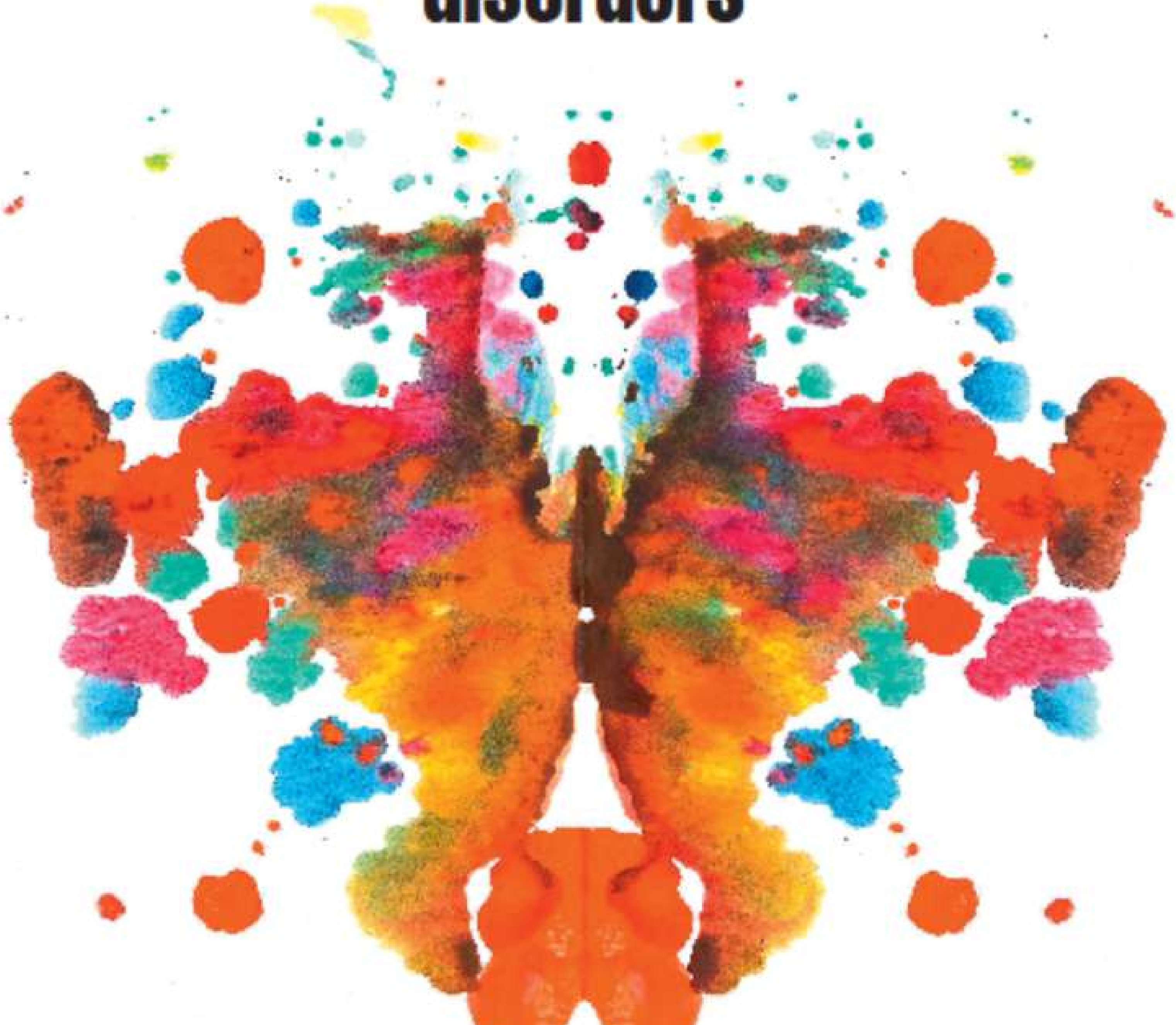
Hallucinated and spoken linguistic patterns as markers of psychiatric disorders

Corona Hernández, Hugo

DOI:
10.33612/diss.872402816

Publication date:
2024

Hallucinated and spoken linguistic patterns as markers of psychiatric disorders



Hugo Corona Hernandez

University of Groningen

Hallucinated and spoken linguistic patterns as markers of psychiatric disorders

Hugo Corona Hernandez

The research presented in this thesis was financially supported by the Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), the University Medical Center Groningen (UMCG), the University of Groningen (RUG), and the Research School of Behavioral and Cognitive Neurosciences (BCN).

The publication of this thesis was financially supported by the University Medical Center Groningen (UMCG), the University of Groningen (RUG), and the Research School of Behavioral and Cognitive Neurosciences (BCN).

Layout inside: Hugo Corona Hernandez and Ridderprint

Cover design: Hugo Corona Hernandez and Ridderprint

Printed by: Ridderprint

university of
groningen

Hallucinated and spoken linguistic patterns as markers of psychiatric disorders

PhD thesis

to obtain the degree of PhD of the
University of Groningen
on the authority of the
Rector Magnificus Prof. J.M.A. Scherpen
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Monday 19 February 2024 at 11.00 hours

by

Hugo Corona Hernandez

born on 18 May 1991
in Mexico City, Mexico

Supervisor
Prof. I.E.C. Sommer

Co-supervisors
Dr. S.G. Brederoo
Dr. J.N. de Boer

Assessment Committee
Prof. S.A.E. Kotz
Prof. J. van Os
Prof. G.J.M. van Noord

TABLE OF CONTENTS

CHAPTER 1

General introduction

Where do my interests come from?

I cannot identify the precise moment I became enthusiastic about the study of linguistic phenomena in individuals with psychiatric disorders. However, through the years, I have pinpointed a number of important events that shaped my awe of the human language, the human mind, and their physical and cognitive correlates (i.e., their measurable counterparts), an awe that at some vague point also intertwined with a broad curiosity about disorders affecting the brain.

I believe that several facts had a large influence on my interest for the human language. First, I have memories of the language impairments that my grandmother had to face as a consequence of the removal of her entire left temporal lobe after she suffered a stroke. Second, when I was a high school student, I slowly fell in love with poetry, both its reading and writing. Third, as a bachelor's student, I was introduced to biolinguistics, psycholinguistic, neurolinguistics, and natural language processing, offering me an overview of the complexity of studying the human language. As for the brain functioning and the human mind, I can count the unexpected advice (received when I was around 22 years old) that I was in need for psychiatric care, and the still ongoing struggle with my own mental health too. Last and fortunately, I can mention as well my growing interest in meditation techniques and methods to study subjective experiencing (e.g., micro-phenomenology and thinking at the edge), whose constant practice has allowed me to become a bit more acquainted with my own mind.

Only when I became a research master student I really started learning about the scientific work (mainly that from the 20th and 21st centuries) on the composition and functioning of the human language, the human mind, the human brain, and their interrelationships. I quickly learned that the study of these phenomena and this organ is an endeavor so complicated that it requires collaboration between researchers from multiple disciplines. I became genuinely surprised when I found out that empirical research merging my interests had already been circulating through the scientific community even before I was born (e.g., Hoffman, 1986). Even more, it was clear that the human language, the human mind, and the human brain and its diseases were not new in any form as topics of research, since pioneering thinkers such as von Humboldt in the 19th century AD (von Humboldt, 1836/1999), Descartes in the 17th century AD (see Westphal, 2016), and Aristotle in the 4th century BC (see Gross, 1995) had already ventured into their study, respectively.

During the last year of my master studies, I was hungrily googling, struggling to find a potential research group that I might aspire to be part of, and doubtful about whether I, as a *linguist, might have any chance to carry out research in the medical sciences. Memorably, I was able to track a study by de Boer and colleagues (2016), titled "A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it". I could not have guessed it then, but, in retrospect, this was a breakthrough moment in my career path. Having found this article provided me with the starting guidance I had been looking for. It propelled me into the venture that finally led me to obtain a PhD position in the Netherlands, at the University Medical Center Groningen, where I started studying two phenomena that language plays a core role in, and that, as I was about to learn, even if they can be present in non-clinical populations, they are particularly burdensome for individuals seeking and/or receiving mental-health care: auditory verbal hallucinations and disorganized speech.*

Auditory verbal hallucinations, a.k.a., voices

Auditory verbal hallucinations (AVHs) can be loosely defined as "a diverse phenomenological experience, which may involve single and/or multiple *voices* [emphasis added], who may be known and/or unknown, speaking sequentially and/or simultaneously, in the first, second, and/or third person and which may give commands, comments, insults, or encouragement" (Jones, 2010). Despite nowadays the terms *AVHs* and *voices* are still sometimes used interchangeably (as it is the case in this thesis), it is worth noting that, over almost the last 2500 years of human history, both terms were coined in different contexts and have received distinct conceptualizations.

In 1838, with the aim of distinguishing between illusions and hallucinations, the French psychiatrist Étienne Esquirol offered a concise and broad, medical description of what experiencing a *hallucination* is like: "Un homme qui a la conviction intime d'une sensation actuellement perçue, alors que nul objet extérieur propre à exciter cette sensation n'est à portée de ses sens, est dans un état d'hallucination"[1] (Esquirol, 1838). In 1892, further contending that language plays an important role in their characterization (Brémaud, 2016), the French psychiatrist Jules Séglas proposed the distinction between psychomotor hallucinations, visual verbal hallucinations, and *auditory verbal hallucinations*, stressing that, in the experience of AVHs, "les paroles ou les phrases (…) semblent venir du dehors, de l'extérieur, et sont perçues par le sujet absolument de la même façon que si elles étaient réellement émises en sa présence

[1] "A person who has the conviction of an actual sensation, whereas there is no external object capable of leading to this sensation, is therefore in a hallucinating state". [Free translation]

par un interlocuteur et venaient frapper son oreille"[2] (Séglas, 1892). Thereby, as a medical expression, the use of the term *AVHs* began its expansion in the late 19th century (Brémaud, 2016; Peyroux & Franck, 2013). Of note, the conceptualization of the term *AVHs* remained contentious throughout the 20th century (Brémaud, 2016), and it is still a matter of conceptual debate in these days (Parnas et al., 2023; Waters & Jardri, 2015).

In contrast of the extended use of the term *AVHs* in the scientific literature of the 19th, 20th, and 21st centuries, the term *voices* had already appeared long before in the accounts describing the experiences of individuals who heard verbal units (e.g., independent words or sentences) in the absence of actual auditory, linguistic stimuli (McCarthy-Jones, 2012; Peyroux & Franck, 2013). For instance, as pointed out by Powell (2022), in *VITA SANCTI GUTHLACI* (The life of Saint Guthlac) (Felix, 715 - 716/1985), it is mentioned that Guthlac, a hermit who lived between the 7th and 8th medieval centuries, encountered "inmundorum spirituum catervis (…) faucibus tortis, labro lato, *vocibus horrisonis* [emphasis added], comis obustis, buccula crassa, pectore arduo…"[3] (Felix, 715 - 716/1985, p. 102).

Not surprisingly, it has been largely acknowledged that the term *voices* likewise remains both phenomenologically and theoretically problematic. It does not fully capture the range of hallucinations with linguistic content that are experienced by individuals (Humpston & Broome, 2016; Waters & Jardri, 2015; Woods et al., 2015) and it can further refer to distinct categorical phenomena (Wilkinson & Krueger, 2022), namely, either to the speech-like quality of the hallucinations with linguistic content or to "the speaker behind the voice" (i.e., a specific agent that is different from the actual individual experiencing the hallucinated voices) (Deamer & Wilkinson, 2015). Despite this, since the 1980's and 1990's (Romme et al., 1992; Romme & Escher, 1989), the term *voices* has increasingly been promoted by researchers and individuals who hear voices to emphasize that, for some individuals, the voice-hearing experience is not necessarily linked to a mental disorder (Corstens et al., 2014). In fact, studies have repeatedly shown that AVHs can be found in both non-clinical (Baumeister et al., 2017; Linszen et al., 2022; Maijer et al., 2018) and clinical adult populations (Upthegrove et al., 2016).

In non-clinical populations, voices have been found to occur in individuals free from a psychiatric diagnosis and/or without current psychiatric treatment (Sommer et al., 2010). Voices have been reported by spiritualists and psychic mediums, i.e., individuals contending

[2] "The words or phrases (…) seem to come from outside, and are perceived by the subject in absolutely the same way as if they were actually emitted by an interlocutor and they came to strike her/his ear". [Free translation]
[3] "Horrible troops of foul spirits (…) [with] twisted jaws, thick lips, *strident voices* [emphasis added], singed hair, fat cheeks, pigeon breasts…". [Translated by Colgrave (Felix, 715 - 716/1985, pp. 101, 103)]

that they can communicate with deceased people (Andrew et al., 2008; Taylor & Murray, 2012). Members of Evangelical Christians groups (Davies et al., 2001), traditional health practitioners (van der Zeijst et al., 2021), and individuals grieving the loss of their loved ones (Sabucedo et al., 2020) have been found to experience AVHs too. Moreover, voices can be induced in non-clinical populations by experimental procedures (Rogers et al., 2021; Wackermann et al., 2008) and/or by the influence of recreational drug use (van der Weijden-Germann et al., 2023).

In clinical populations, AVHs can be present in a variety of psychiatric disorders, including, for instance, schizophrenia-spectrum disorders (SSD) (Hoffman, 1986; Lorente-Rovira et al., 2020; Nayani & David, 1996) and borderline personality disorder (Beatson et al., 2019; Hayward et al., 2022; Slotema et al., 2019) (for a systematic review across diagnoses, see Waters & Fernyhough, 2017). Further, AVHs can manifest in neurodegenerative diseases as well (Eversfield & Orton, 2019). Even if classification systems for psychiatric disorders still consider AVHs as a core symptom for reaching a categorical diagnosis (Waters et al., 2018), nowadays the sole occurrence of AVHs in clinical populations is no longer strictly associated with schizophrenia (Waters et al., 2018). Rather, it has been argued that AVHs can be considered to be either a dissociative experience (Moskowitz & Corstens, 2007) or a transdiagnostic, non-specific psychotic symptom (Waters et al., 2018; Waters & Fernyhough, 2017).

Assuming that AVHs are a transdiagnostic psychotic symptom, epidemiological models of the dimensional (i.e., psychosis-continuum) occurrence of AVHs across clinical and non-clinical populations have been proposed (Baumeister et al., 2017; van Os et al., 2009). Broadly, these models posit that an existing distribution of AVHs-proneness across the whole general population makes it possible that AVHs can present in both clinical and non-clinical populations alike (Baumeister et al., 2017; van Os et al., 2009). These models contend that the occurrence of the AVHs alone is not necessarily related to distress or need for care, with these latter rather resulting from an exposure to risk factors, from a higher frequency of AVHs experiences, and/or from the co-occurrence of AVHs and further psychological difficulties (Baumeister et al., 2017; van Os et al., 2009). Discontinuous (i.e., discrete or non-dimensional) models on AVHs have been proposed too (Baumeister et al., 2017; David, 2010; Linscott & van Os, 2010), and it is currently unclear whether both or just one type of model could better account for AVHs from an epidemiological perspective (Baumeister et al., 2017; Linscott & van Os, 2010).

Along with their epidemiology, accounting for the neural and cognitive mechanisms that might underlie the experience of AVHs is necessary to achieve a comprehensive theory of them and to allow the development of efficient pharmacological and/or psychological treatments for AVHs. Recent meta-analytic evidence suggests that not a single, but a plurality of neurocognitive mechanisms underlying AVHs might exist (Rollins et al., 2019). Given different underlying mechanisms, it might be expected that AVHs' phenomenological characteristics are heterogeneous (McCarthy-Jones et al., 2013). In fact, it has been consistently reported that AVHs phenomenology can largely vary within and across populations (Larøi, 2006; Melvin, Crossley, et al., 2021; Melvin, Rollins, et al., 2021; Nayani & David, 1996; Woods et al., 2015). Thus, hypothetically, any correspondence between specific phenomenological characteristics and distinct underlying mechanisms would suggest the existence of AVHs subtypes (McCarthy-Jones et al., 2013; McCarthy-Jones, Thomas, et al., 2014).

Considering their putative "verbal" nature, it is not surprising that previous studies that subtyped AVHs by phenomenological features considered some dependent linguistic variables as well for analysis (e.g., McCarthy-Jones, Trauer, et al., 2014; Stephane et al., 2003). The problem, though, is that this halts the possibility to know whether AVHs' linguistic features alone might lead to the identification of linguistic subtypes of AVHs. Thus, it has been posited that AVHs' phenomenology might comprise information associated with different AVHs' underlying neurocognitive mechanisms (McCarthy-Jones et al., 2013; McCarthy-Jones, Thomas, et al., 2014) and that this might guide differential diagnosis and/or treatment for AVHs (Cancel et al., 2018; Larøi, 2006; Lowe, 1973; McCarthy-Jones, Thomas, et al., 2014). Yet, whether the case might be such for linguistic subtyping of AVHs remains largely unknown.

In addition to AVHs subtyping, linguistic analysis of AVHs might contribute to current debates over the role of specific AVHs' characteristics and their relation to voice-hearing distress. For instance, distress related to voice-hearing has been posited to occur as a result of metacognitive processes (i.e., beliefs) that influence the experience of voice hearing (Chadwick et al., 2000; Chadwick & Birchwood, 1994), rather than arising from AVHs' characteristics themselves, such as negative content (Baumeister et al., 2022; Larøi et al., 2019; Peters et al., 2012; Silver et al., 2023; Waite et al., 2019). Of note, a controversy still exists between this account and the possibility that negative content in itself leads to AVHs-related distress (Larøi et al., 2019). Moreover, defining what "negative content" is remains problematic (Larøi et al., 2019).

Disorganized speech

Different accounts about how speech can express atypical/anomalous connections between thoughts can be backtracked at least to the 19th century. During that time, a series of terms were already used to (putatively) refer to different psychiatric symptoms (e.g., "*incoherence des idées*" in French, meaning "incoherence of ideas", and "*formale Abweichungen*" in German, meaning "formal deviations") (for more examples, see Jerónimo et al., 2018). In the early 20th century, Eugen Bleuler (1911/1950) and Emil Kraepelin (1919) maintained too that there is a link between speech and disturbances of thought (Hart & Lewine, 2017; Jerónimo et al., 2018). By the second half of the 20th century, Andreasen (1979) emphasized that, despite the contentious interchangeable use of the terms *disorganized speech* (DS) and *formal thought disorder* (FTD), the term DS might be preferred to the term FTD "since disorganized speech is a more accurate term for the behaviors they [i.e., clinicians] are observing". By the end of the same century, the term *DS* was actually incorporated into the Diagnostic and Statistics Manual IV (American Psychiatric Association, 1994), but nowadays researchers and clinicians still find the distinction between DS and FTD a matter of debate (Cohen et al., 2017; Covington et al., 2005; Hart & Lewine, 2017; Jerónimo et al., 2018; Kuperberg, 2010). Of note, yet, cognitive neurosciences and medical sciences have arguably favored the construct FTD over the construct of DS (Kircher et al., 2018), while computational and language sciences have embraced the later (Corcoran et al., 2020; Corcoran & Cecchi, 2020; Covington et al., 2005).

