between linguistic features can be found in Appendix A

Lexical. Our analysis of lexical variation across symptom annotations revealed significant differences across symptoms for token count, type count, TTR, and syllables per word. We observed that texts annotated with *Suicide/Self-extermination* (16) and *Physical symptom* (5) differed significantly from other symptoms in terms of the number of tokens and types and their ratio (TTR). Notably, *Physical symptom* (5) texts tended to be shorter and less lexically diverse, whereas *Suicide/Self-*

Symp	Lexical	Syntactic	Psycholinguistic
1	syllables ↑ (1)	–	–
2	–	1 Sing ↑ (1); 3 Sing ↓ (1)	–
3	syllables ↓ (1)	–	anxiety ↓ (1)
4	tokens ↑ (1); types ↑ (1); ttr ↓ (1)	NOUN ↑ (1)	anxiety ↓ (1)
5	tokens ↑ (6); types ↑ (6); ttr ↓ (2)	NOUN ↑ (1)	anxiety ↑ (2)
6	ttr ↓ (2); tokens ↑ (1)	NOUN ↓ (7); PRON ↑ (2); ADV ↑ (1)	positive ↑ (1); anxiety ↓ (1); death ↓ (1)
7	–	ADV ↑ (1); advmod ↑ (1)	–
8	tokens ↓ (1); types ↓ (1)	NOUN ↑ (1); 1 Sing ↑ (1); 3 Sing ↓ (1)	sadness ↑ (3); anxiety ↓ (1)
9	–	ADV ↑ (1); NOUN ↓ (1); PROPN ↓ (1)	anxiety ↓ (1); death ↓ (1)
10	tokens ↑ (2); types ↑ (2); ttr ↓ (1)	1 Sing ↓ (12); ADV ↓ (10); advmod ↓ (10)	positive ↓ (3); anxiety ↓ (2); sadness ↓ (2)
11	tokens ↓ (1); types ↓ (1)	PRON ↓ (4); NOUN ↑ (2); ADV ↑ (1)	anxiety ↓ (1); death ↓ (1)
12	tokens ↑ (1); types ↑ (1)	ADV ↑ (1); PROPN ↓ (1); advmod ↑ (1)	sadness ↑ (3); anxiety ↓ (1); death ↓ (1)
13	tokens ↑ (1); types ↑ (1)	advmod ↑ (2); ADV ↑ (1); NOUN ↓ (1)	anxiety ↓ (1); death ↓ (1)
14	tokens ↓ (1); types ↓ (1)	NOUN ↑ (6); PRON ↓ (2); ADP ↑ (1)	anxiety ↑ (13); sadness ↓ (2); death ↓ (1)
15	tokens ↑ (1); types ↑ (1)	NOUN ↓ (4); PRON ↑ (3); case ↓ (2)	affect total ↑ (1); positive ↑ (1); anxiety ↓ (1)
16	tokens ↓ (8); types ↓ (7); ttr ↑ (4)	ADV ↑ (1); NOUN ↓ (1); PRON ↑ (1)	death ↑ (10); anxiety ↓ (2); sadness ↓ (2)
17	tokens ↑ (1); tokens ↓ (1); types ↑ (1)	NOUN ↓ (2); ADV ↑ (1); PRON ↑ (1)	positive ↑ (1); anxiety ↓ (1); death ↓ (1)
18	tokens ↓ (2); types ↓ (2); ttr ↑ (2)	NOUN ↑ (2); ADV ↑ (1); advmod ↑ (1)	anxiety ↓ (1); death ↓ (1)

Table 6: Top 3 dominant feature-direction combinations per symptom. The numbers in brackets indicate the number of symptoms that are statistically different.

extermination (16) texts were comparatively longer and less lexically concentrated. The difference in the number of syllables appears only between *Attention/memory deficit* (2) and *Alteration in weight/eating habits* (3), with 3 tending to have words with slightly more syllables (1.77 vs 1.73).

