Schizophrenia Research: Cognition 40 (2025) 100343

Contents lists available at ScienceDirect

Schizophrenia Research: Cognition

journal homepage: www.elsevier.com/locate/scog

Research Paper

Use of reference markers in the speech of people with schizophrenia spectrum disorders: Evidence from two referential communication tasks manipulating common ground with the interaction partner

Amélie M. Achim [a,*], Dominique Knutsen [b], Marc-André Roy [c], Souleymane Gadio [d], Marion Fossard [e]

[a] Département de Psychiatrie et neurosciences, Université Laval, Centre de recherche VITAM and Centre de recherche CERVO, 2480, de la Canardière, Québec, QC G1J 2G3, Canada
[b] Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France
[c] Département de Psychiatrie et neurosciences, Université Laval, Centre de recherche CERVO, 2525, chemin de la Canardière, Québec, QC G1J 2G3, Canada
[d] Centre de recherche VITAM, 2480, de la Canardière, Québec, QC G1J 2G3, Canada
[e] Institut des sciences logopédiques, Université de Neuchâtel, Rue Pierre-à-Mazel 7, CH-2000 Neuchâtel, Switzerland

ARTICLE INFO

Keywords:
Communication
Speech
Language
Schizophrenia
Psychosis
Definite
Indefinite
Pronouns
Common ground
Community membership

ABSTRACT

Introduction: People with schizophrenia spectrum disorders present with language dysfunctions, yet we know little about their use of reference markers (indefinite markers, definite markers, pronouns or names), a fundamental aspect of efficient speech production.
Methods: Twenty-five (25) participants with a recent-onset schizophrenia spectrum disorder (SZ) and 25 healthy controls (HC) completed two referential communication tasks. The tasks involved presenting to an interaction partner a series of movie characters (character identification task) and movie scenes composed of six images (narration task). A manipulation was introduced such that half of the movies could be considered as Likely-Known by the interaction partner, whereas the other half was Likely-Unknown. The analyses focused on the reference markers used to present the movie characters during the tasks.
Results: During the character identification task, the SZ group used fewer names and more pronouns than the HC. During the narration task, the SZ group used fewer names and more definite references when initially introducing the main story characters, while no group effect emerged for subsequent mentions of the characters. The observed effects of conditions were generally present across both groups, except for a lesser adjustment in the use of definite markers when introducing the story characters.
Conclusions: While some group differences emerged, people with SZ were generally sensitive to the manipulation regarding their interaction partner's likely knowledge of the characters. A better understanding of the conditions in which speech production is affected in SZ could help promote more efficient communication.

1. Introduction

People with schizophrenia present with a range of symptoms and deficits, including language dysfunctions. Language production deficits (or speech deficits) have long been recognized in schizophrenia, with multiple reports of disorganized or ambiguous speech (Harvey, 1983; Docherty et al., 1996; de Sousa et al., 2019; Covington et al., 2005). Disorganized speech has been largely studied using the thought disorder approach, in which language disorders are targeted as a way of examining disorganized thinking. Thought disorder symptoms are a hallmark feature of schizophrenia (Andreasen, 1979a; Andreasen, 1979b; Breier and Berg, 1999) and studies focusing on these symptoms have revealed a negative association with social functioning (Marggraf et al., 2020; Bowie and Harvey, 2008), highlighting the importance of identifying and addressing these speech dysfunctions to promote recovery.

While studies focusing on thought disorder symptoms usually rely on subjective ratings based on the global impressions of patient's speech

* Corresponding author.
E-mail addresses: amelie.achim@fmed.ulaval.ca (A.M. Achim), dominique.knutsen@univ-lille.fr (D. Knutsen), marc-andre.roy@fmed.ulaval.ca (M.-A. Roy), souleymane.gadio.ciussscn@ssss.gouv.qc.ca (S. Gadio), marion.fossard@unine.ch (M. Fossard).

https://doi.org/10.1016/j.scog.2024.100343
Received 11 July 2024; Received in revised form 20 December 2024; Accepted 26 December 2024

observed during clinical or research interviews, other approaches have been proposed to more systematically examine speech dysfunctions in people with schizophrenia (Docherty et al., 1996; Rochester and Martin, 1979). For example, the Communication Disturbance Index (CDI) (Docherty et al., 1996) proposes a method in which all the ambiguous words or segments present in transcribed speech samples can be linked to 6 categories of communication disturbances (e.g. vague references, missing information references, etc.), allowing to calculate either a global rate of communication disturbances per 100 words (Docherty et al., 1996), or rates for the six separate categories. Several studies using the CDI have reported higher rates of communication disturbances in the speech of people with schizophrenia as compared to healthy controls (Docherty et al., 1996; Rubino et al., 2011; Docherty, 2005; Docherty et al., 2013), with all six categories from the CDI being affected (e.g. (Docherty et al., 1996). Overall, these results indicate that the speech of people with schizophrenia contains more words or word segments that fail to clearly identify the intended referent (i.e. the entity being discussed).

Reference choices are a fundamental aspect of speech production (Gundel et al., 1993; Fossard et al., 2018; Arnold, 2010), and these choices can be examined with an even finer granularity using theoretic models and methods derived from the psycholinguistic literature. Many characteristics of the references can be examined, and one for which we have good theoretical background is the choices of reference markers, i.e. the different types of linguistic expressions that can be employed to identify the intended referent (the entity being referred to). For example, when referring to a person, we could choose to say "a man" (indefinite marker), "the man" (definite marker), "he" (pronoun), or the name of that person if we know it (name).

From the literature on discourse production, we know that the use of different reference markers is highly influenced by whether or not the speaker believes that the referent is known by their partner (Achim et al., 2017), i.e. whether it belongs to their common ground (Clark, 1996). Common ground is defined as the information that two interaction partners share and are aware of sharing, and according to Clark and Marshall (Clark and Marshall, 1981), it can be established based on *dialogue history* (the referent was mentioned earlier during the dialogue or during previous interactions between the same dialogue partners), based on *visual copresence* (the dialogue partners can see the referent as they speak) or based on *shared community membership* (what the members of a given community can be expected to already know). For instance, even if we did not previously talk about the DSM-5 in this paper, I would assume that most people interested in schizophrenia research would already be familiar with the DSM-5 and that we share some common ground regarding that construct. Importantly, common ground influences speakers' choices of reference markers, with several studies reporting that indefinite references (e.g., "a cat") are favored when referents do not belong to the common ground with the interaction partner while definite references are favored when referents do belong to the common ground (e.g., "the cat") (Fossard et al., 2018; Achim et al., 2017; Champagne-Lavau et al., 2009; Clark and Wilkes-Gibbs, 1986). In other words, dialogue partners rely on their common ground to adjust their choices of reference markers as they interact.

