Linguistic markers of schizophrenia: a case study of Robert Walser

Ivan Nenchev [1,2], Tatjana Scheffler [3], Marie de la Fuente [4], Heiner Stuke [1], Benjamin Wilck [5], Sandra Anna Just [1], Christiane Montag [1]

[1] Department of Psychiatry and Psychotherapy, Charité Campus Mitte, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health,

[2] Berlin Institute of Health at Charité – Universitätsmedizin Berlin,

[3] German Studies, Ruhr-Universität Bochum,

[4] Department of Linguistics, University of Potsdam,

[5] The Hebrew University of Jerusalem, Martin Buber Society of Fellows in the Humanities and Social Sciences

ivan.nenchev@charite.de

Abstract

We present a study of the linguistic output of the German-speaking writer Robert Walser using NLP. We curated a corpus comprising texts written by Walser during periods of sound health, and writings from the year before his hospitalization, and writings from the first year of his stay in a psychiatric clinic, all likely attributed to schizophrenia. Within this corpus, we identified and analyzed a total of 20 linguistic markers encompassing established metrics for lexical diversity, semantic similarity, and syntactic complexity. Additionally, we explored lesser-known markers such as lexical innovation, concreteness, and imageability. Notably, we introduced two additional markers for phonological similarity for the first time within this context. Our findings reveal significant temporal dynamics in these markers closely associated with Walser's contemporaneous diagnosis of schizophrenia. Furthermore, we investigated the relationship between these markers, leveraging them for classification of the schizophrenic episode.

1 Introduction

Schizophrenia is a heterogeneous psychiatric disorder characterized by diverse symptoms impacting a person's perception, cognition, language and motor functions. The disorder displays variable courses; some patients undergo circumscribed episodes with psychotic symptoms and either complete or incomplete remission, while others follow a chronic course with persistent symptoms at a relatively stable level. Typically, there is a prodromal period, ranging from several weeks to several years, that precedes the first psychotic episode. Symptoms in schizophrenia can be broadly categorized into two groups: positive symptoms, such as hallucinations, delusions, and certain formal thought disorders like derailment and word salad; and negative symptoms, including poverty of speech, alogia, anhedonia, and social withdrawal (Andreasen, 1990). In clinical psychiatry, the diagnosis of schizophrenia is established based on interviews and diagnostic manuals that provide comprehensive descriptions of symptoms. Despite advances in modern medicine, the absence of concrete (bio)markers for diagnosis and individualized treatment for schizophrenia persists. Concurrently, patients exhibit fluctuating alterations in language production and comprehension that correlate with the illness's dynamics and severity. Despite the long history of study of language in connection to mental illness (the first accounts on recognizable linguistic patterns are indeed older than the term "schizophrenia" and even its predecessor "dementia praecox" (Griesinger, 1845; Brosius, 1857), only the recent advances of computational linguistics and NLP provide the necessary tools and technology to analyze substantial linguistic datasets and extract linguistic features in an objective and replicable manner (Hitczenko et al., 2021; Crema et al., 2022). Several authors identify the potential of linguistic features extracted by means of NLP as possible biomarkers of psychosis (Corcoran et al., 2020; de Boer et al., 2020). Palaniyappan (2021) emphasizes that linguistic production not only reflects biological processes but also incorporates social aspects. Consequently, he contends that language can be regarded as a biosocial marker. Our stance is that schizophrenia (and mental illness in general) should not be oversimplified to mere biology. We view the concept of biomarkers as an analogy, comparable to blood sugar levels or blood pressure in somatic medicine. Notably, a significant gap exists in the current state of research, with no established population norms for the NLP features. This contrasts sharply with

Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2024), pages 41–60

the majority of biomarkers in traditional medicine. Addressing this gap could involve adopting a personalized methodology, observing changes in linguistic output over time in individual subjects.

Nowadays, NLP methodologies are commonly used to demonstrate significant differences in the linguistic production between participants suffering from schizophrenia or high-risk individuals compared with neurotypical participants. Furthermore, it seems that computational linguistic features at least partially correlate with the severity of psychotic symptoms. A promising recent trend in the field is to produce and analyse longitudinal datasets and explore the stability or the dynamics of the linguistics markers within schizophrenia. As schizophrenia is characterized by a highly heterogenous expression of symptoms, it is not surprising that there is evidence of a substantial difference in the linguistic markers within the schizophrenia group (Liebenthal et al., 2023). Given the advancements in individualized medicine, an important question arises: Can variations in linguistic markers be identified within each individual's language output, enabling a truly personalized method for diagnosing relapses and monitoring the disease? If so, NLP could furnish tools for tailoring individualized detection algorithms, thereby aiding in the prevention of future psychotic episodes.

In this study, we adopt such a longitudinal approach using a single-subject design. Our study is based on a linguistic corpus comprising short literary texts authored by the German-speaking writer Robert Walser who probably suffered from schizophrenia. From this corpus, we extracted established linguistic features for lexical diversity and lexical innovation, syntactic complexity, and semantic similarity. Additionally, we introduce markers for phonological similarity. We present evidence of significant marker dynamics temporally linked to Walser's diagnosis of schizophrenia and subsequent hospitalization.

In this work, we contribute to the intersection of NLP and psychiatry through multiple avenues. Primarily, we explore various NLP findings at the individual level. Furthermore, we introduce a novel linguistic marker, phonological similarity, warranting future investigation. Additionally, our utilization of a corpus from a German-speaking writer enhances the linguistic diversity within this domain.

2 Background

2.1 Linguistic markers of schizophrenia

One extensively studied linguistic aspect in psychotic language through NLP is semantic coherence, addressing the relatedness between word chunks or sentences, aiming to capture formal thought disorders (disorganisation, tangentiality, derailment and poverty of speech) in schizophrenia. In coherence analysis, words and sentences are commonly represented as vector embeddings in a multidimensional semantic space, with relatedness gauged via cosine similarity between these vectors. Currently, there is no consensus on a best practice approach regarding segmentation. Some studies focus on the semantic similarity between chunks of 5 or 10 tokens (called coherence-5 and coherence-10), other studies examine the similarity between sentences (first- and second-order coherence, measuring semantic similarity with a sentence's first or second neighbour(Parola et al., 2023)). Furthermore, there is also no consensus on the preferred type of embeddings. Studies reveal substantial differences in the semantic coherence when comparing patients with schizophrenia to neurotypical controls. The majority of studies suggest reduced semantic coherence in schizophrenia patients, notably derived from analyses based on word2vec, GloVe, and fastText embeddings (Corona-Hernández et al., 2023; Voleti et al., 2023; Iter et al., 2018; Morgan et al., 2021; Voppel et al., 2021; Just et al., 2020; Parola et al., 2023). The findings have encountered challenges. Alonso-Sánchez et al. (2022) revealed an increase in semantic coherence among a cohort experiencing the first episode of psychosis, countering previous assumptions. Moreover, Tang et al. (2021), employing BERT embeddings, yielded inconclusive coherence outcomes in inter-group comparisons. Intriguingly, only second-order coherence demonstrates potential for cross-language generalization (Parola et al., 2023).

Another frequently employed set of linguistic markers revolves around gauging the lexical diversity within language samples. It reflects the variety and richness of vocabulary within a text. The Type-Token Ratio (TTR) specifically quantifies the ratio of unique words to the total words in a text. However, due to its sensitivity to text length variations, various other markers have emerged to address this limitation and offer a more nuanced understanding of lexical richness. Among these markers, the Measure of Textual Lexical Diversity

(MTLD) and Mean Average TTR (MATTR) stand out (McCarthy and Jarvis, 2010). Several studies have investigated the differences in lexical diversity between subjects with schizophrenia and neurotypical subjects. Voleti et al. (2023) report lower lexical diversity, while Ziv et al. (2021) observe the opposite trend. Lundin et al. (2023) and Schneider et al. (2023) report a negative result on MTLD and TTR respectively. Notably, Bambini et al. (2022) utilize TTR for the identification of clusters of individuals with schizophrenia. Additionally, Pavy et al. (1969) report significantly higher TTR for individuals with acute schizophrenia compared to those with a chronic condition. In schizophrenia patients, lexical diversity measures seem to vary based on clinical symptoms. Some individuals demonstrate an increase, while others display a decrease in these metrics.

The words used in a text can be further examined beyond their mere counts. Despite the availability of comprehensive linguistic norms across languages, semantic norms related to concreteness and imageability are seldom applied in analyzing linguistic output from individuals affected by schizophrenia. Concreteness refers to the extent to which a word signifies something tangible, specific, and easily perceivable through the senses; imageability refers to the potential of words to evoke vivid mental images. Oertel et al. (2009) and Sack et al. (2005) have observed that individuals with schizophrenia and their first-degree relatives exhibit heightened vividness of mental imagery, as assessed through a standardized questionnaire. Notably, because of the absence of a correlation between vividness scores and symptoms, the authors interpret this phenomenon as indicative of a trait marker in schizophrenia. The investigation of mental imagery within in the linguistic production of individuals affected by schizophrenia remains unexplored in current research. The adoption of a vocabulary characterized by lower concreteness and imageability, possibly resulting in a more abstract linguistic style, might reflect reminiscences of psychopathological symptoms such as poverty of content of speech and stilted speech. Conversely, a discourse marked by high concreteness and vivid imagery also appears plausible. Minor et al. (2019) report that concreteness is not connected to neurocognitive, socialcognitive or metacognitive deficits in schizophrenia. A more recent study of Minor et al. (2023) examined the test-retest reliability of concreteness using the Coh-Metrix tool and reported a good intraclass-correlation.

