How can a corpus analysis tool help describe mentalizing skills in the speech of individuals with schizophrenia?

https://doi.org/10.48040/PL.2023.1.1

Individuals with schizophrenia reportedly exhibit severe speech and mentalizing difficulties. The study of schizophrenic speech is a multifaceted research field, including a functional linguistic approach. In order to understand the nature of patients’ linguistic dysfunction, the primary task is to identify the occurrence of linguistic disturbances during mentalizing processes. The study being part of an interdisciplinary research is based on guided interviews related to Hemingway’s short story entitled The End of Something, and it primarily focuses on identifying and classifying mental state language reflecting mentalizing processes. The present corpus includes 20 guided patient interviews which were digitally recorded and transcribed. The qualitative analysis targeting the description and classification of recurring mental state language patterns related to mentalizing skills was performed with the help of the Sketch Engine corpus analysis tool. The program assisted in the identification and classification of collocations associated with the interviewees’ mental processes. The results suggest that the language use of individuals with schizophrenia reflects mild microlinguistic impairments but more severe dysfunctions at the macrolinguistic level. Hopefully, the findings can offer some possible indications for psychotherapists on how to detect and interpret characteristic linguistic impairments and improve mentalizing capacities in this patient group.

Keywords: conversation analysis, discourse analysis, mentalization, mental state language, schizophrenic speech

Introduction

Since the second half of the 1990s, a large number of studies investigating the mentalizing skills of individuals with schizophrenia have indicated that these skills are reportedly impaired or underdeveloped (Thibaudeau et al., 2017) in this group; this suggests that these individuals are likely to face difficulties detecting and interpreting others’ mental states (Dodell-Feder et al., 2013). These deficits can create substantial handicaps for these individuals on the level of communication and interpersonal skills. As a consequence, individuals with schizophrenia cannot adequately take part in social life, their standard of living is low, and they are exposed to marginalization in all aspects of life. Hence, communication impairments associated with schizophrenia come into the focus as a central issue to investigate in order to help these individuals optimize their quality of life and functioning in society on both personal and professional levels (Joyal et al., 2016).

Recent studies aim to establish appropriate speech and language therapy (SLT) as part of rehabilitation for patients with schizophrenia by improving their pragmatic or expressive discursive skills. Evidence shows that certain areas of language can be treated through therapy (Joyal et al., 2016). However, it is difficult to define the type of approach that should be implemented to treat language impairments in schizophrenia because the linguistic characterization of schizophrenia depends on which aspects of speech are affected (Joyal et al., 2016). The impairment of speech production typically refers to formal thought disorder, anomalies in speech perception are associated with verbal hallucinations, and impaired speech content may point to delusions (Hinzen–Rosselló, 2015). Previous studies have outlined that schizophrenia can be clearly indicated by certain communication impairments, and more
specifically in the areas of pragmatics and discourse understanding (Joyal et al., 2016). Furthermore, pragmatic deficits in schizophrenia have been associated with impairment in the theory of mind (ToM) (Brüne–Bodenstein, 2005; Mazza et al. 2007 in Joyal et al, 2016).

Mentalizing skills and language use

Theory of mind (also known as mentalizing skills, mindreading, or mentalization) refers to the capacity and skill to impute mental states (beliefs, knowledge, emotions and intentions) to oneself and others. Mentalizing skills have two major types: cognitive and affective skills, which can be either implicit (automatic) or explicit (deliberative) (Turner et al., 2017). Cognitive mentalizing skills are understood as the capacity to draw inferences from others’ beliefs and intentions, whereas affective mentalizing skills imply drawing inferences from others’ emotions (Sebastian et al., 2012). The success of communication – the pragmatic aspect of language – depends on the extent to which one is able to adequately infer the conversation partner’s beliefs and intentions. A significant part of communication can be successfully realised with the help of linguistic tools. Multiple studies have maintained that improving mentalizing skills cannot be separated from language use, these capacities at least partially depend on it (Hinzen–Rosselló, 2015).

Language is a crucial instrument in the development of mentalizing skills. The linguistic field of pragmatics studies the interlocutors’ intentions and language use, the way utterances are interpreted. Reflecting on others’ mental states and the capacity of adequate language use are closely related processes; hence, the investigation of mentalizing skills and pragmatics may also yield some corresponding results (Herold, 2005). Former studies have shown evidence that language and mentalizing skills are closely related phenomena, which explains how the low mentalizing performance of individuals with schizophrenia is linked to their linguistic-pragmatic skills (Herold, 2005:16). As language is closely related to mindreading capacity, impaired mentalizing skills are likely to present in impaired linguistic expression.