As a matter of fact, with the advent of natural language processing (NLP) and spoken language processing (SLP) tools and techniques, objectively measuring and analyzing DS has become a main aim in the field (Corcoran et al., 2020; Corcoran & Cecchi, 2020; de Boer et al., 2020; Hitczenko et al., 2021). Interestingly, even though DS has been studied in both non-clinical (Barrera et al., 2015; Hain et al., 1995) and clinical populations (Cohen et al., 2017; Corcoran et al., 2020; Corcoran & Cecchi, 2020; Covington et al., 2005; de Boer et al., 2020; Hitczenko et al., 2021), studies with a focus on clinical populations have particularly accumulated a large body of evidence about the linguistic characteristics of DS. Specifically, despite its recognition as a transdiagnostic psychiatric symptom (Cohen et al., 2017), DS has been mainly studied in either at-risk-of-psychosis populations or in individuals with a diagnosis of SSD (Corcoran et al., 2020; Corcoran & Cecchi, 2020; de Boer et al., 2020).

Using Andreasen's categories (1979) as constructs for analysis, linguistic phenomena related to DS such as derailment/loose associations (i.e. "progressively moving off topic"), tangentiality (i.e., "oblique or irrelevant answers"), and incoherence (i.e., "incomprehensible speech") have already been studied using multiple NLP techniques and tools (see reviews by Corcoran et al., 2020; Corcoran & Cecchi, 2020; Hitczenko et al., 2021). Of note, even if derailment/loose associations, tangentiality, and incoherence could arguably be analyzed across different linguistic levels (i.e., phonetics, morphology, syntax, semantics, and pragmatics), most previous NLP studies on these phenomena have implemented computational "semantic" approaches (e.g., semantic space models[4]) to assess the extent to which the meaning of the speech produced by patients with either a SSD or first-episode psychosis can be used to distinguished them from control participants, from patients with different psychiatric disorders, or to predict illness transitions (Corcoran et al., 2020; Corcoran & Cecchi, 2020; Hitczenko et al., 2021). As yet, promising results have been obtained, and the assessment of DS and its subconstructs using NLP-based features along with AI algorithms to accomplish a series of prediction tasks is thought to become a viable digital marker for psychosis in the near future.

Along with encouraging findings, several challenges and limitations for the validity and reliability of computational semantic-based AI algorithms to assess DS have been noticed. For instance, computational shortcomings include sociodemographic biases in the developed AI algorithms and semantic space models, standardization and harmonization issues comprise inconsistencies in speech elicitation techniques and data preprocessing steps, and generalizability obstacles entail the poor cross-linguistic robustness of NLP-based semantic measures to assess DS (for details, see Corcoran et al., 2020; Corcoran & Cecchi, 2020; Hitczenko et al., 2021; Parola et al., 2022).

Surprisingly, the problem of interrelating a refined linguistic definition of DS or any of its subconstructs (e.g., incoherence) and its corresponding NLP operationalization has been largely overlooked. For instance, NLP-based assessments of semantic coherence have been indeed carried out using a variety of calculations, such as similarity mean (i.e., the "average semantic similarity of each word to the immediately preceding word"), the so-called first-order coherence (i.e., the "similarity of consecutive phrase vectors"), and the so-called second-order coherence (i.e., the "similarity between phrases separated by another intervening phrase"), among others (Parola et al., 2022). However, all these different options to operationalize semantic coherence both ignore and conflate two different linguistic components that semantic coherence relies upon, namely, thematic continuity (i.e., the shifting preservation of semantic content across speech or texts) and grammatical connectivity (i.e., the use of linguistic explicit markers to organize the sequence of semantic content) (Givón, 2020).

Arguably, no previous study has attempted to disentangle thematic continuity from grammatical connectivity in analyzing semantic coherence in speech from individuals with

[4] These computational models are thought to numerically represent semantic features and distributional properties of words or phrase, for instance. A core assumption of any NLP analysis based on these models is that, within a given semantic space model, it is possible to "locate" the representations of, let's say, two different words, in turn allowing to calculate the cosine similarity (or "semantic proximity") between them. For an introduction, see Jurafsky and Martin (Online draft 3rd edition).

either SSD or first-episode psychosis. This is remarkably unexpected considering that Andreasen (1979) had already highlighted that connectives (i.e., explicit grammatical markers that hierarchical organize semantic content) are important elements to asses incoherence in SSD.

Incoherence is only one among several constructs for which NLP/SLP-based digital markers of psychosis and other psychiatric disorders are currently under extensive examination and development. To mention a few more, syntactic complexity, poverty of content, referentiality, metaphors, prosody, and lexical abnormalities might also be used as digital markers in psychiatry (Corcoran et al., 2020; Hitczenko et al., 2021). A main goal related to the study of speech- and texts-derived NLP/SLP-based digital markers for psychiatric disorders is that, in the near future, their inclusion in daily clinical practice may improve the psychiatric care provided to patients. Making such an aim a reality comprises a huge complexity that intertwines clinical, cross-linguistic, cultural, economic, ethical, inter-disciplinary, legal, and technological challenges.

Outline of this thesis

AVHs/voices

Chapter 2 explores whether a data-driven clustering analysis on linguistic features alone can be used to distinguish putative linguistic subtypes of AVHs, and whether the resulting AVHs-clusters might associate with AVHs' phenomenology. Moreover, this chapter discusses how putative linguistic subtypes of AVHs might be accounted by different AVHs neurocognitive models, and what potential clinical use might arise from the characterization of those linguistic subtypes of AVHs.

Chapter 3 analyzes whether the linguistic computational approach called sentiment analysis (i.e., the determination of positive, neutral, and negative linguistic valence) can be used to reliably operationalize and quantify linguistic negative content based on transcripts of AVHs from clinical and non-clinical voice-hearers. Possible associations between those sentiment/valence computational linguistic measures and AVHs-related distress were examined. Results were further discussed in light of current debates about negative content of AVHs, highlighting how these findings can inform psychological interventions for AVHs.

(Disorganized) speech

Chapter 4 proposes an approach to reliably assess (in)coherence based on semantic similarity measures calculated using linguistic connectives and their surrounding words in speech-derived transcriptions of individuals with SSD. Also, this chapter examines the proportion of use of different types of connectives and assesses whether connectives-derived measures and proportion of connectives can be used to accurately distinguish individuals with a diagnosis of SSD from control participants.

Going beyond (in)coherence alone and its assessment for (differential) diagnosis purposes, **chapter 5** broadens the discussion of NLP/SLP-based digital markers of psychosis and other psychiatric disorders. This chapter defines ten clinical priorities for which NLP/SLP-based digital markers are currently under rigorous and intensive development, emphasizing, yet, that the endeavor of creating such digital markers faces multidimensional challenges that require inter-institutional and international collaboration in order to be solved.

General discussion

Chapter 6 starts by summarizing the main findings of the studies presented in **chapters 2, 3, 4** and **5**, which is followed by an interpretation of those findings in light of the strengths and limitations of the studies. In **chapter 6**, an exercise is made to delineate what is still "unknown" and worth investigating to understand what links and what sets apart AVHs and disorganized speech, how multidisciplinary studies could approach those gaps in our knowledge, and how the delivery of clinical care might be enhanced by filling those gaps.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4th ed.). American Psychiatric Association.

Andreasen, N. C. (1979). Thought, Language, and Communication Disorders: I. Clinical Assessment, Definition of Terms, and Evaluation of Their Reliability. *Archives Of General Psychiatry*, *36*(12), 1315–1321.

Andrew, E. M., Gray, N. S., & Snowden, R. J. (2008). The relationship between trauma and beliefs about hearing voices: A study of psychiatric and non-psychiatric voice hearers. *Psychological Medicine*, *38*(10), 1409–1417. https://doi.org/10.1017/S003329170700253X

Barrera, Á., Handel, A., Kondel, T. K., & Laws, K. R. (2015). Formal thought disorder: Self-report in non-clinical populations. *International Journal of Psychology & Psychological Therapy*, *15*(1), 155–167.

Baumeister, D., Peters, E., Pruessner, J., Howes, O., & Chadwick, P. (2022). The effects of voice content on stress reactivity: A simulation paradigm of auditory verbal hallucinations. *Schizophrenia Research*, *243*, 225–231. https://doi.org/10.1016/j.schres.2019.07.019

Baumeister, D., Sedgwick, O., Howes, O., & Peters, E. (2017). Auditory verbal hallucinations and continuum models of psychosis: A systematic review of the healthy voice-hearer literature. *Clinical Psychology Review*, *51*, 125–141.

Beatson, J. A., Broadbear, J. H., Duncan, C., Bourton, D., & Rao, S. (2019). Avoiding Misdiagnosis When Auditory Verbal Hallucinations Are Present in Borderline Personality Disorder. *The Journal of Nervous and Mental Disease*, *207*(12), 1048–1055. https://doi.org/10.1097/NMD.0000000000001073

Bleuler, E. (1950). *Dementia Praecox or the Group of Schizophrenias* (J. Zinkin, Ed.). International Universities Press. (Original work published 1911)

Brémaud, N. (2016). Repères historiques sur les hallucinations verbales psychotiques. *Annales Médico-Psychologiques, Revue Psychiatrique*, *174*(5), 359–365. https://doi.org/10.1016/j.amp.2014.03.015

Cancel, A., Rousseau, P.-F., & Jardri, R. (2018). Peut-on guider le traitement des hallucinations sur la base de la sémiologie et de la phénoménologie ? *Annales Médico-Psychologiques, Revue Psychiatrique*, *176*(8), 819–823. https://doi.org/10.1016/j.amp.2018.08.008

Chadwick, P., & Birchwood, M. (1994). The omnipotence of voices. A cognitive approach to auditory hallucinations. *The British Journal of Psychiatry: The Journal of Mental Science*, *164*(2), 190–201. https://doi.org/10.1192/bjp.164.2.190

Chadwick, P., Sambrooke, S., Rasch, S., & Davies, E. (2000). Challenging the omnipotence of voices: Group cognitive behavior therapy for voices. *Behaviour Research and Therapy*, *38*(10), 993–1003. https://doi.org/10.1016/s0005-7967(99)00126-6

Cohen, A. S., Le, T. P., Fedechko, T. L., & Elvevåg, B. (2017). Can RDoC Help Find Order in Thought Disorder? *Schizophrenia Bulletin*, *43*(3), 503–508. https://doi.org/10.1093/schbul/sbx030

Corcoran, C. M., & Cecchi, G. A. (2020). Using Language Processing and Speech Analysis for the Identification of Psychosis and Other Disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 770–779.

Corcoran, C. M., Mittal, V. A., Bearden, C. E., E. Gur, R., Hitczenko, K., Bilgrami, Z., Savic, A., Cecchi, G. A., & Wolff, P. (2020). Language as a biomarker for psychosis: A natural language processing approach. *Schizophrenia Research*, *226*, 158–166.

Corstens, D., Longden, E., McCarthy-Jones, S., Waddingham, R., & Thomas, N. (2014). Emerging Perspectives From the Hearing Voices Movement: Implications for Research and Practice. *Schizophrenia Bulletin*, *40*(Suppl_4), S285–S294. https://doi.org/10.1093/schbul/sbu007

Covington, M. A., He, C., Brown, C., Naçi, L., McClain, J. T., Fjordbak, B. S., Semple, J., & Brown, J. (2005). Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research*, *77*(1), 85–98. https://doi.org/10.1016/j.schres.2005.01.016

David, A. S. (2010). Why we need more debate on whether psychotic symptoms lie on a continuum with normality. *Psychological Medicine*, *40*, 1935–1942. https://doi.org/10.1017/S0033291710000188

Davies, M. F., Griffin, M., & Vice, S. (2001). Affective reactions to auditory hallucinations in psychotic, evangelical and control groups. *The British Journal of Clinical Psychology*, *40*(4), 361–370. https://doi.org/10.1348/014466501163850

de Boer, J. N., Brederoo, S. G., Voppel, A. E., & Sommer, I. E. C. (2020). Anomalies in language as a biomarker for schizophrenia: *Current Opinion in Psychiatry*, *33*(3), 212–218.

de Boer, J. N., Heringa, S. M., van Dellen, E., Wijnen, F. N. K., & Sommer, I. E. C. (2016). A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it. *Brain and Language*, *162*, 10–18.

Deamer, F., & Wilkinson, S. (2015). The speaker behind the voice: Therapeutic practice from the perspective of pragmatic theory. *Frontiers in Psychology*, *6*.

Esquirol, E. (1838). Des hallucinations. In *Des maladies mentales: Considérées sous les rapports médical, hygiénique et médico-légal: Vol. Tome 1* (pp. 159–201). J. B. Baillière.

Eversfield, C. L., & Orton, L. D. (2019). Auditory and visual hallucination prevalence in Parkinson's disease and dementia with Lewy bodies: A systematic review and meta-analysis. *Psychological Medicine*, *49*(14), 2342–2353. https://doi.org/10.1017/S0033291718003161

Felix. (1985). *Felix's Life of Saint Guthlac: Texts, Translation and Notes* (B. Colgrave, Ed. & Trans.). Cambridge University Press. (Original work published 715 - 716)

Givón, T. (2020). *Coherence*. John Benjamins Publishing Company.

Gross, C. G. (1995). Aristotle on the Brain. *The Neuroscientist*, *1*(4), 245–250. https://doi.org/10.1177/107385849500100408

Hain, C., Maier, W., Hoechst-Janneck, S., & Franke, P. (1995). Subclinical thought disorder in first-degree relatives of schizophrenic patients. Results from a matched-pairs study with the Thought Disorder Index. *Acta Psychiatrica Scandinavica*, *92*(4), 305–309. https://doi.org/10.1111/j.1600-0447.1995.tb09587.x

Hart, M., & Lewine, R. R. J. (2017). Rethinking Thought Disorder. *Schizophrenia Bulletin*, *43*(3), 514–522. https://doi.org/10.1093/schbul/sbx003

Hayward, M., Jones, A.-M., Strawson, W. H., Quadt, L., Larsson, D. E. O., Silva, M., Davies, G., Fielding-Smith, S., Hazell, C. M., Critchley, H. D., & Garfinkel, S. N. (2022). A cross-sectional study of auditory verbal hallucinations experienced by people with a diagnosis of borderline personality disorder. *Clinical Psychology & Psychotherapy*, *29*(2), 631–641. https://doi.org/10.1002/cpp.2655

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

Hoffman, R. E. (1986). Verbal hallucinations and language production processes in schizophrenia. *Behavioral and Brain Sciences*, *9*(3), 503–517.

Humpston, C. S., & Broome, M. R. (2016). The spectra of soundless voices and audible thoughts: Towards an integrative model of auditory verbal hallucinations and thought insertion. *Review of Philosophy and Psychology*, *7*(3), 611–629. https://doi.org/10.1007/s13164-015-0232-9

Jerónimo, J., Queirós, T., Cheniaux, E., & Telles-Correia, D. (2018). Formal Thought Disorders–Historical Roots. *Frontiers in Psychiatry*, *9*, 572. https://doi.org/10.3389/fpsyt.2018.00572

Jones, S. R. (2010). Do We Need Multiple Models of Auditory Verbal Hallucinations? Examining the Phenomenological Fit of Cognitive and Neurological Models. *Schizophrenia Bulletin*, *36*(3), 566–575.

Jurafsky, D., & Martin, J. H. (Online draft 3rd edition). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. https://web.stanford.edu/~jurafsky/slp3/ed3book_jan72023.pdf

Kircher, T., Bröhl, H., Meier, F., & Engelen, J. (2018). Formal thought disorders: From phenomenology to neurobiology. *The Lancet. Psychiatry*, *5*(6), 515–526. https://doi.org/10.1016/S2215-0366(18)30059-2

Kraepelin, E. (1919). *Dementia Praecox and Paraphrenia*. Chicago Medical Book Co.

Kuperberg, G. R. (2010). Language in Schizophrenia Part 1: An Introduction. *Language and Linguistics Compass*, *4*(8), 576–589.

Larøi, F. (2006). The phenomenological diversity of hallucinations: Some theoretical and clinical implications. *Psychologica Belgica*, *46*(1–2), 163–183. https://doi.org/10.5334/pb-46-1-2-163

Larøi, F., Thomas, N., Aleman, A., Fernyhough, C., Wilkinson, S., Deamer, F., & McCarthy-Jones, S. (2019). The ice in voices: Understanding negative content in auditory-verbal hallucinations. *Clinical Psychology Review*, *67*, 1–10.

Linscott, R. J., & van Os, J. (2010). Systematic reviews of categorical versus continuum models in psychosis: Evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI, and DSM-VII. *Annual Review of Clinical Psychology*, *6*, 391–419. https://doi.org/10.1146/annurev.clinpsy.032408.153506

Linszen, M. M. J., de Boer, J. N., Schutte, M. J. L., Begemann, M. J. H., de Vries, J., Koops, S., Blom, R. E., Bohlken, M. M., Heringa, S. M., Blom, J. D., & Sommer, I. E. C. (2022). Occurrence and phenomenology of hallucinations in the general population: A large online survey. *Schizophrenia*, *8*(1), Article 1. https://doi.org/10.1038/s41537-022-00229-9

Lorente-Rovira, E., Grasa, E., Ochoa, S., Corripio, I., Peláez, T., López-Carrillero, R., Morano-Guillén, M., Gutiérrez-Gea, Á., Villagrán, J. M., Llácer, B., & Sanjuán, J. (2020). Positive and Useful Voices in Patients With Schizophrenia: Prevalence, Course, Characteristics, and Correlates. *The Journal of Nervous and Mental Disease*, *208*(8), 587–592. https://doi.org/10.1097/NMD.0000000000001183

Lowe, G. R. (1973). The Phenomenology of Hallucinations as an Aid to Differential Diagnosis. *The British Journal of Psychiatry*, *123*(577), 621–633. https://doi.org/10.1192/bjp.123.6.621

Maijer, K., Begemann, M. J. H., Palmen, S. J. M. C., Leucht, S., & Sommer, I. E. C. (2018). Auditory hallucinations across the lifespan: A systematic review and meta-analysis. *Psychological Medicine*, *48*(6), 879–888.

McCarthy-Jones, S. (2012). *Hearing Voices: The Histories, Causes and Meanings of Auditory Verbal Hallucinations*. Cambridge University Press. https://doi.org/10.1017/CBO9781139017534

McCarthy-Jones, S., Krueger, J., Larøi, F., Broome, M., & Fernyhough, C. (2013). Stop, look, listen: The need for philosophical phenomenological perspectives on auditory verbal hallucinations. *Frontiers in Human Neuroscience*, *7*(127), 1–9.

McCarthy-Jones, S., Thomas, N., Strauss, C., Dodgson, G., Jones, N., Woods, A., Brewin, C. R., Hayward, M., Stephane, M., Barton, J., Kingdon, D., & Sommer, I. E. (2014). Better Than Mermaids and Stray Dogs? Subtyping Auditory Verbal Hallucinations and Its Implications for Research and Practice. *Schizophrenia Bulletin*, *40*(Suppl 4), S275–S284.

