REVIEW PAPER Curr Probl Psychiatry, Vol. 25(2024):127-135

DOI:10.12923/2353-8627/2024-0013

Czasopismo indeksowane na liście MNiSW - 70 pkt.

The Significance of Natural Language Processing and Machine Learning in Schizophasia Description. Identification of Research Trends and Perspectives in Schizophrenia Language Studies

Michał Mazur ABCDEF, https://orcid.org/0000-0003-4969-2792,

Paweł Krukow ABCDEF, https://orcid.org/0000-0001-9497-2713,

Department of Clinical Neuropsychiatry, Medical University of Lublin,Poland

Abstract

Introduction: Language and speech serve as significant biomarkers for psychiatric disorders, including schizophrenia. The linguistic features associated with schizophasia have been a focal point since the early descriptions of schizophrenia. Over the past twenty-five years, scientific reflection on language in mental illnesses has dynamically provided new data identifying the complex phenomenon of speech pathology in schizophrenia.

Material and methods: A bibliometric analysis was conducted using SCOPUS data, focusing on word co-occurrence patterns in schizophrenia research. VOSviewer was employed for visualization, and semantic relationships between words were explored.

Results: An analysis has revealed trends and gaps in research on schizophasia. Integrating temporal and spatial visualizations of metadata has allowed for the identification of currently employed measures of incoherence in schizophatic texts across various levels of linguistic organization. Keyword modeling has demonstrated a growing interest in utilizing artificial intelligence techniques to develop linguistic biomarkers for schizophrenia and other mental disorders.

Conclusions: The harmonization of computational methods for measuring narrative, dialogic, and prosodic coherence holds promise, particularly in cross-validation studies involving other neuroindicators of mental disorders. Developing linguistic biomarkers using broadly understood artificial intelligence requires multidisciplinary research teams integrating experts from psychiatry, neurolinguistics, neurologopedics, and AI engineering. Clear domain-specific regulations are essential to ensure accurate conclusions and ethical considerations. The study of schizophasia prospects is particularly evident at the lexical, semantic, and syntactic levels, along with affective and neurophysiological variables.

Keywords: language, schizophasia, speech pathology, bibliomeric analysis, formal thought disorder

Streszczenie

Wstęp: Język i mowa są istotnymi biomarkerami zaburzeń psychicznych, w szczególności schizofrenii. Zaburzenia mowy o typie schizofazji stanowią jedne z osiowych objawów psychozy, przywoływanych od początku badań nad schizofrenią. W ciągu ostatnich dwudziestu pięciu lat naukowa refleksja nad językiem w chorobach psychicznych dynamicznie dostarczała nowych danych identyfikujących złożone zjawisko patologii mowy w schizofrenii.

Materiał i metody: Przeprowadzono analizę bibliometryczną z wykorzystaniem danych SCOPUS, koncentrując się na wzorcach współwystępowania słów kluczowych w metadanych badań nad schizofazją. Do wizualizacji zgromadzonych danych w formie map wykorzystano VOSviewer.

Dyskusja: Analiza ujawniła trendy i luki w badaniach nad schizofazją. Zintegrowanie czasowych i przestrzennych wizualizacji metadanych pozwoliło zidentyfikować współcześnie podejmowane miary niespójności tekstów schizofatycznych na różnych poziomach organizacji językowej. Modelowanie słów kluczowych wykazało rosnące zainteresowanie wykorzystywaniem technik sztucznej inteligencji w opracowywaniu lingwistycznych biomarkerów schizofrenii i innych chorób psychicznych.

Wnioski: Harmonizacja metod obliczeniowych pomiaru spójności narracyjnej, dialogowej i prozodycznej jest obiecująca, szczególnie w badaniach krzyżowych obejmujących inne neurowskaźniki zaburzeń psychicznych. Opracowywanie biomarkerów językowych z wykorzystaniem szeroko pojętej sztucznej inteligencji wymaga zaangażowania multidyscyplinarnych zespołów badawczych, integrujących ekspertów z zakresu psychiatrii, neurolingwistyki, neurologopedii i inżynierii sztucznej inteligencji. Jasne regulacje dotyczące poszczególnych dziedzin są niezbędne, aby zapewnić trafne wnioski i względy etyczne.

Perspektywa badań schizofazji jest szczególnie widoczna na poziomie opisów leksykalnych, semantycznych i syntaktycznych, przy jednoczesnym uwzględnieniu zmiennych afektywnych i neurofizjologicznych

Słowa kluczowe: język, schizofazja, zaburzenia mowy, analiza bibliometryczna, formalne zaburzenia myślenia

Introduction

Formal thought disorders in schizophrenia are diagnosed based on available linguistic material, and language-communication dysfunctions, referred to as schizophasia, constitute one of the fundamental diagnostic criteria [1]. The disorganization of the language system has often been cited as a central symptom of mental disorders. In-depth understanding of the specifics of speech dysfunction, understood as distruption of "the set of activities that a person performs using language to perceive reality and convey its interpretation to other participants in social life" [2] is a significant challenge. This concept is closely associated with the notion of language behavior and has been explored by psychiatrists, psychologists, neurolinguists, and speech therapists. Scientific reflection consistently emphasizes the diagnostician's ability to describe and assess linguistic aspects of thought organization, both in monologues and during discourse [3].