Individuals with schizophrenia often exhibit lexical innovations, termed neologisms, encompassing words absent from the general lexicon. Surprisingly, the exploration of such lexical innovations in schizophrenia through NLP methodologies remains limited. To date, Just et al. (2020) stand as the sole instance using semi-automated neologism detection effectively, distinguishing schizophrenia individuals from control groups.

Syntactic complexity addresses the intricacy and sophistication of the grammatical structures in a text. It appears reduced in individuals at high risk for developing schizophrenia (Bedi et al., 2015; Corcoran et al., 2018). Schneider et al. (2023) report a significantly reduced syntactic complexity for individual suffering from schizophrenia compared to both controls and patients with depression. Haas et al. (2020) report a negative correlation between negative symptoms and syntactic complexity in clinically high risk individuals. Silva et al. (2023) analyse various indices of syntactic complexity in individuals with first episode psychosis and report that the majority of indices remain stable over a period of 6 months. Voleti et al. (2023) used Yngve scoring to analyse the syntactic complexity of transcribed interviews from individuals suffering from schizophrenia or bipolar disorder and healthy controls. Although syntactic complexity seems lower in the schizophrenia group this marker was not selected for the development of prediction models by the authors.

A commonly replicated linguistic feature of individuals with schizophrenia is the extensive use of first person singular pronouns as a marker of focus on the self (Ziv et al., 2021; Tang et al., 2021; Birnbaum et al., 2017; Lundin et al., 2023; Fineberg et al., 2015). The prominence of increased first-person singular pronoun use extends beyond schizophrenia and has been observed in diverse mental health conditions (Brockmeyer et al., 2015; Edwards and Holtzman, 2017; Lyons et al., 2018). An increased use of second person singular pronouns (Watson et al., 2012) and decreased use of first person plural pronouns (Lundin et al., 2023) have also been reported in schizophrenia.

This study not only delves into established linguistic markers but also introduces phonological similarity as a promising marker in NLP associated with linguistic output in schizophrenia. While drawing from established methodologies for measuring string similarity, this study stands as the pioneering exploration of this approach within the context of linguistic output in schizophrenia. High phonological similarity may be associated with clanging - a rare symptom observed in some patients with schizophrenia which involves using words based on their sound similarity rather than their meaning, e.g. "I'm trying to make sense out of sense. I'm not making sense [cents] anymore. I have to make dollars." (Andreasen, 1986).

2.2 Longitudinal studies

Currently, only a limited number of studies address the question whether linguistic markers represent stable longitudinal traits or capture dynamic shifts in psychological states. Research by Bedi et al. (2015) demonstrated a decrease in semantic coherence via Latent Semantic Analysis (LSA) preceding psychosis, complemented by a decline in markers of syntactic complexity. Corcoran et al. (2018) corroborated these findings. Birnbaum et al. (2019) analyzed Facebook posts, identifying linguistic alterations preceding psychotic relapses, notably an upsurge in first- and second-person pronouns. Minor et al. (2023) observed satisfactory test-retest reliability in speech content and organization over 6 months to a year using LIWC and Coh-Metrix. Alonso-Sánchez et al. (2022) highlighted an augmented semantic similarity in a picture description task over 6 months, correlating with increased negative symptoms. Silva et al. (2023) examined the syntactic complexity of individuals with first episode psychosis and report that over a period of 6 months the majority of the examined markers remain stable.

2.3 Literature in the study of mental illness

The NLP analysis of authors with mental illness includes studies on Iris Murdoch's reduced lexical diversity due to Alzheimer's dementia (Hirst and Wei Feng, 2012; Le et al., 2011; Garrard et al., 2005; Pakhomov et al., 2011). Edgar Allan Poe's works are scrutinized to illuminate aspects of his enigmatic death (Dean and Boyd, 2020). Additionally, research examines linguistic patterns associated with bipolar disorder (Rentoumi et al., 2017) or suicidality in poems or in diaries (Stirman and Pennebaker, 2001; Fernández-Cabana et al., 2013; Baddeley et al., 2011). To our knowledge, a systematic NLP-based analysis of the literary works authored by an individual diagnosed with schizophrenia is still missing. In the current study, we extract NLP features from a corpus of a single individual and observe significant within-subject variations, which could be associated with psychosis.

3 Methods

3.1 Robert Walser

Robert Walser (1878–1956) is a German-speaking writer from the early 20th century, who played a significant role in European literary modernism. Throughout his lifetime, Walser created an extensive body of work, encompassing several novels, numerous short pieces of prose, and poetry. In 1929, he was institutionalized, remaining in psychiatric care for nearly 27 years until his passing. During this period, he received a diagnosis of schizophrenia. Several detailed accounts shed light on Walser's hospitalization and his prolonged stay in psychiatric clinics (Wernli, 2014; Partl et al., 2011). Upon admission to the Waldau psychiatric clinic, Walser exhibited auditory verbal hallucinations, probably persecutory delusions, anxiety, and suicidal thoughts. Subsequently, in the Herisau asylum from 1933 onwards, he consistently reported experiencing commenting and dialogizing voices, as noted by the attending psychiatrist (Vannette, 2020). Through an extensive presentation of Walser's medical records (Wernli, 2014) and a detailed exploration of his family history, which includes multiple instances of schizophrenia or depression among family members (Gisi, 2018), the possibility of his diagnosis becomes evident. However, posthumously, the diagnosis of schizophrenia has faced challenges from various scholars. Lyons and Fitzgerald (2004) suggest, for instance, that Walser might have been experiencing high-functioning autism instead. Other scholars reject any psychiatric diagnosis altogether and assert that Walser's stay in psychiatric institutions was solely due to socio-economic reasons. This study acknowledges that verifying or refuting Walser's psychiatric diagnosis falls outside its scope, particularly given the impossibility of such an assessment for a person who passed away over 60 years ago. Nevertheless, considering the current efforts to identify NLP markers of schizophrenia, Walser's extensive body of work and medical history presents an intriguing case. His extensive body of work presents a compelling opportunity for NLP research due to the substantial volume of text he generated in the decades and years prior to and just

Period	Texts	Tokens (total)	Tokens (mean)
1903-1907	70	50102	715.74
1915-1919	79	80408	1017.82
1928-1929	40	28688	717.20
Total	189	159198	995.56

Table 1: Linguistic corpus

before his hospitalization, enabling comprehensive comparative analysis.

3.2 Corpus

For our analysis, we assembled a corpus consisting of short prose texts authored by Walser for various periodicals or published in collections. Considering the texts share the same genre and have comparable lengths, we consider them comparable for the purposes of our study. This corpus included all currently available texts (n=189) from three distinct timeframes: 70 texts published between 1903 and 1907, sourced from the volume Fritz Kocher's Aufsätze (Walser, 2023a), kleine Prosa (Walser, 2023b), as well as publications in Neue Rundschau (Walser, 2017a), Schaubühne (Walser, 2015), and Berliner Tageblatt (Walser, 2013a). Another set of 79 texts, spanning 1915 to 1919, originated from Neue Züricher Zeitung (Walser, 2013b), Neue Rundschau (Walser, 2017a), Prosastücke (Walser, 2017b), and Poetenleben (Walser, 2014). Notably, Walser's biography does not attest to mental suffering or illness during these periods. Lastly, 40 texts written in 1928 and 1929, form the third part of the corpus (Walser, 2019, 2013b,a). Walser was admitted to the psychiatric clinic in Waldau in January 1929. The assumption that he had been in a state of psychosis in the months (and likely years) leading up to this admission has been expressed by his biographers (Mächler and Seelig, 1992) as cited in (Vannette, 2020). We decided to exclude both poetry and novels to ensure maximal homogeneity in the corpus, facilitating comparisons across different time periods. Walser's three major novels were composed between 1907 and 1909. The draft of a fourth novel, *The Robber*, was found posthumously and dated back to 1925. A detailed examination of Walser's poetry or letters is reserved for future work.

3.3 Linguistic markers

After constructing the corpus and preprocessing, we extracted a number of linguistic markers.

To gauge semantic coherence, we utilized pretrained word2vec embeddings from the Python library spaCy[1]. After removing the stop words, we computed the cosine similarity between chunks of 5 and 10 tokens and averaged these values to derive one score per text, generating coherence-5 and coherence-10 scores for chunks with 5 and 10 tokens respectively. To assess first- and second-order sentence similarity, we employed a pretrained Hugging Face transformer model[2] using the sentence transformers Python library (Reimers and Gurevych, 2019) we calculated cosine similarity between sentence BERT-embeddings, then averaging these values to produce a single score for each text.

To estimate the lexical diversity of the texts we used the python library lexical richness (Shen et al., 2023) and calculated TTR, MLTD and MATTR for both 25 and 50 token windows.