Studying schizophrenic speech requires a multidisciplinary approach involving psychology, psychiatry, neurology, sociolinguistics, and applied linguistics, particularly pragmatics. In a previous study, Joyal et al. (2016) suggested that data on linguistic-pragmatic disturbances in individuals with schizophrenia are scarce. More specifically, few studies have aimed at establishing effective methods to improve speech and language skills in individuals with schizophrenia.

A new method to improve mentalizing skills via reading literary fiction

The research group of the Department of Psychiatry and Psychotherapy (Fekete et al., 2020) adopted a short story task (Dodell-Feder et al., 2013), and later this interdisciplinary research was joined by the present author to contribute to the study from a linguistic aspect. Recent studies have proposed that providing short story tasks for individuals with schizophrenia can be considered one possible and highly effective method for improving mentalizing skills. Reading literary fiction provides the opportunity to ameliorate empathy, and broaden one’s knowledge about others’ thoughts, emotions and intentions. While reading literary fiction, the reader is called to activate mentalizing processes including mindreading and imaging or interpreting specific short story characters (Kidd–Castano, 2013).

The present study is based on Ernest Hemingway’s short story The End of Something. There are specific reasons why the research study applied by Dodell-Feder et al. (2013) relied on Hemingway and this specific short story. On the one hand, this short story has a direct style and a language that is easy to understand: it relates a conflict between a romantic couple, more exactly the end of their relationship as the title suggests. Nick, the male protagonist has a dispute
with his girlfriend, Marjorie, and eventually breaks up with her. On the other hand, Hemingway is recognized for his unique style also referred to as the “iceberg theory”, meaning that he does not explicitly uncover the characters’ inner mental world. Hence, the readers are required to draw mentalizing inferences of the characters’ inner thoughts, emotional states or intentions, necessary for interpreting the story (Dodell-Feder et al., 2013). The linguistic representation of mental states can be realised via expressions linked with perception (e.g. describing experiences), desires (e.g. likes), basic emotions (e.g. angry, scared), social emotions (e.g. envious, guilty), intentions (e.g. wants), or other cognitive terms (e.g. knows) (Langdon et al., 2017).

The present study with a linguistic–pragmatic focus undertakes the linguistic analysis of the language use of individuals with schizophrenia by targeting patients’ mental state talk, which is a set of verbs related to cognition used for attributing cognitive state (e.g., knowing, thinking, understanding, remembering, forgetting, etc.) to others and has commonly been implemented to measure several aspects of mentalizing capacities. (Pinto et al., 2017). The analysis primarily focuses on these mental state expressions which may reflect cognitive and/or affective mentalizing processes suggesting the patients’ underlying mentalizing deficiencies.

Material

Corpus

The corpus of the functional linguistic research involved guided interviews between a PhD student of psychology and individuals with schizophrenia displaying mentalizing difficulties treated at the Department of Psychiatry of the University of Pécs, and separate guided interviews with controls. The interviews were conducted in Hungarian and centred around Hemingway’s short story entitled The End of Something.

The entire corpus consists of 95 guided interviews including 47 individuals with schizophrenia and 48 controls. However, the subcorpus regroups the matched pairs of 20 patients and 20 controls upon the recommendations of a professional psychiatrist. Hence, the criteria of selection consisted of the socio-demographic data including age, gender and education level. As a first step of the functional linguistic analysis, the focus of the research has been narrowed down to the patient group counting 13 female and 7 male patients. The age of the patients ranged from 18–70. In terms of education level, the patients’ educational attainment covered the span of elementary school to postgraduate qualification. The patient group showed varied properties in all evaluation criteria.

Methodological procedure

Continuing the research conducted by Fekete et al. (2020), the present study contributed to the interdisciplinary research by pursuing further functional linguistic analysis after digitally transcribing the interviews in Hungarian in a Word format. The subcorpus constituted by the 20 patient interviews was created for the purposes of classifying and describing cognitive expressions associated with cognitive mentalizing skills. The patients’ language use was analysed in order to gain insight into the extent to which it reflects linguistic impairments suggesting mentalizing deficiencies. The present study applied a bottom-up approach combining quantitative and qualitative aspects when analysing the conversation between the experimenter and the patient.