Speech, considered a biological construct and a reflection of cognitive and neurophysiological processes, becomes a crucial biomarker for psychiatric disorders, particularly schizophrenia. During this time, linguistic competence is partially or completely lost, and surviving functions often experience disorganization.

Speech, as an activity reflecting mental states, represents an abstract entity of great complexity. Approaches to its interpretation, especially in clinical contexts, require advanced knowledge that integrates linguistic, biomedical, psychological, and pedagogical sciences. The intricate structure of speech and the even greater complexity of potential disorganization have long fascinated researchers across these disciplines. This has led to a growing body of research on language in schizophrenia. Despite the significant contributions of these studies to clinical practice, they also present challenges in monitoring the ever-evolving knowledge about schizophasia and schizophrenia. Modern bibliometric research methods allow us to analyze the entire field of speech disorders in schizophrenia, identify research trends, track shifts in scientific directions, address knowledge gaps, and create visualizations of research topics over time. These insights are essential for informed and targeted management of the development of schizophasia-related research.

The question of whether modern natural language processing (NLP) methods can fully replace or even eliminate human involvement in information indexing and retrieval is a central point of investigation. By leveraging automated methods, which sometimes outperform classical cognitive methods used by humans, computers can potentially take over these tasks. The automation of this process offers clear benefits in terms of cost savings and time efficiency associated with creating appropriate profiles. Therefore, research that explores the practical implementation of such capabilities is valuable and useful.

This article aims to provide insights, based on existing literature, into the current state and future prospects of utilizing natural language in information retrieval processes within search systems, including the Internet. The answer to this question could serve as a starting point for broader discussions and engagement of Polish information scientists in this extensive research and application domain.

Methodology

The primary objective of this study was to identify trends and research gaps in contemporary scientific works related to the topic of schizophasia. This goal will be achieved through the following operational objectives:

1. Logically inductive analysis of network connections based on bibliographic data from studies on language and speech in schizophrenia.
2. Comparison of quantitative variables used in this analysis with studies on language in general psychiatric disorders.
3. Development of visualizations depicting the research structure on schizophasia.
4. Identification of the currently burgeoning sub-areas of research on schizophasia in the context of 1999-2024.

The hypothesis posits that the advancement of research methods based on artificial intelligence influences the growth rate of original scientific contributions regarding language in schizophrenia. This language constitutes the fundamental material for describing formal thought disorders. As an extension of this hypothesis, it is assumed that integrating knowledge

from clinical psychiatry, linguistics, and artificial intelligence holds the potential to describe highly complex linguistic constructs. This includes not only isolated levels of structural language organization but also phenomena that require combining selected aspects from different layers of the language and speech system.

To test the above hypothesis, bibliometric sampling focused on mapping and clustering is employed. Mapping enables characterization of relationships between objects in Euclidean scientific space, while clustering determines object membership within specific categories. The integration of these two approaches in describing the structure of scientific disciplines was effectively pioneered by Ludo Waltman and Nees Jan van Eck [4]. Their proposed approach enables a rational combination of diverse mapping and clustering ideas, representing the sole proven perspective for analyzing scientific trends and identifying research gaps [5].

To conduct this review, the research tool developed by these scholars, VOSviewer, is utilized [4].

For this investigation, bibliographic data (bibliographic coupling) and information derived from natural language analysis used in describing metadata of scientific articles (e.g., title, keywords, abstract) are utilized. The goal of this review is achievable through analyses of keyword co-occurrence. Specifically, we will employ co-occurrence analyses (also known as node analyses), focusing on their focused variant [6]. The decision to focus on schizophasia stems from the need to analyze this narrower aspect within the broader field of schizophrenia research. Additionally, the evolution of concepts related to language functioning in individuals with schizophrenia is explored.

The fundamental material for the analysis will be bibliographic metadata records from the SCOPUS database. Filtering of scientific reports was conducted using the following query:

(TITLE-ABS-KEY (schizophasia) OR TITLE-ABS-KEY (language AND shizophrenia) OR TITLE-ABS-KEY (language AND psychosis) OR TITLE-ABS-KEY (language AND competence AND shizophrenia)) AND PUBYEAR > 1999 AND PUBYEAR < 2025.

The time scope for including works in the analysis was proposed based on a prior examination of the general trend of increasing publications that meet the remaining criteria. The last twenty-five years have been particularly notable in this regard. Consequently, the analysis encompasses 2701 scientific works addressing the topic of schizophasia from various scientific and research perspectives worldwide.

Expanding the conditional instruction used for searching scientific works in the SCOPUS database to the following form allowed for the preliminary identification of language organization levels most frequently described:

(TITLE-ABS-KEY("schizophasia") OR TITLE-ABS-KEY("language AND schizophrenia") OR TITLE-ABS-KEY("language AND psychosis") OR TITLE-ABS-KEY("language AND competence AND schizophrenia") OR TITLE-ABS-KEY("language AND phonetics") OR TITLE-ABS-KEY("language AND phonology") OR TITLE-ABS-KEY("language AND semantics") OR TITLE-ABS-KEY("language AND lexicon") OR TITLE-ABS-KEY("language AND syntax") OR TITLE-ABS-KEY("language AND clause") OR TITLE-ABS-KEY("language AND narration") OR TITLE-ABS-KEY("language AND morphology") OR TITLE-ABS-KEY("language AND discourse") OR TITLE-ABS-KEY("language AND pragmatics")) AND PUBYEAR > 1999 AND PUBYEAR < 2025

Furthermore, it was assumed that only phrases meeting the condition of at least 15 iterations (calculated using the binary method) for the first command and 5 iterations for the second command would qualify for quantitative analysis aimed at creating maps using VOSviewer. To avoid keyword duplication, only author keywords were considered in the map generation process.