Another branch of the literature on discourse production has revealed that the use of different types of reference markers is also strongly influenced by the accessibility or salience of the referent in the discourse model, which evolves throughout the verbal interaction (Gundel et al., 1993; Ariel, 1996). Indefinite markers are thus preferentially used for referents with low salience in the discourse, notably when introducing a new referent that has not been previously mentioned (Fossard et al., 2018). Definite markers are typically used when later reintroducing a referent that was previously mentioned, and pronouns are largely favored when maintaining a very salient referent in focus (ex: "A man entered the room, and he then took a seat") (Fossard et al., 2018; Achim et al., 2017; Knutsen et al., 2022). Using the right reference markers is thus crucial to ensure clarity of the discourse and effective communication.

While it is well established that people with schizophrenia produce higher rates of unclear references relative to healthy participants (Docherty et al., 1996; Rochester and Martin, 1979; Rubino et al., 2011; Docherty, 2005; Docherty et al., 2013), surprisingly few studies have examined the use of reference markers in schizophrenia. The first study of which we are aware is that of Champagne-Lavau et al. (Champagne-Lavau et al., 2009), that examined the effect of common ground, and more specifically dialogue history, on the choices of references markers in people with schizophrenia. That study relied on a referential communication paradigm (Isaacs and Clark, 1987), in which the participant holds a set of images in a predetermined order, the interaction partner holds the same images in a random order, and the participant's task is to describe each image in turn so that the interaction partner can place their images in the same order. The participants and their interaction partner (a research assistant) did the task five times with the same five abstract images, with a different image order for each turn. This introduced a manipulation of the turn, with a first turn for which we would expect a greater proportion of indefinite references (as the referents are being introduced for the first time and do not yet belong to the partners' common ground) and four subsequent turns for which we would expect a greater proportion of definite references (as participants are now describing referents that have already been mentioned before and are thus part of the partners' common ground). The analyses revealed that people with schizophrenia and controls did not significantly differ on the first turn of the task, but differed on the subsequent turns, such that the healthy participants used a greater proportion of definite references, whereas participants with schizophrenia continued to use a greater proportion of indefinite references. This suggests that the choices of indefinite versus definite references was not properly modulated by the common ground that was established with their interaction partner on the previous turns of the task. As far as we know, in the literature in schizophrenia, no study has yet examined the effect of other factors from which common ground can be established (visual copresence and community membership), nor studied the impact of the relative salience that the referents can have at different point during the discourse (e.g. very salient when just mentioned as the main topic of the previous utterance, and less salient if not recently mentioned).

There are however a few studies that examined the general rates of definite and indefinite references per number of words in the speech of people with schizophrenia relative to healthy controls, based on more classical linguistic tasks that do not involve a joint activity with an interaction partner. In Çokal et al. (Çokal et al., 2018), participants told the story depicted in an eight-picture comic strip, and the participants with schizophrenia with more severe thought disorder symptoms showed a reduced use of definite references in their verbal productions relative the healthy control group, while people with schizophrenia with less sever thought disorder symptoms did not differ significantly from the control group. In another study by Çokal et al. (Çokal et al., 2023), participants had to talk about a series of individual pictures and people with schizophrenia with or without thought disorder symptoms showed a reduced use of definite expressions in their verbal productions, whereas the subgroup with thought disorder also showed a reduced use of indefinite expressions. These two studies also reported that people with schizophrenia made their definite references more often ambiguous to interpret according to trained raters (Çokal et al., 2018; Çokal et al., 2023).

Overall, there are indications that the use of reference markers can be affected in people with schizophrenia, yet many factors known to influence reference choice based on the literature in healthy participants have not yet been examined in people with schizophrenia.

In the current study, we aimed to investigate the use of reference markers in people with schizophrenia as a function of common ground determined by community membership (i.e. what the interaction partner can be expected to already know about the topic being discussed). For this study, we focused not only on the presentation of individual

referents (i.e. separate images, as in Champagne-Lavau et al. (Champagne-Lavau et al., 2009)), but also on a continuous narrative discourse where a strong effect of the accessibility or salience of the different referents can be expected at different discourse stages (e.g. when introducing, maintaining in focus or later reintroducing a referent; (Fossard et al., 2018; Achim et al., 2017). We hypothesized that people with schizophrenia may show a reduced effect of common ground based on community membership on their choices of reference markers (i.e. not adjusting as much as the control group based on their interaction partner's likely knowledge of the referents), especially for the use of definite references or proper names for which such affects have been reported in healthy participants (Achim et al., 2017).

2. Methods

2.1. Participants

The data was drawn from a language corpus used for prior studies in healthy participants or in schizophrenia (Achim et al., 2015, 2017, 2022, 2023). No prior studies with this corpus had however examined the use of reference markers in people with schizophrenia. Twenty-five (25) participants with a schizophrenia spectrum disorder (27.0 years, 22 men[1]) and 25 matched healthy controls (25.1 years, 20 men) were included in this study. The schizophrenia group (SZ) was recruited from the Clinique Notre-Dame des Victoires, a specialized clinic for recent-onset psychosis in Quebec City, Canada. The mean duration of illness was of 2.18 years (range 5 to 67 months). The diagnoses included schizophrenia ($N = 17$), schizoaffective disorder ($N = 6$), and delusional disorder ($N = 2$) according to the DSM-IV criteria (Association, 1994). As further detailed in Table 1, all participants from the schizophrenia group were taking an antipsychotic medication at the time of the testing.

The healthy controls (HC) were selected to best match the schizophrenia group in terms of age, gender and parental socioeconomic status. They reported no history of a psychiatric diagnosis and were taking no psychoactive medication.

The Research Ethics Board of the CIUSSS-CN - neuroscience and mental health division approved the study (project #182-2007) and all participants provided informed consent.

2.2. Referential communication tasks

The protocol included two distinct referential communication tasks, during which the participant and the interaction partner were seated in front of each other, separated by an opaque panel to limit non-verbal communication. As illustrated in Fig. 1, on each trial, the participant held a series of images, and the interaction partner held the same images as the participant but in a random order. The instructions and the images differed between the two tasks (see below), but in both cases the images were drawn from movies that could be considered as either Likely-known or Likely-unknown to the interaction partner (a woman in her twenties) as established based on a series of surveys (the validation surveys are presented in details in (Achim et al., 2017)). For both tasks, the role of the interaction partner was held by the research assistant administering the tasks, always a woman in her twenties who knew all the Likely-Known movies and was trained to act as if she did not know any of the Likely-Unknown movies. A concealment strategy was used to lead the participants to think that the assistant did the two tasks with different images for each participant (see (Achim et al., 2017) for more details).