McCarthy-Jones, S., Trauer, T., Mackinnon, A., Sims, E., Thomas, N., & Copolov, D. L. (2014). A new phenomenological survey of auditory hallucinations: Evidence for subtypes and implications for theory and practice. *Schizophrenia Bulletin*, *40*(1), 225–235. https://doi.org/10.1093/schbul/sbs156

Melvin, K., Crossley, J., & Cromby, J. (2021). The feeling, embodiment and emotion of hallucinations in first episode psychosis: A prospective phenomenological visual-ecological study using novel multimodal unusual sensory experience (MUSE) maps. *EClinicalMedicine*, *41*, 101153. https://doi.org/10.1016/j.eclinm.2021.101153

Melvin, K., Rollins, C. P. E., Cromby, J., Crossley, J., Garrison, J. R., Murray, G. K., & Suckling, J. (2021). Arts-based methods for hallucination research. *Cognitive Neuropsychiatry*, 1–20. https://doi.org/10.1080/13546805.2021.1993807

Moskowitz, A., & Corstens, D. (2007). Auditory hallucinations: Psychotic symptom or dissociative experience? *Journal of Psychological Trauma*, *6*(2–3), 35–63. https://doi.org/10.1300/J513v06n02_04

Nayani, T. H., & David, A. S. (1996). The auditory hallucination: A phenomenological survey. *Psychological Medicine*, *26*(1), 177–189. https://doi.org/10.1017/s003329170003381x

Parnas, J., Yttri, J.-E., & Urfer-Parnas, A. (2023). Phenomenology of auditory verbal hallucination in schizophrenia: An erroneous perception or something else? *Schizophrenia Research*, S0920-9964(23)00139-1. https://doi.org/10.1016/j.schres.2023.03.045

Parola, A., Lin, J. M., Simonsen, A., Bliksted, V., Zhou, Y., Wang, H., Inoue, L., Koelkebeck, K., & Fusaroli, R. (2022). Speech disturbances in schizophrenia: Assessing cross-linguistic generalizability of NLP automated measures of coherence. *Schizophrenia Research*, S0920-9964(22)00274-2. https://doi.org/10.1016/j.schres.2022.07.002

Peters, E. R., Williams, S. L., Cooke, M. A., & Kuipers, E. (2012). It's not what you hear, it's the way you think about it: Appraisals as determinants of affect and behaviour in voice hearers. *Psychological Medicine*, *42*(7), 1507–1514. https://doi.org/10.1017/S0033291711002650

Peyroux, E., & Franck, N. (2013). An Epistemological Approach: History of Concepts and Ideas About Hallucinations in Classical Psychiatry. In R. Jardri, A. Cachia, P. Thomas, & D. Pins (Eds.), *The Neuroscience of Hallucinations* (pp. 3–20). Springer.

Powell, H. (2022). Voices in Psychosis: A Medieval Perspective. In A. Woods, B. Alderson-Day, & C. Fernyhough (Eds.), *Voices in Psychosis Interdisciplinary Perspectives* (pp. 93–100). Oxford University Press.

Rogers, S., Keogh, R., & Pearson, J. (2021). Hallucinations on demand: The utility of experimentally induced phenomena in hallucination research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, 20200233. https://doi.org/10.1098/rstb.2020.0233

Rollins, C. P. E., Garrison, J. R., Simons, J. S., Rowe, J. B., O'Callaghan, C., Murray, G. K., & Suckling, J. (2019). Meta-analytic Evidence for the Plurality of Mechanisms in Transdiagnostic Structural MRI Studies of Hallucination Status. *EClinicalMedicine*, *8*, 57–71.

Romme, M. A., & Escher, A. D. (1989). Hearing Voices. *Schizophrenia Bulletin*, *15*(2), 209–216.

Romme, M. A., Honig, A., Noorthoorn, E. O., & Escher, A. D. (1992). Coping with hearing voices: An emancipatory approach. *The British Journal of Psychiatry*, *161*, 99–103. https://doi.org/10.1192/bjp.161.1.99

Sabucedo, P., Evans, C., & Hayes, J. (2020). Perceiving those who are gone: Cultural research on post-bereavement perception or hallucination of the deceased. *Transcultural Psychiatry*. https://doi.org/10.1177/1363461520962887

Séglas, J. (1892). *Des troubles du langage chez les aliénés*. J. Rueff et Cie.

Silver, J. H., Lewton, M., & Lewis, H. W. (2023). Mediators of negative content and voice-related distress in a diverse sample of clinical and non-clinical voice-hearers. *The British Journal of Clinical Psychology*, *62*(1), 96–111. https://doi.org/10.1111/bjc.12396

Slotema, C. W., Bayrak, H., Linszen, M. M. J., Deen, M., & Sommer, I. E. C. (2019). Hallucinations in patients with borderline personality disorder: Characteristics, severity, and relationship with schizotypy and loneliness. *Acta Psychiatrica Scandinavica*, *139*(5), 434–442. https://doi.org/10.1111/acps.13012

Sommer, I. E. C., Daalman, K., Rietkerk, T., Diederen, K. M., Bakker, S., Wijkstra, J., & Boks, M. P. M. (2010). Healthy individuals with auditory verbal hallucinations; who are they? Psychiatric assessments of a selected sample of 103 subjects. *Schizophrenia Bulletin*, *36*(3), 633–641. https://doi.org/10.1093/schbul/sbn130

Stephane, M., Thuras, P., Nasrallah, H., & Georgopoulos, A. P. (2003). The internal structure of the phenomenology of auditory verbal hallucinations. *Schizophrenia Research*, *61*(2–3), 185–193.

Taylor, G., & Murray, C. (2012). A qualitative investigation into non-clinical voice hearing: What factors may protect against distress? *Mental Health, Religion & Culture*, *15*(4), 373–388. https://doi.org/10.1080/13674676.2011.577411

Upthegrove, R., Broome, M. R., Caldwell, K., Ives, J., Oyebode, F., & Wood, S. J. (2016). Understanding auditory verbal hallucinations: A systematic review of current evidence. *Acta Psychiatrica Scandinavica*, *133*(5), 352–367.

van der Weijden-Germann, M., Brederoo, S. G., Linszen, M. M. J., & Sommer, I. E. C. (2023). Recreational Drug Use and Distress From Hallucinations in the General Dutch Population. *Schizophrenia Bulletin*, *49*(Supplement_1), S41–S47. https://doi.org/10.1093/schbul/sbac190

van der Zeijst, M., Veling, W., Makhathini, E. M., Susser, E., Burns, J. K., Hoek, H. W., & Susser, I. (2021). Ancestral calling, traditional health practitioner training and mental illness: An ethnographic study from rural KwaZulu-Natal, South Africa. *Transcultural Psychiatry*, *58*(4), 471–485. https://doi.org/10.1177/1363461520909615

van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. (2009). A systematic review and meta-analysis of the psychosis continuum: Evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychological Medicine*, *39*(2), 179–195. https://doi.org/10.1017/S0033291708003814

von Humboldt, W. (1999). *On Language: On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species* (M. Losonsky, Ed.; P. Heath, Trans.; 2nd Edition). Cambridge University Press. https://www.cambridge.org/nl/academic/subjects/philosophy/philosophy-texts/humboldt-language-diversity-human-language-construction-and-its-influence-mental-development-human-species-2nd-edition, https://www.cambridge.org/nl/academic/subjects/philosophy/philosophy-texts (Original work published 1836)

Wackermann, J., Pütz, P., & Allefeld, C. (2008). Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex*, *44*(10), 1364–1378. https://doi.org/10.1016/j.cortex.2007.05.003

Waite, F., Diamond, R., Collett, N., Chadwick, E., Bold, E., Teale, A.-L., Taylor, K. M., Kirkham, M., Twivy, E., Causier, C., Carr, L., Bird, J. C., Černis, E., Isham, L., & Freeman, D. (2019). The comments of voices on the appearance of patients with psychosis: “the voices tell me that I am ugly”. *BJPsych Open*, *5*(5), e86. https://doi.org/10.1192/bjo.2019.66

Waters, F., Blom, J. D., Jardri, R., Hugdahl, K., & Sommer, I. E. C. (2018). Auditory hallucinations, not necessarily a hallmark of psychotic disorder. *Psychological Medicine*, *48*(4), 529–536. https://doi.org/10.1017/S0033291717002203

Waters, F., & Fernyhough, C. (2017). Hallucinations: A Systematic Review of Points of Similarity and Difference Across Diagnostic Classes. *Schizophrenia Bulletin*, *43*(1), 32–43.

Waters, F., & Jardri, R. (2015). Auditory Hallucinations: Debunking the Myth of Language Supremacy. *Schizophrenia Bulletin*, *41*(3), 533–534.

Westphal, J. (2016). *The Mind-Body Problem*. MIT Press.

Wilkinson, S., & Krueger, J. (2022). The Phenomenology of Voice-Hearing and Two Concepts of Voice. In A. Woods, B. Alderson-Day, & C. Fernyhough (Eds.), *Voices in Psychosis: Interdisciplinary Perspectives* (pp. 127–133). Oxford University Press. https://library.oapen.org/handle/20.500.12657/58176

Woods, A., Jones, N., Alderson-Day, B., Callard, F., & Fernyhough, C. (2015). Experiences of hearing voices: Analysis of a novel phenomenological survey. *The Lancet Psychiatry*, *2*(4), 323–331.

CHAPTER 2

A data-driven linguistic characterization of hallucinated voices in clinical and non-clinical voice-hearers

H. Corona-Hernández[1], S. G. Brederoo[1,2], J. N. de Boer[1,3], I. E. C. Sommer[1,2]

[1] Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, University of Groningen, The Netherlands
[2] Department of Psychiatry, University Medical Center Groningen, University of Groningen, The Netherlands
[3] Department of Psychiatry, University Medical Center Utrecht, Utrecht University & Brain Center Rudolf Magnus, Utrecht, The Netherlands

Schizophrenia research. 2022; 241: 210–217.

ABSTRACT

Background: Auditory verbal hallucinations (AVHs) are heterogeneous regarding phenomenology and etiology. This has led to the proposal of AVHs subtypes. Distinguishing AVHs subtypes can inform AVHs neurocognitive models and also have implications for clinical practice. A scarcely studied source of heterogeneity relates to the AVHs linguistic characteristics. Therefore, in this study we investigate whether linguistic features distinguish AVHs subtypes, and whether linguistic AVH-subtypes are associated with phenomenology and voice-hearers' clinical status.

Methods: Twenty-one clinical and nineteen non-clinical voice-hearers participated in this study. Participants were instructed to repeat verbatim their AVHs just after experiencing them. AVH-repetitions were audio-recorded and transcribed. AVHs phenomenology was assessed using the Auditory Hallucinations Rating Scale of the Psychotic Symptom Rating Scales. Hierarchical clustering analyses without a priori group dichotomization were performed using quantitative measures of sixteen linguistic features to distinguish sets of AVHs.

Results: A two-AVHs-cluster solution best partitioned the data. AVHs-clusters significantly differed in linguistic features ($p<.001$); AVHs phenomenology ($p<.001$); and distribution of clinical voice-hearers ($p<.001$). The "expanded-AVHs" cluster was characterized by more determiners, more prepositions, longer utterances (all $p<.01$), and mainly contained non-clinical voice-hearers. The "compact-AVHs" cluster had fewer determiners and prepositions, shorter utterances (all $p<.01$), more negative content, higher degree of negativity (both $p<.05$), and predominantly came from clinical voice-hearers.

Discussion: Two voice-speech clusters were recognized, differing in syntactic-grammatical complexity and negative phenomenology. Our results suggest clinical voice-hearers often hear negative, "compact-voices", understandable under Broca's right hemisphere homologue and memory-based mechanisms. Conversely, non-clinical voice-hearers experience "expanded-voices", better accounted by inner speech AVHs models.

Keywords
Auditory Verbal Hallucinations; Clustering; Linguistics; Phenomenology; Schizophrenia; Voice-hearers

1. Introduction

Auditory verbal hallucinations (AVHs) are understood as the experience of hearing voices in absence of corresponding stimuli. Nowadays, their presence across psychiatric disorders is well recognized, as is the fact that they also occur in non-clinical populations, with estimated lifetime frequencies in the 5-15% range (Beavan et al., 2011; Maijer et al., 2018). In recent years, the appreciation of heterogeneous features in AVHs has increased, both within and across different populations (Larøi et al., 2012; Waters and Fernyhough, 2017; Woods et al., 2015). For instance, while loudness and number of voices are similar between clinical and non-clinical voice-hearers, these two groups differ in frequency of and control over their AVHs (Daalman et al., 2011). It has been suggested that, to fully understand AVHs and their origin, this heterogeneity may require the study of potential underlying subtypes of AVHs (Jones, 2010; McCarthy-Jones et al., 2014a, 2014b; Sommer et al., 2018). In previous research, some support was found for five AVHs subtypes, namely hypervigilant subtype, autobiographical memory subtype, inner speech subtype, an epileptic subtype, and a deafferentation subtype (McCarthy-Jones et al., 2014a). So far, data-driven support for this subdivision is lacking. In addition, negative emotional content and form (e.g., commands) were identified as dimensional constructs that vary across subtypes (McCarthy-Jones et al., 2014a). Altogether, distinguishing subtypes of AVHs can have important implications for both clinical practice and research, such as developing treatments for AVHs and informing neurocognitive models of AVHs (David, 2010; McCarthy-Jones et al., 2014a; Sommer et al., 2018).

Until now, studies on AVHs subtypes have mostly relied on phenomenological and etiological features, scarcely including linguistic characteristics (Chang et al., 2015, 2009; McCarthy-Jones et al., 2014a, 2014b; Stephane et al., 2003). Moreover, most studies on AVHs subtypes have focused exclusively on clinical AVHs (Chang et al., 2009; McCarthy-Jones et al., 2014b; Stephane et al., 2003). Therefore, to arrive at a more comprehensive understanding of AVHs subtypes, a linguistic approach to AVH heterogeneity in both clinical and non-clinical voice-hearers is warranted.

The language of AVHs or “voice-speech” from clinical voice-hearers presumably has linguistic characteristics that distinguish it from other registers of speech (Tovar et al., 2019). Specifically, it has been suggested that clinical voice-speech displays unpleasant and recurrent semantic content, short utterances lacking syntactical errors or grammatical connectivity, and a low use of the grammatical first person (Frank et al., 1980; Hoffman et al., 1994; Tovar et al., 2019; Turkington et al., 2019). Importantly, when comparing voice-speech between clinical and non-clinical voice-hearers, both similarities and dissimilarities in linguistic features have been found (de Boer et al., 2016). Dissimilarities include shorter mean length of utterance, lower verb complexity, and more verbal abuses and perseverations in the voice-speech of individuals with a clinical status (de Boer et al., 2016). This raises the question whether different linguistic subtypes of AVH can be identified, and whether these subtypes are present in both clinical and non-clinical voice-hearers. Therefore, the main aim of the present study was to investigate how linguistic features can be used to distinguish subtypes of AVHs. Specifically, we studied how AVHs linguistic subtypes might be characterized in terms of phenomenology, and whether they are associated with participants' clinical status. To achieve this, we used a data-driven approach to overcome possible limitations of a priori dichotomization of clinical and non-clinical voice-hearers.

2. Methods

2.1. Participants

A total of 21 clinical and 19 non-clinical voice-hearers participated in this study. The majority of this sample was previously described in de Boer et al. (2016). Four clinical voice-hearers were added in the current study. Inclusion criteria for all participants were: (a) being a native Dutch speaker, (b) experiencing verbal hallucinations at least once per month, (c) at least three months free of alcohol or drugs abuse, and (d) absence of a chronic somatic disorder. Patients were recruited via the University Medical Center Utrecht. Non-clinical voice-hearers were recruited via a website (www.verkenuwgeest.nl), and were required to pass an online screening about hallucinations, a telephone interview about the inclusion criteria, and the face-to-face psychiatric screening. All participants were screened for a psychiatric disorder using the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992) and the Structured Clinical Interview for Personality Disorder (SCID-II) (First et al., 1995). Participants were classified as non-clinical voice-hearers when they did not meet the criteria for a psychiatric disorder, and as clinical voice-hearers when they did.

2.2. Procedures

Registrations of the AVHs were collected from all participants using the shadowing procedure. This consisted of instructing each participant to "shadow" (i.e., repeat verbatim) her/his AVHs just after experiencing them (de Boer et al., 2016). The verbatim repetitions were recorded on a voice recording device. This procedure was repeated three times, each recording lasting a minimum of one minute, resulting in a total of at least three minutes of "shadows" per participant. Similar to previous reports (de Boer et al., 2016; Tovar et al., 2019), the recording time required for obtaining the sound recordings spanned between a couple of minutes and half an hour, depending on the frequency and duration of the hallucinations. All sound recordings of the “shadows” were orthographically transcribed by consensus rating of three Dutch native, linguistics graduate students. They successfully identified blurry sounds as either verbal or non-verbal units, disentangled specific words, and differentiated “shadows” from self-talk (see also Linguistic data preprocessing in supplementary materials). The Auditory Hallucinations Rating Scale (AHRS) and the Psychotic Symptom Rating Scales (PSYRATS) (Haddock et al., 1999) were used to assess eleven phenomenological features of the AVHs. Procedures were approved by the ethical committee of the University of Utrecht, and participants provided written informed consent before participation. Declaration of Helsinki’s principles were followed throughout all steps of the research.

2.3. Linguistic features

Sixteen features were analyzed: two types of pronouns (nominative first-person singular, and relatives), three verbal time expressions (simple past, present, and future tenses), three content-and-structure measures (mean length of utterance or MLU, mean word length, and moving-average type-token ratio), four function-word classes (definite and indefinite articles, prepositions, and subordinating conjunctions), and four content-word classes (attributive adjectives, locative adverbs, plural and singular nouns) (see definitions and examples in Table 1). This choice was informed by work about spoken Dutch from Grieve et al. (2017). Based on it, and following Biber’s procedure (1988) to retain only salient linguistic variables for analysis, these 16 features were identified as the most suited to explore differentiating patterns in spoken Dutch.

Relative frequencies of the linguistic features were calculated by dividing each absolute frequency by the total number of words of the corresponding shadow file and then multiplying the quotient by 10,000. In order to prevent absolute frequencies of zero from remaining zero in relative frequencies, one unit was added to all relative frequencies. The Moving-Average Type–Token Ratio (MATTR) was computed by means of the Quantitative Analysis of Textual Data tool version 2.0.1 (Benoit et al., 2018) implementing logarithm with base ten and a moving window with a size of ten words as parameters. MATTR (Covington and McFall, 2010) was chosen over Type-Token Ratio (TTR) (Richards, 1987) as it is more robust in dealing with the influence of text length on the ratio calculation (Brezina, 2018; Covington and McFall, 2010).

Table 1. Linguistic features for analysis in the AVHs.