The quantitative results obtained from the aforementioned steps was complemented and visualized through a qualitative literature review. This comprehensive analysis aims to delve into the intricate aspects of language description in schizophrenia. The discussion was conducted based on a literature review. The structure of this discussion and review took the following form: after identifying general research trends in the analytical part, they were discussed based on the most frequently described schizotypal features, as distinguished in Thought, Language and Communication (TLC) according to Andreasen [7]. Selected symptoms of schizophrenia, chosen due to the number of new publications addressing them, were presented in the context of currently appreciated and prospective research methods.

Results

The co-occurrence of keywords in the selected set of publications (Figure 1) allowed for the identification of 76 items organized into 7 thematically diverse clusters. The arrangement of keywords within these clusters clearly indicates a focus on language characterization related to describing the essence of schizophrenia and various psychoses. Notably, the cluster containing the keyword “language” ranks third in terms of richness, while also housing keywords with the highest Total Link Strength (TLS): “schizophrenia” (TLS = 834) and “psychosis” (TLS = 817). The keyword “language” itself ranks third in terms of both Occurrences and TLS.

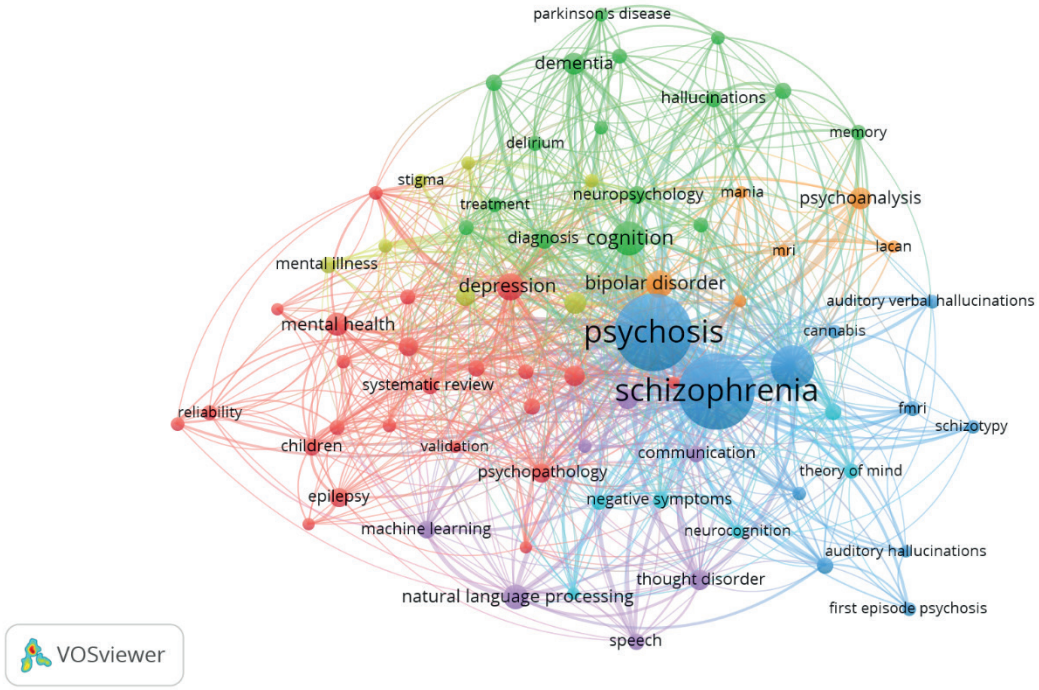

Fig. 1 Cluster and keyword connections map in the Network Visualization view.
Source: The figure was generated using VOSviewer software based on the database collected from SCOPUS.