2.2.1. Character identification task

The first task was a character identification task, in which the participants had to present a single series of 10 images, each depicting a single movie character (all male characters), with five characters from the Likely-Known condition and five characters from the Likely-Unknown condition, organized in a pseudo-random order. Participants were not informed of the conditions. They were simply asked to present each character one after the other so that the interaction partner, who held the same images in a random order, could place her images in the correct order. While the participant presented the different characters, the interaction partner was allowed to ask for clarifications when needed or to provide feedback such as “ok” or “hum-hum”.

2.2.2. Narration task

The second task was a narration task, in which the participants were asked to narrate six different movie scenes, each presented on a series of six images. Each of the six movie scenes involved two different male characters, a main character of interest as well as a secondary character that became visually in focus in images 3 and 4. Having this second character ensured that the three different stage of narrative elaboration be represented in the narrations, including segments where the characters are introduced (introduction stage), segments where the characters are then maintained in focus in the discourse (maintaining stage)

Table 1
Information about the participants.

	Schizophrenia	Healthy controls	Between-group effects
Age (mean, SD)	27.0 (4.6)	25.1 (5.8)	t(48) = 1.28, NS
Gender (Male/Female)	22/3	20/5	$X^2(1)$ = 0.60, NS
Parental socioeconomic status (Mean, SD)	3.5 (1.0)	3.2 (0.9)	t(47) = 1.10, NS
Number of Likely-Known movies that the participants had seen (/5 from the identification task)	4.2 (0.9)	4.6 (0.7)	t(48) = 1.69, NS
Number of Likely-Unknown movies that the participants had seen (/5 from the identification task)	2.8 (1.3)	2.8 (1.6)	t(48) = −1.90, NS
Number of Likely-Known movies that the participants had seen (/3 from the narration task)	2.5 (0.6)	2.6 (0.06)	t(48) = 0.69, NS
Number of Likely-Unknown movies that the participants had seen (/3 from the narration task)	1.6 (1.1)	1.4 (1.0)	t(48) = −0.69, NS
PANSS 5 factors (mean, SD)			
PANSS Positive	14.6 (4.9)	–	–
PANSS Negative	16.0 (6.3)	–	–
PANSS Cognitive/ Disorganisation	9.4 (3.3)	–	–
PANSS Depression/Anxiety	8.3 (2.6)	–	–
PANSS Excitability/Hostility	5.8 (2.3)	–	–
PANSS Total (mean, SD)	61.6 (16.9)	–	–
SOFAS (mean, SD)	56.5 (11.9)	–	–
Antipsychotic medication			
Aripiprazole	N = 9	–	–
Quetiapine XR	N = 5	–	–
Risperidone or Risperidone consta	N = 5	–	–
Clozapine	N = 3	–	–
Paliperidone	N = 1	–	–
Olanzapine	N = 1	–	–
Combination Aripiprazole + Quetiapine	N = 1	–	–
Other medications			
Benzodiazepines	N = 8	–	–
Antidepressant	N = 5	–	–
Lamotrigine	N = 5	–	–
Valproic acid	N = 1	–	–
Lithium	N = 1	–	–

NS = Not statistically significant; SD = Standard deviation; PANSS = Positive and Negative Symptoms Scale (Kay et al., 1987); SOFAS = Social and occupational functioning assessment scale (Association, 1994).

[1] The high proportion of male in the sample (22/25 = 88 %) is very similar to that of our prior studies with participants from the same clinic.

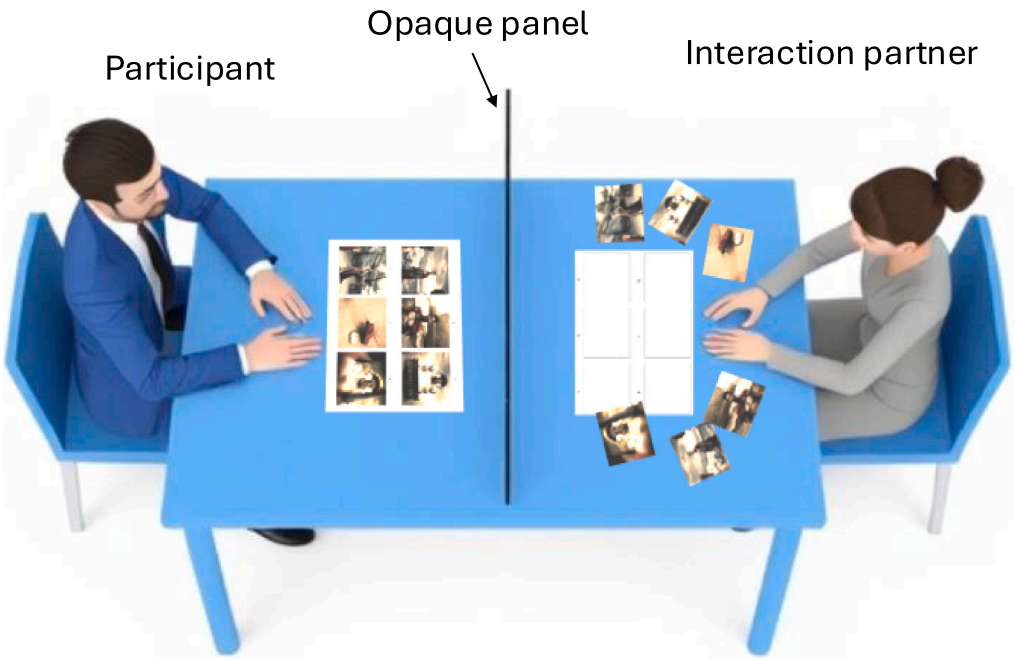

Fig. 1. Illustration of the referential communication procedure used for this study, here with a trial from the narration task.

and segments where the characters are reintroduced after not being the focus in the previous discourse segment (reintroduction stage). Having two characters of the same gender in the different movie scenes was used to increase the referential complexity of the stories, given that the same pronoun (i.e. "he") or noun phrases (e.g. "the guy") could refer to either one of the two male characters, depending on the discourse context.