Syntactic. The symptom *Irritation/Aggressiveness* (10) systematically has less adverbial usage (PoS ADV: and DEP advmod) compared with other symptoms. This might suggest that a trigger for identifying symptom 10 is the verb modification. We also observe a difference in the proportion of noun usage in some symptoms. The *Difficulty in decision-making* (6) showed the highest proportion of nouns, while *Worry/Fear/Anxiety* (14) and *Irritation/Aggressiveness* (10) had the smallest proportion. These findings indicate that symptom clusters differ in the degree of nominal elaboration and referential density. Proper noun usage differed significantly across symptoms ($p = .0001$), with repeated contrasts involving *Irritation/Aggressiveness* (10). This may reflect differences in narrative framing. This might indicate a tendency to indicate references to named entities, people or institutions targeted.

Looking in more detail, targeting the pronoun usage, we also observe a reduction in the use of *first-person singular* in the *Irritation/Aggressiveness* (10) symptom.

Psycholinguistc. At the symptom level, significant differences emerged for several affective and cognitive dimensions.

A global difference in *affect* was observed across symptoms ($p = 0.0062$). Pairwise comparisons indicated that *Irritation/Aggressiveness* (10) differed significantly from *Feeling of worthlessness/Low self-esteem* (15). Inspection of group means suggests that 10 presented one of the highest affective loads, whereas 15 showed comparatively lower values. This indicates heterogeneity in overall affective expression across symptom categories.

Positive emotion varied significantly by symptom. Differences primarily involved *Irritation/Aggressiveness* (10), which showed higher positive emotion levels. Overall, decision-related narratives exhibited the highest positive emotion proportion, whereas irritability-related narratives showed among the lowest.

Anxiety showed the strongest symptom-level effect ($p < 0.0001$). As expected, *Worry/Fear/Anxiety* (14) presented the highest mean anxiety proportion (0.0168), significantly exceeding multiple other symptom categories. This confirms that anxiety-related symptomatology is linguistically reflected in increased anxiety lexicon. To illustrate that Figure 1 show the word cloud for the text spans annotated with *Worry/Fear/Anxiety* (211 instances), where we can see the highlighted words *medo* (fear), with 76 occurrences, and *ansiedade* (anxiety), with 74 occurrences.

Sadness also differed across symptoms ($p = 0.0005$). Smallest sadness proportions were observed in *Tiredness/Discouragement/Fatigue*, 12, (0.016). This suggests that symptoms related to fatigue and lack of enjoyment are more strongly associated with the explicit lexicon of sadness.

Death-related words showed the clearest differentiation ($p < 0.001$). *Suicide/Self-extermination* (16) exhibited substantially higher death lexicon proportions (0.0214) than all other symptom cate-

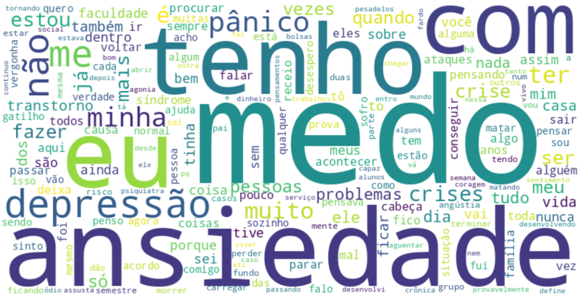

Figure 1: Word cloud for the text spans annotated with *Worry/Fear/Anxiety*

gories. This pattern is theoretically coherent and supports the construct validity of the symptom annotation.

4.3 Symptoms and PDP (RQ1 & RQ2)

Now that we identified the link between the linguistic marker related to the PDP and the symptoms, we explore a possible interplay between RQ1 and RQ2. In other words, can the linguistic characteristics of symptoms differ between PDP and non-PDP?