In addition to this, we calculated the concreteness and imageability values for each text as the mean of the concreteness and imageability norms for the individual words used in the texts. We used the word norms from the newly developed GLEAN dataset for German (Lüdtke and Hugentobler).

We employed a German reference corpus[3] comprising 249 million tokens extracted from texts between 1465 and 1969 to detect neologism. This corpus, spanning Walser's lifetime, was chosen due to its written content. After preprocessing involving stopword removal and lemmatization, we determined the relative frequency of tokens/lemmata in texts that were out of scope of the German reference corpus.

After POS-tagging, we estimated syntactic complexity by calculating the mean Yngve score, sentence length, and number of clauses per sentence in each text. The Yngve Score measures the depth of the parsing tree, with higher scores indicating a more complex syntactic structure (Yngve, 1960). Furthermore, we included the mean frequency of 1st and 2nd person singular and also 1st person plural pronouns per text in our report.

To assess the phonological similarity among

[1] github.com/explosion/spacy-models/releases/tag/de_core_news_lg-3.7.0

[2] huggingface.co/aari1995/German_Semantic_STS_V2

[3] www.dwds.de/r/lexdb/dta/lex

strings, we analyzed the raw un-preprocessed text. Initially, we transformed the graphemes into phonemes using the Python library Epitran (Mortensen et al., 2018). Subsequently, we calculated the Jaro-Winkler similarity between a string and each token from the consecutive 10- and 20-token sequences. We selected these window sizes based on the established capacity of working memory, commonly regarded as 7±2 tokens (Miller, 1956), but acknowledging potential individual differences. Following this, we derived a mean similarity score for each text.

3.4 Analysis

Comparison between periods. The texts from the three distinct periods were compared for differences in linguistic markers via a one-way ANOVA analysis and a subsequent Tukey test through the statsmodels Python library (Seabold and Perktold, 2010). In total we estimated 20 linguistic markers. To mitigate Type I errors, a Bonferroni correction was implemented to adjust for the multitude of linguistic markers assessed ($p < 0.0025$ for 20 markers). In assessing the practical significance of observed distinctions, we provide effect sizes for individual markers within our analysis. Our evaluation indicates that small effects align with η^2 around 0.01, medium effects with η^2 around 0.06 and large effects with η^2 around 0.14 (Cohen, 1988). We proceeded to analyze the relationships between the linguistic markers by computing their correlations.

Classification. In our study, we applied sklearn's (Pedregosa et al., 2011) classification algorithms – logistic regression, SVM, random forests, and Naïve Bayes – to categorize texts based on extracted linguistic markers. We assigned "healthy" to texts from the initial two periods and "ill" to those from the last period. The robust 10-fold cross-validation technique notably bolstered the models' reliability and ability to generalize effectively across the diverse entries within the text corpus.

4 Results

Statistically significant differences between time periods were observed for the majority of linguistic markers after Bonferroni correction (Table 2). Subsequent post-hoc Tukey tests identified these differences to reside mainly in the third period. Specifically, markers related to lexical richness, concreteness, imageability, neologism frequency, 1st and 2nd order coherence, and phonological similarity exhibited large effect sizes. Other markers such as syntactic complexity displayed moderate effect sizes. Coherence-10 and the frequency of personal pronouns did not show significant differences. Correlation analysis (table and hierarchically-clustered heatmap in the Appendix21) unveiled generally moderate to weak correlations among linguistic markers across different domains. For instance, coherence-5 and coherence-10 demonstrated high mutual correlation but exhibited weak associations with measures of lexical diversity, syntactic complexity, or phonological similarity. Within our analysis involving 20 linguistic markers, the Naïve Bayes classifier demonstrated superior performance in distinguishing texts written in the third period from those in earlier periods (Table 3). In the context of a personalized approach, this classification algorithm can only be applied to the linguistic output of the person on whose corpus it was trained.

4.1 Figures and tables

In this section, we present scattered boxplots annotated for significance using the Tukey test ($p > 0.0025$ not significant, $p < 0.0025$ '*', $p < 0.001$ '**', $p < 0.0001$ '***') for MLTD, neologisms, concreteness, phonological similarity (20-token window), and coherence-5 (Fig 1-5). Plots for the remaining markers can be found in the Appendix A. In Table 2, we present the descriptive results and outcomes from the ANOVA comparisons. In Table 3, we summarize the results from the classification. In Table 5, we provide examples of text exhibiting high phonological similarity.

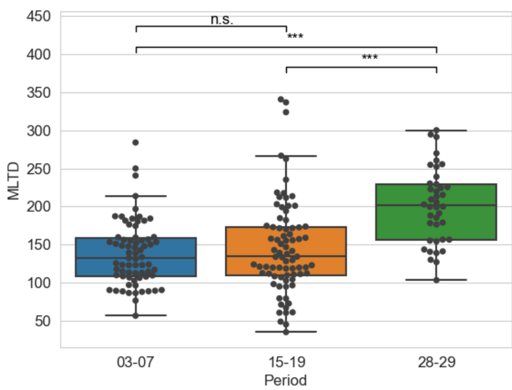

Figure 1: MLTD

	1903-1907 (n=70)	**1915-1919 (n=79)**	**1928-1929 (n=40)**	**ANOVA**	η^2
1.Coherence-5	0.465 (0.044)	0.456 (0.034)	0.426 (0.041)	**F=12.351**	0.117
2.Coherence-10	0.59 (0.041)	0.593 (0.038)	0.569 (0.042)	F=5.224	0.053
3.1st order coherence	0.622 (0.034)	0.647 (0.029)	0.653 (0.036)	**F=15.603**	0.143
4.2dn order coherence	0.604 (0.033)	0.63 (0.032)	0.637 (0.037)	**F=16.091**	0.147
5.TTR	0.561 (0.091)	0.527 (0.09)	0.593 (0.057)	**F=8.468**	0.083
6.MTLD	138.665 (42.5)	146.376 (61.9)	201.478 (48.9)	**F=19.942**	0.298
7.MATTR 25	0.907 (0.022)	0.901 (0.027)	0.931 (0.012)	**F=24.047**	0.205
8.MATTR 50	0.847 (0.028)	0.838 (0.037)	0.88 (0.018)	**F=26.196**	0.219
9.Mean token length	7.392 (0.457)	7.554 (0.479)	8.389 (0.548)	**F=57.264**	0.381
10.Out of scope tokens	0.081 (0.033)	0.084 (0.026)	0.143 (0.032)	**F=63.81**	0.406
11.Concreteness	2.053 (1.057)	1.51 (0.809)	0.447 (0.823)	**F=39.541**	0.298
12.Imageability	1.702 (0.874)	1.185 (0.77)	0.219 (0.745)	**F=43.193**	0.317
13.Yngve score	4.669 (1.056)	5.418 (1.206)	5.442 (1.075)	**F=9.996**	0.097
14.Mean sentence length	15.945 (6.92)	20.023 (6.349)	19.982 (7.013)	**F=8.093**	0.080
15.Clauses	1.959 (1.021)	2.413 (0.735)	2.556 (0.764)	**F=7.911**	0.078
16.1st person singular pronouns	0.02 (0.023)	0.03 (0.022)	0.032 (0.022)	F=5.619	0.0569
17.2nd person singular pronouns	0.003 (0.012)	0.005 (0.012)	0.004 (0.007)	F=0.53	0.005
18.1st person plural pronouns	0.006 (0.009)	0.003 (0.005)	0.002 (0.002)	F=4.715	0.0482
19.Phonological similarity 20	0.29 (0.013)	0.285 (0.015)	0.305 (0.013)	**F=28.961**	0.237
20.Phonological similarity 10	0.289 (0.014)	0.285 (0.015)	0.305 (0.013)	**F=28.459**	0.234

Table 2: Desriptive statistics and results from the ANOVA: means and standard deviations (in brackets) are listed for each group and category. Bold font indicates significant results after Bonferroni correction for multiple testing.