The guided interviews were recorded in .wma files by the PhD student of psychology (investigator). They were originally recorded for the purpose of previous research conducted by psychiatric professionals (Fekete et al., 2020) targeting the assessment of cognitive mentalizing skills of patients with schizophrenia contrasted with healthy participants with no
diagnosis of mental disorder. The study design was approved by the Committee on Medical Ethics, University of Pécs (ethical permit number: 6539). In line with previous research methodology (Dodell-Feder et al., 2013), the research material was translated and adapted to Hungarian conditions. Before reading the short story, participants received the following instructions:

“You are going to read a short story called The End of Something. The story is only a few pages, but take your time reading it. Try to get a sense of what happens and what the relationships are between the characters. After you’ve finished, I’m going to ask you some questions and tape-record your responses. Do you have any questions before we begin?” (Dodell-Feder et al., 2013: 4)

Participants were asked to read the short story in Hungarian prior to the interview. Next, the experimenter asked 14 open-ended questions in a structured format (see Appendix 1). Participants were allowed to consult the text in case they needed to check specific details. Before starting the guided interview, the participants were instructed as follows:

“Now I’m going to ask you some questions about the story. Here is a copy of the questions I’ll be asking so you can read along. For most of the questions, there are no right or wrong answers and the questions can be answered with short responses. We’re also interested in the character’s thoughts, feelings and intentions when it applies to the question.” (Dodell-Feder et al., 2013: 4).

The interviews were structured and guided in the sense that participants were given 14 questions related to the literary work targeting the assessment of their cognitive mentalizing performance. 5 questions were related to comprehension, 1 question investigated spontaneous mental state reasoning (participants were invited to give a brief summary of the plot), and 8 questions were focused on explicit mental state reasoning (inferring the characters’ thoughts, emotions and intentions) (Dodell-Feder et al., 2013).

Quantitative method

The subcorpus incorporating the 20 patients’ interviews was evaluated by Sketch Engine online text analysis tool. After loading the subcorpus selected, the program generated quantitative data under the corpus info tab detecting the type-token ratio, that is counts of tokens (all instances of words occurring in the corpus) and words (distinct types of words).

The frequency of each qualitative category (Word Sketch, Wordlist and Concordance) also displayed quantitative data. However, due to the special focus of the research targeting the description and evaluation of mental state terms, the quantitative aspect of the analysis was proportionally less emphasized than the qualitative profile.

Qualitative method

Mental state terms were categorized and analysed based on three specific tools displayed on the dashboard of Sketch Engine tool that we found most useful for the purposes of the functional linguistic analysis, namely, the Word Sketch, the Wordlist and the Concordance tools.

The Word Sketch tool allowed us to identify collocations and word combinations. It was used for summarizing a word’s grammatical and collocational behaviour. The results were organized into categories, called grammatical relations, such as words that serve as an object of the verb, words that serve as a subject of the verb, words that modify the word, etc. Given a specific expression such as the verb ‘tud’ “knows”, Word Sketch generated a list of its modifiers, its pronominal objects and subjects, and its usage patterns including, for example, the present

1 https://www.sketchengine.eu/
tense form, singular or plural form, or conjugated forms in the first-second and third person singular or plural variants.

With the help of the Wordlist, it was possible to establish the frequency of certain expressions. Practically, in this function, we could filter for words (multiple conjugated forms or a specific word, e.g. the lemma ‘tud’ “knows” could have ‘tudja’ “she knows it” or ‘tudta’ “she knew it” as different word forms), lemmas (basic forms) or specific parts of speech such as nouns, verbs, adjectives, and adverbs. We could also filter for words beginning, ending, or containing certain characters. In the present analysis, the criteria of selection were words and verbs.

Finally, in the Concordance tool, we could search for examples of use in the context. The search always started from the beginning of the corpus and the concordance lines were displayed in the order in which they were found in the corpus. To test the Concordance function, “the negative epistemic marker claiming insufficient knowledge” (Doehler, 2016) ‘nem tudom’ (“I don’t know”, hereinafter referred to as IDK token) was selected as the point of reference and analysis criterion because this expression showed the highest occurrence from the point of view of word frequency in the corpus. The excerpts were based on the Hungarian transcripts and were also provided in English maintaining the conventions regarding the orthography of the transcripts. IDK tokens of the selected transcript excerpts were labelled in bold.