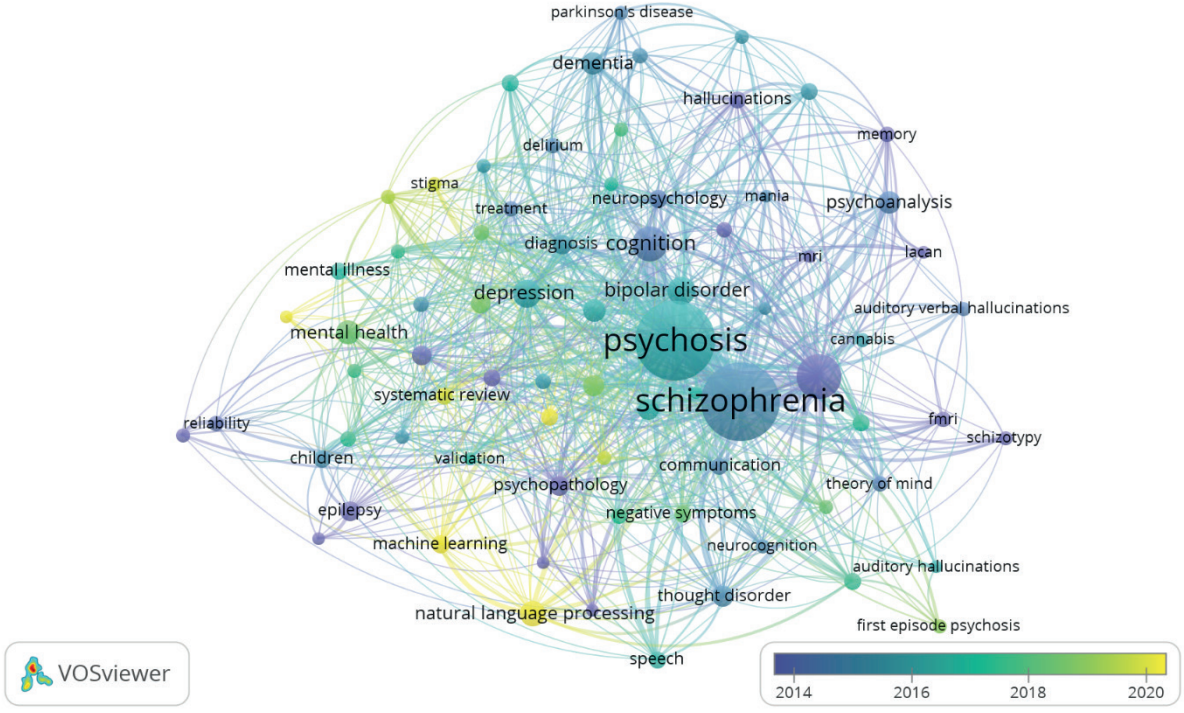

Fig. 2 Keyword co-occurrence map considering the peak interest over time in the Overlay Visualization view.
Source: The figure was generated using VOSviewer software based on the database collected from SCOPUS.

The strong connections between the keyword "language" and other keywords are evident, particularly within the fifth cluster. These connections include derivatives of the keyword, such as "speech" and "communication." All of these terms are closely linked to descriptions of language research methodologies, stemming from the broader field of artificial intelligence. Notably, natural language processing (NLP) and machine learning (ML), despite their shorter history compared to other keywords, stand out on the map with significant TLS and Occurrences values: "NLP" (TLS = 102, Occurrences = 56) and "ML" (TLS = 70, Occurrences = 31).

Schizophasia, characterized by language competence breakdown in various mental disorders (not limited to schizophrenia alone), is most prominently associated with schizophrenia. This alignment is reflected in

the map, where advanced statistical methods for describing language behaviors are primarily utilized in characterizing schizophrenia and psychosis — especially during the first episode of psychoses. Expanding the observation reveals that NLP and ML are also significantly employed in language research related to depression and bipolar affective disorder.

Applying a time filter to the map of keywords related to schizophasia (Overlay Visualization) identifies trends specific to recent years in schizophrenia language research (Figure 2). Over the past five years, NLP and ML have emerged as prominent trends. These methods serve both as methodological tools in ongoing research and as subjects of investigation regarding their potential for describing linguistic biomarkers of mental disorders using artificial intelligence.

Quantitative neuro- and psycholinguistic analysis of specific speech disorders within the schizophrenia group, closely aligned with real-time observations, is now feasible thanks to NLP methods. Advances in natural language understanding and processing enable efficient sampling of language material and identification of variables relevant to various aspects of central nervous system activity and cognition. Consequently, objective linguistic biomarkers for schizophrenia and other mental disorders can be developed. Interest in utilizing NLP and ML for schizophasia description extends beyond technical advancements and aligns with socio-clinical needs [8].

The examination of keyword nodes related to schizophrenia, linguistics, artificial intelligence, and data science reveals an increasing interest in quantifying semantic, grammatical, and acoustic language properties. Researchers from interdisciplinary backgrounds strive to propose objective methods for identifying linguistic indicators of schizophrenia while considering their potential clinical applications.

Over the past 25 years, research on lexicon and semantics has remained consistently relevant. Alongside syntactic studies (after thorough analysis of co-occurring keywords and broader qualitative semiotics), these topics constitute some of the most extensively explored aspects related to schizophasia. While monologue narration research has decreased, investigations into individual levels of structural language organization, their interactions, and critical discourse analysis have gained prominence. The former primarily focuses on schizophrenia, while discourse analyses serve both as research subjects and methodological tools. Additionally, the growing influence of philosophy and language logic in schizophrenia research is justified not only by methodological considerations but also by ethical concerns related to autonomy and linguistic self-determination in individuals with schizophrenia, especially in communication contexts. Logic, as a fundamental discipline, plays a crucial role in language, statistics, and data science research.