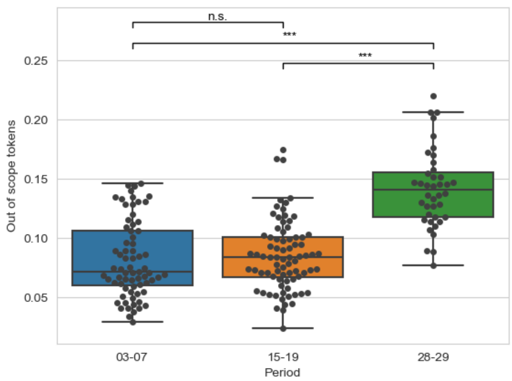

Figure 2: Out of scope tokens

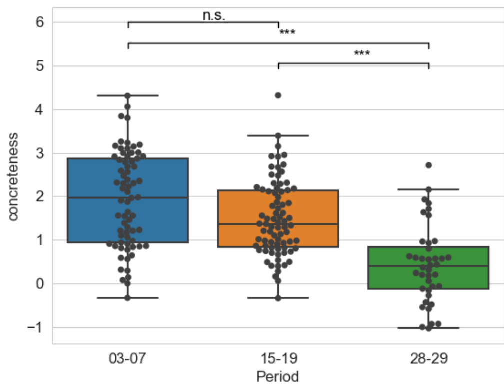

Figure 3: Concreteness

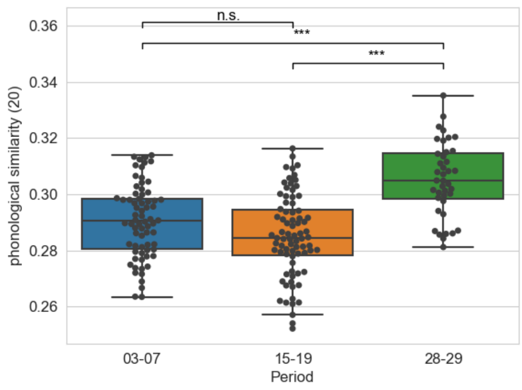

Figure 4: Phonological similarity (20-token window)

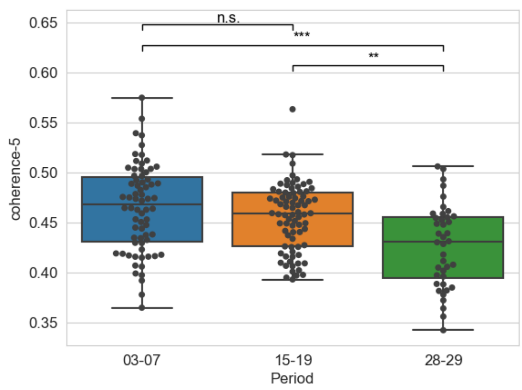

Figure 5: Coherence-5

5 Discussion

In this study, we analyzed the linguistic production of the German-speaking author Robert Walser, who probably suffered from schizophrenia, using NLP. Walser's case is particularly intriguing, given that he produced a substantial body of literary texts over a period of more than 30 years, and towards its end he was hospitalized in a psychiatric clinic with a diagnosis of schizophrenia. To construct our corpus, we compiled short prose from three distinct timeframes: 1903 – 1907, 1915 – 1919, and 1928 – 1929, the latter coinciding with Walser's hospitalization for schizophrenia in January 1929. Biographical notes suggest his experience of psychotic symptoms before hospitalization, indicating that the texts from the third period were likely composed during a psychotic state. Examining several established NLP features, we investigate their temporal association with Walser's psychosis. Additionally, we introduce two novel markers which aim at capturing phonological similarity and imagiability. Our analysis delineates a significant linguistic shift temporally linked to the onset of schizophrenia.

Model	Acc.	Prec.	Rec.	F1
Logistic regression	0.85	0.76	0.60	0.60
Random Forest	0.90	0.81	0.69	0.71
SVM	0.87	0.79	0.70	0.67
Naïve Bayes	**0.90**	**0.72**	**0.88**	**0.75**

Table 3: Results from the classification

Walser's texts from 1928 – 1929 showcase a significant lexical expansion, characterized by heightened lexical diversity, increased usage of out of scope words, and generally longer words. Metrics including MLTD, MATTR 25, MATTR 50, and out of scope token frequency demonstrate substantial effect sizes. Notably, Ziv et al. (2021) align with our findings on lexical diversity, diverging from Voleti et al. (2023), Schneider et al. (2023), and Lundin et al. (2023). Bambini et al. (2022) suggests an association between lexical diversity and psychopathological symptoms, demonstrating higher TTR in patients with pronounced symptoms. Additionally, our study replicates Just et al. (2020) observation of increased neologism use. Remarkably, a significant moderate correlation between overall lexical diversity and the usage of out of scope tokens is evident in Walser's case, potentially linked to schizophrenic symptomatology. While yielding significant results, our algorithm for the detection of lexical innovation and neologisms requires further refinement. It currently captures not only true neologisms (e.g. "Unbewusstheitsabwesenheit", "humorentfremden", "Shakespearehaftigkeit" and "Shakespearesch", "Schwalbenessay")[4] but also tokens written in Swiss German or tokens with deviant orthography.

Furthermore, Walser's later texts feature a notable decrease in both concreteness and imageability, resulting in a more abstract and ambiguous tone. The high correlation between imageability and concreteness suggests a shared underlying phenomenon. Minor et al. (2023) demonstrate sufficient test-retest reliability for word concreteness over a 6-month period, indicating its stability within that timeframe. Our analysis spans a much

[4] e.g. "unconsciousness absence", "humor alienation", "Shakespeareanism" and "Shakespearean", "Swallow essay"

longer period (1903 to 1929), allowing for a comprehensive comparison. To date, only Minor et al. (2019). reported a null association between social cognition or metacognition and concreteness so that the relationship between these markers and psychopathological symptoms remains to be studied. Overall, concreteness and imageability emerge as promising markers warranting further investigation.

In our analysis, we introduced a novel measure of phonological similarity, utilizing Jaro-Winkler similarity for the first time in schizophrenia-related NLP research. Specifically, we assessed similarity between a token (as phoneme) and a subsequent window of 10 or 20 tokens, yielding two highly correlated values likely measuring the same phenomenon. In Table 5, we provide some examples of text exhibiting high phonological similarity. ANOVA results indicated significant differences among the three samples, with post-hoc Tukey tests pinpointing distinctions in the third time period. Notably, this finding exhibited a high effect size and demonstrated moderate correlation with measures of lexical diversity. We believe this marker holds promise and warrants further investigation. As for its correlation with clanging or other psychopathological symptoms, this remains an open question that requires additional exploration.

The analysis of semantic coherence yielded contrasting outcomes. With pretrained word2vec embeddings, a significant reduction in semantic similarity for 5-token chunks emerged in the third period, showcasing a moderate effect size. These results align with previous findings (Corona-Hernández et al., 2023; Voleti et al., 2023; Iter et al., 2018; Morgan et al., 2021; Voppel et al., 2021). However, semantic similarity for 10-word chunks did not attain significance post-Bonferroni correction. Notably, these markers exhibited high correlations solely among themselves, distinct from other dataset markers. Conversely, utilizing pretrained BERT embeddings revealed increased cosine similarity between consecutive sentences in the third period compared to the first period. Additionally, 1st and 2nd order coherence exhibited significant positive correlations with measures of syntactic complexity. It seems that in the case of Walser the cosine similarity measured from the pretrained BERT model could be associated with syntax. Interestingly, the correlation between semantic similarity values from word2vec and BERT embeddings did not demonstrate significance, suggesting that they capture of distinct underlying phenomena.

There were also significant differences in the syntactic features although their effect was much less pronounced compared to the linguistic markers already described. Notably we could not find the reduction of syntactic complexity which has been described by Bedi et al. (2015) and Corcoran et al. (2018). Since this reduction of complexity has been associated with depression and negative symptoms, we can speculate that Walser was not showing these symptoms at the time before his schizophrenia diagnosis.

The observation regarding the lack of significant differences in the use of pronouns in Walser's texts is intriguing. This finding seemingly contradicts prior research (Ziv et al., 2021; Tang et al., 2021; Lundin et al., 2023). However, it's important to note that the increased use of 1st person pronouns, associated with schizophrenia, might not exclusively indicate this condition but might be observed across various mental health conditions (Lyons et al., 2018). Therefore, this particular marker might not hold substantial promise for schizophrenia detection.

Our findings offer an intriguing perspective on the concept of biomarkers or biosocial markers in schizophrenia. Unlike traditional biomarkers in somatic medicine, we observe an individual constellation that may not always align with the typical profile for the disease in the population. For instance, an increase in the usage of personal pronouns, a linguistic feature typically found in schizophrenia, does not manifest in Walser's later texts. Conversely, we observe a decrease in concreteness and imageability, along with an increase in phonological similarity. Considering the pressing demand for markers for schizophrenia in clinical practice, an NLP-driven approach shows promise. Its strength lies in its capacity for personalized analysis, identifying individual markers with significant predictive power. Utilizing these markers holds potential for predicting relapses and enhancing tailored interventions.

Limitations

Our study delineates several limitations that shape the scope and interpretation of our findings. Foremost, the exclusive focus on a singular individual restrains the generalizability of our outcomes. It's imperative to acknowledge that Walser was diagnosed with schizophrenia in 1929, almost a century preceding our study. This historical context bears significance, as diagnostic criteria and the conceptualization of psychiatric disorders have substantially evolved since that period. The diagnostic classifications prevalent during Walser's era might not seamlessly align with contemporary diagnostic manuals, potentially impacting the interpretation and contextualization of clinical data within modern psychiatric frameworks. As previously mentioned, the diagnosis itself has also faced challenges from scholars posthumously. Furthermore, the absence of standardized ratings for psychopathology in Walser's case introduces a pivotal gap. Relying solely on clinical records devoid of standardized assessments markedly curtails the depth of psychopathological insights. In addition, we recognize that alternative explanations may exist for the observed changes in linguistic style across the third time period. There are several alternative explanations of the observed results which cannot be easily addressed in the current single case design. Considering Walser's vocation as a writer, linguistic shifts in his work may reflect deliberate adaptations and conscious development in his literary style or changes in the topics he addressed. Additionally, linguistic changes associated with aging are less probable, as at the time of his hospitalization at around 50 years old, typical cognitive changes related to older age do not seem likely. The current corpus is limited to literary prose texts, posing a potential limitation. Future research should consider including additional sources of text, such as Walser's personal letters, to enhance the breadth and depth of analysis.