Lexical. Overall, lexical complexity differences between PDP and non-PDP are not uniform across symptoms. However, symptoms such as *suicide, despair, and low self-esteem* show a consistent pattern of structural compression and increased readability in PDP contexts. This suggests that depressive profile status does not merely increase symptom frequency but is associated with subtle shifts in linguistic organization.

Syntactic. The analysis revealed a robust and structured pattern of syntactic differentiation across several grammatical domains. Regarding POS distribution, significant interaction effects were observed.[8] Crucially, these effects were not uniform but driven by specific symptom–PDP combinations. For instance, *verb* usage differed significantly between *Suicide/Self-extermination-PDP* and *Helplessness/Social harm/Loneliness-non-PDP* ($p = 0.0239$), as well as between *Feeling of worthlessness/Low self-esteem-PDP* and *Helplessness/Social harm/Loneliness-non-PDP* ($p = 0.0208$). *Adverbial* usage showed a consistent pattern: *Irritation/Aggressiveness-non-PDP* differed from multiple PDP-marked symptom groups (e.g., *Sadness/Depressed mood*, *Despair* and *Feeling of guilt* in PDPs; $p \leq 0.0389$), suggesting that adverbial modulation may differentiate emotionally reactive non-PDP irritation from more internally focused symptom constellations. *Noun* and *preposition* distributions further distinguished PDP-marked low self-worth (*Feeling of worthlessness/Low self-esteem-PDP*) from anxiety-related non-PDP cases (*Worry/Fear/Anxiety-non-PDP*), indicating structural shifts in nominal framing and relational marking.

Dependency-level analyses reinforced these findings.[9] The *adverbial modifier* pattern mirrored the POS-level *adverb* effect, with *Irritation/Aggressiveness-non-PDP* contrasting systematically with multiple PDP symptom groups. Additionally, *case* and *mark* dependencies showed differences primarily between *Feeling of worthlessness/Low self-esteem-PDP* and *Worry/Fear/Anxiety-non-PDP*, suggesting alterations in clausal embedding and argument structure associated with depressive self-referential symptoms. These results indicate that depressive profiles are not characterized solely by lexical shifts but by reorganizations in clause architecture and modifier integration.

Verb-person morphology exhibited one of the strongest interaction effects. *First-person singular* verbs showed significant differences ($p < 0.0001$), with *Irritation/Aggressiveness-non-PDP* differing from nearly all PDP-marked symptom groups.

Similarly, *third-person singular* verbs ($p < 0.0001$) distinguished fatigue- and apathy-related PDP groups from several non-PDP symptom categories. These patterns suggest that symptom configurations interact with PDP status in shaping perspectival anchoring: PDP contexts tend to intensify first-person singular marking, while non-PDP irritation shows comparatively distinct distributional behavior. Pronoun-level interactions were more selective. *Third-person singular pronouns* ($p = 0.0006$) differentiated several PDP-marked symptoms (e.g., *Tiredness/Discouragement/Fatigue-PDP*, *Suicide/Self-extermination-PDP*) from *Sadness/Depressed mood-non-PDP*. *First-person singular pronouns* ($p = 0.0005$) also exhibited interaction effects, reinforcing the verb-based

[8] We found significant interaction effects for verbs ($p = 0.001$), adverbs ($p = 0.001$), nouns ($p < 0.0001$), prepositions ($p = 0.0083$), subordinating conjunction ($p = 0.0326$), pronouns ($p = 0.0011$), and proper nouns ($p < 0.0001$).

[9] Significant interaction effects emerged for *adverbial modifier* ($p = 0.0001$), *case* ($p = 0.0027$), *nominal dependents* ($p = 0.0231$), *mark* ($p = 0.0118$), *auxiliary* ($p = 0.0124$), and *indirect object* ($p = 0.0136$).

self-referential pattern. Oblique pronouns showed no significant interaction effects, indicating that argument-level object marking remains relatively stable across symptom–PDP combinations.