Linguistic feature	Description
Attributive adjective	Word conveying properties of or adding features to a given noun, e.g., “pretty”.
Definite article	Determiner for identifiable entities, e.g., “the”.
Indefinite article	Determiner for non-identifiable entities, e.g., “a” or “an”.
Locative adverb	Word for details of place or position, e.g., “homeward”.
Mean length of utterance	The average number of words forming an utterance.
Mean word length	The average number of letters forming a word.
Moving-Average Type-Token Ratio	A ratio that expresses the number of different word forms relative to the total number of words.
Nominative first-person singular pronoun	Word standing for the grammatical first person functioning as singular subject, i.e., “I”.
Plural noun	Word standing for concrete or abstract entities, e.g., “tables”.
Preposition	Word usually placed before a noun phrase and typically indicating spatial or temporal relations, e.g., “at” and “in”.
Relative pronoun	Word that depends on an antecedent and that both introduces and plays a role in a new sentence, e.g., “who”.
Simple future	Grammatical value of a later time, e.g., “will”.
Simple past	Grammatical value of a previous time, e.g., “did”.
Simple present	Grammatical value of a current time, e.g., “have”.
Singular noun	Word standing for a concrete or abstract entity, e.g., “table”.
Subordinating conjunction	Word linking an independent clause to a dependent clause, e.g., “because”.

2.4. Statistical analysis

We used hierarchical clustering analyses to classify the participants into groups based on linguistic characteristics of their AVH. Specifically, different subgroups of AVHs were distinguished by grouping according to standardized linguistic features, using Canberra distance and Ward’s method. A two-step procedure was conducted to assess AVHs-cluster validity implementing both relative and internal criteria. First, the number of AVHs-clusters was estimated by means of the R ‘NbClust’ package (Charrad et al., 2014). Secondly, a Sequential Minimal Optimization algorithm (SMO) (Platt, 1999) was implemented along with polynomial kernel and ten-fold cross-validation in order to evaluate the AVHs-clusters’ partitions as labels for classification categories. The Waikato Environment for Knowledge Analysis software (Weka) (Witten et al., 2016) was used only for the classification task.

Chi-square tests of independence without continuity correction were carried out to test for differences in the distribution of clinical status and sex. Two-tailed independent t-tests were performed to test for differences in age and years hearing AVHs. One-way non-parametric multivariate analysis of variance (MANOVA) (Burchett et al., 2017) was done independently to test for the presence of significant differences on both the combined linguistic variables and the combined phenomenological features between AVHs-clusters. Non-parametric two-tailed independent Wilcoxon rank sum tests with continuity correction were carried out for analyzing possible differences in both individual linguistic quantitative measures and individual phenomenological features between AVHs-clusters. Non-parametric two-tailed Spearman bivariate correlations were conducted in order to assess possible relations between individual linguistic features and individual phenomenological features. In both Wilcoxon and correlation tests, Holm correction for multiple comparisons was applied to control for false positive results. This method was used because it is suited for exploratory studies (Menyhart et al., 2021).

For all analyses, after correcting for multiple comparisons, statistical results with p-values $<.05$ were considered to be significant. All analyses were performed in RStudio version 1.2.5019 (RStudio Team, 2019) running R version 4.0.2 (R Core Team, 2020).

3. Results

Clinical voice-hearers were diagnosed either with a schizophrenia-spectrum (95%) or a bipolar (5%) disorder. Clinical and non-clinical voice-hearers were similar in age (ranging from 21 to 75) and sex (both $p>.05$), but differed in years hearing AVHs ($p<.001$) (Table 2). General characteristics of their AVHs are presented in Table 3.

Table 2. Demographic characteristics of participants.

Characteristic	Clinical voice-hearers (n=21)	Non-clinical voice-hearers (n=19)	Statistic value	*P*-value
Age in years, mean (SD)	43.2 (11.38)	50.7 (16.19)	$t\,(38) = 1.70$	.09
Diagnosis, n (%)				
Bipolar disorder	1 (5%)			
Psychosis not otherwise specified	5 (24%)			
Schizoaffective disorder	3 (14%)			
Schizophrenia	12 (57%)			
Females, n (%)	8 (38%)	8 (42%)	$\chi^2\,(1) = .06$	.79
Medication [a], n (%)				
None	2 (11%)			
Only antipsychotic	9 (50%)			
Both antipsychotic and antidepressant	7 (39%)			
Years hearing AVHs [b], mean (SD)	19.3 (15.2)	40.4 (18.56)	$t\,(33) = 3.86$	<.001

[a] Data available only for 18 clinical voice hearers.
[b] Data available only for 18 clinical and 17 non-clinical voice-hearers.
n = sample size, SD = standard deviation.

The data-driven hierarchical clustering procedure showed that a two-AVHs-cluster solution best partitioned the data. AVHs-clusters' validation showed that the percentage of correctly predicted instances was 92.5% (Cohen's kappa coefficient=0.85). There were no significant differences between the two AVHs-clusters regarding age, sex, and years hearing AVHs. The AVHs-clusters differed significantly in terms of distribution of participants with and without a psychiatric disorder (see Table 4), with one AVHs-cluster consisting mainly of non-clinical voice-hearers, and the other predominantly of clinical voice-hearers (p<.001).

Table 3. Details of the AVHs' shadows per group of voice-hearers.

Feature	Voice-hearers	
	Clinical (n=21)	**Non-clinical (n=19)**
Frequency of AVHs, mean (SD) [a]	3.2 (0.89)	2.4 (1.28)
Description [b]	Voices occur at least once an hour	Voices occur at least once a day
Number of different word forms (*types*)		
Mean (SD)	77.4 (45.21)	97.1 (53.56)
Total	690	795
Number of running words (*tokens*)		
Mean (SD)	176.2 (122.96)	199.6 (139.14)
Total	3701	3794
Recording time		
Mean (SD)	5m 47s (2m 54s)	3m 19s (1m 38s)
Total	2hrs 1m 35s	1hr 3m 4s

[a] This information was available only for 35 out of 40 participants.
[b] According to the Psychotic Symptom Rating Scales (PSYRATS) (Haddock et al., 1999).

Multivariate analysis of variance indicated the presence of significant differences on the combined linguistic variables between AVHs-clusters, $F(9.05,340.34)=7.42$, $p<.001$. Follow-up comparisons showed that the AVHs-clusters differed on four linguistic features. Compared to the other cluster, AVHs in the cluster with mainly clinical voice-hearers were characterized by fewer definite and indefinite articles, less prepositions and shorter utterances. Henceforth, the AVHs from this cluster will be called "compact-AVHs". In contrast, the AVHs from the cluster made out of mainly non-clinical voice hearers, being richer in articles and prepositions and showing longer utterances, will henceforth be called "expanded-AVHs" (see Table 5). Illustrative fragments of compact-AVHs and expanded-AVHs are given in Figure 1.

Furthermore, compact-AVHs and expanded-AVHs differed significantly in terms of phenomenology, $F(4.28,141.19)=5.34$, $p<.001$. Compact-AVHs showed a larger amount of negative content and a higher degree of negativity than expanded-AVHs (Table 6). There were no significant correlations between linguistic and phenomenological variables (all *rho*(ρ)<0.4, all $p>.05$).

Table 4. Characteristics of the participants per cluster of AVHs.

Characteristic	**Cluster with n=18**	**Cluster with n=22**	**Statistic value**	***P*-value**	**Cramer's V effect size**
	Mean (SD) or n (%)	**Mean (SD) or n (%)**			
Age in years	47.2 (14.70)	46.5 (14.14)	$t\,(38) = .14$	.88	
Clinical voice-hearers	4 (22.22%)	17 (77.27%)	$\chi^2\,(1) = 12.03$	<.001	.54
Diagnosis					
Bipolar disorder		1 (4.54%)			
Psychosis not otherwise specified	1 (5.5%)	4 (18.18%)			
Schizoaffective disorder	1 (5.5%)	2 (9.09%)			
Schizophrenia	2 (11.11%)	10 (45.45%)			
Medication [a]					
None	1 (5.5%)	1 (4.54%)			
Only antipsychotic	3 (16.6%)	6 (27.27%)			
Both antipsychotic and antidepressant		7 (31.81%)			
Sex (female)	6 (33.33%)	10 (45.45%)	$\chi^2\,(1) = .60$	.43	
Years hearing AVHs [b]	33 (18.5)	26.3 (19.64)	$t\,(33) = 1.02$	.31	

[a] Data available only for 18 clinical voice hearers.
[b] Data available only for 18 clinical and 17 non-clinical voice-hearers.
n = sample size, SD = standard deviation.

As the possibility exists that AVHs in patients with schizophrenia-spectrum disorders are different from those in another psychiatric disorder, we duplicated our analyses with the exclusion of the patient who was diagnosed with bipolar disorder. Compared to the above-mentioned results, this did not lead to substantial changes (see supplementary Tables S1-S3).

Table 5. Linguistic variables across the two clusters of AVHs.

Linguistic variable	Expanded-AVHs cluster (n=18), mean (SD)	Compact-AVHs cluster (n=22), mean (SD)	Statistic		*r* effect size
			W	Adj. *p*-value	
Attributive adjectives	150.1 (141.56)	134 (168.32)	225.5	.98	
Definite articles	545.3 (187.24)	224.7 (191.85)	356	<.001	.68
Indefinite articles	207.7 (140.26)	46.8 (64.63)	346	<.001	.64
Locative adverbs	190.3 (214.43)	109.3 (179.63)	273.5	.42	
Mean length of utterance	8.6 (4.44)	5.3 (2.74)	323.5	.008	.53
Mean word length	4.4 (.49)	4.3 (.66)	249.5	.98	
Moving-Average Type-Token Ratio	.9 (.02)	.9 (.09)	245	.98	
Nominative first-person singular pronoun	254.3 (289.92)	333.7 (272.15)	158	.98	
Plural nouns	229.6 (207.74)	139.3 (173.78)	278	.36	
Prepositions	718 (304.43)	292.2 (236.70)	366	<.001	.72
Relative pronouns	38.1 (44.33)	12.2 (22.68)	266	.42	
Simple future	49.4 (80.68)	13.8 (38.06)	249.5	.54	
Simple past	99.6 (207.28)	143.6 (204.01)	150.5	.98	
Simple present	1623.8 (447.25)	1770.6 (391.56)	128.5	.54	
Singular nouns	1305.3 (601.49)	1027 (503.72)	253	.96	
Subordinating conjunctions	397.5 (323.62)	290.5 (246.76)	239	.98	

P-values were adjusted using Holm correction.
n = sample size, SD = standard deviation.

Compact-AVHs excerpts

Example 1

•Dutch	*Ja*	*het*	*wordt*	*toch*	*niks.*
•Gloss	Yes	it	become.3SG	still	not.
•Feature of interest			simple present		
•Translation	**"Yes, it won't work".**				

Example 2

•Dutch	*Allemaal*	*negatieve*	*dingen.*
•Gloss	All	negative	things.
•Feature of interest		attributive adjective	plural noun
•Translation	**"All negative things".**		

Expanded-AVHs excerpts

Example 3

•Dutch	*Ik*	*mag*	*niets*	*zeggen*	*maar*	*dat*	*kind*	*van*	*hun*	*heeft*	*toch*	*wel*	*een*	*behoorlijke*	*invloed*	*op*	*hun*	*relatie.*
•Gloss	NOM.1SG	to be allowed to.1SG	nothing	say.INF	but	that	child	of	them	have.3SG	still	indeed	a	considerable	influence	on	their	relationship.
•Feature of interest	first-person pronoun	simple present					singular noun	preposition		simple present			indefinite article	attributive adjective	singular noun	preposition		singular noun
•Translation	**"I should not say anything, but that child of theirs surely has a considerable influence on their relationship".**																	

Example 4

•Dutch	*Dan*	*hoef*	*je*	*er*	*niet*	*meer*	*uit*	*te*	*zien*	*als*	*die*	*oude*	*vrouw*	*die*	*toen*	*in*	*dat*	*verpleeghuis*	*lag.*
•Gloss	Then	have.2SG	you	PARTITIVE	not	more	PREP.PREF	PAR	look.INF	like	that	old	woman	who	at that time	in	that	nursing home	lay.3SG.IMP.P
•Feature of interest		simple present										attributive adjective	singular noun	relative pronoun		preposition		singular noun	Simple past
•Translation	**"Then you no longer have to look like that old woman who laid in that nursing home at that time".**																		

Figure 1. Examples of both compact-AVHs and expanded-AVHs. The excerpts belong to a single participant per cluster. IMP=imperfect aspect; INF=infinitive verb; NOM=nominative case; P=past tense; PAR=particle; PREP=prepositional; PREF=prefix; SG=singular; 1=first grammatical person; 2=second grammatical person; 3=third grammatical person.

Table 6. Phenomenological features of the two clusters of AVHs.

PSYRATS item	Phenomenological feature of the AVHs	Expanded-AVHs cluster, description of median [median]	Compact-AVHs cluster, description of median [median]	Expanded-AVHs (n=18), mean (SD)[a]	Compact-AVHs (n=22), mean (SD)[b]	Statistic		r effect size
						W	Adj. *p*-value	
1	Frequency	Voices occur at least once an hour [3]	Voices occur continuously or almost continuously [4]	2.5 (1.23)	3.1 (1.04)	111.5	.70	
2	Duration	Voices last for several minutes [2]	Voices last for at least one hour [3]	2.2 (1.03)	2.8 (1.20)	110	.70	
3	Location	Voices outside the head, but close to ears or head [2]	Voices outside the head, but close to ears or head [2]	1.9 (1.03)	2.1 (1.04)	134	.99	
4	Loudness	About same loudness as own voice [2]	About same loudness as own voice [2]	1.6 (0.70)	1.9 (0.87)	125	.97	
5	Origin	More than 50% but less than 100% of voices originate from external causes [3]	Less than 50% of voices originate from external causes [2]	2.7 (1.25)	2.6 (1.29)	162.5	.99	
6	Amount of negative content	No unpleasant content [0]	More than 50% of voice content is unpleasant or negative [3]	.8 (1.36)	2.5 (1.47)	62	.01	-.52
7	Degree of negative content	Not unpleasant or negative [0]	Personal verbal abuse relating to self-concept [3]	.8 (1.36)	2.6 (1.20)	55	.009	-.56
8	Amount of distress	Voices not distressing at all [0]	More than 50% but less than 100% of distressing voices [3]	1.1 (1.69)	2.8 (1.47)	73	.05	
9	Intensity of distress	Voices not distressing at all [0]	Voices are distressing to a moderate degree [2]	.9 (1.34)	2 (1.19)	80.5	.09	
10	Disruption to life	No disruption to life [0]	Voices cause moderate amount of disruption to life [2]	.7 (1.25)	2 (1.28)	75.5	.05	
11	Control	Some control most of the time [1]	Only occasional control [3]	1.7 (1.72)	2.8 (1.34)	96	.31	

[a] This information was available only for 17 of 18 participants.
[b] This information was available only for 18 of 22 participants.
P-values were adjusted using Holm correction.
n = sample size, SD = standard deviation.

4. Discussion

In this study we investigated whether linguistic features can be used to recognize different subtypes of AVHs, and whether this links to AVHs phenomenology and presence or absence of a clinical diagnosis. The cluster analysis revealed a two-cluster solution. The expanded-AVHs cluster mainly contained non-clinical voice-hearers' AVHs, while the compact-AVHs cluster mostly had clinical voice-hearers' AVHs. In comparison with expanded-AVHs, compact-AVHs had fewer determiners and prepositions, as well as shorter MLU. Regarding phenomenology, compact-AVHs showed a larger amount of negative content and higher degree of negativity. Linguistic features and phenomenology were not correlated, emphasizing the importance of a role of language in characterizing our sample of AVHs.

4.1. Linguistic characterization of the AVHs

Compared to expanded-AVHs, compact-AVHs had fewer determiners overall. Determiners identify referents, either as concrete (e.g., *the* country you live in) or abstract entities (e.g., *the* topic of this text) (Giacalone Ramat and Andorno, 2006; Juvonen, 2006). This smaller number of determiners was unlikely due to fewer noun phrases in compact-AVHs, since both singular and plural nouns did not differ significantly between AVHs-clusters. Post-hoc qualitative assessment of our results gave rise to a few possible explanations. First, compact-AVHs were found to have a larger proportion of singular proper nouns (see supplementary Table S4), which never require a determiner in Dutch (Hanks, 2006; Oosterhoff, 2015). Second, the Dutch indefinite article "*een*" can form adverbial constructions of degree (e.g., *een beetje*; "a bit") (Klein, 1998), and the proportion of this type of constructions was smaller in compact-AVHs (see supplementary Table S5). Third, compact-AVHs had a larger proportion of countable indefinite plural nouns (e.g., *Ø stemmen*; "voices"), which always lack a determiner (Oosterhoff, 2015) (see supplementary Table S6).

Furthermore, our results suggest that prepositional constructions were more frequent in expanded-AVHs compared to compact-AVHs. Prepositions typically form locational/positional, temporal or directional/movement constructions (Kurzon, 2006; Svenonius, 2007). Importantly, prepositional constructions often specify information in connection with verbs (Svenonius, 2007; Talmy, 2000). For instance, someone can either "live" *in a house*, "eat" *at lunchtime*, or "walk" *towards a door*. Indeed, a post-hoc exploratory evaluation of our data showed that expanded-AVHs had a larger proportion of verbs of location/position, time, and direction/movement (see supplementary Table S7).

MLU was also significantly different between the AVHs-clusters, with compact-AVHs consisting of shorter MLU than expanded-AVHs. It can be reasonably assumed that fewer determiners and prepositions underlie this shorter MLU of compact-AVHs. However, correlations between MLU and both determiners and prepositions did not remain significant after correction for multiple comparisons, leaving this hypothesis unconfirmed. Notwithstanding, our finding that compact-AVHs had shorter MLU and were more often experienced by individuals with a psychiatric disorder is in line with previous studies that found a shorter MLU in clinical AVHs (de Boer et al., 2016) and reduced grammatical connectivity in clinical voice-hearers' AVHs (Tovar et al., 2019).

Interestingly, our results showed that the occurrence of the nominative first-person singular pronoun was not different between AVHs-clusters. This is intriguing, since this feature has been shown to be characteristic of clinical AVHs (McCarthy-Jones et al., 2014b). When including other linguistic functions and also counting its plural forms, the grammatical first person could still typify clinical AVHs (Tovar et al., 2019). The pronoun "*ik*" (i.e., "I") represents only one of several forms and functions in which the grammatical first person can be indicated in Dutch. Therefore, our results would suggest that, rather than the nominative first-person singular pronoun, other forms and functions of the grammatical first person could be more important in distinguishing linguistic sets of AVHs across clinical and non-clinical voice hearers.