Ethics Statement

The current study is part of a broader project on the linguistic production of individuals with schizophrenia, which has received approval from the ethics board at the Charité – Universitätsmedizin Berlin. The texts from Robert Walser used in this research are published open access by the Robert Walser Archive and the Critical Walser Edition. Walser's medical history has been previously published. The author died 67 years ago. We acknowledge that the findings of this research should not be utilized for the initial diagnosis of schizophrenia. However, we recognize the potential of employing NLP in detecting relapses among individuals already diagnosed.

Acknowledgements

We express our gratitude to the Robert Walser Archive for providing us with the texts. Dr. Ivan Nenchev is participant in the BIH Charité Digital Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin, and the Berlin Institute of Health at Charité (BIH).

References

Maria Francisca Alonso-Sánchez, Sabrina D. Ford, Michael MacKinley, Angélica Silva, Roberto Limongi, and Lena Palaniyappan. 2022. Progressive changes in descriptive discourse in First Episode Schizophrenia: a longitudinal computational semantics study. *Schizophrenia*, 8(1):1–9. Number: 1 Publisher: Nature Publishing Group.

N. C. Andreasen. 1990. Methods for assessing positive and negative symptoms. *Modern Problems of Pharmacopsychiatry*, 24:73–88.

Nancy C. Andreasen. 1986. Scale for the assessment of thought, language, and communication (TLC). *Schizophrenia Bulletin*, 12(3):473–482.

Jenna L. Baddeley, Gwyneth R. Daniel, and James W. Pennebaker. 2011. How Henry Hellyer's Use of Language Foretold His Suicide. *Crisis*, 32(5):288–292. Publisher: Hogrefe Publishing.

Valentina Bambini, Federico Frau, Luca Bischetti, Federica Cuoco, Margherita Bechi, Mariachiara Buonocore, Giulia Agostoni, Ilaria Ferri, Jacopo Sapienza, Francesca Martini, Marco Spangaro, Giorgia Bigai, Federica Cocchi, Roberto Cavallaro, and Marta Bosia. 2022. Deconstructing heterogeneity in schizophrenia through language: a semi-automated linguistic analysis and data-driven clustering approach. *Schizophrenia*, 8(1):1–12. Number: 1 Publisher: Nature Publishing Group.

Gillinder Bedi, Facundo Carrillo, Guillermo A. Cecchi, Diego Fernández Slezak, Mariano Sigman, Natália B. Mota, Sidarta Ribeiro, Daniel C. Javitt, Mauro Copelli, and Cheryl M. Corcoran. 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ schizophrenia*, 1:15030.

M. L. Birnbaum, S. K. Ernala, A. F. Rizvi, E. Arenare, A. R. Van Meter, M. De Choudhury, and J. M. Kane. 2019. Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *npj Schizophrenia*, 5(1):1–9. Number: 1 Publisher: Nature Publishing Group.

Michael L. Birnbaum, Sindhu Kiranmai Ernala, Asra F. Rizvi, Munmun De Choudhury, and John M. Kane. 2017. A Collaborative Approach to Identifying Social Media Markers of Schizophrenia by Employing Machine Learning and Clinical Appraisals. *Journal of Medical Internet Research*, 19(8):e289.

Timo Brockmeyer, Johannes Zimmermann, Dominika Kulessa, Martin Hautzinger, Hinrich Bents, Hans-Christoph Friederich, Wolfgang Herzog, and Matthias Backenstrass. 2015. Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. *Frontiers in Psychology*, 6.

Casper M. Brosius. 1857. Über die Sprache der Irren. *Allgemeine Zeitschrift für Psychiatrie und psychisch-gerichtliche Medizin*, 1.:37 – 64.

Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2 edition. Routledge, New York.

Cheryl M. Corcoran, Facundo Carrillo, Diego Fernández-Slezak, Gillinder Bedi, Casimir Klim, Daniel C. Javitt, Carrie E. Bearden, and Guillermo A. Cecchi. 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World psychiatry: official journal of the World Psychiatric Association (WPA)*, 17(1):67–75.

Cheryl M. Corcoran, Vijay A. Mittal, Carrie E. Bearden, Raquel E Gur, Kasia Hitczenko, Zarina Bilgrami, Aleksandar Savic, Guillermo A. Cecchi, and Phillip Wolff. 2020. Language as a biomarker for psychosis: A natural language processing approach. *Schizophrenia Research*, 226:158–166.

H. Corona-Hernández, J. N. de Boer, S. G. Brederoo, A. E. Voppel, and I. E. C. Sommer. 2023. Assessing coherence through linguistic connectives: Analysis of speech in patients with schizophrenia-spectrum disorders. *Schizophrenia Research*, 259:48–58.

Claudio Crema, Giuseppe Attardi, Daniele Sartiano, and Alberto Redolfi. 2022. Natural language processing in clinical neuroscience and psychiatry: A review. *Frontiers in Psychiatry*, 13.

Janna N. de Boer, Sanne G. Brederoo, Alban E. Voppel, and Iris E. C. Sommer. 2020. Anomalies in language as a biomarker for schizophrenia. *Current Opinion in Psychiatry*, 33(3):212–218.

Hannah J. Dean and Ryan L. Boyd. 2020. Deep into that darkness peering: A computational analysis of the role of depression in Edgar Allan Poe's life and death. *Journal of Affective Disorders*, 266:482–491.

To'Meisha Edwards and Nicholas S. Holtzman. 2017. A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality*, 68:63–68.

M. Fernández-Cabana, A. García-Caballero, M. T. Alves-Pérez, M. J. García-García, and R. Mateos. 2013. Suicidal traits in Marilyn Monroe's Fragments: an LIWC analysis. *Crisis*, 34(2):124–130.

S. K. Fineberg, S. Deutsch-Link, M. Ichinose, T. McGuinness, A. J. Bessette, C. K. Chung, and P. R. Corlett. 2015. Word use in first-person accounts of schizophrenia†. *The British Journal of Psychiatry*, 206(1):32–38.

Peter Garrard, Lisa M. Maloney, John R. Hodges, and Karalyn Patterson. 2005. The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. *Brain*, 128(2):250–260.

Lucas Marco Gisi, editor. 2018. *Robert Walser-Handbuch*. J.B. Metzler, Stuttgart.

Wilhelm Griesinger. 1845. *Die Pathologie und Therapie der psychischen Krankheiten, für Aerzte und Studirende*, 1. auflage edition. Krabbe, Stuttgart.

S. S. Haas, G. E. Doucet, S. Garg, S. N. Herrera, C. Sarac, Z. R. Bilgrami, R. B. Shaik, and C. M. Corcoran. 2020. Linking language features to clinical symptoms and multimodal imaging in individuals at clinical high risk for psychosis. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 63(1):e72.

Graeme Hirst and Vanessa Wei Feng. 2012. Changes in Style in Authors with Alzheimer's Disease. *English Studies*, 93(3):357–370. Publisher: Routledge _eprint: https://doi.org/10.1080/0013838X.2012.668789.

Kasia Hitczenko, Vijay A Mittal, and Matthew Goldrick. 2021. Understanding Language Abnormalities and Associated Clinical Markers in Psychosis: The Promise of Computational Methods. *Schizophrenia Bulletin*, 47(2):344–362.

Dan Iter, Jong Yoon, and Dan Jurafsky. 2018. Automatic Detection of Incoherent Speech for Diagnosing Schizophrenia. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 136–146, New Orleans, LA. Association for Computational Linguistics.

Sandra A. Just, Erik Haegert, Nora Kořánová, Anna-Lena Bröcker, Ivan Nenchev, Jakob Funcke, Andreas Heinz, Felix Bermpohl, Manfred Stede, and Christiane Montag. 2020. Modeling Incoherent Discourse in Non-Affective Psychosis. *Frontiers in Psychiatry*, 11. Publisher: Frontiers.

Xuan Le, Ian Lancashire, Graeme Hirst, and Regina Jokel. 2011. Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. *Literary and Linguistic Computing*, 26(4):435–461.

Einat Liebenthal, Michaela Ennis, Habiballah Rahimi-Eichi, Eric Lin, Yoonho Chung, and Justin T. Baker. 2023. Linguistic and non-linguistic markers of disorganization in psychotic illness. *Schizophrenia Research*, 259:111–120.

Nancy B. Lundin, Henry R. Cowan, Divnoor K. Singh, and Aubrey M. Moe. 2023. Lower cohesion and altered first-person pronoun usage in the spoken life narratives of individuals with schizophrenia. *Schizophrenia Research*, 259:140–149.

Minna Lyons, Nazli Deniz Aksayli, and Gayle Brewer. 2018. Mental distress and language use: Linguistic analysis of discussion forum posts. *Computers in Human Behavior*, 87:207–211.

Viktoria Lyons and Michael Fitzgerald. 2004. The case of Robert Walser (1878-1956). *Irish Journal of Psychological Medicine*, 21(4):138–142. Publisher: Cambridge University Press.