For each trial, the participant was given a sheet with the six images of the movie scene in a predetermined order while the interaction partner held six separate image cards with the images of the scene, given to them in a random order. While the participant narrated the story, the interaction partner had to place the different images in the correct order. Again, the interaction partner was allowed to ask for clarifications or provide feedback such as "ok" or "hum-hum". Three of the movie scenes were drawn from Likely-Known movies and the other three from Likely-Unknown movies, but the participants were not informed of this manipulation and were simply asked to narrate the movie scenes so that their interaction partner could place the image cards in the correct order.

2.2.3. Participant's knowledge of the movies

Prior to performing the referential communication tasks, participants filled a questionnaire in which they indicated which movies they had previously seen (to retain only the movies they had seen for the analyses, see below).

2.3. Data processing

The tasks were audio-recorded and the verbal productions were transcribed verbatim to facilitate the coding. For the current project, we focused specifically on the reference markers used to present the characters and coded whether it corresponded to a name (e.g. "Leonidas"), an indefinite expression (e.g. "a warrior"), a definite expression (e.g. "the warrior") or a pronoun (e.g. "he"). For the identification task, we extracted a single reference marker per character, focusing on the first mention of the character in the transcript. For example, "The next one is Leonidas from the movie 300" would be coded as a name, whereas "The next one is a warrior named Leonidas" would be coded as an indefinite reference.

For the narration task, the analyses focused strictly on the main character, who was introduced at the beginning of each movie scene (Introduction stage, which occurred once per story). The main character could however be maintained in focus in several subsequent clauses (Maintaining stage) and could also be reintroduced any number of times during the narration (Reintroduction stage). For each mention of the character during the narration, we thus had to identify whether it corresponded to the introduction stage, the maintaining stage or the reintroduction stage. The procedure used to identify the stages of narrative elaboration (Introduction, Maintaining and Reintroduction) is presented in detail in Achim et al. (Achim et al., 2017). Briefly, the narrative speech was first divided into clausal units, and for each clausal unit we determined whether the main character was perceived as being in focus or not by three different research assistants. More specifically, these assistants were asked to rate their subjective appreciation of whether the main character was the most salient (was in focus) in each of the clausal units, and any discrepancies were thereafter resolved by an independent judge to obtain the final ratings. The Introduction stage corresponded to the first mention of the main character for each story. The Maintaining stage corresponded to the cases in which the main character was considered in focus after also being in focus in the immediately preceding clausal unit. The Reintroduction stage corresponded to the cases in which the main character was considered in focus after *not* being in focus in the immediately preceding clausal unit.

2.4. Analyses

For the analyses, we retained only the references produced for the movies that the participants had seen to really focus on the effect of the interaction partner's likely knowledge while controlling for the participant's own knowledge (which did not differ between groups, see Table 1). The analyses were performed independently for both tasks (identification and narration, with the three narration stages also analyzed separately) and for each of our four categories of reference markers (names, indefinite markers, definite markers and pronouns). Logistic mixed models were performed using SAS On Demand for Academics. For each model, the dependent variable (DV) was the probability of producing the targeted reference marker. The independent variables included group membership (between subject: SZ or HC) and condition (within subject: Likely-Known or Likely-Unknown).

Following the recommendations by Barr et al. (Barr et al., 2013), mixed models should include the maximal random effect structure (random intercepts and random slopes) justified by the experimental design. In the current design, this involved including both participants and items (films) as random effects (i.e., by-participant random intercepts, by-participant random slopes corresponding to condition, by-item random intercepts and by-item random slopes corresponding to condition and group membership). Models that include the maximal structure however often fail to properly converge. In these cases, SAS identifies the effects that cause convergence problems, which can then be removed before performing the final analyses (Kiernan et al., 2012).

For all analyses, the Satterthwaite correction (Satterthwaite, 1946) was used to estimate the degrees of freedom, which was required given the uneven number of observations for the different cells of the design (different number of movies that the participant had seen, different number of maintaining and reintroduction following the initial introduction in the narrative task, etc.)

3. Results

3.1. Character identification task

The results for the identification task are illustrated in Fig. 2, and the model parameters are presented in Table A1 in Appendix 1.

For the use of *names*, a significant effect of group emerged ($F(1, 49) = 5.92$, $p = .019$), such that the SZ group used fewer names than the HC group. The effect of condition ($F(1, 7) = 1.61$, $p = .247$) and the interaction ($F(1, 356) = 0.83$, $p = .363$) were not statistically significant.

For the use of *indefinite markers*, none of the effects were statistically significant, including the effect of group ($F(1, 17) = 0.41$, $p = .529$), the effect of condition ($F(1, 8) = 0.00$, $p = .977$) and the interaction between group and condition ($F(1, 9) = 1.64$, $p = .232$).

For the use of *definite markers*, none of the effects were statistically significant, including the effect of group ($F(1, 15) = 0.77$, $p = .393$), the effect of condition ($F(1, 8) = 1.39$, $p = .274$) and the interaction between group and condition ($F(1, 8) = 0.94$, $p = .360$).

For use of *pronouns*, a significant effect of group emerged ($F(1, 47) = 4.37$, $p = .042$), such that the SZ group used more pronouns than the HC group. An effect of condition also emerged, ($F(1, 28) = 6.80$, $p = .015$), such that pronouns were more often used to present the Likely-Unknown than the Likely-Known characters. The interactions did not reach statistical significance ($F(1, 28) = 1.06$, $p = .312$).

3.2. Narration task

The results for the narration task are illustrated in Fig. 3, and the model parameters are presented in Table A2-A4 in Appendix 1.

3.2.1. Introduction stage

For the use of *names*, there was a significant effect of group ($F(1, 55) = 6.05$, $p = .017$), such that the SZ used fewer names than the HC group. There was also a significant effect of condition ($F(1, 6) = 31.85$, $p = .001$), such that names were more often used to introduce Likely-Known than Likely-Unknown characters. There was no significant interaction between group and condition ($F(1,202) = 0.20$, $p = .658$).

For the use of *indefinite markers*, there was a significant effect of condition ($F(1, 202) = 10.48$, $p = .001$), such that indefinite markers were more often used to introduce Likely-Unknown than Likely-Known characters. The effect of group ($F(1, 45) = 1.16$, $p = .288$) and the interaction ($F(1, 202) = 0.31$, $p = .578$) were not significant.

For the use of *definite markers*, there was a significant effect of group ($F(1, 108) = 4.42$, $p = .038$), a significant effect of condition ($F(1, 6) = 14.14$, $p = .009$) and a significant interaction between group and condition ($F(1, 125) = 7.36$, $p = .008$). The decomposition of the interaction revealed that the effect of condition was significant in the HC group ($t(40) = -3.80$, $p = .003$), with more definite markers for the Likely-Unknown than the Likely-Known characters, while the effect of condition was not significant for the SZ group ($t(5) = -1.27$, $p = 1.000$).