Overall, the interaction results demonstrate that depressive linguistic profiles are not reducible to either symptom type or PDP status alone. Rather, specific symptom configurations acquire distinct syntactic signatures when co-occurring with PDP. The most consistent interaction markers involve (i) verbal density and person marking, (ii) adverbial modification, and (iii) nominal and relational structuring. These findings support the hypothesis that PDP amplifies or reshapes the grammatical realization of certain symptom, particularly self-reference, cognitive evaluation and affective modulation.

Psycholinguistc. The clearest descriptive amplification of PDP effects occurs within *Suicide/Self-extermination* (16). PDP participants in this category show markedly higher *death* lexicon proportion (0.0223) compared to non-PDP individuals with the same symptom (0.0100). Similar, though smaller, PDP-related increases appear in *Feeling of worthlessness/Low self-esteem* (15) and *Sadness/Depressed mood* (18).

Within *Feeling of guilt* (9), PDP individuals show higher *negative emotion* (0.0477) compared to non-PDP (0.0346). A similar pattern appears in *Sadness/Depressed mood* (18) and *Suicide/Self-extermination* (16), suggesting that PDP intensifies negative affective expression.

PDP participants generally show higher *negation* and *absolutist* usage within most symptom categories (e.g., *Feeling of worthlessness/Low self-esteem* (15), *Loss/Diminishment of pleasure/libido* (8)), suggesting that diagnostic status amplifies cognitively rigid and negatively framed language beyond symptom effects alone.

4.4 Effects of symptoms on PDP (RQ3)

All chains converged satisfactorily ($\hat{R} \approx 1.00$ for all parameters). The subject-level variance parameter ($\sigma_{\text{subject}} = 0.339$) indicates moderate heterogeneity across individuals.

Table 7 summarizes posterior estimates for symptom coefficients (β), their 94% Highest Density Intervals (HDI), and the corresponding odds ratios (OR = e^{β}).

Suicidal ideation (16) exhibited the strongest association with PDP ($\beta = 1.516$), corresponding to an OR of approximately 4.55. This indicates that a one-standard-deviation increase in this symptom multiplies the odds of PDP by four. *Despair* (13) also showed a strong positive effect (OR ≈ 2.30), doubling the odds of PDP. Core affective symptoms such as sadness (18), despair (15) and functional impairment (11) nearly doubled the odds as well.

Irritation/Aggressiveness (10) showed a negative association ($\beta = -0.511$), with an OR of 0.60. This suggests that, controlling for other symptoms, higher levels of irritability were associated with lower odds of PDP in this dataset.

Symptom ID	β	94% HDI	OR
16	1.516	[1.064; 2.021]	4.55
13	0.835	[0.385; 1.313]	2.30
18	0.762	[0.466; 1.037]	2.14
15	0.579	[0.185; 0.959]	1.78
11	0.568	[0.248; 0.908]	1.76
8	0.439	[0.104; 0.805]	1.55
17	0.437	[0.154; 0.719]	1.55
7	0.435	[-0.037; 0.953]	1.54
1	0.375	[-0.011; 0.786]	1.45
2	0.333	[-0.036; 0.755]	1.40
12	0.171	[-0.156; 0.509]	1.19
5	0.106	[-0.132; 0.374]	1.11
9	0.098	[-0.188; 0.407]	1.10
14	-0.015	[-0.259; 0.225]	0.99
4	-0.111	[-0.463; 0.225]	0.89
3	-0.115	[-0.406; 0.142]	0.89
6	-0.210	[-0.497; 0.079]	0.81
10	-0.511	[-0.772; -0.228]	0.60

Table 7: Posterior estimates for symptom effects on PPD.