Jana Lüdtke and Katharina G. Hugentobler. Using emotional word ratings to extrapolated norms for valence, arousal, imagebility and concreteness: The German list of extrapolated affective norms (GLEAN). *Proceedings of KogWis2022, the 5th Biannual Conference of the German Society for Cognitive Science.*

Philip M. McCarthy and Scott Jarvis. 2010. MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2):381–392.

George A. Miller. 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97.

K. S. Minor, J. A. Willits, M. P. Marggraf, M. N. Jones, and P. H. Lysaker. 2019. Measuring disorganized speech in schizophrenia: automated analysis explains variance in cognitive deficits beyond clinician-rated scales. *Psychological Medicine*, 49(3):440–448. Publisher: Cambridge University Press.

Kyle S. Minor, Nancy B. Lundin, Evan J. Myers, Aitana Fernández-Villardón, and Paul H. Lysaker. 2023. Automated measures of speech content and speech organization in schizophrenia: Test-retest reliability and generalizability across demographic variables. *Psychiatry Research*, 320:115048.

Sarah E. Morgan, Kelly Diederen, Petra E. Vértes, Samantha H. Y. Ip, Bo Wang, Bethany Thompson, Arsime Demjaha, Andrea De Micheli, Dominic Oliver, Maria Liakata, Paolo Fusar-Poli, Tom J. Spencer, and Philip McGuire. 2021. Natural Language Processing markers in first episode psychosis and people at clinical high-risk. *Translational Psychiatry*, 11(1):1–9. Number: 1 Publisher: Nature Publishing Group.

David R. Mortensen, Siddharth Dalmia, and Patrick Littell. 2018. Epitran: Precision G2P for Many Languages. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Robert Mächler and Carl Seelig. 1992. *Das Leben Robert Walsers: eine dokumentarische Biographie*. Suhrkamp. Google-Books-ID: czFcAAAAMAAJ.

Viola Oertel, Anna Rotarska-Jagiela, Vincent van de Ven, Corinna Haenschel, Michael Grube, Ulrich Stangier, Konrad Maurer, and David E. J. Linden. 2009. Mental imagery vividness as a trait marker across the schizophrenia spectrum. *Psychiatry Research*, 167(1):1–11.

Serguei Pakhomov, Dustin Chacon, Mark Wicklund, and Jeanette Gundel. 2011. Computerized assessment of syntactic complexity in Alzheimer's disease: a case study of Iris Murdoch's writing. *Behavior Research Methods*, 43(1):136–144.

Lena Palaniyappan. 2021. More than a biomarker: could language be a biosocial marker of psychosis? *npj Schizophrenia*, 7(1):1–5. Number: 1 Publisher: Nature Publishing Group.

Alberto Parola, Jessica Mary Lin, Arndis Simonsen, Vibeke Bliksted, Yuan Zhou, Huiling Wang, Lana Inoue, Katja Koelkebeck, and Riccardo Fusaroli. 2023. Speech disturbances in schizophrenia: Assessing cross-linguistic generalizability of NLP automated measures of coherence. *Schizophrenia Research*, 259:59–70.

S. Partl, B. Pfuhlmann, B. Jabs, and G. Stöber. 2011. „Meine Krankheit ist eine Kopfkrankheit, die schwer zu definieren ist". *Der Nervenarzt*, 82(1):67–78.

D. Pavy, L. Grinspoon, and R. I. Shader. 1969. Word frequency measures of verbal disorders in schizophrenia. *Diseases of the Nervous System*, 30(8):553–556.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. ArXiv:1908.10084 [cs].

Vassiliki Rentoumi, Timothy Peters, Jonathan Conlin, and Peter Garrard. 2017. The acute mania of King George III: A computational linguistic analysis. *PLoS ONE*, 12(3):e0171626.

Alexander T Sack, Vincent G van de Ven, Simone Etschenberg, Dietmar Schatz, and David E. J Linden. 2005. Enhanced vividness of mental imagery as a trait marker of schizophrenia? *Schizophrenia Bulletin*, 31(1):97–104.

Katharina Schneider, Katrin Leinweber, Hamidreza Jamalabadi, Lea Teutenberg, Katharina Brosch, Julia-Katharina Pfarr, Florian Thomas-Odenthal, Paula Usemann, Adrian Wroblewski, Benjamin Straube, Nina Alexander, Igor Nenadić, Andreas Jansen, Axel Krug, Udo Dannlowski, Tilo Kircher, Arne Nagels, and Frederike Stein. 2023. Syntactic complexity and diversity of spontaneous speech production in schizophrenia spectrum and major depressive disorders. *Schizophrenia*, 9(1):1–10. Number: 1 Publisher: Nature Publishing Group.

Skipper Seabold and Josef Perktold. 2010. Statsmodels: Econometric and Statistical Modeling with Python.

Proceedings of the 9th Python in Science Conference, pages 92–96. Conference Name: Proceedings of the 9th Python in Science Conference.

Lucas Shen, David Lesieur, Christophe Bedetti, Earl Brown, Garrett Hurst, and Pip coder. 2023. LSYS/LexicalRichness: v0.5.0.

Angelica M. Silva, Roberto Limongi, Michael MacKinley, Sabrina D. Ford, Maria Francisca Alonso-Sánchez, and Lena Palaniyappan. 2023. Syntactic complexity of spoken language in the diagnosis of schizophrenia: A probabilistic Bayes network model. *Schizophrenia Research*, 259:88–96.

S. W. Stirman and J. W. Pennebaker. 2001. Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63(4):517–522.

Sunny X. Tang, Reno Kriz, Sunghye Cho, Suh Jung Park, Jenna Harowitz, Raquel E. Gur, Mahendra T. Bhati, Daniel H. Wolf, João Sedoc, and Mark Y. Liberman. 2021. Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. *NPJ schizophrenia*, 7(1):25.

Charles Vannette. 2020. *Robert Walser: Unmoored*. Peter Lang, Bern, Switzerland.

Rohit Voleti, Stephanie M Woolridge, Julie M Liss, Melissa Milanovic, Gabriela Stegmann, Shira Hahn, Philip D Harvey, Thomas L Patterson, Christopher R Bowie, and Visar Berisha. 2023. Language Analytics for Assessment of Mental Health Status and Functional Competency. *Schizophrenia Bulletin*, 49(Supplement_2):S183–S195.

Alban E. Voppel, Janna N. de Boer, SG Brederoo, HG Schnack, and Iris E. C. Sommer. 2021. Quantified language connectedness in schizophrenia-spectrum disorders. *Psychiatry Research*, 304:114130.

Robert Walser. 2013a. *Drucke im Berliner Tageblatt*. Schwabe Verlag.

Robert Walser. 2013b. *Drucke in der Neuen Zürcher Zeitung*. Schwabe Verlag.

Robert Walser. 2014. *Poetenleben - Kritische Edition der Erstausgabe*. Schwabe Verlag.

Robert Walser. 2015. *Drucke in der Schaubühne/ Weltbühne*. Schwabe Verlag.

Robert Walser. 2017a. *Drucke in der «Neuen Rundschau»*. Schwabe Verlag.

Robert Walser. 2017b. *Prosastücke - Kleine Prosa - Der Spaziergang*. Schwabe Verlag.

Robert Walser. 2019. *Drucke in der Prager Presse*. Schwabe Verlag.

Robert Walser. 2023a. *Fritz Kocher's Aufsätze - Kritische Edition und Reprint der Erstausgabe*. Schwabe Verlag.

Robert Walser. 2023b. *Kleine Dichtungen*. Schwabe Verlag.

Andrew R. Watson, Cağla Defterali, Thomas H. Bak, Antonella Sorace, Andrew M. McIntosh, David G. C. Owens, Eve C. Johnstone, and Stephen M. Lawrie. 2012. Use of second-person pronouns and schizophrenia. *The British Journal of Psychiatry: The Journal of Mental Science*, 200(4):342–343.

Martina Wernli. 2014. *Schreiben am Rand: Die »Bernische kantonale Irrenanstalt Waldau« und ihre Narrative (1895-1936)*. transcript Verlag. Publication Title: Schreiben am Rand.

Victor H. Yngve. 1960. A model and an hypothesis for language structure. Accepted: 2004-03-02T23:07:37Z Publisher: Massachusetts Institute of Technology, Research Laboratory of Electronics.

Ido Ziv, Heli Baram, Kfir Bar, Vered Zilberstein, Samuel Itzikowitz, Eran V. Harel, and Nachum Dershowitz. 2021. Morphological characteristics of spoken language in schizophrenia patients - an exploratory study. *Scandinavian Journal of Psychology*.