Results

Quantitative results

The subcorpus incorporating the 20 patient interviews was evaluated with Sketch Engine online text analysis tool for quantitative data. Based on the software, the corpus analyses included 2,959 tokens (all instances of words occurring in the corpus) and 2,923 words (distinct types of words). In the selected corpus, in total, 76 IDK tokens could be identified. The frequency of each qualitative category (Word Sketch, Wordlist and Concordance) also displayed quantitative data.

Qualitative results

This section presents the results of the three categories (Wordlist, Word sketch and Concordance) investigated as part of the qualitative analyses.

The Wordlist tool was tested primarily for the categories of words and verbs. When searching for the frequency of words, the word ‘nem’ “no” showed the highest occurrence (180 instances) which, when investigating its example of use in the context in the Concordance tool, was most commonly followed by the verb ‘tud’ “knows” in 74 instances. After excluding non-words such as the definite articles ‘a/az” “the” or conjunction ‘hogy” “that” or demonstrative pronouns ‘az/ez, azt/ezt” “this/that”, expressions related to cognitive mentalizing processes such as ‘szerintem” “in my opinion” (n=32), ‘lehet” “maybe” (n=23), ‘talan’ “perhaps” (n=16), ‘konkrétan” “precisely” (n=11) and ‘valószínűleg” “probably” (n=10) showed the highest incidence. In case of filtering for verbs, the most frequent items were ‘tud” “knows” (n=102), ‘gondol” “thinks” (n=41), ‘emlékszik” “remembers” (n=13), ‘akar” “wants” (n=13), ‘érez” “feels” (n=10), ‘lát” “can see” (n=9), ‘szeret” “loves” (n=8), ‘ért” “understands” (n=7), ‘un” “is bored with” (n=6), and ‘hisz” “believes” (n=5).

As a next step, in Word Sketch, after typing in the cognitive mentalizing verb ‘tud” “knows” in the search field, the tool generated four categories regrouping the distinct modifiers of ‘tud’, the pronominal object of ‘tud’, the pronominal subjects of ‘tud’, and most usefully its usage patterns. Among the most common modifiers, it could be observed that the verb ‘tud” was most commonly displayed in its negated form ‘nem tudom” “I don’t know”, (n=74)
complemented by the conjunction word ‘hogy’ “what/if”, ‘pontosan’ “exactly”, and ‘miért’ “why”. Regarding usage patterns, a special emphasis was placed on comparing the frequency of 1st person singular and 3rd person singular forms. Altogether 85 items of 1st person singular forms were detected in contrast to only 10 items of 3rd person singulars.

Finally, in Concordance, of 74 instances, distinctive uses and functions of ‘nem tudom’ (IDK tokens) were identified, based on the context.

Example 1.

**nem tudom** hogy mire gondol hogy nem kapnak
‘I don’t know what he means when he says “They [the fish] aren’t striking”’.

Example 2.

**nem tudom** miért ül háttal
‘I don’t know why she is sitting with her back toward him’

Example 3.

**nem tudom** miért nem megy gyalog
‘I don’t know why he doesn’t go by foot’

Example 4.

**nem tudom** hogy raktak belőle tűzet
‘I don’t know how they made a fire’

Example 5.

**nem tudom** pontosan hogy hölgy vagy férfi
‘I don’t know exactly if it’s a he or a she’

Example 6.

ezt **nem tudom** jó menjünk tovább
‘well then [I don't know] let's move on’

Example 7.

**nem tudom** mert **nem tudom**
‘I don’t know because I don’t know’.

Example 8.

DR: Miért nem mer Nick Marjoriere nézni?
Why is Nick afraid to look at Marjorie?

PT: Ezt **nem tudom**
I don’t know [this]

Example 9.

DR: Mit érez Nick amikor azt mondja hogy “Ó menj el Bill menj el egy kis időre”?
‘What is Nick feeling when he says, “Oh, go away, Bill! Go away for a while!”?’

PT: Talán hogy gondolkodjon és és azért hogy **nem tudom** ezt nem nem értem ezt a Bill hogy megjelenik
‘Maybe to think it over and and also to I don’t know [I don’t understand] that this Bill shows up’
Discussion

The present study introduced a functional linguistic methodology assisted by Sketch Engine corpus analysis tool targeting the analysis of linguistic mentalizing impairment in persons with schizophrenia within guided interviews structured around a specific literary work. The Wordlist, Word Sketch and Concordance functions proved to be potentially suitable for identifying and classifying mental state expressions in the transcribed patient interviews. The term ‘nem tudom’ (IDK) was noticeably one of the most commonly used phrases. The examples illustrated the relation between mentalizing deficits and linguistic disturbances in the above categories.