5 Conclusions

This paper focused on characterising the language of possible depressive profile narratives, looking at lexical, syntactic and psycholinguistic clues in texts and in the symptoms in them. For RQ1 and RQ2 the results confirm that there are relevant signals, but more investigation is needed for their links with PDP and their interplay (§ 4.1 and 4.2). For RQ3, we found symptoms closely associated with PDP (§ 4.4). These findings shed light on how symptoms of depression appear on social media, opening up new avenues in a range of fields, such as clinical practice and the development of more accurate computational tools, including LLMs that can warn of potential cases where immediate intervention is needed. Future work includes extending the characterisation to include analyses of the incidence of figurative language and also replicate this analysis on English language since AIM-Health project is an international research cooperation between UK and Brazil.

Ethical Considerations

The data for this research comprises anonymous text posts automatically collected from a public social media platform. These posts are part of a corpus made available to researchers upon request at no cost. No information that could identify the authors of the posts is present in the texts, thereby protecting their privacy and minimizing potential harm. Researchers were informed that the posts' content may include sensitive material related to mental health. All analyses were conducted using established, widely recognized methods.

Limitations

This analysis reflects the contents of the corpus and any biases derived from the corpus construction regarding the social network (Facebook) selection, the keywords and protocols applied. The annotation may also be biased, even though it was carried out by psychology, psychiatry and occupational therapy students after training with experienced mental health professionals and with their review. However, it is important to emphasize that for the purpose of the work, a false positive is better than a false negative; that is, it is better to find an indication of depression that is not exactly depression and invite the person to talk than to miss an indication of depression in someone who really should be invited for a conversation. Another limitation of this work is the small corpus size. However, although small, the collection contains rich data that should be considered representative of the problem: depression among university students.

Acknowledgments

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Process Number #2024/10233-7 (AIM-Health project). This project is also financed by UKRI/MRC – UK Research and Innovation / Medical Research Council. We thank the Amive project team for giving us access to the corpus. The authors thank the support received from the World University Network (WUN) RDF.

References

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022.

Margaret M Bradley and Peter J Lang. 1999. Affective norms for english words (anew): Instruction manual and affective ratings. Technical report, Technical report C-1, The center for research in psychophysiology, University of Florida.

Vinicius Casani, Alinne C. Correa Souza, Rafael G. Mantovani, and Francisco Carlos M. Souza. 2021. DP-symptom-identifier: uma estratégia para classificar sintomas de depressão utilizando um conjunto de dados textuais na língua portuguesa. In *Annals of the XIII Brazilian Symposium of Information Tecnology and Human Language (STIL 2021)*. SBC.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media*, volume 7, pages 128–137.

Antonia Errazuriza, Dalia Avello-Vegab, Juan P. Ramirez-Mahalufa, Rafael Torresa, Nicolas A. Crossleya, Eduardo A. Undurragac, and Peter B. Jones. 2023. Prevalence of depressive disorder in the adult population of latin america: a systematic review and meta-analysis. *The Lancet Regional Health – Americas*, 26(100587).

Pedro P. Balage Filho, Thiago Alexandre Salgueiro Pardo, and Sandra M. Aluísio. 2013. An evaluation of the Brazilian Portuguese LIWC dictionary for sentiment analysis. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*.

Global Burden of Disease Collaborative Network. 2025. *Global Burden of Disease Study 2023 (GBD 2023) Covariates 1980-2023*. Institute for Health Metrics and Evaluation (IHME), Seattle, United States of America.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Shaoxiong Ji, Celina Ping Yu, Sai-fu Fung, Shirui Pan, Guodong Long, and Gao Cong. 2018. Supervised learning for suicidal ideation detection in online user content. *Complex.*, 2018.

Paulo Mann, Aline Paes, and Elton H. Matsushima. 2020. See and read: Detecting depression symptoms in higher education students using multimodal social media data. *Proceedings of the International AAAI Conference on Web and Social Media*, 14:440–451.

Paulo Mann, Matheus Yasuo Ribeiro Utino, Elton Hiroshi Matsushima, and Aline Paes. 2025. The topics of depression on social networking sites. *Journal of the Brazilian Computer Society*, 31(1):771–806.