A Appendix

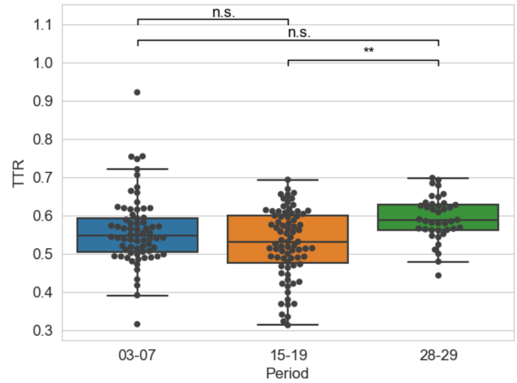

Figure 6: TTR

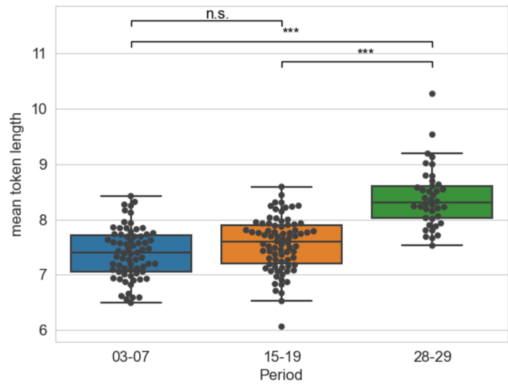

Figure 9: Mean token length

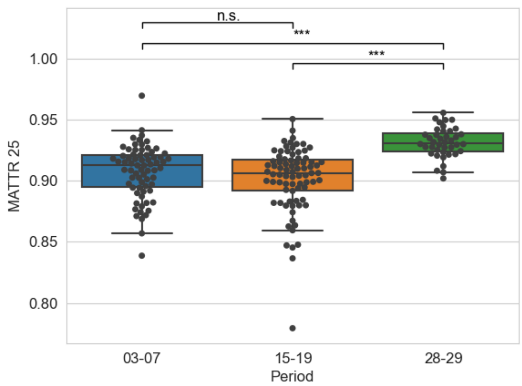

Figure 7: MATTR (25-token window)

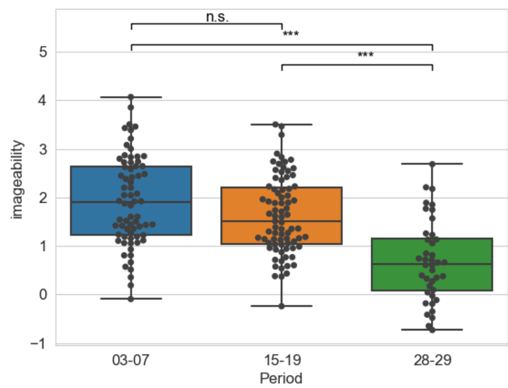

Figure 10: Imageability

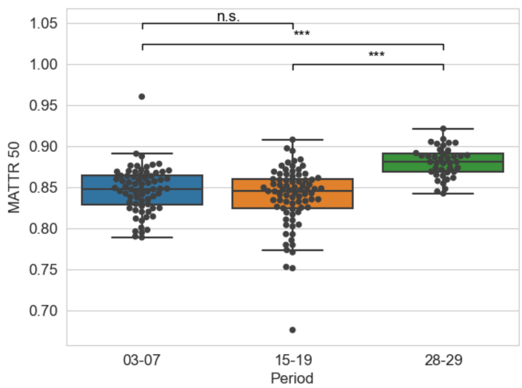

Figure 8: MATTR (50-token window)

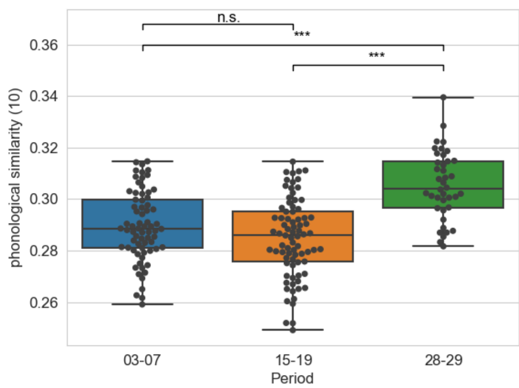

Figure 11: Phonological similarity (10-token window)