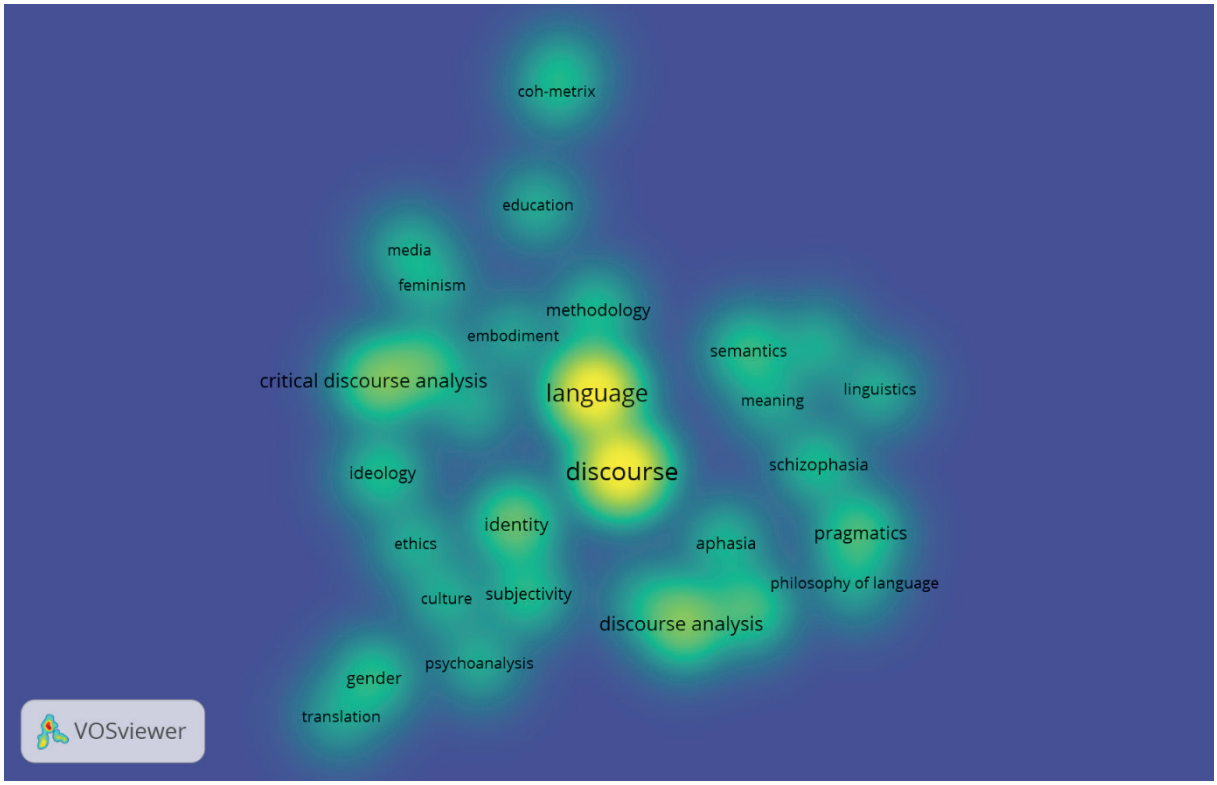

Fig. 3 Density map of key terms representing linguistic properties in the Density Visualization view.
Source: The figure was generated using VOSviewer software based on the database collected from SCOPUS.

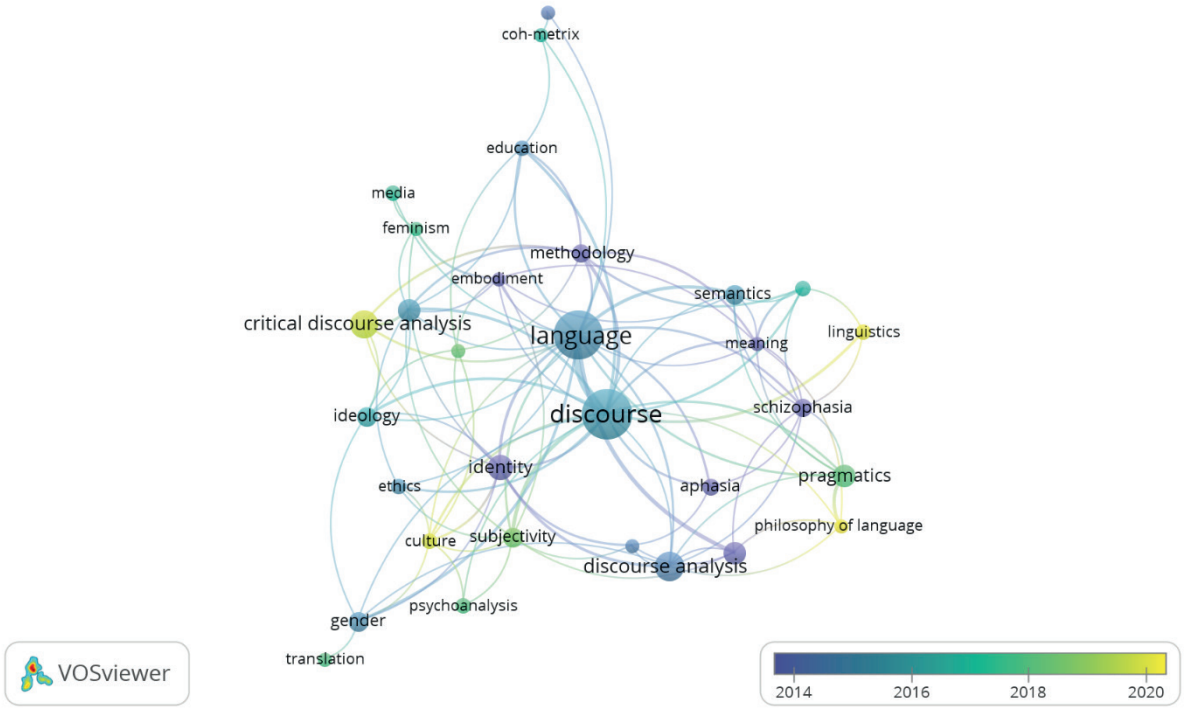

Fig. 4 Keyword co-occurrence map representing linguistic properties, considering the peak interest over time in the Overlay Visualization view.
Source: The figure was generated using VOSviewer software based on the database collected from SCOPUS.