Teresa BF Martins, Claudete M Ghiraldelo, Maria das Graças Volpe Nunes, and Osvaldo Novais de Oliveira Junior. 1996. Readability formulas applied to textbooks in brazilian portuguese.

Augusto R. Mendes and Helena Caseli. 2024. Identifying fine-grained depression signs in social media posts. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 8594–8604, Torino, Italia. ELRA and ICCL.

Priya Nambisan, Zhihui Luo, Akshat Kapoor, Timothy B. Patrick, and Ron A. Cisler. 2015. Social media, big data, and public health informatics: Ruminating behavior of depression revealed through twitter. In *2015 48th Hawaii International Conference on System Sciences*, pages 2906–2913.

J.W. Pennebaker, R.L. Boyd, K. Jordan, and K. Blackburn. 2015. The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin.

Wesley Ramos dos Santos, Rafael Lage de Oliveira, and Ivandré Paraboni. 2023. SetembroBR: a social media corpus for depression and anxiety disorder prediction. *Language Resources and Evaluation*, pages 1–28.

Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.

Raluca Nicoleta Trifu, Bogdan Nemeș, Dana Cristina Herta, Carolina Bodea-Hategan, Dorina Anca Talaș, and Horia Coman. 2024. Linguistic markers for major depressive disorder: a cross-sectional study using an automated procedure. *Frontiers in Psychology*, Volume 15 - 2024.

World Health Organization. 2021. Comprehensive mental health action plan 2013–2030. Technical report, World Health Organization.

Shweta Yadav, Jainish Chauhan, Joy Prakash Sain, Krishnaprasad Thirunarayan, Amit Sheth, and Jeremiah Schumm. 2020. Identifying depressive symptoms from tweets: Figurative language enabled multitask learning framework. *arXiv preprint arXiv:2011.06149*.

Amir Hossein Yazdavar, Hussein S Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-supervised approach to monitoring clinical depressive symptoms in social media. In *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, pages 1191–1198.

A Symptom-Level Pairwise Linguistic Contrasts

A.1 Lexical contrasts

A.2 Syntactic contrasts

A.3 Psycholinguistic contrasts

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	–		syllables↑															
3	syllables↓		–															
4				–												tk↑, ttr↓, ty↑		
5					–			tk↑, ty↑			tk↑, ty↑			tk↑, ty↑		tk↑, ttr↓, ty↑	tk↑, ty↑	tk↑, ttr↓, ty↑
6						–										tk↑, ttr↓		ttr↓
8					tk↓, ty↓			–										
10										–						tk↑, ttr↓, ty↑		tk↑, ty↑
11					tk↓, ty↓						–							
12												–				tk↑, ty↑		
13													–			tk↑, ty↑		
14					tk↓, ty↓									–				
15															–	tk↑, ty↑		
16				tk↓, ttr↑, ty↓	tk↓, ttr↑, ty↓	tk↓, ttr↑				tk↓, ttr↑, ty↓		tk↓, ty↓	tk↓, ty↓		tk↓, ty↓	–	tk↓, ty↓	
17					tk↓, ty↓											tk↑, ty↑	–	
18					tk↓, ttr↑, ty↓	ttr↑				tk↓, ty↓								–

Table 8: Lexical symptom contrasts. Cells indicate significant differences; arrows denote direction (row relative to column).