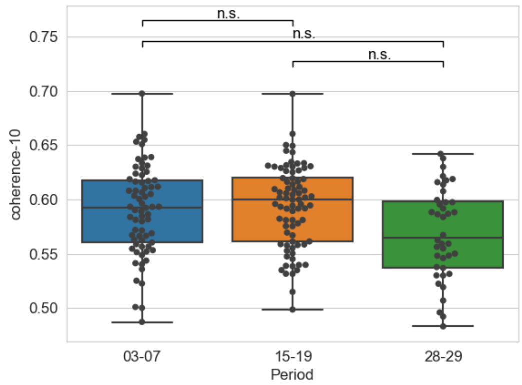

Figure 12: Coherence-10

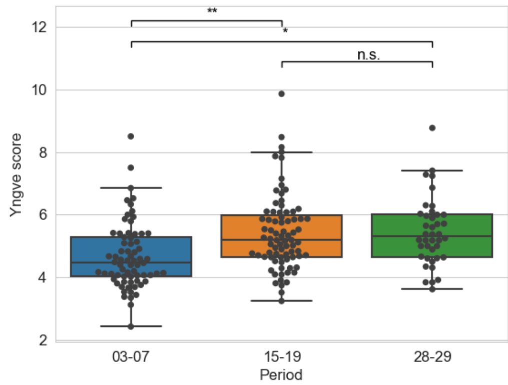

Figure 15: Yngve score

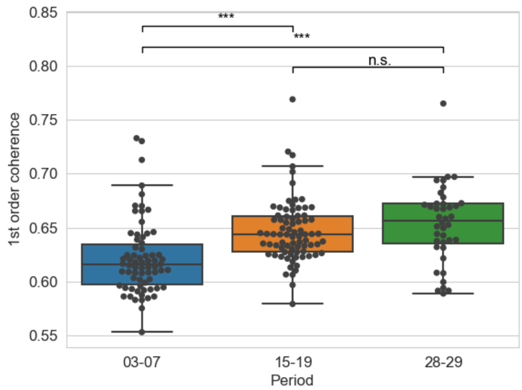

Figure 13: First order coherence

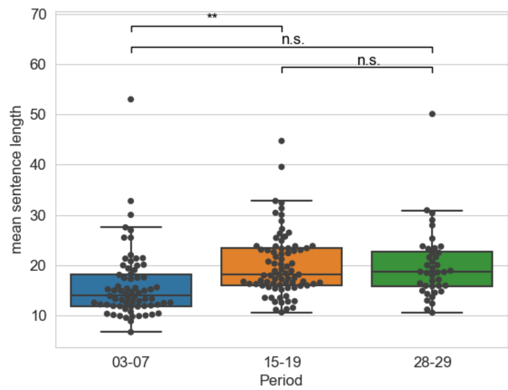

Figure 16: Mean sentence length

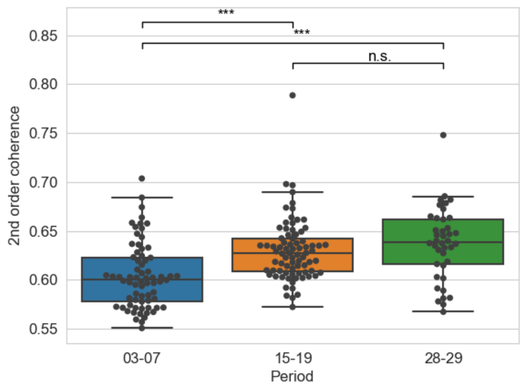

Figure 14: Second order coherence

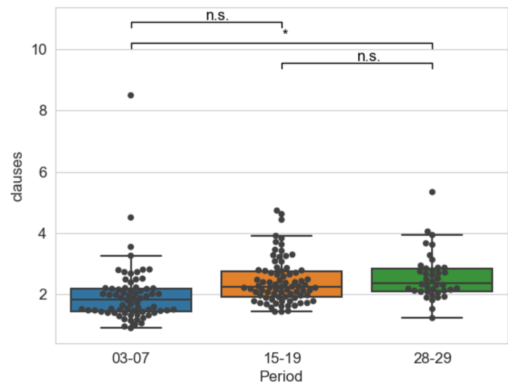

Figure 17: Clauses per sentence

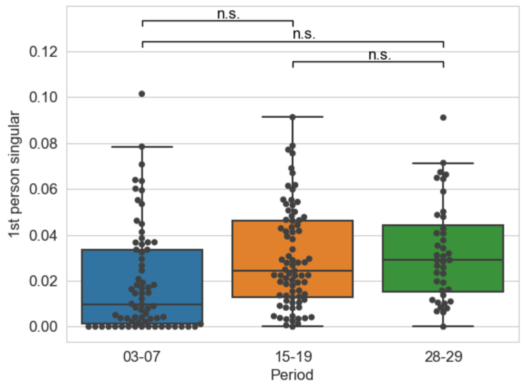

Figure 18: First person singular pronouns

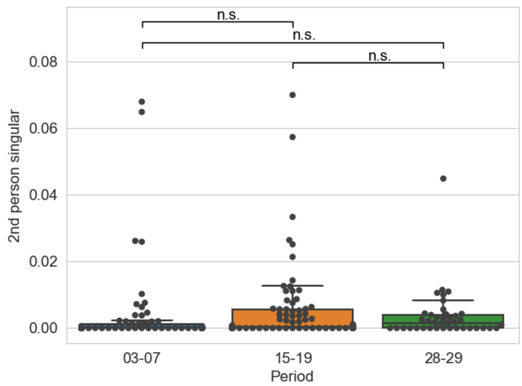

Figure 19: Second person singular pronouns

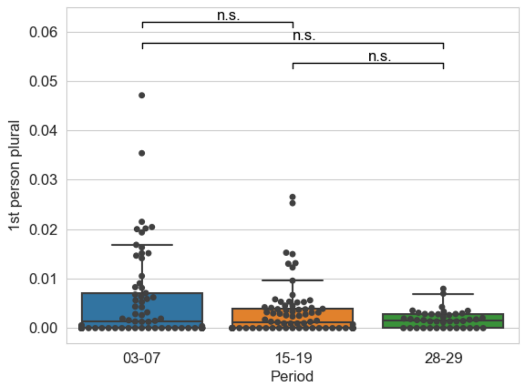

Figure 20: First person plural pronouns

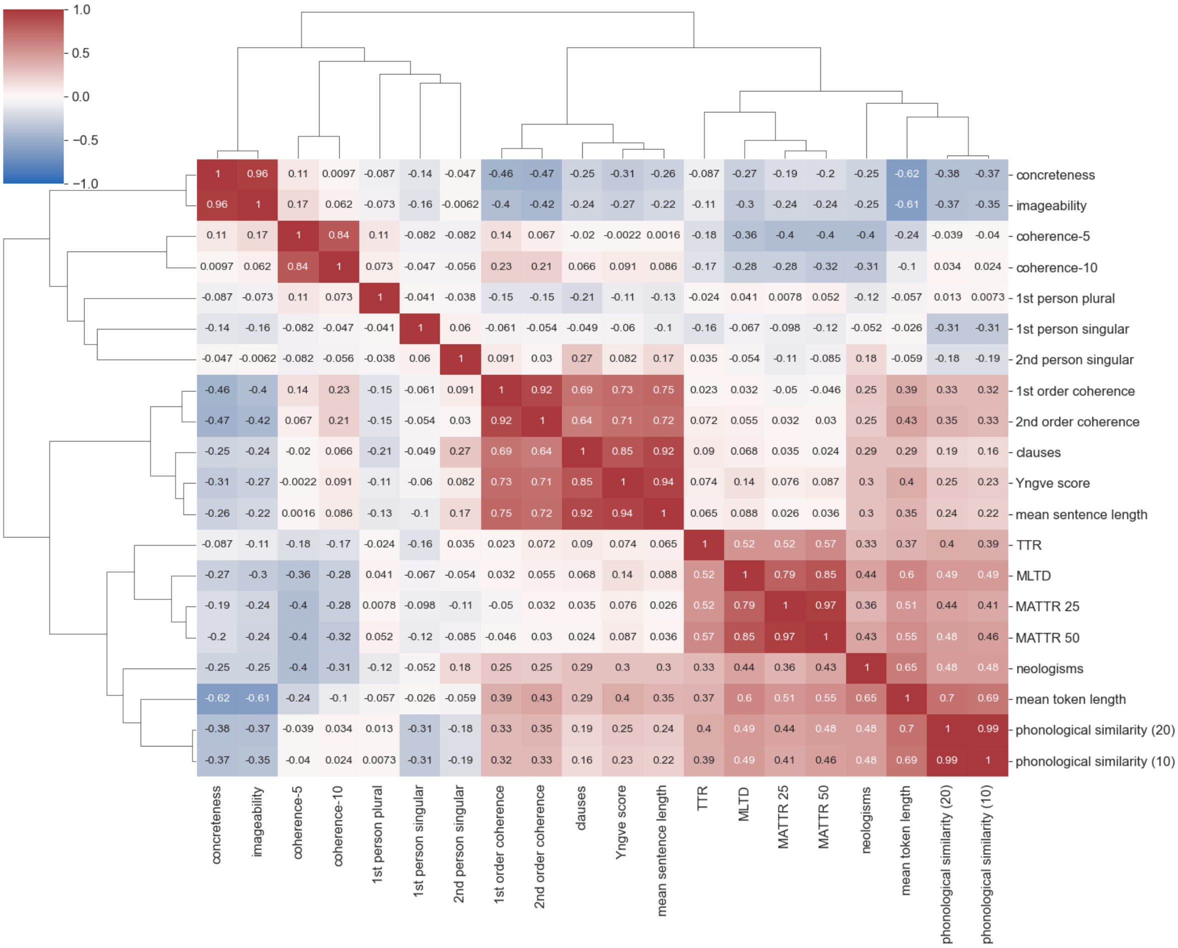

Figure 21: Hierarchically-clustered heatmap of the 20 linguistic markers.

	coh-5	coh-10	1st coh	2nd coh	MATTR 25	MATTR 50	MLTD	TTR	out of scope tokens
mean token length	-0.24***	-0.1	0.39***	0.43***	0.51***	0.55***	0.6***	0.37***	0.65***
coherence-5		0.84***	0.14	0.07	-0.4***	-0.4***	-0.36***	-0.18*	-0.4***
coherence-10			0.23**	0.21**	-0.28***	-0.32***	-0.28***	-0.17*	-0.31***
1st order coherence				0.92***	-0.05	-0.05	0.03	0.02	0.25***
2nd order coherence					0.03	0.03	0.06	0.07	0.25***
MATTR 25						0.97***	0.79***	0.52***	0.36***
MATTR 50							0.85***	0.57***	0.43***
MLTD								0.52***	0.44***
TTR									0.33***
out of scope tokens									
concreteness									
imageability									
Yngve score									
mean sentence length									
clauses									
1st person singular									
phonological similarity (20)									
phonological similarity (10)									
1st person plural									

Table 4: Table demonstrating the correlations between the linguistic marker, ($p > 0.0025$ not significant, $p < 0.05$ '*', $p < 0.01$ '**', $p < 0.001$ '***')

	concr	image	Yngve	msl	clauses	1st sing	phon 20	phon 10	1st pl	2nd sing
mean token length	-0.62***	-0.61***	0.4***	0.35***	0.29***	-0.03	0.7***	0.69***	-0.06	-0.06
coherence-5	0.11	0.17*	-0.0	0.0	-0.02	-0.08	-0.04	-0.04	0.11	-0.08
coherence-10	0.01	0.06	0.09	0.09	0.07	-0.05	0.03	0.02	0.07	-0.06
1st order coherence	-0.46***	-0.4***	0.73***	0.75***	0.69***	-0.06	0.33***	0.32***	-0.15*	0.09
2nd order coherence	-0.47***	-0.42***	0.71***	0.72***	0.64***	-0.05	0.35***	0.33***	-0.15*	0.03
MATTR 25	-0.19**	-0.24***	0.08	0.03	0.03	-0.1	0.44***	0.41***	0.01	-0.11
MATTR 50	-0.2**	-0.24***	0.09	0.04	0.02	-0.12	0.48***	0.46***	0.05	-0.08
MLTD	-0.27***	-0.3***	0.14	0.09	0.07	-0.07	0.49***	0.49***	0.04	-0.05
TTR	-0.09	-0.11	0.07	0.07	0.09	-0.16*	0.4***	0.39***	-0.02	0.03
out of scope tokens	-0.25***	-0.25***	0.3***	0.3***	0.29***	-0.05	0.48***	0.48***	-0.12	0.18*
concreteness		0.96***	-0.31***	-0.26***	-0.25***	-0.14*	-0.38***	-0.37***	-0.09	-0.05
imageability			-0.27***	-0.22**	-0.24**	-0.16*	-0.37***	-0.35***	-0.07	-0.01
Yngve score				0.94***	0.85***	-0.06	0.25***	0.23**	-0.11	0.08
mean sentence length					0.92***	-0.1	0.24***	0.22**	-0.13	0.17*
clauses						-0.05	0.19**	0.16*	-0.21**	0.27***
1st person singular							-0.31***	-0.31***	-0.04	0.06
phonological similarity (20)								0.99***	0.01	-0.18*
phonological similarity (10)									0.01	-0.19*
1st person plural										-0.04

	Source	Example in German and English translation
Phonological similarity	Mondschein-geschichte, 1928	Noch nie, solange ich dichte, dichtete ich eine schlichtere Geschichte, wie die, worin ich berichte,... Never, as long as I have been writing poetry, have I written a simpler story than the one in which I report...
	Freiheits-aufsatz, 1928	Sie ist eine Freie und infolgedessen eine Feine, die jede Unfeinheit aufs feinste empfindet, mit anderen Worten, die jede Freiheit, die man sich ihr gegenüber herausnimmt, als etwas Unfeines betrachtet... She is a free woman and, as a result, a fine woman who feels every impurity in the finest way, in other words, who regards every freedom taken towards her as something impure...
	Ein dummer Junge, 1928	Auffallend viele Menschen, die einen Namen haben, einen Wert auf den Achseln tragen, feiern in diesen Tagen ihren sechzigsten Geburtstag. A conspicuous number of people who have a name, a value on their armpits, are celebrating their sixtieth birthday these days.

Table 5: Examples of text exhibiting high phonological similarity