For the use of *pronouns*, the model failed to converge due to the lack of any pronouns in healthy controls for the Likely-Known condition.

3.2.2. Maintaining stage

For the use of *names*, none of the effects were statistically significant, including the effect of group ($F(1, 277) = 0.03$, $p = .865$), the effect of condition ($F(1, 5) = 6.07$, $p = .056$) and the interaction between group and condition ($F(1, 769) = 0.45$, $p = .505$).

For the use of *indefinite markers*, none of the effects were statistically significant, including the effect of group ($F(1, 72) = 0.01$, $p = .941$), the effect of condition ($F(1, 4) = 0.78$, $p = .425$) and the interaction between group and condition ($F(1, 72) = 0.50$, $p = .480$).

For the use of *definite markers*, there was a significant effect of condition ($F(1, 7) = 10.38$, $p = .015$), such that definite markers were more often used to maintain Likely-Unknown than Likely-Known characters in focus. The effect of group ($F(1, 9) = 1.39$, $p = .268$) and the interaction ($F(1, 7) = 1.25$, $p = .300$) were not significant.

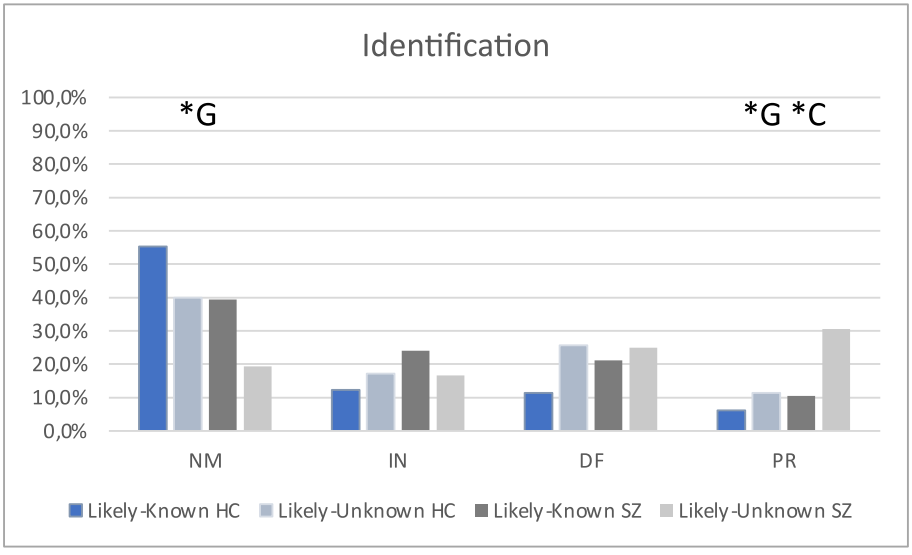

Fig. 2. Proportions of trials in which the different reference markers were employed during Task 1.
NM = Names; IN = Indefinite markers; DF = Definite markers; PR = Pronouns; *G = Significant effect of group; *C = Significant effect of condition.

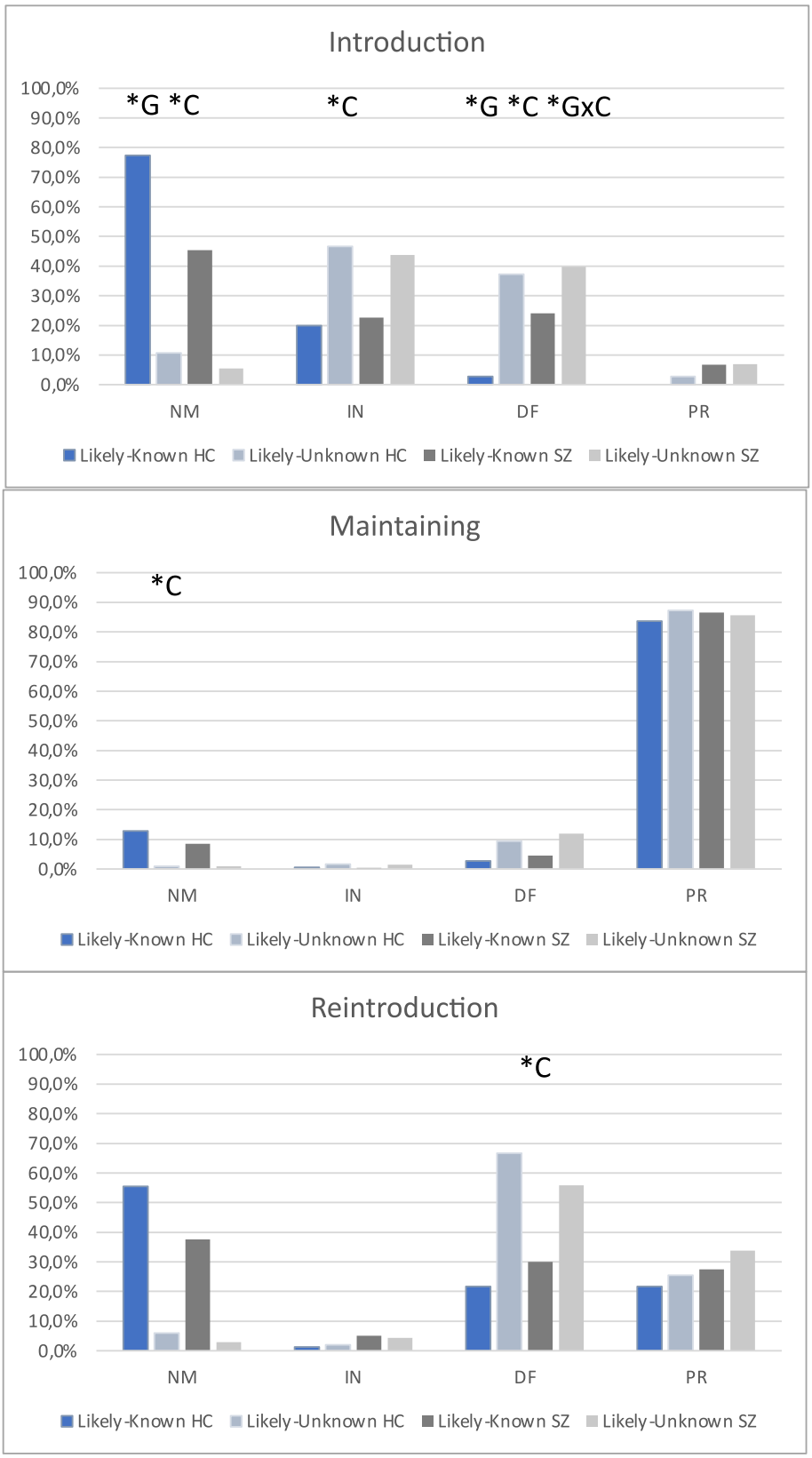

Fig. 3. Proportions of trials in which the different reference markers were employed during the different stages of Task 2.
NM = Names; IN = Indefinite markers; DF = Definite markers; PR = Pronouns; *G = Significant effect of group; *C = Significant effect of condition; *GxC = Significant interaction between group and condition.