	2	4	5	6	7	8	9	10
2	–							1 Sing↑, 3 Sing↓
4		–		NOUN↑				
5			–	NOUN↑				
6		NOUN↓	NOUN↓	–		NOUN↓		1 Sing↑, ADV↑, NOUN↓, advmod↑
7					–			ADV↑, advmod↑
8				NOUN↑		–		1 Sing↑, 3 Sing↓
9							–	1 Sing↑, 3 Sing↓, ADV↑, PROPN↓, advmod↑
10	1 Sing↓, 3 Sing↑			1 Sing↓, ADV↓, NOUN↑, advmod↓	ADV↓, advmod↓	1 Sing↓, 3 Sing↑	1 Sing↓, 3 Sing↑, ADV↓, PROPN↑, advmod↓	–
11				NOUN↑, PRON↓				1 Sing↑, 3 Sing↓, ADV↑, advmod↑
12								1 Sing↑, ADV↑, PROPN↓, advmod↑
13								1 Sing↑, 3 Sing↓, ADV↑, PROPN↓, advmod↑
14				NOUN↑, PRON↓			NOUN↑	1 Sing↑
15								1 Sing↑, 3 Sing↓, ADV↑, NOUN↓, PRON↑, PROPN↓, advmod↑, case↓
16								1 Sing↑, 3 Sing↓, ADV↑, PROPN↓, advmod↑
17								1 Sing↑, 3 Sing↓, ADV↑, NOUN↓, PROPN↓, advmod↑
18				NOUN↑				1 Sing↑, ADV↑, advmod↑

Table 9: Syntactic symptom contrasts (from symptoms 1 to 10). Cells indicate significant differences; arrows denote direction (row relative to column).

	11	12	13	14	15	16	17	18
2								
4								
5								
6	NOUN↓, PRON↑			NOUN↓, PRON↑				NOUN↓
7								
8								
9				NOUN↓				
10	1 Sing↓, 3 Sing↑, ADV↓, advmod↓	1 Sing↓, ADV↓, PROPN↑, advmod↓	1 Sing↓, 3 Sing↑, ADV↓, PROPN↑, advmod↓	1 Sing↓	1 Sing↓, 3 Sing↑, ADV↓, NOUN↑, PRON↓, PROPN↑, advmod↓, case↑	1 Sing↓, 3 Sing↑, ADV↓, PROPN↑, advmod↓	1 Sing↓, 3 Sing↑, ADV↓, NOUN↑, PROPN↑, advmod↓	1 Sing↓, ADV↓, advmod↓
11	–				NOUN↑, PRON↓	PRON↓	PRON↓	
12		–						
13			–	NOUN↓, advmod↑				
14			NOUN↑, advmod↓	–	ADP↑, NOUN↑, PRON↓, case↑	NOUN↑	NOUN↑	
15	NOUN↓, PRON↑			ADP↓, NOUN↓, PRON↑, case↓	–			NOUN↓
16	PRON↑			NOUN↓		–		
17	PRON↑			NOUN↓			–	
18					NOUN↑			–

Table 10: Syntactic symptom contrasts (from symptoms 11 to 18). Cells indicate significant differences; arrows denote direction (row relative to column).

	3	4	5	6	8	9	10	11	12	13	14	15	16	17	18
3	–										anxiety↓				
4		–									anxiety↓				
5			–				anxiety↑						anxiety↑		
6				–			positive↑				anxiety↓		death↓		
8					–		sadness↑				anxiety↓, sadness↑		sadness↑		
9						–					anxiety↓		death↓		
10			anxiety↓	positive↓	sadness↓		–		sadness↓		anxiety↓	affect↓, positive↓	death↓	positive↓	
11								–			anxiety↓		death↓		
12							sadness↑		–		anxiety↓, sadness↑		death↓, sadness↑		
13										–	anxiety↓		death↓		
14	anxiety↑	anxiety↑		anxiety↑	anxiety↑, sadness↓	anxiety↑	anxiety↑	anxiety↑	anxiety↑, sadness↓	anxiety↑	–	anxiety↑	anxiety↑, death↓	anxiety↑	anxiety↑
15							affect↑, positive↑				anxiety↓	–	death↓		
16			anxiety↓	death↑	sadness↓	death↑	death↑	death↑	death↑, sadness↓	death↑	anxiety↓, death↑	death↑	–	death↑	death↑
17							positive↑				anxiety↓		death↓	–	
18											anxiety↓		death↓		–

Table 11: Psycholinguistic symptom contrasts. Cells indicate significant differences; arrows denote direction (row relative to column).