4.2. Phenomenological characterization of the AVHs

We found that compact-AVHs, which had a higher proportion of AVHs from clinical voice-hearers, were associated with a larger amount and a higher degree of negative content. This is consistent with previous research showing that AVHs are experienced to be more negative by clinical voice-hearers, although non-clinical voice-hearers also report some negative AVHs (Baumeister et al., 2017; Daalman et al., 2011; Larøi et al., 2019; Nayani and David, 1996). It is noteworthy that we failed to find a relation between the amount and degree of negative content on the one hand, and linguistic features on the other hand, as both differed between the two AVHs-clusters. However, this absence of an association is not surprising, since emotional valence of words is mainly grounded in adjectives, nouns, and verbs, as shown by affective norms for words (e.g., Moors et al., 2013), rather than in determiners, prepositions, or MLU that emerged from our linguistic analysis.

4.3. Neurocognitive models of AVHs

Inner speech/self-monitoring AVH models (Frith and Done, 1988; Jones and Fernyhough, 2007; McGuire et al., 1995) suggest that AVHs arise from inner speech, which is substantiated by neuroimaging studies (Allen et al., 2008, 2007). In comparing characteristics of inner speech to our subtypes of AVHs, a few points are of note. Fernyhough (2004) describes two types of inner speech, namely expanded and condensed inner speech. Expanded inner speech represents an internalization of overt dialog (Fernyhough, 2004). As such, it might commonly display locational, temporal, and directional information (Yule, 1996), much like expanded-AVHs. In contrast, condensed inner speech retains few of these features (Fernyhough, 2004), similarly to compact-AVHs. However, compact-AVHs differ from condensed inner speech in their negative phenomenology. This is in line with previous research showing that clinical voice-speech contains more unpleasant and more controlling words than inner verbal thoughts (Turkington et al., 2019). While expanded-AVHs might instantiate what Fernyhough (2004) calls "expanded inner speech", our results suggest that this is not the case for all AVHs. Specifically, the inner speech-model seems to accommodate some linguistic characteristics, but not the negative valence of compact-AVHs.

Meta-analytic evidence consistently relates activation of Broca's area right homologue to experiencing AVHs (Zmigrod et al., 2016). Importantly, it is likely the source of AVHs with negative phenomenology and relatively simple linguistic features (de Boer et al., 2016; Larøi et al., 2019; Sommer and Diederen, 2009). Since compact-AVHs displayed features like these, it can be hypothesized that they are triggered by activation of Broca's homologue in the right hemisphere.

Another possibility is that compact-AVHs are in fact memories of previously heard speech, which is no longer recognized as such (Waters et al., 2006). Deficits in biding contextual cues (Waters et al., 2006) might also underlie the scarce identification through determiners and the few locational, temporal, and directional information shown by compact-AVHs. Moreover, misattributed recalled memories might trigger negative affect (Waters et al., 2006), and this would be reflected in compact-AVHs' negative phenomenology.

Alternatively, compact-AVHs' linguistic features might arise from disruptions in generative circuits of language-related mechanisms (Brown and Kuperberg, 2015), which tallies with the observation that referential systems are disrupted in schizophrenia-spectrum disorders (van Schuppen et al., 2019; Zimmerer et al., 2017). Çokal et al. (2018) and Sevilla et al. (2018) showed that speech of patients with schizophrenia-spectrum disorders and formal thought disorder is characterized by aberrant use of definite, but not of indefinite articles. This suggests that referentiality deficits in formal thought disorder may be linked to linguistic definiteness (Çokal et al., 2018). Our results are surprising in this respect, as both definite and indefinite articles were less frequent in compact-AHVs than in expanded-AVHs. As mentioned earlier, this does not necessarily imply anomalous referentiality in compact-AVHs, since these had more proper nouns, which by definition are referential words (Hanks, 2006). However, to date no direct comparison has been made between linguistic characterizations of spontaneous speech of voice-hearers, on the one hand, and the content of their voices, on the other. For this reason, it remains an open question to what extent our diverging findings may be explained by referentiality differences in spontaneous speech compared to AVHs.

4.4. Future directions

It has been pointed out that shared mechanisms may account for different subtypes of AVHs (Keshavan, 2013; McCarthy-Jones et al., 2014a). This possibility is partially supported by a recent study showing that, in non-clinical voice-hearers, different AVHs show both common and distinct brain activation (Lin et al., 2020). This is in line with the claim that multiple mechanisms could be at the root of AVHs subtypes (Cicchetti and Rogosch, 1996; McCarthy-Jones et al., 2014a), which is related to meta-analytic evidence on plurality of mechanisms in AVHs (Rollins et al., 2019). It remains an open question whether expanded-AVHs and compact-AVHs can be explained within the same neurocognitive models, or whether different models underlie the emergence of these different linguistic subtypes of AVHs. Future research must address this controversy, taking into account that both linguistic features and phenomenology of AVHs subtypes should be consistently integrated with findings on other levels of explanation (Hugdahl and Sommer, 2018).

Our findings on expanded-AVHs and compact-AVHs can benefit personalized psychological therapies for AVHs. For example, expanded-AVHs displayed more identification through determiners, and richer locational, temporal, and directional information. These features are important for conversation (Yule, 1996). Thus, the Voice Dialogue method (Stone and Stone, 1989) and derivatives thereof (e.g., Corstens et al., 2012) may be indicated for voice-hearers with expanded-AVH. Those features were less present in compact-AVHs, and these AVHs were also shorter. In treating compact-AVHs, dialogue therapies would then face obstacles similar to those of trying to talk to someone who violates communicative principles (Clark, 2004; Grice, 1975). Besides, compact-AVHs had overall negative phenomenology. Considering all this, a psychological treatment for compact-AVHs might rather take advantage of voice-hearers' metalinguistic skills (Basturkmen et al., 2002; Bialystok and Ryan, 1985;

Gombert, 1993) in judging linguistic properties of their voice-speech. Just as people display attitudes toward speakers' features of speech (Dragojevic et al., 2020), voice-hearers with compact-AVHs can be taught to do so with their voices. Research shows that perceived power and superiority of the AVHs plays a large role in the resulting distress (Chadwick and Birchwood, 1994). Hence, metalinguistic therapy reflecting on the "poor" linguistic quality of compact-AVHs could help voice-hearers to counter with the perceived status of the voices. This might also deviate their attention from the negative phenomenology, possibly alleviating distress.

Finally, as some languages do not have determiners (e.g., Finnish) nor prepositions (e.g., Pilagá, spoken in Argentina) (Dryer, 2013a, 2013b), part of our results may not be replicable in those languages. In that case, the analysis will have to be focused on elements or mechanisms that may fulfill the corresponding typological linguistic functions of determiners, prepositions and/or referential units.

4.5. Limitations

A first limitation of the present study is that we constrained our analyses to 16 linguistic variables, while of course more aspects of language could have been assessed. A related matter is that our sample size was relatively small, and there was no homogeneity in clinical voice-hearers' diagnosis nor in non-clinical voice-hearers' phenotype (Baumeister et al., 2017). This might have reduced statistical power and added variation to the data, increasing the difficulty of finding linguistic patterns in our relatively small sample. On the other hand, this heterogeneity does best reflect clinical practice and could therefore make the generalization of our results easier, since the differences we found between compact-AVHs and expanded-AVHs in both linguistic and phenomenological features had either a medium or a large effect size.

Secondly, there is no way of knowing whether the "shadows" of the participants' AVHs are a perfect reflection of the AVHs they heard. For instance, due to embarrassment of the AVHs content, some participants might have been hesitant in fully repeating their AVHs. Moreover, personal (e.g., mood and cognitive abilities) and situational factors (e.g., distractions in the room) might have influenced participants' performance in repeating their AVHs verbatim. In parallel, AVHs data from clinical voice-hearers might have been constrained by their verbal repetition skills, since results from a simulated shadow-task showed that patients with schizophrenia or schizoaffective disorder have poorer performance when compared to controls (Fuentes-Claramonte et al., 2019). Furthermore, we did not control for medication use. As antipsychotic medication can influence language production (de Boer et al., 2020; Salomé et al., 2000), the possibility that medication effects underlie some characteristics of the clinical voice-hearers' AVHs as repeated by them cannot be ruled out.

5. Conclusions

To conclude, using a data-driven approach with linguistic features, two clusters of voice-speech could be recognized. Linguistically, these AVHs-clusters, which we named "compact" and "expanded", mainly differed in their use of referential information and syntactic complexity. Phenomenologically, the amount of negative content and degree of negativity were the most important differences between the AVHs-clusters. Our data show that, compared to expanded-AVHs, compact-AVHs were mainly experienced by clinical voice-hearers. These findings can inform neurocognitive models of AVHs, and also be useful in developing treatments for people with a specific subtype of AVHs.

Role of the funding source

The founding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CRediT author statement

H. Corona-Hernández: Conceptualization, Data curation, Formal analysis, Methodology, Writing-Original draft. **S. G. Brederoo:** Conceptualization, Methodology, Writing-Review & Editing. **J. N. de Boer:** Conceptualization, Methodology, Writing-Review & Editing. **I. E. C. Sommer:** Funding Acquisition, Supervision, Writing-Review & Editing.

Conflict of interest

All authors declare that they have no conflicts of interest.

Acknowledgements

This project was supported by ZonMW mental health (GGZ) call, project number 636310010. HCH was additionally supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604).

Supplementary materials

Linguistic data preprocessing

Throughout the whole inter-rater agreement process, the graduate students were blinded to the clinical status of all participants. Neither of these students had been involved in the making of the preliminary orthographical transcripts. Furthermore, phonetic and grammatical patterns presumably due to spoken language characteristics were also assessed. An example of a phonetic phenomenon of sounds elision is "(ee)n" (corresponding to the indefinite article "a/an" in English), meaning that, although "(ee)n" was written down in the transcript, just the sound [n] was actually produced by the participant. An instance of a grammatical contraction found in the transcripts is "tis", standing for two different words ("het" and "is", corresponding to the English words "it" and "is", respectively) that were contracted and produced as just one articulatory unit. In cases such as these, the underlying (abstract) form of the words remained in the transcripts. This means that, considering the examples given above, "een" and "het is" were the words that actually remained in the individual plain-text transcripts. This decision was supported by two arguments. First, in cases of either sound elision or grammatical contraction, it was not possible to know whether the non-full form of the word had already been perceived as such in the AVHs or whether they were just instances of the participant's actual speech production characteristics. Secondly, non-contracted word forms were sufficient for carrying out both lexical-semantic tagging and frequency extraction. For the protection of personal data, a pseudonymization procedure was also applied.

Table S1. Participants' characteristics by AVHs cluster after excluding the subject with bipolar disorder.

Characteristic	Cluster with n=20	Cluster with n=19	Statistic value	*P*-value	Cramer's V effect size
	Mean (SD) or n (%)	Mean (SD) or n (%)			
Age in years	48.5 (14.96)	44.63 (13.72)	t (37) = .84	.40	
Clinical voice-hearers	4 (20%)	16 (84.21%)	χ^2 (1) = 16.08	<.001	.64
Diagnosis					
Psychosis not otherwise specified	1 (5%)	4 (21.05%)			
Schizoaffective disorder	1 (5%)	2 (10.53%)			
Schizophrenia	2 (10%)	10 (52.63%)			
Medication [a]					
None	1 (5%)	1 (5.26%)			
Only antipsychotic	3 (15%)	6 (31.58%)			
Both antipsychotic and antidepressant		7 (36.84%)			
Sex (female)	8 (40%)	7 (36.84%)	χ^2 (1) = .04	.83	
Years hearing AVHs [b]	33.7 (18.20)	25.2 (19.61)	t (33) = 1.32	.19	

[a] Data available only for 18 clinical voice hearers.
[b] Data available only for 18 clinical and 17 non-clinical voice-hearers.
n = sample size, SD = standard deviation.

Table S2. Linguistic characteristics by AVHs cluster after excluding the subject with bipolar disorder.

Linguistic variable	Expanded-AVHs cluster (n=20), mean (SD)	Compact-AVHs cluster (n=19), mean (SD)	Statistic		*r* effect size
			W	Adj. *p*-value	
Attributive adjectives	144.4 (135.64)	145.3 (177.70)	209	.99	
Definite articles	513.6 (202.40)	220.2 (206.22)	323	.002	.59
Indefinite articles	192.5 (141.78)	44.2 (66.86)	322	.002	.60
Locative adverbs	187.4 (203.75)	105.6 (192.36)	275.5	.13	
Mean length of utterance	8.3 (4.31)	4.7 (1.51)	323.5	.002	.60
Mean word length	4.4 (0.47)	4.3 (0.70)	226	.99	
Moving-Average Type-Token Ratio	0.9 (0.02)	0.8 (0.09)	247	.78	
Nominative first-person singular pronoun	289.3 (295.72)	285.3 (259.09)	186	.99	
Plural nouns	221.2 (198.22)	144.5 (185.66)	263	.32	
Prepositions	681.2 (309.51)	278.6 (252.28)	341	<.001	.67
Relative pronouns	38.1 (42.24)	8.7 (21.76)	278	.06	
Simple future	50.5 (76.44)	9.6 (37.60)	263	.08	
Simple past	107.6 (197.69)	143.4 (219.37)	170.5	.99	
Simple present	1608.2 (426.23)	1830.0 (386.48)	95.5	.08	
Singular nouns	1293.3 (570.22)	1035.0 (530.16)	242	.88	
Subordinating conjunctions	389.9 (308.74)	274.3 (257.33)	236	.99	

P-values were adjusted using Holm correction.
n = sample size, SD = standard deviation.

Table S3. Phenomenological features by AVHs cluster after excluding the subject with bipolar disorder.

PSYRATS item	Phenomenological feature of the AVHs	Expanded-AVHs cluster, description of median [median]	Compact-AVHs cluster, description of median [median]	Expanded-AVHs (n=20), mean (SD)[a]	Compact-AVHs (n=19), mean (SD)[b]	Statistic		r effect size
						W	Adj. *p*-value	
1	Frequency	Voices occur at least once an hour [3]	Voices occur continuously or almost continuously [4]	2.6 (1.23)	3.1 (1.05)	121.5	.83	
2	Duration	Voices last for several minutes [2]	Voices last for at least one hour [3]	2.2 (1.00)	2.8 (1.21)	106	.53	
3	Location	Voices outside the head, but close to ears or head [2]	Voices outside the head, but close to ears or head [2]	2.0 (1.11)	2.0 (0.96)	150	.99	
4	Loudness	Quieter than own voice, whispers [1]	About same loudness as own voice [2]	1.6 (0.69)	2.0 (0.86)	114.5	.69	
5	Origin	More than 50% but less than 100% of voices originate from external causes [3]	Less than 50% of voices originate from external causes [2]	2.7 (1.21)	2.5 (1.32)	163.5	.99	
6	Amount of negative content	No unpleasant content [0]	More than 50% of voice content is unpleasant or negative [3]	0.8 (1.33)	2.6 (1.36)	51	.005	-.59
7	Degree of negative content	Not unpleasant or negative [0]	Personal verbal abuse relating to self-concept [3]	0.8 (1.33)	2.7 (1.03)	43.5	.002	-.63
8	Amount of distress	Voices not distressing at all [0]	Voices always distressing [4]	1.0 (1.66)	3.0 (1.32)	61.5	.01	-.53
9	Intensity of distress	Voices not distressing at all [0]	Voices are distressing to a moderate degree [2]	0.8 (1.32)	2.1 (1.11)	69	.03	-.48
10	Disruption to life	No disruption to life [0]	Voices cause moderate amount of disruption to life [2]	0.7 (1.22)	2.1 (1.21)	65.5	.01	-.51
11	Control	Some control most of the time [1]	Only occasional control [3]	1.6 (1.72)	3.0 (1.17)	82	.09	

[a] This information was available only for 18 of 20 participants.
[b] This information was available only for 17 of 19 participants.
P-values were adjusted using Holm correction.
n = sample size, SD = standard deviation.

Table S4. Proportion of common and proper singular nouns per cluster of AVHs.

Characteristic	Expanded-AVHs (n=18)		Compact-AVHs (n=22)	
	Raw frequency	Percentage	Raw frequency	Percentage
Singular common nouns	276	84.66%	212	78.52%
Singular proper nouns	50	15.34%	58	21.48%
Total singular nouns	326	100%	270	100%

Table S5. Proportion of *een* functions per cluster of AVHs.

Characteristic	Expanded-AVHs (n=18)		Compact-AVHs (n=22)	
	Raw frequency	Percentage	Raw frequency	Percentage
Adverbial constructions with *een*	24	36.36%	7	33.33%
Doubtful cases of *een*	3	4.55%	1	4.77%
Use of *een* as indefinite article	38	57.57%	13	61.90%
Use of *een* as hedge	1	1.52%	0	
Total occurrences of *een*	66	100%	21	100%

Table S6. Proportion of indefinite plural nouns that are countable per cluster of AVHs.

Characteristic	Expanded-AVHs (n=18)		Compact-AVHs (n=22)	
	Raw frequency	Percentage	Raw frequency	Percentage
Countable indefinite plural nouns	37	46.84%	17	58.62%
Other plural nouns	42	53.16	12	41.38%
Total plural nouns	79	100%	29	100%

Note: Nouns were first automatically annotated for countability; this was followed by manual inspection.

Table S7. Proportion of locational/positional, temporal, and movement verbs per cluster of AVHs.

Characteristic	Expanded-AVHs (n=18)		Compact-AVHs (n=22)	
	Raw frequency	Percentage	Raw frequency	Percentage
Verbs of direction/movement	48	5.94%	38	4.68%
Verbs of location/position	45	5.57%	27	3.32%
Verbs of time	4	0.49%	1	0.12%
Other verbs	711	88%	747	91.88%
Total verbs	808	100%	813	100%

Note: Verbs were first automatically annotated for semantic category; this was followed by manual inspection.

References

Allen, P., Aleman, A., McGuire, P.K., 2007. Inner speech models of auditory verbal hallucinations: Evidence from behavioral and neuroimaging studies. Int. Rev. Psychiatry 19, 409–417.

Allen, P., Larøi, F., McGuire, P.K., Aleman, A., 2008. The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. Neurosci. Biobehav. Rev. 32, 175–191.

Andreasen, N.C., Flaum, M., Arndt, S., 1992. The Comprehensive Assessment of Symptoms and History (CASH): An Instrument for Assessing Diagnosis and Psychopathology. Arch. Gen. Psychiatry 49, 615–623.

Basturkmen, H., Loewen, S., Ellis, R., 2002. Metalanguage in Focus on Form in the Communicative Classroom. Lang. Aware. 11, 1–13.

Baumeister, D., Sedgwick, O., Howes, O., Peters, E., 2017. Auditory verbal hallucinations and continuum models of psychosis: A systematic review of the healthy voice-hearer literature. Clin. Psychol. Rev. 51, 125–141.

Beavan, V., Read, J., Cartwright, C., 2011. The prevalence of voice-hearers in the general population: A literature review. J Ment Health 20, 281–292.

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., Matsuo, A., 2018. quanteda: An R package for the quantitative analysis of textual data. J. Open Source Softw. 3, 774.

Bialystok, E., Ryan, E.B., 1985. Toward a Definition of Metalinguistic Skill. Merrill-Palmer Q. 31, 229–251.