For use of *pronouns*, none of the effects were statistically significant, including the effect of group ($F(1, 95) = 0.00$, $p = .951$), the effect of condition ($F(1, 5) = 0.10$, $p = .761$) and the interaction between group and condition ($F(1, 95) = 0.18$, $p = .669$).

3.2.3. Reintroduction stage

For the use of *names*, there was a significant effect of condition ($F(1, 5) = 7.52$, $p = .043$), such that names were more often used to reintroduce Likely-Known than Likely-Unknown characters. The effect of group ($F(1, 77) = 0.49$, $p = .487$) and the interaction ($F(1, 193) = 0.05$, $p = .818$) were not significant.

For the use of *indefinite markers*, none of the effects were statistically significant, including the effect of group ($F(1, 14) = 0.15$, $p = .703$), the effect of condition ($F(1,13) = 0.01$, $p = .940$) and the interaction between group and condition ($F(1, 13) = 0.68$, $p = .426$).

For the use of *definite markers*, none of the effects were statistically significant, including the effect of group ($F(1, 5) = 0.60$, $p = .475$), the effect of condition ($F(1, 5) = 6.07$, $p = .058$) and the interaction between group and condition ($F(1, 4) = 0.14$, $p = .727$).

For use of *pronouns*, none of the effects were statistically significant, including the effect of group ($F(1, 51) = 0.02$, $p = .883$), the effect of condition ($F(1, 4) = 1.21$, $p = .329$) and the interaction between group and condition ($F(1, 45) = 0.12$, $p = .725$).

4. Discussion

While it is clearly established that people with schizophrenia (SZ) produce higher rates of unclear references in their speech relative to healthy controls (HC) (Docherty et al., 1996; Rubino et al., 2011), we know little about their choices of reference markers, an important aspect of efficient speech production. This study examined the production of four types of reference markers (names, indefinite markers, definite markers and pronouns) in people with SZ and HC during the performance of two distinct referential communication tasks. Both tasks manipulated the interaction partner's likely knowledge of the material being presented.

During the character identification task, the SZ group used fewer names and more pronouns than the HC, but the groups did not significantly differ in their use of indefinite or definite markers. During the second task (narration task), group differences were observed when the participants first introduced the main character at the beginning of each narration, such that the SZ group used fewer names and more definite references than the HC group, the SZ also showing a lesser adjustment in their use of definite markers as a function of the condition (Likely-Known versus Likely-Unknown movie characters), resulting in a significant group by condition interaction. No significant effect involving the group emerged when the characters were subsequently mentioned during the narration, be it when maintaining a character in focus in the discourse of when later reintroducing a character that was not previously in focus. Overall, these results indicate that the use of reference markers in the speech of people with schizophrenia is mainly affected when introducing new referents (here movie characters), be it when these referents are being listed (first task) or when new referents are introduced at the onset of a narration (second task).

Importantly, people with SZ did show some adjustments to the likely knowledge of their interaction partner, as evidenced by several instances where significant effects of conditions were observed across both groups (i.e. without a significant interaction with group; see Fig. 3). For instance, both groups used indefinite references more often to introduce the Likely-Unknown than the Likely-Known characters in the introduction stage of the narration task, and both groups used definite references more often when later reintroducing a Likely-Unknown than a Likely-Known character. Overall, our results suggest that people with schizophrenia were generally sensitive to the likely knowledge of their interaction partner, hence being able to take common ground based on *community membership* into account, which is an important finding from this study. These results contrast with those of a prior study in which people with SZ showed difficulties adjusting their choices of references markers based on *dialogue history* (Champagne-Lavau et al., 2009), suggesting that different forms of common ground can be differentially affected.

While our study was the first one to examine the effect of the interaction partner's likely knowledge of the referents (common ground based on community membership) on the choices of reference *markers* in people with schizophrenia, a prior study with the same dataset examined the *content* of the referential expressions used during the identification task (Achim et al., 2022). That prior study found that healthy controls used more elements of information to present Likely-Unknown than to present Likely-Known characters (e.g. “It is Leonidas from the movie 300, he has a red cape” versus “ It's Harry Potter”, respectively) whereas people with schizophrenia did not adjust the number of elements of information that they used as a function of their interaction partner's likely knowledge, leading to a significant group-by-condition interaction. In fact, people with schizophrenia tended to present the characters using only descriptive information (e.g. “the next one wears a red cape”) in that study, even if they knew the movies according to our movie knowledge questionnaire, a strategy that was significantly less used by the healthy participants. Taken together, the results from Achim et al. (Achim et al., 2022) and the current results indicate that people with schizophrenia do present some difficulties in adjusting their choices of reference *content* as a function of the likely knowledge of their interaction partner, whereas they display no significant difficulties adjusting their choices of references *markers* (except for the use of definite references during the introduction stage of the narration task).

Beyond these effects linked to the interaction partner's likely knowledge of the characters, the main effects of group observed in this study also deserves further attention, especially the group effects observed for the use of the character's names that were significant both during the character identification task (first task) and during the introduction stage of the narration task (second task). One possible interpretation for this result is that people with schizophrenia may not have remembered the names of the characters as well as did the healthy controls, which would be consistent with the memory deficits often reported in people with schizophrenia (Aleman et al., 1999). In future studies, it would thus be interesting to use a more complete movie knowledge questionnaire in which participants are asked to state the name of the character if they can remember it. Such a strategy would be necessary to disentangle whether people with SZ did not use the names because they did not remember them, or if they remembered the names yet did not use them.