The primary themes within individual components of the language system include semantics, syntax, and pragmatics. Notably, research is increasingly effective in exploring objects at the intersection of at least two levels of language organization. Investigations into morphotactic and morphophonological rules, connotation and denotation issues, and the relationship between complex linguistic constructs and acoustic features in schizophrenia are successfully undertaken within psychology, psychiatry, and speech therapy.

The development of computational language analysis methods has directly impacted our understanding of verbal auditory hallucinations in schizophrenia. Qualitative analyses of keyword associations in this area reveal a trend toward dynamic research growth, particularly in early psychosis relapse detection.

The analysis of keyword nodes from the seventh and fifth clusters of the initial filtering query, with particular emphasis on the chronological view, revealed a specific interest in the following linguistic symptoms of psychiatric disorders, especially schizophrenia: discourse coherence, poverty of content, syntactic disturbances, features of metaphorical language content, referential coherence, echolalia, and other phonetic-acoustic symptoms.

Discussion

The model of integrating multidisciplinary research focused on characterizing language disorders in mental illnesses has equipped logopedics, psychiatry, and neuropsychology with historical insights into pathological language phenomena in schizophrenia over the past decades [9]. The adoption and development of Eugen Bleuler's and Emil Kraepelin's thoughts, which considered associative disturbances as significant core symptoms of schizophrenia [10], led to their inclusion in contemporary classifications and manuals (ICD, DSM). Expanding language analysis in the context of schizophrenia from a linguistic-medical perspective, utilizing various linguistic tools, has contributed to the ongoing development of a set of mechanisms and symptomatic sequences related to various hebefrenias.

The advancement of individual components of artificial intelligence over the past decades now allows for the creation of advanced computational methods capable of investigating linguistic phenomena. These methods include structuralist-defined rules, inter-word semantic relationships, narrative analysis, and partially prosody-expressed intentions. Leveraging corpora specific to different mental disorders enables precise recognition and categorization of linguistic features. Identifying linguistic indicators in individuals with mental illnesses becomes more effective through natural language processing (NLP), which employs advanced statistical methods and computational principles.

Implementing self-improving algorithms based on experience, fine-tuned using advanced mathematical and informatics models, opens up possibilities for extensive research on the organization of language production

at each level. Prototyping mathematical models based on existing linguistic research, correlated with binary system translations representing their structural-semantic values, is feasible. Developing an initial algorithm using neurolinguistic-clinical-programmatic experiences seems to be the most appropriate approach to address contemporary research needs. Such interdisciplinary collaboration could lead to software automating data acquisition and analysis, with the ability to refine its own system through carefully guided initial input data selection (the so-called "training set"). While this approach may raise concerns about the objectivity of machine learning-based research, it also safeguards against generating results that are incomprehensible or inconsistent with fundamental scientific-practical observations. It is essential to consider the specific form of language system breakdown in mental disorders.

The ability of self-learning machines to dynamically modify themselves and tackle complex problems while considering interdependencies is a fundamental characteristic that justifies their application in modeling linguistic features. These features are particularly relevant in the context of specific language-related disorders, such as the entire spectrum of mental illnesses. Consequently, classification methods that focus on creating new terms and explaining relationships between them play a crucial role. Some of these methods include rule sets, Bayesian approaches, decision trees, and a combination of inductive and analytical learning, all aimed at representing results in various forms: taxonomic, finite automaton transition functions, formal grammar, or visualized semantic network structures.

In the design of neurolinguistic research on language in mental disorders, one critical decision involves choosing a computationally understandable representation of lexemes. This representation should be suitable for self-learning machines during processing and classification of natural language components [11]. Alongside classical methods like assigning words to consecutive integers (commonly found in dictionaries) and one-hot vectors (which encode words as binary sequences), "word embeddings" (WE) have gained prominence. WE's primary potential lies in eliminating the distance between synonymous or semantically related words and reducing the number of zeros, directly enhancing computational and memory efficiency. This approach aligns with distributional semantics theory, developed in linguistic research during the mid-20th century by John Firth. According to this theory, we can determine semantic similarity between words based on the contexts in which they appear [12]. The key argument supporting the use of WE in translation and classification processes is its non-binary structure. In this approach, lexemes are represented as complex vectors of real numbers (typically 100- or 300-dimensional), with the final size defined by the researcher. Neurolinguistic studies focused on clinically analyzing language behaviors, utilizing WE vectors, are made possible, in part, by the word2vec algorithm developed in 2013. Based on input data provided to the machine, this algorithm iteratively learns word representations using WE [13]. As iterations progress, vector quality improves, making it highly valuable for clinical research. Importantly, from a linguistic perspective, it also enables automatic recognition of the fact and degree of semantic (or pragmatic) similarity between word pairs, leveraging shallow neural networks. Tuning the weights of hidden layers in the network used by WE becomes the foundation for this utility [14].

Applying this algorithmic model without modifications poses the risk of research errors due to its limited ability to interpret polysemy or words not included in the training dataset. Attempts to address such issues have been made repeatedly over the past five years, and the choice of a solution model may directly depend on research objectives and the clinical-linguistic context influencing studies on schizophasia and/or language [15]. In the case of analyzing narrative structures with "split" language competence, architectural solutions like transformers (e.g., BERT from 2019) can be relevant [16].