Biber, D., 1988. Variation across speech and writing. Cambridge University Press, New York, NY.

Brezina, V., 2018. Statistics in corpus linguistics: a practical guide. Cambridge University Press, New York, NY.

Brown, M., Kuperberg, G.R., 2015. A Hierarchical Generative Framework of Language Processing: Linking Language Perception, Interpretation, and Production Abnormalities in Schizophrenia. Front. Hum. Neurosci. 9, 1–23.

Burchett, W.W., Ellis, A.R., Harrar, S.W., Bathke, A.C., 2017. Nonparametric Inference for Multivariate Data: The R Package npmv. J. Stat. Softw. 76, 1–18.

Chadwick, P., Birchwood, M., 1994. The Omnipotence of Voices: A Cognitive Approach to Auditory Hallucinations. Br. J. Psychiatry 164, 190–201.

Chang, J.S., Kim, Y., Kim, S.H., Hwang, S., Kim, J., Chung, I.-W., Kim, Y.S., Jung, H.-Y., 2015. Differences in the internal structure of hallucinatory experiences between clinical and nonclinical populations. Psychiatry Res. 226, 204–210.

Chang, J.S., Yi, J.S., Ahn, Y.M., Kim, J.H., Kim, Y.S., 2009. Stabilization of the Internal Structure of Persistent Auditory Verbal Hallucinations in Schizophrenia. Aust. N. Z. J. Psychiatry 43, 244–251.

Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. J. Stat. Softw. 61, 1–36.

Cicchetti, D., Rogosch, F.A., 1996. Equifinality and multifinality in developmental psychopathology. Dev. Psychopathol. 8, 597–600.

Clark, H.H., 2004. Pragmatics of Language Performance, in: Horn, L.R., Ward, G. (Eds.), The Handbook of Pragmatics. Blackwell, Malden, Massachusetts, pp. 365–382.

Çokal, D., Sevilla, G., Jones, W.S., Zimmerer, V., Deamer, F., Douglas, M., Spencer, H., Turkington, D., Ferrier, N., Varley, R., Watson, S., Hinzen, W., 2018. The language profile of formal thought disorder. Npj Schizophr. 4, 1–8.

Corstens, D., Longden, E., May, R., 2012. Talking with voices: Exploring what is expressed by the voices people hear. Psychos. Psychol. Soc. Integr. Approaches 4, 95–104.

Covington, M.A., McFall, J.D., 2010. Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). J. Quant. Linguist. 17, 94–100.

Daalman, K., Boks, M.P.M., Diederen, K.M., de Weijer, A.D., Blom, J.D., Kahn, R.S., Sommer, I.E.C., 2011. The Same or Different? A Phenomenological Comparison of Auditory Verbal Hallucinations in Healthy and Psychotic Individuals. J. Clin. Psychiatry 72, 320–325.

David, A.S., 2010. Why we need more debate on whether psychotic symptoms lie on a continuum with normality. Psychol. Med. 40, 1935–1942.

de Boer, J.N., Heringa, S.M., van Dellen, E., Wijnen, F.N.K., Sommer, I.E.C., 2016. A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it. Brain Lang. 162, 10–18.

de Boer, J.N., Voppel, A.E., Brederoo, S.G., Wijnen, F.N.K., Sommer, I.E.C., 2020. Language disturbances in schizophrenia: the relation with antipsychotic medication. Npj Schizophr. 6, 1–9.

Dragojevic, M., Fasoli, F., Cramer, J., Rakić, T., 2020. Toward a Century of Language Attitudes Research: Looking Back and Moving Forward. J. Lang. Soc. Psychol. 1–20.

Dryer, M.S., 2013a. Order of Adposition and Noun Phrase, in: Dryer, M.S., Haspelmath, M. (Eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dryer, M.S., 2013b. Definite Articles, in: Dryer, M.S., Haspelmath, M. (Eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Fernyhough, C., 2004. Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. New Ideas Psychol. 22, 49–68.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1995. The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part I: Description. J. Personal. Disord. 9, 83–91.

Frank, S.M., Rendon, M.I., Siomopoulous, G., 1980. Language in Hallucinations of Adolescent Schizophrenics, in: Rieber, R.W. (Ed.), Applied Psycholinguistics and Mental Health. Springer, Boston, MA, pp. 115–126.

Frith, C.D., Done, D.J., 1988. Towards a Neuropsychology of Schizophrenia. Br. J. Psychiatry 153, 437–443.

Fuentes-Claramonte, P., Soler, J., Hinzen, W., Ramiro-Sousa, N., Rodriguez-Martinez, A., Sarri-Closa, C., Sarró, S., Larrubia, J., Padilla, P.P., McKenna, P.J., Pomarol-Clotet, E., 2019. The interfering effects of frequent auditory verbal hallucinations on shadowing performance in schizophrenia. Schizophr. Res. 208, 488–489.

Giacalone Ramat, A., Andorno, C.M., 2006. Referential Relations in Spoken Discourse, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 450–457.

Gombert, J.E., 1993. Metacognition, Metalanguage and Metapragmatics. Int. J. Psychol. 28, 571–580.

Grice, H.P., 1975. Logic and conversation, in: Cole, P., Morgan, J.L. (Eds.), Syntax and Semantics 3: Speech Acts. Academic Press, New York, NY, pp. 41–58.

Grieve, J., Ruette, T., Speelman, D., Geeraerts, D., 2017. Social Functional Linguistic Variation in Conversational Dutch, in: Friginal, E. (Ed.), Studies in Corpus-Based Sociolinguistics. Routledge, New York, NY, pp. 253–272.

Haddock, G., McCarron, J., Tarrier, N., Faragher, E.B., 1999. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). Psychol. Med. 29, 879–889.

Hanks, P., 2006. Proper Names: Linguistic Status, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 134–137.

Hoffman, R.E., Oates, E., Hafner, J., Hustig, H.H., McGlashan, T.H., 1994. Semantic organization of hallucinated ‘‘voices” in schizophrenia. Am. J. Psychiatry 151, 1229–1230.

Hugdahl, K., Sommer, I.E.C., 2018. Auditory Verbal Hallucinations in Schizophrenia From a Levels of Explanation Perspective. Schizophr. Bull. 44, 234–241.

Jones, S.R., 2010. Do We Need Multiple Models of Auditory Verbal Hallucinations? Examining the Phenomenological Fit of Cognitive and Neurological Models. Schizophr. Bull. 36, 566–575.

Jones, S.R., Fernyhough, C., 2007. Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. Conscious. Cogn. 16, 391–399.

Juvonen, P., 2006. Articles, Definite and Indefinite, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 484–487.

Keshavan, M.S., 2013. Nosology of psychoses in DSM-5: Inches ahead but miles to go. Schizophr. Res. 150, 40–41.

Klein, H., 1998. Adverbs of Degree in Dutch and Related Languages. John Benjamins Publishing Company, Amsterdam, The Netherlands.

Kurzon, D., 2006. Adpositions, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 63–66.

Larøi, F., Sommer, I.E., Blom, J.D., Fernyhough, C., ffytche, D.H., Hugdahl, K., Johns, L.C., McCarthy-Jones, S., Preti, A., Raballo, A., Slotema, C.W., Stephane, M., Waters, F., 2012. The Characteristic Features of Auditory Verbal Hallucinations in Clinical and Nonclinical Groups: State-of-the-Art Overview and Future Directions. Schizophr. Bull. 38, 724–733.

Larøi, F., Thomas, N., Aleman, A., Fernyhough, C., Wilkinson, S., Deamer, F., McCarthy-Jones, S., 2019. The ice in voices: Understanding negative content in auditory-verbal hallucinations. Clin. Psychol. Rev. 67, 1–10.

Lin, X., Zhuo, C., Li, G., Li, J., Gao, X., Chen, C., Jiang, D., 2020. Functional brain alterations in auditory hallucination subtypes in individuals with auditory hallucinations without the diagnosis of specific neurological diseases and mental disorders at the current stage. Brain Behav. 10, 1–12.

Maijer, K., Begemann, M.J.H., Palmen, S.J.M.C., Leucht, S., Sommer, I.E.C., 2018. Auditory hallucinations across the lifespan: a systematic review and meta-analysis. Psychol. Med. 48, 879–888.

McCarthy-Jones, S., Thomas, N., Strauss, C., Dodgson, G., Jones, N., Woods, A., Brewin, C.R., Hayward, M., Stephane, M., Barton, J., Kingdon, D., Sommer, I.E., 2014a. Better Than Mermaids and Stray Dogs? Subtyping Auditory Verbal Hallucinations and Its Implications for Research and Practice. Schizophr. Bull. 40, S275–S284.

McCarthy-Jones, S., Trauer, T., Mackinnon, A., Sims, E., Thomas, N., Copolov, D.L., 2014b. A New Phenomenological Survey of Auditory Hallucinations: Evidence for Subtypes and Implications for Theory and Practice. Schizophr. Bull. 40, 231–235.

McGuire, P.K., David, A.S., Murray, R.M., Frackowiak, R.S.J., Frith, C.D., Wright, I., Silbersweig, D.A., 1995. Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. The Lancet 346, 596–600.

Menyhart, O., Weltz, B., Győrffy, B., 2021. MultipleTesting.com: A tool for life science researchers for multiple hypothesis testing correction. PLOS ONE 16, e0245824. https://doi.org/10.1371/journal.pone.0245824

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., De Schryver, M., De Winne, J., Brysbaert, M., 2013. Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. Behav. Res. Methods 45, 169–177.

Nayani, T.H., David, A.S., 1996. The auditory hallucination: a phenomenological survey. Psychol. Med. 26, 177–189.

Oosterhoff, J., 2015. Modern Dutch Grammar: A Practical Guide. Routledge, New York, NY.

Platt, J.C., 1999. Fast training of support vector machines using sequential minimal optimization, in: Schölkopf, B., Burges, C.J.C., Smola, A.J. (Eds.), Advances in Kernel Methods: Support Vector Learning. MIT Press, Cambridge, Mass, pp. 185–208.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, B., 1987. Type/Token Ratios: what do they really tell us? J. Child Lang. 14, 201–209.

Rollins, C.P.E., Garrison, J.R., Simons, J.S., Rowe, J.B., O’Callaghan, C., Murray, G.K., Suckling, J., 2019. Meta-analytic Evidence for the Plurality of Mechanisms in Transdiagnostic Structural MRI Studies of Hallucination Status. EClinicalMedicine 8, 57–71.

RStudio Team, 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.

Salomé, F., Boyer, P., Fayol, M., 2000. The effects of psychoactive drugs and neuroleptics on language in normal subjects and schizophrenic patients: a review. Eur. Psychiatry 15, 461–469.

Sevilla, G., Rosselló, J., Salvador, R., Sarró, S., López-Araquistain, L., Pomarol-Clotet, E., Hinzen, W., 2018. Deficits in nominal reference identify thought disordered speech in a narrative production task. PLOS ONE 13, e0201545.
Sommer, I.E., Diederen, K.M., 2009. Language production in the non-dominant hemisphere as a potential source of auditory verbal hallucinations. Brain 132, e124.
Sommer, I.E.C., Kleijer, H., Hugdahl, K., 2018. Toward personalized treatment of hallucinations. Curr. Opin. Psychiatry 31, 237–245.
Stephane, M., Thuras, P., Nasrallah, H., Georgopoulos, A.P., 2003. The internal structure of the phenomenology of auditory verbal hallucinations. Schizophr. Res. 61, 185–193.
Stone, H., Stone, S., 1989. Embracing our selves: The voice dialogue training manual. Nataraj Publishing, Novato, California.
Svenonius, P., 2007. Adpositions, particles and the arguments they introduce, in: Reuland, E.J., Bhattacharya, T., Spathas, G. (Eds.), Argument Structure. John Benjamins, Amsterdam, The Netherlands, pp. 63–103.
Talmy, L., 2000. Toward a Cognitive Semantics. The MIT press, Cambridge, Mass.
Tovar, A., Fuentes-Claramonte, P., Soler-Vidal, J., Ramiro-Sousa, N., Rodriguez-Martinez, A., Sarri-Closa, C., Sarró, S., Larrubia, J., Andrés-Bergareche, H., Miguel-Cesma, M.C., Padilla, P.P., Salvador, R., Pomarol-Clotet, E., Hinzen, W., 2019. The linguistic signature of hallucinated voice talk in schizophrenia. Schizophr. Res. 206, 111–117.
Turkington, G.D., Lawrence, T., Stuart, W., 2019. Linguistic analysis of the valence, arousal and dominance of auditory hallucinations and internal thoughts in schizophrenia: Implications for psychoeducation and CBT. Cogent Psychol. 6, 1–5.
van Schuppen, L., van Krieken, K., Sanders, J., 2019. Deictic Navigation Network: Linguistic Viewpoint Disturbances in Schizophrenia. Front. Psychol. 10, 1–8.
Waters, F., Badcock, J., Michie, P., Maybery, M., 2006. Auditory hallucinations in schizophrenia: Intrusive thoughts and forgotten memories. Cognit. Neuropsychiatry 11, 65–83.
Waters, F., Fernyhough, C., 2017. Hallucinations: A Systematic Review of Points of Similarity and Difference Across Diagnostic Classes. Schizophr. Bull. 43, 32–43.
Witten, I.H., Frank, E., Hall, M.A., Pal, C.J., 2016. The WEKA workbench, in: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, Cambridge, Mass, pp. 553–571.
Woods, A., Jones, N., Alderson-Day, B., Callard, F., Fernyhough, C., 2015. Experiences of hearing voices: analysis of a novel phenomenological survey. Lancet Psychiatry 2, 323–331.
Yule, G., 1996. Pragmatics. Oxford University Press, New York, NY.
Zimmerer, V.C., Watson, S., Turkington, D., Ferrier, I.N., Hinzen, W., 2017. Deictic and Propositional Meaning-New Perspectives on Language in Schizophrenia. Front. Psychiatry 8, 1–17.
Zmigrod, L., Garrison, J.R., Carr, J., Simons, J.S., 2016. The neural mechanisms of hallucinations: A quantitative meta-analysis of neuroimaging studies. Neurosci. Biobehav. Rev. 69, 113–123.

CHAPTER 3

Negative content in auditory verbal hallucinations: a natural language processing approach

J. N. de Boer[a,b], H. Corona-Hernández[a], F. Gerritse[b], S. G. Brederoo[a,c], F. N. K. Wijnen[d] & I. E. C. Sommer[a]

[a] Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, University of Groningen, The Netherlands
[b] Department of Psychiatry, University Medical Center Utrecht, Utrecht University & Brain Center Rudolf Magnus, Utrecht, The Netherlands
[c] Department of Psychiatry, University Medical Center Groningen, University of Groningen, The Netherlands
[d] Utrecht Institute of Linguistics OTS, Utrecht University, Utrecht, The Netherlands

Cognitive neuropsychiatry. 2022; 27(2-3): 139–149.

ABSTRACT

Introduction: Negative content of auditory verbal hallucinations (AVH) is a strong predictor of distress and impairment. This paper quantifies emotional voice-content in order to explore both subjective (i.e., perceived) and objectively (i.e., linguistic sentiment) measured negativity and investigates associations with distress.

Methods: Clinical and non-clinical participants with frequent AVH (n= 40) repeated and recorded their AVH verbatim directly upon hearing. The AVH were analyzed for emotional valence using Pattern, a rule-based sentiment analyzer for Dutch. The AVH of the clinical individuals were compared to those of non-clinical voice-hearers on emotional valence and associated with experienced distress.

Results: The mean objective valence of AVH in patients was significantly more negative than those of non-clinical voice-hearers. In the clinical individuals a larger proportion of the voice-utterances was negative (34.7% versus 18.4%) in objective valence. The linguistic valence of the AVH showed a significant, strong association with the perceived negativity, amount of distress and disruption of life, but not with the intensity of distress.

Conclusions: Our results indicate that AVH of patients have a more negative linguistic content than those of non-clinical voice-hearers, which is associated with the experienced distress. Thus, patients not only perceive their voices as more negative, objective analyses confirm this.

Keywords

Hallucinations; Language; Psychosis; Schizophrenia; Sentiment

Introduction

Auditory verbal hallucinations (AVH) are a cardinal feature of psychosis and one of the most common positive symptoms in schizophrenia (Baethge et al., 2005). They also occur in individuals without a psychiatric or neurological disorder, with median reported prevalences around 9.6% (Maijer et al., 2018). A recent population study found that up to 29.4% of the general population reported the experience of AVH over the course of a month (Linszen et al., in press) when a sensitive questionnaire is used (Schutte et al., 2018). AVH in non-clinical and clinical individuals are similar in terms of loudness, personification and number of voices heard, but the perceived emotional content differs with a tendency towards negative valence content in patients (Daalman et al., 2011). Negative voice-content appears to be one of the major differences between clinical and non-clinical voice-hearers (Larøi, 2012) and is a strong predictor of experienced distress and impairment in daily functioning (Larøi et al., 2019).

It is not yet clear how we should define "negative" voice-content. Linguistic voice-content assessments, based on emotional valence estimations for individual words, may not lead to valid estimations of the negative content, since words are best interpreted in their context. Personal memories or experiences can give a certain passage a negative meaning, although the meanings of its constituting words might appear neutral or even positive. For example, a patient much detested that AVH called him by his last name, as children who bullied him at school used to do this.

In a previous study on this topic (van der Gaag et al., 2003), voice-content was rated by two independent raters. Their results indicate that both positive and negative voice-content assessed by the raters is interpreted as such by voice-hearers. However, content assessed to be neutral by independent raters could still be interpreted as either positive or negative by the voice-hearers. This finding indeed confirms that seemingly neutral voices can have a personal negative/positive valence, perhaps depending on adverse life experiences or affective processing alterations in clinical voice-hearers (Aleman & Kahn, 2005; Cohen & Minor, 2010; Reiffet al., 2012). This could lead to the hypothesis that clinical and non-clinical individuals with AVH have similar voices in terms of linguistic emotional valence, but differ in the processing or interpretation of the voice-content. The cause for more severe distress from AVH in clinical voice-hearers would then lie in affective processing, rather than in the objective valence of the AVH.

Little is known about the objective linguistic emotional valence of AVH. A recent preliminary study (n= 6) explored the emotional content of AVH compared to general inner thoughts based on linguistic emotional valence, suggesting that AVH were more negative than inner thoughts (Turkington et al., 2019). A previous study by our group (De Boer et al., 2016) showed that the AVH of individuals with a psychotic disorder contained more terms of abuse than AVH of non-clinical individuals.

Given the consistent association between negative emotional content and distress engendered by AVH, reducing negative (interpretations of) voice-content is an often-applied approach for cognitive behavioral therapy (CBT) in patients with distressing hallucinations. To further inform such lines of treatment, detailed knowledge about the emotional content of AVH is essential.

The current study examines the emotional valence of voice-content using linguistic sentiment analysis in clinical and non-clinical voice-hearers. Sentiment analysis is a method in natural language processing that aims to quantify the emotional polarity or valence of text, which can be negative, neutral or positive. Second, we assess the relation between linguistic sentiment and perceived negativity and distress in both clinical and non-clinical voice-hearers. By assessing linguistic emotional valence as well as self-rated perceived negativity we aim to establish whether negative voice-content is objectively more negative in clinical voice-hearers, or whether they process their voice-content in such a way that it leads to a more subjectively negative perception.