Some limitations of this study should be acknowledged. First, the sample was relatively modest in size and composed of a large proportion of men. It would thus be interesting to replicate this study in a larger group of participants and to include a greater proportion of women to improve the generalizability of the results. Second, a major methodological challenge in this study was the convergence failure of some statistical models, particularly those investigating complex group-by-condition interactions. These challenges stem from the modest sample size (50 participants) and the complexity of the random effects structures. Following the recommendations of Barr et al. (Barr et al., 2013), we simplified these structures by removing non-essential random slopes, ensuring convergence while preserving analytical integrity. However, we acknowledge that these adjustments may influence the precision of *p*-values and, consequently, the robustness of some findings. Third, the participants from the schizophrenia group had a recent onset of psychosis, and it is possible that people with a longer illness duration may be more affected in their use of reference markers. Fourth, participants from the SZ group had relatively mild thought disorder symptoms, and more important deficits in the use of reference markers may be observed in people with SZ who present with more important thought disorder symptoms. A fifth limitation is that all the participants from the SZ group were taking an antipsychotic medication, whereas the healthy controls

did not take any psychoactive medication. While this is the case with most studies targeting people with schizophrenia, it nonetheless leaves the door open that part of the findings could be related to the medication rather than the presence of a schizophrenia spectrum disorder. The effect of the medication on the use of reference markers, or on other aspects of narrative speech production, in people with SSD should be further examined in future studies.

Despite these limitations, the current study adds to our knowledge about the use of reference markers by people with schizophrenia and established that these choices are affected principally when people with schizophrenia introduce new referents, be it as part of a list of individual referents (as in the character identification task) or when they first introduce a new referent as part of a narration (as in the narration task). Adequate choices of reference markers are central for efficient speech production, and a better characterization of the difficulties faced by people with SZ could lead to better tools to measure speech dysfunctions, ultimately enabling better management of these difficulties in the future.

CRediT authorship contribution statement

Amélie M. Achim: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Dominique Knutsen:** Writing – review & editing, Methodology. **Marc-André Roy:** Writing – review & editing, Data curation. **Souleymane Gadio:** Formal analysis, Writing – review & editing. **Marion Fossard:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgments

This work was supported by a Social Sciences and Humanity Research Council of Canada (SSHRC) grant (#410-2011-1628), and the Fonds de Recherche du Québec – Santé supported our team through salary awards to AMA.

Appendix 1

Table A1
Coefficients associated with the random and fixed effects for the models targeting the identification task.

	Names		Indefinites		Definites		Pronouns	
Model Parameters	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Random effects								
By-participant intercept	1.82	0.64	1.04	0.55	1.17	0.49	2.24	0.77
By-participant random slope corresponding to condition	–	–	0.39	0.51	–	–	–	–
By-film intercept	3.00	1.80	0.27	0.39	0.43	0.50	–	–
By-film random slope corresponding to group	–	–	0.25	0.37	0.46	0.46	0.02	0.21
Fixed effects								
Intercept	−2.46	0.95	−1.89	0.55	−1.27	0.57	−1.03	0.44
Group: HC	1.54	0.69	0.21	0.70	−0.05	0.70	−1.74	0.74
Group: SZ	0		0		0		0	
Condition: Likely-Known	1.77	1.22	0.55	0.66	−0.29	0.72	−1.47	0.46
Condition: Likely-Unknown	0		0		0		0	
Group*Condition: HC, Likely-Known	−0.60	0.66	−1.07	0.84	−0.85	0.88	0.83	0.81
Group*Condition: HC, Likely-Unknown	0		0		0		0	
Group*Condition: SZ, Likely-Known	0		0		0		0	
Group*Condition: SZ, Likely-Unknown	0		0		0		0	

Note: The "–" denote the effects that had to be removed in order for the model to converge.

Table A2
Coefficients associated with the random and fixed effects for the models targeting the introduction stage from the narration task.

	Names		Indefinites		Definites		Pronouns	
Model Parameters	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Random effects								
By-participant intercept	1.99	0.84	1.82	0.74	0.34	0.51	–	–
By-participant random slope corresponding to condition	–	–	–	–	0.09	0.59	–	–
By-film intercept	0.13	0.27	–	–	0.06	0.21	–	–
By-film random slope corresponding to group	–	–	–	–	–	–	–	–
Fixed effects								
Intercept	−3.05	0.76	−0.10	0.46	−0.57	0.39		
Group: HC	1.30	0.92	−0.80	0.69	0.22	0.52		
Group: SZ	0		0		0			
Condition: Likely-Known	2.96	0.79	−1.45	0.50	−0.63	0.50		
Condition: Likely-Unknown	0		0		0			
Group*Condition: HC, Likely-Known	0.43	0.95	0.43	0.76	−2.52	0.93		

(*continued on next page*)

Table A2 (*continued*)

	Names		Indefinites		Definites		Pronouns	
Model Parameters	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Group*Condition: HC, Likely-Unknown	0		0		0			
Group*Condition: SZ, Likely-Known	0		0		0			
Group*Condition: SZ, Likely-Unknown	0		0		0			

Note: The "–" denote the effects that had to be removed in order for the model to converge. For the Pronouns, the model failed to converge.

Table A3
Coefficients associated with the random and fixed effects for the models targeting the maintaining stage from the narration task.

	Names		Indefinites		Definites		Pronouns	
Model Parameters	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Random effects								
By-participant intercept	0.69	0.32	–	–	0.42	0.58	–	–
By-participant random slope corresponding to condition	–	–	0.09	2.18	0.26	0.61	0.49	0.23
By-film intercept	0.59	0.64	0.36	1.00	–	–	0.24	0.22
By-film random slope corresponding to group	–	–	–	–	0.20	0.35	–	–
Fixed effects								
Intercept	−4.10	0.88	−4.03	0.80	−2.19	0.46	1.84	0.43
Group: HC	−0.26	1.07	−0.69	1.24	−0.07	0.66	0.11	0.47
Group: SZ	0		0		0		0	
Condition: Likely-Known	1.70	0.99	−1.51	1.33	−1.01	0.61	−0.03	0.57
Condition: Likely-Unknown	0		0		0		0	
Group*Condition: HC, Likely-Known	0.72	1.08	1.24	1.75	−1.08	0.96	−0.25	0.57
Group*Condition: HC, Likely-Unknown	0		0		0		0	
Group*Condition: SZ, Likely-Known	0		0		0		0	
Group*Condition: SZ, Likely-Unknown	0		0		0		0	

Note: The "–" denote the effects that had to be removed in order for the model to converge.

Table A4
Coefficients associated with the random and fixed effects for the models targeting the reintroduction stage from the narration task.