The Wordlist function helped us identify which word forms or verbs showed the highest incidence. In case of word forms, the negation word ‘nem’ “no” was the most common expression, followed by terms related to cognitive mentalizing processes, expressing the speaker’s uncertainty or hesitation. Interestingly, when filtering for verbs, the list predominantly represented cognitive and/or affective mentalizing verbs such as tud “knows”, gondol “thinks”, emlékzik “remembers”, akar “wants”, érez “feels”, lát “sees”, szeret “loves”, ért “understands”, un “is bored with”, and hisz “believes”. These results suggest that the patients made an effort to infer the characters’ inner mental states or refer to their own mental states, although not correctly, by showing signs of uncertainty or lack of knowledge.

An important finding in the Word Sketch function was the significant disproportion between usage patterns of IDK in the 1st person singular or 3rd person singular forms. This result indicates that patients may have severe difficulty inferring others’ mental states, but have minor difficulties when reflecting on their own inner mental world.

Sketch Engine tool provided the most useful assistance in the Concordance function when investigating the diverse uses and functions of the phrase ‘nem tudom’ (IDK) in context. As the examples suggest, IDK is a recurring feature in patients’ language use as they indeed have difficulty interpreting the characters’ thoughts (Ex. 1), motivations (Ex. 2), intentions (Ex. 3), do not understand the process of events (Ex. 4), are unable to identify the characters’ identity or specific role in the short story (Ex. 5). These instances may indicate that patients’ mentalizing skills are impaired when attributing mental state to others, as was described by Fekete et al., 2020.

Furthermore, IDK was also used by the patients to express their intention to move on, skip the question and steer away or definitively avoid the topic (Ex. 6, 7, 8), hence employing an avoidance or resistance strategy possibly in order to conceal their uncertainty or lack of knowledge. In Ex. 8, the patient answers a question related to explicit mental state reasoning Why is Nick afraid to look at Marjorie?, which is directed to the character’s feelings. In this example, the patient either could not or did not want to respond to this question, but this may also indicate the patient’s inability to attribute mental states to others.

The preliminary analyses assisted by Sketch Engine text analysis tool confirmed the results of the short story task assessing theory of mind capacities in adults with schizophrenia (Dodell-Feder et al., 2013) in that cognitive mentalizing skills are clearly reflected in language use, for example, the patients expressed uncertainty on several occasions, which was linguistically realized with the help of cognition terms applied in negation forms (e.g. nem tudom miért ül neki háttal ‘I don’t know why she is sitting with her back toward him’). As a result, the patients’ language use expressed hesitation and uncertainty, combined with a less confident content knowledge (Ex. 9). Moreover, patients may have poorer capacities to draw logical inferences of the characters’ thoughts, intentions, beliefs, desires and other mental state contents.

In addition to cognitive mentalizing impairment, affective dullness (Ex. 8 and 9), and further typical features characterizing the speech of individuals with schizophrenia could be identified, including derailment, that is loose associations, steering away from the topic (e.g.
Ex. 6, 7), and deficiency of content of speech and thought (Ex. 1–9) (McKenna – Oh, 2005). Therefore, when patients use IDK expressions, they presumably try to compensate their impaired cognitive mentalizing skills, or employ avoidance strategies to conceal their insufficient knowledge, uncertainty or deficient thought when they do not know, do not understand or do not remember details related to the characters’ inner mental states (thoughts, beliefs, intentions, or desires).

Future research should investigate this aspect on an extended corpus, including controls as well. Additionally, the quantitative results of the token and word counts may be significant to describe vocabulary variability.

Conclusion

This research shows the necessity for further research in order to confirm the preliminary results and to be able to describe mental state language associated with patients’ mentalizing skills in a more nuanced way with the help of Sketch Engine corpus analysis tool applied to an extended corpus involving healthy controls. The findings may contribute to a deeper understanding of the language use of patients with mental disorders from a pragmatic aspect as well as to the success of psychotherapeutic counselling.

References


**Online sources**

Sketch Engine. [https://www.sketchengine.eu/](https://www.sketchengine.eu/)