Based on previous work by our group (Daalman et al., 2011; De Boer et al., 2016), we hypothesize that both objective and subjective voice-content in clinical voice-hearers is more negative than in non-clinical voice-hearers. We further expect objective voice-content to be predominantly negative in the clinical voice-hearers, and predominantly positive in the non-clinical group. Finally, we expect objective voice-content to be strongly associated with subjective negativity, distress and disruption of life in both groups.

Methods

Participants

All participants experienced persistent AVHs (i.e., at least once a month for over a year). A total of 40 participants were included: 21 patients with a psychotic disorder and 19 non-clinical participants who experience AVH. Participants were included if they heard voices at least daily. Non-clinical participants were recruited through a Dutch website; for full methodology see previous reports on this sample (Daalman et al., 2011; De Boer et al., 2016). The non-clinical voice-hearers were screened for the absence of a psychiatric disorder by psychiatrists using the Comprehensive Assessment of Symptoms and History (CASH) interview (Andreasen et al., 1992) and the Structured Clinical Interview for Personality Disorder (SCID-II) (First et al., 1997). Non-clinical voice-hearers were excluded if (1) they had a diagnosis or treatment for psychiatric disorders other than depressive or anxiety disorders in complete remission; (2) they had a history of alcohol or drug abuse in the past 3 months. The Psychotic Symptom Rating Scale (PSYRATS) for auditory hallucinations was applied for the phenomenological characteristics of the hallucinations (Haddock et al., 1999). All procedures were approved by the Ethical Review Board of the University Medical Center Utrecht. All participants gave written informed consent.

Procedures

Shadowing procedure

The shadowing procedure was conducted at the University Medical Center Utrecht. Participants were instructed to repeat out loud their AVH verbatim directly upon hearing them for the duration of one minute. They were further instructed to repeat their AVH with the same intonation, loudness, and pronunciation as the voice(s) they perceived. Their verbatim repetitions were recorded using a voice-recording device. Voice recording started with the onset of the participants' repetition of the AVH and was stopped after one minute. This procedure was repeated three times per participant in the same session, resulting in three minutes of recorded voice-speech. Some participants experience AVH almost continuously, whereas others had less frequent AVH on the day of the recording. The procedure lasted between10 and 30 min, depending on the frequency of the AVH. Recordings were saved as .wav files.

Language analyses

The shadowing audio files were transcribed using CLAN software, according to the CHILDES manual (MacWhinney, 2000). All transcriptions were made by trained linguistics students who were native speakers of Dutch, and were blinded to the presence of a clinical disorder. The transcriptions were divided into utterances. Utterance boundaries were determined on the basis of prosodic and semantic coherence. Sentiment analyses were performed using Pattern (https://github.com/clips/pattern), an open-source Python package for natural language processing. The Dutch submodule contains a rule-based sentiment analyzer, which is based on a lexicon of about 4000 Dutch lemmas. The algorithm takes into account downtoners, amplifiers and negations. Downtoners are adverbs that diminish the sentiment of an adjective (e.g., "nearly dark"), whereas amplifiers strengthen it (e.g., "very dark"), and negations assert that something is not the case (e.g., "not dark"). Following previous research (Nazareth et al., 2019), this lexicon was expanded using Moors lexicon (Moors et al., 2013), which contains valence scores for approximately 4300 Dutch words. These valence scores were rescaled to the [−1; 1] range Pattern uses and were added to the lexicon, along with their corresponding part of speech (POS) tags. The final lexicon contained 8218 different Dutch nouns, verbs, adverbs and adjectives. Pattern's "parse" and "split" functions were used to annotate words with their POS tags. This lexicon is a selection of the Dutch language vocabulary, which is estimated to consist of at least 1 million words. The "sentiment" function was used to calculate the sentiment of each utterance. Mean valence scores were calculated per participant by averaging over all utterances. The variance of valence was calculated as the standard deviation of the valence of all utterances per participant. The minimal and maximal valence scores were calculated as the utterance with respectively the lowest and highest valence per participant. Valence scores in the [−.03; .03] range were considered neutral, conforming to previous research (Nazareth et al., 2019). On average, 31 utterances were obtained per participant, of which 17 received a valence score that was used in the analyses. One clinical voice-hearer only heard English hallucinations during the shadowing procedure and was therefore excluded from the analyses.

Statistics

All statistical analyses were run in IBM SPSS Statistics version 25.0.0.2. Participant characteristics were compared between groups using an analysis of variance (ANOVA) for continuous values, and χ^2 tests for categorical values. ANOVAs were used to compare the linguistic emotional valence characteristics between groups. The grouping variable was the presence/absence of a psychotic disorder. Mann–Whitney *U* tests were used to assess differences in the phenomenological characteristics of AVHs between the two groups. The phenomenological outcome measures were derived from the PSYRATS. A χ^2 test was used to test differences in distributions across groups. Bivariate Pearson's correlations were used to assess the association between linguistic valence and the characteristics of AVHs. Alpha was set at .05 for all analyses.

Results

Demographic characteristics of the clinical and the non-clinical voice-hearers are presented in Table 1. Clinical and non-clinical voice-hearers did not differ in age or sex. The non-clinical voice-hearers had a younger age of onset of AVH than the clinical voice-hearers. None of the non-clinical voice-hearers had a history of depression or anxiety disorders or used psychotropic medication. One of the clinical voice-hearers had a comorbid borderline personality disorder. The proportion of voice-utterances that were scored using Pattern differed between groups (F (1, 38) = 8.46, Partial η^2 = .181, p = .006). Pattern recognized a greater proportion of the non-clinical voice-utterances than of the clinical voice-utterances. The mean valence of the voice-utterances significantly differed between the groups (see Table 1).

Table 1. Demographics and AVH valence characteristics.

	Non-clinical voice-hearers (n = 19)	Clinical voice-hearers (n = 20)	Statistics	Effect size	p-value
Age, mean years (SD)	46.8 (12.80)	41.6 (10.18)	F(1,37) = 2.01	Partial η^2 = .051	.165
Age at onset of AVH, mean years (SD) [range]	10.8 (14.16) [2-43]	21.9 (10.14) [7-49]	F(1,33) = 7.28	Partial η^2 = .181	.011*
Time since onset of AVH, mean years (SD)	34.9 (15.14)	17.9 (15.07)	F(1,33) = 11.1	Partial η^2 = .252	.002**
Sex, M (%)	11 (57.9)	13 (61.9)	χ^2 (1, n = 39) = .067	φ=.041	.796
Diagnosis, n (%)					
Schizophrenia		12 (57.1)			
Schizoaffective disorder		3 (14.3)			
Bipolar I disorder		1 (4.8)			
Psychotic disorder NOS		5 (23.8)			
Mean valence (SD)	.16 (.124)	.04 (.154)	F(1,37) = 7.33	Partial η^2 = .165	.010*
Variance of valence (SD)	.25 (.065)	.32 (.113)	F(1,37) = 4.28	Partial η^2 = .104	.046*
Minimal valence	-.37(.228)	-.53 (.280)	F(1,37) = 3.76	Partial η^2 = .092	.060
Maximal valence	.56 (.243)	.53 (.232)	F(1,37) = .218	Partial η^2 = .006	.643

Note: SD: standard deviation; n: sample size; NOS: not otherwise specified.
* p-value <.05.
** p-value <.01.

Phenomenological characteristics, including perceived (negativity) of the voices, are presented in Table 2. The voice-utterances (clinical n = 338, non-clinical n = 310) showed significantly different objective (i.e., linguistic sentiment) valence distributions in the clinical versus the non-clinical voice-hearers (χ^2 (2, n = 648) = 23.76, φ = .192, p < .0001), see Figure 1. In the clinical individuals, 34.7% of the voice-utterances were objectively negative, 6.8% were neutral and 58.5% were positive. In the non-clinical voice-hearers 18.4% of the voice-utterances were objectively negative, 4.9% were neutral and 76.7% were positive. Post-hoc analyses revealed that the distribution of objectively positive versus negative voice-utterances differed between groups (χ^2 (1, n = 606) = 23.26, φ= .196, p < .0001), whereas the distribution of objectively positive versus neutral or negative versus neutral voice-utterances was not significantly different between groups (p > .05).

Table 2. Comparison of AVHs in clinical and non-clinical voice-hearers.

PSYRATS items	Non-clinical voice-hearers, description of median	Clinical voice-hearers, description of median	Median non-clinical voice-hearers [range]	Median clinical voice-hearers [range]	Mann-Whitney *U*	*p*-value
1. Frequency	At least once daily	Continuously / nearly continuously	4 [2-6]	6 [4-6]	209.0	.067
2. Duration	Several minutes	Several hours on end	2 [1-4]	4 [1-4]	219.5	.027*
3. Location	Inside head near ears	Inside head near ears	2 [1-4]	2 [1-4]	161.0	.807
4. Loudness	Same loudness as own voice	Same loudness as own voice	2 [1-3]	2 [1-4]	166.5	.660
5. Explanation of origin	More than 50% external	50% external / More than 50% external	3 [1-4]	2 / 3[a] [1-4]	134.0	.546
6. Amount of negative content	No negative content	Most content is negative	0 [0-4]	3 [0-4]	279.0	<.001**
7. Degree of negative content	No negative content	Negative voices directed at self-concept	0 [0-3]	3 [1-4]	284.0	<.001**
8. Amount of distress	No distress	Always distressing	0 [0-4]	4 [1-4]	292.0	<.001**
9. Intensity of distress	No distress	Severely distressing	0 [0-4]	3 [1-4]	282.5	<.001**
10. Disruption of life	No disruption	Moderate / severe disruption	0 [0-3]	2 / 3[a] [1-4]	292.5	<.001**
11. Controllability	Mostly some control	No control	1 [0-4]	4 [1-4]	262.0	<.001**

Note: PSYRATS: Psychotic Symptom Rating Scale for auditory hallucinations.
[a] Two scores were tied as median.
* *p*-value <.05.
** *p*-value <.01.

When looking at the linguistic sentiment distribution of voice-utterances over the participants, our results indicate that 65% of the clinical voice-hearers predominantly heard positive voices, 25% predominantly heard negative voices and 10% heard an equal amount of positive and negative voices. Of the non-clinical voice-hearers, 90% heard predominantly positive voices, 5% heard negative voices and 5% heard an equal amount of positive and negative voices. An example of one of the clinical voice-utterances classified as positive was "once your turn will come" (translated from the original Dutch "eens kom je aan de beurt"), whereas a negative utterance was "that bitch must die" (translated from the original Dutch "dat wijf moet dood").

The mean objective valence was strongly associated with the amount and intensity of perceived (subjective) negativity ($r = -.619$, $p = .001$, $r = -.474$, $p = .005$ respectively), amount of distress ($r = -.579$, $p = .004$), and disruption of life ($r = -.409$, $p = .016$), whereas no significant association was found between objective valence and the intensity of distress ($r = -.295$, $p = .090$).

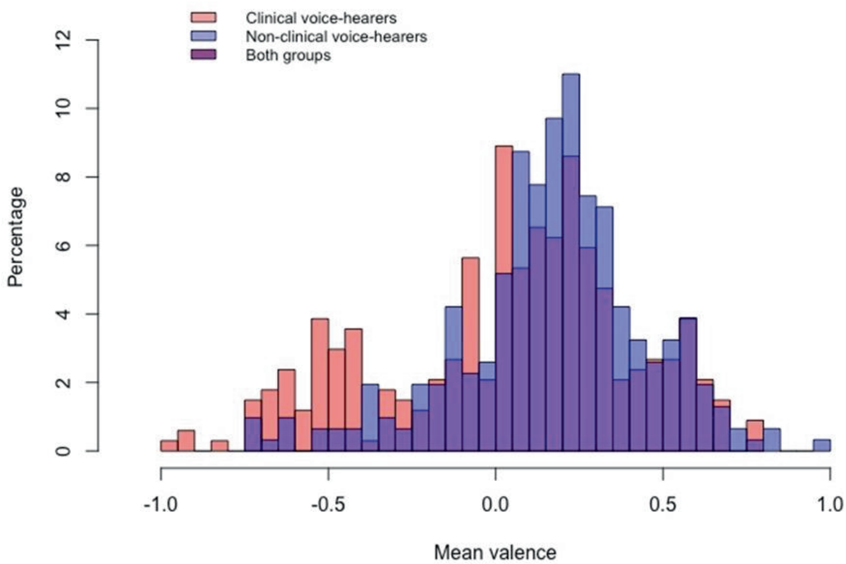


Figure 1. Distribution of sentiment of the AVH in clinical and non-clinical voice-hearers. Mean emotional valence of the AVH range from −1 to 1, where 1 indicates highly positive valence, and −1 highly negative valence. Absolute frequencies are displayed. Valence scores of 0 are considered neutral.

Discussion

In line with our expectations, we found that the AVH utterances of patients with a psychotic disorder had a more negative linguistic emotional valence than those of non-clinical voice-hearers. Our findings are in line with previous research that shows a preponderance towards negative self-rated voice-content in patients (Daalman et al., 2011; Larøi, 2012; Larøi et al., 2019). We extend these findings by showing that this tendency remains when objectively quantified, in the absence of information on linguistic context. Moreover, in contrast to expectations, most clinical and non-clinical voice-hearers predominantly heard objectively positive voices, yet the proportion of positive versus negative and neutral voices was smaller in the patients. The perceived negativity, amount of distress from the voices and the disruption of life by the voices was strongly associated with the mean linguistic emotional valence of the voices, whereas the intensity of distress from the voices was not.

Our study has both scientific and therapeutic implications. First, we have shown that even in the absence of linguistic context, patients' AVH contain more objectively negative content than AVH of non-clinical voice-hearers. This suggests that AVH language in patients is more often negative (objectively), independent of potential alterations in emotional processing, personal memories or negative associations that may additionally affect the perceived negativity. A prominent pathophysiological model for explaining negative voice-content suggests that AVH result from activation of the right hemisphere Broca's area homologue, which is associated with the production of swear words (Sommer et al., 2008; Sommer & Diederen, 2009). However, since we did not test swear words in this project, we were unable to assess whether our results are in line with this framework.

Second, CBT for AVH is currently focused on changing voice-hearers' beliefs about their voices, based on the cognitive model of hallucinations (Chadwick & Birchwood, 1994) which suggests that "distress and behavioral repertoire in voice-hearers is most closely tied to beliefs about voices, irrespective of content" (Larøi et al., 2019; Peterset al., 2012, p. 1507). Whereas we do not deny the importance of a person's beliefs in the generation of distress, our findings show that distress is closely tied to content, even when beliefs are left out of the equation. Solely focusing on beliefs about the voices might therefore not be sufficient to alleviate the distress. Indeed, whereas CBT has proven effective for AVH, effect sizes are relatively small and there is no evidence that CBT changes the perceived malevolence of voices (Sommer et al., 2012; van der Gaaget al., 2014). Our findings may contribute to developing additional angles for CBT. For example, our results show that although patients hear more objectively negative voices than non-clinical voice-hearers, both groups also hear objectively positive voices. Therefore, an additional aim in CBT could be to shift some of the attentional weight from the negative towards the positive voice-utterances, in an attempt to relieve some of the distress. This could be achieved by training a person's metalinguistic skills (Bialystok & Ryan, 1985; Tunmer et al., 1988), which can enhance their ability to focus on their positive or negative valence, rather than on the content itself. Other metalinguistic approaches to hallucinations include focusing on the grammatical structure of the voices (Corona Hernández et al., 2022), or reducing negative associations by replacing them with positive word associations (Moritz et al., 2007; Moritz & Jelinek, 2011; Moritz & Russu, 2013).

This study has some limitations. First, the sample size is small. Second, although we extended the emotional valence tool with an additional set of words, only~60% of all voice-utterances were recognized by Pattern. A word category that is not included in Pattern is swear words. Previous work by our group (De Boer et al., 2016) indicates that the voices of patients contain more swear words, yet these were not rated by Pattern. This likely affected our results and may explain in part why also the clinical voice-hearers predominantly heard positive voices and why a smaller proportion of the utterances of the clinical voice-hearers were recognized by Pattern. Third, all patients were collected through the "voices clinic" of the UMC Utrecht, which is an outpatient clinic for patients with chronic AVH. This may have led to a selection bias since only patients that regard their voices as distressing come to the clinic for treatment. Fourth, although we did not exclude lifetime diagnoses of depression or anxiety disorders in the non-clinical voice-hearers, none of the non-clinical participants had a history of mental illness. This may have influenced the perceived distress or characteristics of the voices. Finally, the AVH were obtained using the shadowing procedure, which has several limitations. Shadow recordings started with the onset of AVH and stopped after one minute. Some participants experienced AVH for the full duration of the recording, whereas others did not. This may have influenced the amount of AVH captured on record. Further, the participant is trusted to repeat the contents of their hallucinations correctly, which makes the recordings subjective. Also, as a result of the use of this method, the participant focuses on the AVH which could result in a change in cognitive processing of the hallucination. This could result in a recording that is not representative of the AVH that are generally experienced by the participant. However, it has to be acknowledged that there is no more direct way to gain access to the content of AVH, as this is a strictly private experience. In future studies, this may be checked by asking participants to rate the resemblance of the recorded AVH and the AVH they generally hear on a Likert scale.

It is important to note that sentiment analysis, in general, has its limitations. It is, for example, less capable of dealing with highly complex sentences and performs less accurately in new domains (Astya, 2017). Replication studies are required to establish whether sentiment analyses are indeed accurate at capturing negative voice-content. Future research should also focus on more in-depth linguistic analyses of the differences between clinical and non-clinical voices since, for example, the use of power or politeness by the voices can shed a light on the relationship individuals have with their voices (Demjen et al., 2020), which could provide additional angles for CBT.

In conclusion, we have shown that both clinical and non-clinical voice-hearers predominantly hear positive voices, yet the proportion of objectively negative versus positive voices is larger in patients. The linguistic emotional valence of voices is strongly associated with the perceived distress and disruption of life, irrespective of context or personal memories. This has important implications for additions to current CBT regimes since current models are based on the idea that distress in voice-hearers is caused by their beliefs about the voices, irrespective of their content. Instead, our findings suggest the opposite is also true, namely, distress is closely tied to the content of the voices, irrespective of personal beliefs.

Acknowledgments

We would like to thank those who participated in the study and lab members for their help in processing and collecting data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Note on Contributors

Janna de Boer: Conceptualization, Methodology, Formal analysis, Writing–Original draft. **Hugo Corona Hernández**: Writing–Review & Editing. **Frank Gerritse**: Methodology, Software. **Sanne Brederoo**: Writing–Review & Editing. **Frank Wijnen**: Writing–Review & Editing, Supervision. **Iris Sommer**: Writing–Review & Editing, Supervision, Funding acquisition.