	Names		Indefinites		Definites		Pronouns	
Model Parameters	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Random effects								
By-participant intercept	1.97	0.87	0.90	1.01	0.43	0.34	0.53	0.54
By-participant random slope corresponding to condition	–	–	–	–	–	–	0.14	0.59
By-film intercept	1.20	1.13	–	–	0.11	0.40	0.17	0.25
By-film random slope corresponding to group	–	–	0.31	1.02	0.38	0.54		
Fixed effects								
Intercept	−3.09	1.09	−3.74	1.12	0.31	0.57	−0.88	0.50
Group: HC	0.36	1.18	0.41	1.57	−0.22	0.87	0.20	0.65
Group: SZ	0		0		0		0	
Condition: Likely-Known	2.83	1.27	0.74	1.27	−1.22	0.75	−0.42	0.62
Condition: Likely-Unknown			0		0		0	
Group*Condition: HC, Likely-Known	0.28	1.21	−1.63	1.98	−0.38	1.02	−0.27	0.77
Group*Condition: HC, Likely-Unknown	0		0		0		0	
Group*Condition: SZ, Likely-Known	0		0		0		0	
Group*Condition: SZ, Likely-Unknown	0		0		0		0	

Note: The "–" denote the effects that had to be removed in order for the model to converge.

Data availability

The data that support the findings of this study are available on request from the corresponding author [AMA].

References

Achim, A.M., et al., 2015. Adjustment of speaker's referential expressions to an addressee's likely knowledge and link with theory of mind abilities. Front. Psychol. 6, 823.

Achim, A.M., Achim, A., Fossard, M., 2017. Knowledge likely held by others affects speakers' choices of referential expressions at different stages of discourse. Language, Cognitiion and Neuroscience 32 (1), 21–36.

Achim, A.M., Achim, A., Fossard, M., 2022. Referential communication in people with recent-onset schizophrenia-spectrum disorders. Front. Psych. 13, 971256.

Achim, A.M., Roy, M.A., Fossard, M., 2023. The other side of the social interaction: theory of mind impairments in people with schizophrenia are linked to other people's difficulties in understanding them. Schizophr. Res. 259, 150–157.

Aleman, A., et al., 1999. Memory impairment in schizophrenia: a meta-analysis. Am. J. Psychiatry 156 (9), 1358–1366.

Andreasen, N.C., 1979a. Thought, language, and communication disorders. II. Diagnostic significance. Arch Gen Psychiatry 36 (12), 1325–1330.

Andreasen, N.C., 1979b. Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. Arch. Gen. Psychiatry 36 (12), 1315–1321.

Ariel, M., 1996. Referring expressions and the +/− coreference distinction. In: Fretheim, T., Gundel, J. (Eds.), Reference and Referent Accessibility. John Benjamins, Amsterdam, pp. 13–25.

Arnold, J.E., 2010. How speakers refer: the role of accessibility. Language and Linguistics Compass 4 (4), 187–203.
Association, A.P., 1994. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, fourth ed. American Psychiatric Association Press, Washington D.C.
Barr, D.J., et al., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J. Mem. Lang. 68 (3).
Bowie, C.R., Harvey, P.D., 2008. Communication abnormalities predict functional outcomes in chronic schizophrenia: differential associations with social and adaptive functions. Schizophr. Res. 103 (1–3), 240–247.
Breier, A., Berg, P.H., 1999. The psychosis of schizophrenia: prevalence, response to atypical antipsychotics, and prediction of outcome. Biol. Psychiatry 46 (3), 361–364.
Champagne-Lavau, M., et al., 2009. Do patients with schizophrenia attribute mental states in a referential communication task? Cogn. Neuropsychiatry 14 (3), 217–239.
Clark, H.H., 1996. Using Language. Cambridge University Press.
Clark, H., Marshall, C.R., 1981. Definite reference and mutual knowledge. In: Joshi, A., Webber, B., Sag, I. (Eds.), Elements of Discourse Understanding. Cambridge University Press, Cambridge.
Clark, H.H., Wilkes-Gibbs, D., 1986. Referring as a collaborative process. Cognition 22 (1), 1–39.
Çokal, D., et al., 2018. The language profile of formal thought disorder. NPJ Schizophr. 4 (1), 18.
Çokal, D., et al., 2023. Referential noun phrases distribute differently in Turkish speakers with schizophrenia. Schizophr. Res. 259, 104–110.
Covington, M.A., et al., 2005. Schizophrenia and the structure of language: the linguist's view. Schizophr. Res. 77 (1), 85–98.
Docherty, N.M., 2005. Cognitive impairments and disordered speech in schizophrenia: thought disorder, disorganization, and communication failure perspectives. J. Abnorm. Psychol. 114 (2), 269–278.
Docherty, N.M., DeRosa, M., Andreasen, N.C., 1996. Communication disturbances in schizophrenia and mania. Arch. Gen. Psychiatry 53 (4), 358–364.
Docherty, N.M., et al., 2013. Effects of social cognitive impairment on speech disorder in schizophrenia. Schizophr. Bull. 39 (3), 608–616.
Fossard, M., et al., 2018. Referential choices in a collaborative storytelling task: discourse stages and referential complexity matter. Front. Psychol. 9, 176.
Gundel, J., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions in discourse. Language 69 (2), 274–307.
Harvey, P.D., 1983. Speech competence in manic and schizophrenic psychoses: the association between clinically rated thought disorder and cohesion and reference performance. J. Abnorm. Psychol. 92 (3), 368–377.
Isaacs, E.A., Clark, H.H., 1987. References in conversations between experts and novices. J. Exp. Psychol. Gen. 116 (1), 26–37.
Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13 (2), 261–276.
Kiernan, K., Tao, J., Gibbs, P., 2012. Tips and strategies for mixed modelling with SAS/STAT procedures. In: 2012 SAS Global Forum. Orlando, FL. In: http://support.sas.com/resources/papers/proceedings12/332-2012.pdf.
Knutsen, D., Fossard, M., Achim, A.M., 2022. Comparing individual and collective management of referential choices in dialogue. Q J Exp Psychol (Hove) 75 (4), 712–729.
Marggraf, M.P., et al., 2020. The link between formal thought disorder and social functioning in schizophrenia: a meta-analysis. Eur. Psychiatry 63 (1), e34.
Rochester, S., Martin, J.R., 1979. Crazy Talk: A Study of the Discourse of Schizophrenic Speakers. Plenum Publishing, New York.
Rubino, I.A., et al., 2011. Referential failures and affective reactivity of language in schizophrenia and unipolar depression. Schizophr. Bull. 37 (3), 554–560.
Satterthwaite, F.E., 1946. An approximate distribution of estimates of variance components. Biometrics 2 (6), 110–114.
de Sousa, P., et al., 2019. Disorganisation, thought disorder and socio-cognitive functioning in schizophrenia spectrum disorders. Br. J. Psychiatry 214 (2), 103–112.