Semantic relationships between words, represented as vectors, can be further specified by calculating the scalar product of at least two vectors. Their analogy is then measured by the cosine of the angle between them. If two vectors are aligned in a similar direction in a multidimensional space, there is a high likelihood of their proportional semantic closeness. The distance from cos(0) to the value 1 for two vectors, according to the adopted model, reflects their similarity.

The abundance of potential data identifying cognitive processes, neurophysiological interactions, and ways of thinking accessible through narrative analysis offers invaluable research opportunities. However, there is also a risk of underutilizing these resources. For analyzing data from large language corpora, an initial approach involves leveraging machine learning methods based on deep and shallow neural networks.

The previously mentioned methods for translating linguistic material into multidimensional matrices (particularly word2vec) hold potential for research endeavors focused on characterizing discourse coherence in schizophrenia [17]. Schizophrenia's language disturbances constitute one of its core features. To describe the coherence of statements, it is reasonable to support the aforementioned method with an analysis of relationships between content and the concepts contained

within. This can be achieved by applying techniques such as latent semantic analysis (LSA) and Global Vectors for Word Representation (GloVe) within the realm of hidden semantic analysis [18, 19]. These techniques would automatically represent each sentence as a vector and compare the closeness of neighboring sentence values using cosine distances [20].

Additionally, descriptions of content poverty can also be based on advanced vector analyses (including those mentioned earlier) [21]. By presenting and analyzing the same data, we can compute semantic density, which involves calculating vectors necessary for reconstructing sentence meanings [22]. Word2vec further enables the labeling of words as literal or metaphorical, leveraging neural networks. This opens up the possibility of describing metaphorical features in language generated by individuals with mental disorders [23]. Gutierrez demonstrated this approach based on the description of a first psychotic episode [24].

Syntax complexity studies pose another significant challenge, and their approach is essential for describing cognitive changes during the breakdown of competence. When describing schizophasia using AI, it becomes necessary to address syntax. Corcoran suggested using part-of-speech (POS) syntactic analysis, building on the previous work by Bedi and colleagues [25, 26]. Automatically measuring phrase structure, sentence length, and part-of-speech class frequencies provides sorted data for further classical linguistic and clinical analyses. POS tagging allows for comparing the number of shared morphological roots across different sentences, and tools like Coh-Metrix could be useful for assessing the referential coherence of schizophrenic texts [27].

The context of research on searching for biomarkers differentiating psychiatric disorders in clinical-scientific practice using NLP has also shown potential beyond the use of acoustic material in neuropsychiatric descriptions and diagnoses. The specific nature of prosody and all suprasegmental features of speech, measurable by acoustic-auditory methods, in the course of schizophrenia is one of the most frequently indicated in studies on schizophasia. It affects not only the emotional expression of the text but also the segmental layer of the speech itself, as evidenced by symptoms like clanging and echolalia, implying the necessity of successive words in a text having a common phonetic pattern or sound resemblance, regardless of the logic of the text. Research on the potential of setting a psychiatric differential diagnosis using NLP has demonstrated the ability to differentiate based on advanced analysis of acoustic features of speech (emotionally regulated) of depression, bipolar disorders, and schizophrenia [28].

Conclusions

The significant increase in research on language in schizophrenia and other mental disorders in recent years is closely tied to the development of broadly understood artificial intelligence, particularly the utilization of advanced computational methods, data science, NLP and ML. These methods find application in both creating research methodologies focused on describing schizophasia and as subjects of investigation themselves.

1. Embracing AI models in linguistic research, aimed at describing and characterizing language behaviors following the breakdown of language competence due to various mental disorders, allows for defining and developing new paradigms in schizophasia analysis.
2. Ethical considerations are crucial, and researchers utilizing ML should be prepared for limitations arising from partial analysis and potentially false interpretations.
3. Clear domain-specific regulations are essential. In formulating final research conclusions, those eager to leverage knowledge from artificial intelligence should also align with research approaches consistent with critical rationalism, as advocated by Popper, emphasizing that inductive reasoning can only assume a falsifiable character.
4. Developing a set of linguistic biomarkers for mental illnesses using broadly understood artificial intelligence requires a multidisciplinary research team, integrating experts from psychiatry, neurolinguistics, neurologopedics, and artificial intelligence engineering. The preparation for conducting such studies directly depends on the language used as the analytical material for the research group.
5. Interest is growing in complex components of the language system that undergo breakdown in schizophrenia, such as disturbances in applying morphonological rules or connotations [29]. Despite significant technological advancements, many studies on mutual lexeme capabilities for opening discourse spaces in both normative and psychopathological contexts have primarily employed linguistic methods for clinical description [29].
6. The development of schizophasia research prospects is particularly evident at the lexical, semantic, and syntactic levels. Additionally, integrating linguistic behavioral research with affective and neurophysiological variables is essential.

Conflict of interest

The authors have declared no conflict of interest.

References

1. Czernikiewicz A. Przewodnik po zaburzeniach językowych w schizofrenii, Warszawa; Instytut Psychiatrii i Neurologii: 2004
2. Grabias S. Teoria zaburzeń mowy. Perspektywy badań, typologie zaburzeń, procedury postępowania logopedycznego, W: Grabias S, Kurkowski M. (red.) Logopedia teoria zaburzeń mowy, Lublin; UMCS: 2012: 15–73.
3. Woźniak T. Zaburzenia mowy w schizofrenii, W: Grabias S, Kurkowski M. (red.) Logopedia teoria zaburzeń mowy, Lublin; UMCS: 2012: 645-678.
4. Van Eck NJ, Waltman L. Vosviewer: A computer program for bibliometric mapping. ERIM Report Series Reference No. ERS-2009-005-LIS. 2009
5. Waltman L, van Eck NJ. Text mining and visualization using VOSviewer. ISSI Newsletter. 2011;7(3)
6. Waltman L, van Exk NJ, Noyons ECM. A unified approach to mapping and clustering of bibliometric networks. Journal of Informatrics. 2010;4(4)
7. Andreasen N, Grove W. Thought, language, and communication in schizophrenia: Diagnosis and prognosis. Schizophr. Bull. 1986;12:348–358
8. de Boer J. Voices: a clinical computational psycholinguistic approach to language and hallucinations in schizophrenia spectrum disorders. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. Groningen 2023. doi. org/10.33612/diss.668440255
9. Czernikiewicz A, Woźniak T. Diagnoza psychologicznych zaburzeń mowy W: Milewski S, Czaplewska E. (red.), Diagnoza logopedyczna, Gdańsk 2012
10. Bleuler E. Dementia praecox oder Gruppe der Schizophrenien. Leipzig: Deuticke; 1911.
11. Hernandez HC. Natural Language Processing Markers for Psychosis and Other Psychiatric Disorders: Emerging Themes and Research Agenda From a Cross-Linguistic Workshop. Schizophr. Bull. 2023;49(2):86-92
12. Firth JR. Applications of general linguistics. Philological Society. 1957;56(1):1-14
13. Cabuk T, et al. Natural Language Processing for defining linguistic features in schizophrenia: A sample from Turkish speakers. Schizophrenia Research. 2024;266:183-189
14. Ehlen F, et al. Linguistic findngs in persons with schizophrenia – a review of the current literature. Frontiers in Psychology. Psychology of Language. 2023;14, doi.org/10.3389/ fpsyg.2023.1287706
15. Jeong L, et al. Exploring the Use of Natural Language Processing for Objective Assessment of Disorganized Speech in Schizophrenia. Psychiatric Research & Clinical Practice. 2023;5(3) doi.org/10.1176/appi.prcp.20230003
16. Lin PY, et al. Deep Learning for schizophrenia classification based on natural langiage processing – A pilot study. Research Square, 2024 PREPRINT doi.org/10.21203/rs.3.rs-3836497/v1
17. Ryazanskaya G, Khudyakova M. Automated Analysis of Discourse Coherence in Schizophrenia: Approximation of Manual Measures. Language Resource and Evaluation Confernce 15-16.05.2020, Proceedings
18. Just SA, et al. Validation of natural language processing methods capturing semantic incoherence in the speech of patients with non-affective psychosis, Sec. Schizophrenia, 2023;14 doi. org/10.3389/fpsyt.2023.1208856
19. Corona-Hernandez H, et al. Assessing coherence throught linguistic connectivities: Analysis of speech in patients with schizophrenia-spectrum disorders Schizophrenia Research. 2023;259:48-58 doi.org/10.1016/j.schres.2022.06.013
20. Corcoran CM, Carillo F, Fernandez-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CF, Cecchi GA. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry. 2018;17:67-75 doi: 10.1002/wps.20491
21. Guillaume BF, et al. Poverty and inequality in real-world schizophrenia: a national study. Sec. Public Mental Health. 2023;11
22. Rezaii N, Walker E, Wolff P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. NPJ Schizophrenia. 2019;5(9) doi:10.7910/DVN/ K9WKPV
23. Glaz LA, et al. Machine Learning and Natural Language Processing in Mental health: Systematic Review. Journal of Medical Internet Research. 2021;23(5) doi: 10.2196/15708
24. Gutierrez ED, Corlett PR, Corcoran CM, Cecchi GA. Using automated metaphor identification to aid in detection and prediction of first episode schizophrenia, in EMNLP 2017 – Conference on Empirical Method in Natural Language Processing, Proceedings, 2017
25. Corcoran CM, Carillo F, Fernandez-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CF, Cecchi GA. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry. 2018;17:67-75
26. Bedi G, Carrillo F, Cecchi GA, Fernandez-Slezak D, Sigman M, Mota NB, Riberto S, Javitt DS, Copelli M, Corcoran CM. Automated analysis of free speech predicts psychosis onset in high-risk youths. NPJ Schizophrenia. 2015:1
27. Gupta T, Hespos SJ, Horton WS, Mittal VA. Automated analysis of written narratives receals abnormalities in referential cohesion in youth at ultra high risk for psychosis. Schizophrenia Research. 2018;192:82-88
28. Huang KU, Wu CH, Su MH, Kuo YT. Detecting unipolar and bipolar depressive disorers from elicited speech responses using latent affective structure model. IEE Trans Affect Comput. 2020;11(3):393-404
29. Lipski W. Zaburzenia konotacji w schizofrenii. Lublin; UMCS: 2018

Corresponding author

Michał Mazur
e-mail: mazur.logopeda@gmail.com
Department of Clinical Neuropsychiatry, Medical University of Lublin,Poland

Otrzymano: 26.04.2024
Zrecenzowano: 10.06.2024
Przyjęto do publikacji: 25.06.2024