References

Aleman, A., & Kahn, R. S. (2005). Strange feelings: Do amygdala abnormalities dysregulate the emotional brain in schizophrenia? *Progress in Neurobiology*, *77*(5), 283–298. https://doi.org/10.1016/j.pneurobio.2005.11.005

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH): an instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry*, *49*(8), 615–623.

Astya, P. (2017). Sentiment analysis: approaches and open issues. *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 154–158.

Baethge, C., Baldessarini, R. J., Freudenthal, K., Streeruwitz, A., Bauer, M., & Bschor, T. (2005). Hallucinations in bipolar disorder: characteristics and comparison to unipolar depression and schizophrenia. *Bipolar Disorders*, *7*(2), 136–145.

Bialystok, E., & Ryan, E. B. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly (1982-)*, 229–251.

Chadwick, P., & Birchwood, M. (1994). The omnipotence of voices. *British Journal of Psychiatry*, *164*(2), 190–201.

Cohen, A. S., & Minor, K. S. (2010). Emotional experience in patients with schizophrenia revisited: Meta-analysis of laboratory studies. *Schizophrenia Bulletin*, *36*(1), 143–150. https://doi.org/10.1093/schbul/sbn061

Corona-Hernández, H., Brederoo, S. G., de Boer, J. N., & Sommer, I. E. C. (2022). A data-driven linguistic characterization of hallucinated voices in clinical and non-clinical voice-hearers. *Schizophrenia research, 241*, 210–217. https://doi.org/10.1016/j.schres.2022.01.055

Daalman, K., Boks, M. P., Diederen, K. M., de Weijer, A. D., Blom, J. D., Kahn, R. S., & Sommer, I. E. (2011). The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *J Clin Psychiatry*, *72*(3), 320–325.

De Boer, J. N., Heringa, S. M., van Dellen, E., Wijnen, F. N. K., & Sommer, I. E. C. (2016). A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it. *Brain and Language*, *162*, 10–18.

Demjen, Z., Marszalek, A., Semino, E., & Varese, F. (2020). *Applying Linguistics in Illness and Healthcare Contexts: Chapter one*. Bloomsbury Publishing.

First, M. B., Gibbon, M., Spitzer, R. L., Benjamin, L. S., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM-IV® Axis II Personality Disorders SCID-II*. American Psychiatric Pub.

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychological Medicine*, *29*(4), 879–889.

Larøi, F. (2012). How do auditory verbal hallucinations in patients differ from those in non-patients? *Frontiers in Human Neuroscience*, *6*, 25.

Larøi, F., Thomas, N., Aleman, A., Fernyhough, C., Wilkinson, S., Deamer, F., & McCarthy-Jones, S. (2019). The ice in voices: Understanding negative content in auditory-verbal hallucinations. *Clinical Psychology Review*, *67*(July 2018), 1–10. https://doi.org/10.1016/j.cpr.2018.11.001

Linszen, M., de Boer, J., Schutte, M., Begemann, M., de Vries, J., Koops, S., Blom, R., Bohlken, M., Heringa, S., Blom, J., & Sommer, I. (2021). Hallucinations in the general population: content and phenomenology. *NPJ Schizophrenia*, *In press*.

MacWhinney, B. (2000). *Tools for Analyzing Talk Part 1: The CHAT Transcription Format* (3rd ed.). Lawrence Erlbaum Associates. https://doi.org/10.1111/1460-6984.12101/abstract

Maijer, K., Begemann, M. J. H., Palmen, S. J. M. C., Leucht, S., & Sommer, I. E. C. (2018). Auditory hallucinations across the lifespan: a systematic review and meta-analysis. *Psychological Medicine*, *48*(6), 879–888.

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., Van Schie, K., Van Harmelen, A.-L., De Schryver, M., De Winne, J., & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177.

Moritz, S., & Jelinek, L. (2011). Further evidence for the efficacy of association splitting as a self-help technique for reducing obsessive thoughts. *Depression and Anxiety*, *28*(7), 574–581.

Moritz, S., Jelinek, L., Klinge, R., & Naber, D. (2007). Fight fire with fireflies! Association splitting: A novel cognitive technique to reduce obsessive thoughts. *Behavioural and Cognitive Psychotherapy*, *35*(5), 631.

Moritz, S., & Russu, R. (2013). Further evidence for the efficacy of association splitting in obsessive-compulsive disorder. An internet study in a Russian-speaking sample. *Journal of Obsessive-Compulsive and Related Disorders*, *2*(2), 91–98.

Nazareth, D. S., Tournier, E., Leimkötter, S., Janse, E., Heylen, D., Westerhof, G. J., & Truong, K. P. (2019). An acoustic and lexical analysis of emotional valence in spontaneous speech: Autobiographical memory recall in older adults. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *2019-Septe*, 3287–3291. https://doi.org/10.21437/Interspeech.2019-1823

Peters, E. R., Williams, S. L., Cooke, M. A., & Kuipers, E. (2012). It's not what you hear, it's the way you think about it: Appraisals as determinants of affect and behaviour in voice hearers. *Psychological Medicine*, *42*(7), 1507–1514. https://doi.org/10.1017/S0033291711002650

Reiff, M., Castille, D. M., Muenzenmaier, K., & Link, B. (2012). Childhood abuse and the content of adult psychotic symptoms. In *Psychological Trauma: Theory, Research, Practice, and Policy* (Vol. 4, Issue 4, pp. 356–369). Educational Publishing Foundation. https://doi.org/10.1037/a0024203

Schutte, M., Linszen, M., Ffytche, D., Koops, S., van Dellen, E., Heringa, S., Slooter, A., Teunisse, R., van den Heuvel, O., Lemstra, A., Foncke, E., Slotema, C., de Jong, J., Hoffman, R., Strauss, C., Thomas, N., Rossell, S., & Sommer, I. (2018). *Development and psychometric validation of a new transdiagnostic instrument: the Questionnaire for Psychotic Experiences (QPE).*

Sommer, I., & Diederen, K. (2009). Language production in the non-dominant hemisphere as a potential source of auditory verbal hallucinations. *Brain*, *132*(10), e124–e124.

https://doi.org/10.1093/brain/awp040

Sommer, I., Diederen, K., Blom, J., Willems, A., Kushan, L., Slotema, K., Boks, M., Daalman, K., Hoek, H., & Neggers, S. (2008). Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain*, *131*(12), 3169–3177.

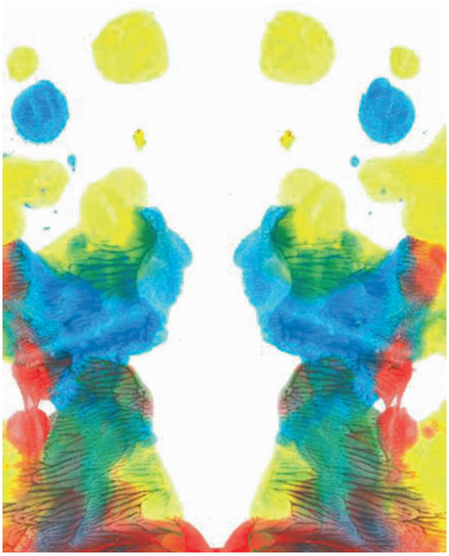
Sommer, I. E. C., Slotema, C. W., Daskalakis, Z. J., Derks, E. M., Blom, J. D., & van der Gaag, M. (2012). The treatment of hallucinations in schizophrenia spectrum disorders. *Schizophrenia Bulletin*, *38*(4), 704–714.

Tunmer, W. E., Herriman, M. L., & Nesdale, A. R. (1988). Metalinguistic abilities and beginning reading. *Reading Research Quarterly*, 134–158.

Turkington, G. D., Lawrence, T., & Stuart, W. (2019). Linguistic analysis of the valence, arousal and dominance of auditory hallucinations and internal thoughts in schizophrenia: Implications for psychoeducation and CBT. *Cogent Psychology*, *6*(1). https://doi.org/10.1080/23311908.2019.1703463

Van Der Gaag, M., Hageman, M. C., & Birchwood, M. (2003). Evidence for a cognitive model of auditory hallucinations. *Journal of Nervous and Mental Disease*, *191*(8), 542–545. https://doi.org/10.1097/01.nmd.0000082183.95853.ec

van der Gaag, M., Valmaggia, L. R., & Smit, F. (2014). The effects of individually tailored formulation-based cognitive behavioural therapy in auditory hallucinations and delusions: a meta-analysis. *Schizophrenia Research*, *156*(1), 30–37.



CHAPTER 4

Assessing coherence through linguistic connectives: analysis of speech in patients with schizophrenia-spectrum disorders

H. Corona-Hernández[1], J.N. de Boer[1,2], S.G. Brederoo[1,3], A.E. Voppel[1], & I.E.C. Sommer[1,3]

[1] Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, University of Groningen, The Netherlands

[2] Department of Psychiatry, University Medical Center Utrecht, Utrecht University & Brain Center Rudolf Magnus, Utrecht, the Netherlands

[3] Department of Psychiatry, University Medical Center Groningen, University of Groningen, The Netherlands

Schizophrenia research. 2023; 259: 48–58.

ABSTRACT

Background: Incoherent speech is a core diagnostic symptom of schizophrenia-spectrum disorders (SSD) that can be studied using semantic space models. Since linguistic connectives signal relations between words, they and their surrounding words might represent linguistic loci to detect unusual coherence in speech. Therefore, we investigated whether connectives' measures are useful to assess incoherent speech in SSD.
Methods: Connectives and their surrounding words were extracted from transcripts of spontaneous speech of 50 SSD-patients and 50 control participants. Using word2vec, two different cosine similarities were calculated: those of connectives and their surrounding words (connectives-related similarity), and those of free-of-connectives words-chunks (non-connectives similarity). Differences between groups in proportion of five types of connectives were assessed using generalized logistic models, and connectives-related similarity was analyzed through non-parametric multivariate analysis of variance. These features were evaluated in classification tasks to differentiate between groups.
Results: SSD-patients used less contingency (e.g., because) (p = .008) and multiclass connectives (e.g., as) (p < .001) than control participants. SSD-patients had higher minimum similarity of multiclass (adj-p = .04) and temporality connectives (e.g., after) (adj-p < .001), narrower similarity-range of expansion (e.g., and) (adj-p = .002) and multiclass connectives (adj-p = .04), and lower maximum similarity of expansion connectives (adj-p = .005). Using connectives' features alone, SSD-patients and controls could be distinguished with 85% accuracy.
Discussion: Our results show that SSD-speech can be distinguished from speech of control participants with high accuracy, based solely on connectives' features. We conclude that including connectives could strengthen computational models to categorize SSD.

Keywords
Classification; Disorganized speech; Grammatical connectivity; Schizophrenia-Spectrum Disorders; Semantic cosine similarity; Word2vec

1. Introduction

Disorganized speech is a core feature of schizophrenia-spectrum disorders (SSD) (American Psychiatric Association, 2013) that has been increasingly assessed using semantic space models (Corcoran et al., 2020; Corcoran and Cecchi, 2020; Hitczenko et al., 2020). Such computational models create *n* dimensions, each standing for an abstract feature of word meaning. The represented meaning (i.e., vector) of a given word can thus be located within the semantic space of *n* dimensions, and it is posited that words with similar meaning are found close to each other within a given semantic space (Landauer et al., 1998; Mikolov et al., 2013b). Using these models, it has been shown that patients with SSD can be distinguished from healthy controls with accuracies between 70% and 93% (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2020; Tang et al., 2021; Voppel et al., 2021), while predicting psychosis onset in at-high-risk individuals has accuracies ranging from 72% to 100% (Bedi et al., 2015; Corcoran et al., 2018; Rezaii et al., 2019).

Disorganization in speech is considered to signal a reduction in the underlying semantic coherence of a given message (Corcoran and Cecchi, 2020; Hitczenko et al., 2020). To be attained, coherence requires *thematic continuity* and *grammatical connectivity* (Givón, 2020). While *thematic continuity* is reflected in the maintenance of semantic content, *grammatical connectivity* refers to the use of explicit markers to hierarchically organize the sequence of the content. Syntactically, *grammatical connectivity* is most clearly instantiated by linguistic connectives, which relate two or more words, clauses, or sentences to each other (Maat and Sanders, 2006; Sanders and Maat, 2006; van der Vliet and Redeker, 2014). Importantly, connectives establish different types of explicit coherence relations in discourse, such as comparative (e.g., this flower is red, *while* that other is white), contingent (e.g., I have to replace this piece *because* it is damaged), expansive (e.g., *besides* being mammals, gorillas are primates), and temporal (e.g., we will go outside *after* the rain stops) (Bourgonje et al., 2018; Stede et al., 2019). Thus, in fine-tuning semantic space models to better quantify disorganized speech, it could be valuable to separately assess *thematic continuity* and *grammatical connectivity*.

General semantic content has been the main focus of interest in previous studies using semantic space models to quantify coherence. Specifically, in speech of patients with SSD, unusual general semantic content has often been assessed across entire interviews or conversations. Typically, the measures for analysis are obtained by averaging series of semantic distances (i.e., cosine similarities) across all words or sentences uttered by patients, one after each other (Hitczenko et al., 2020). This form of assessment is reasonable considering that semantic space models perform better if they are built upon sentences rather than upon speech samples from word-association or verbal-fluency tasks (de Boer et al., 2018). However, this procedure hampers the possibility to quantify the coherence that relates to syntactic markers of connectivity (i.e., connectives).

While the use of connectives has not been separately assessed in semantic space models yet, previous studies have examined their occurrence in speech. Patients with SSD have been found to use less connectives of differentiation in comparison to control participants, (Just et al., 2020), and less causal, contrastive, and logical connectives when compared to adults with a diagnosis of HIV+ (Willits et al., 2018). In contrast, another study showed that untreated first episode psychosis patients with high scores in conceptual disorganization (PANSS Item P2) overall used more connectives than control participants (Mackinley et al., 2021).

These inconsistencies in results currently limit our knowledge about the frequency and coherence with which different types of connectives are used by patients with SSD compared to control participants. Moreover, even though semantic space models have been shown to be reliable tools to assess disorganized speech in patients with SSD, no previous research has specifically focused on grammatical connectivity. Considering this, in the present study we first evaluated whether patients with SSD and control participants use different types of connectives in similar proportions. Second, by calculating cosine similarity between connectives and their surrounding words (i.e., connectives-related similarity), we assessed whether connectives and their surrounding words can be used as linguistic loci to detect unusual coherence in speech of patients with SSD. Third, we tested how automatic classification driven by connectives-related similarity compares to another driven by non-connectives similarity, and how accurately connectives-related similarity and proportions per type of connective together could distinguish patients with SSD from control participants.

2. Methods

2.1. Participants

Fifty individuals with a schizophrenia-spectrum disorder and fifty healthy control participants, all native Dutch speakers, took part in this study. These participants had been previously investigated in Voppel et. al (2021). Their inclusion took place at the University Medical Center Utrecht. Patients' diagnoses were established by a trained physician, and confirmed using either the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992) or the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). Patients' severity of symptoms was assessed with the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987). Control participants were included if they had neither current nor a history of psychiatric disorders. All participants gave written informed consent before obtaining the measurements.

2.2. Speech sampling

Speech was elicited using a semi-structured interview, comprising 60 open-ended questions, from which a subset was presented in a semi-randomized order across all participants. The questions were designed to elicit spontaneous speech but prevent excessive emotional arousal, with topics such as life experiences, current daily habits, hobbies, and hypothetical situations, avoiding health-related and psychopathology-related topics. To prevent participants from adjusting their own speech, they were informed about the research aim only after concluding the interview. All interviews were conducted by trained researchers. The elicited verbal samples were audio-recorded and later transcribed following the CHILDES protocol (MacWhinney, 2000). Transcribers were blind to participants' group. Procedures were approved by the ethical committee of the University Medical Center Utrecht.

2.3. Linguistic features

2.3.1. Data preprocessing

Using the CHILDES transcripts, individual plain text files were derived for each participant. Speech of the interviewers, punctuation marks, metadata, headers, special characters, and markers of events (e.g., &=laughs) were all excluded from these files. In the resulting files, all words were set to lowercase, and grammatical contractions (e.g., *t'is*, standing for *het is*, "it is") were retained. Fillers (e.g., *ehm*) and repetitions have been shown to add noise to semantic similarity calculations (Iter et al., 2018). However, they were also retained in the transcripts used for analysis for three reasons. First, it is currently unknown whether fillers and repetitions might influence connectives-related similarity. Second, despite recent attempts to control for "inadequate repetitions" of speech (Just et al., 2020), there is no standard procedure for avoiding this bias. Third, fillers (Tang et al., 2021) and repetitions (Andreasen, 1979; Hong et al., 2015; Maher, 1972) have been shown to be important to distinguish patients with SSD from control participants.

2.3.2. Selection and occurrence of connectives

Based on Bourgonje et al. (2018), 188 different Dutch connectives were selected for this study. These connectives were originally divided in four broad categories, and we created a fifth category (i.e., multiclass) to be filled in with all connectives that were listed in more than one category, excluding such connectives from their initial lists and being relocated to this new one. Connectives ultimately belonged to only one of the following categories: comparison (*n* = 37), contingency (*n* = 48), expansion (*n* = 44), multiclass (*n* = 23), and temporality (*n* = 36) (see supplementary materials: 1. List of connectives in Dutch).

For each preprocessed transcript, all occurrences of the 188 different Dutch connectives were automatically extracted using the R "quanteda" package (Benoit et al., 2018). Subsequently, all the extracted connectives were automatically given the label of the category they belonged to. For each occurrence, along with the connective, the previous and the following three words were retained, considering them as part of the surrounding context of the connective, resulting in a seven-words fixed window size. This length was chosen for two main reasons: shorter word-windows are less suited to reveal differences in cosine similarity between groups (Elvevåg et al., 2007), and larger word-windows were found to be poorer informative features for classification tasks (Voppel et al., 2021). All cases in which there were fewer words in the surrounding context of the connective were also preserved (i.e., connectives occurring at the end of an interview), leading a subset of instances to have less than seven words.

2.3.3. Semantic space model

A semantic space with 300 dimensions was modeled using the skip-gram method of the word2vec learning algorithm (Mikolov et al., 2013b). It was trained on more than five-million words from The Netherlands' transcripts collection of the *Corpus Gesproken Nederlands* (CGN) (van Eerten, 2007), using the R "word2vec" package (Wijffels, 2020). In this model, each dimension might be taken to represent an abstract feature of word meaning, and the meaning of each word (i.e., its word embedding) is the vector indicating the position of the word relative to the 300 semantic dimensions of the model. Using the skip-gram method, the word2vec algorithm computes each word embedding in a few steps. For each word, first a random embedding is created. Next, using all instances of the word and its surrounding words as constraint, the random embedding is iteratively changed to resemble the embeddings of its surrounding words more, and the embeddings of words which do not appear nearby less (Mikolov et al., 2013b, 2013a). Finally, each word is assigned a fixed and unique embedding, which can then be used to measure the semantic (i.e., cosine) similarity between embeddings (Mikolov et al., 2013b).

2.3.4. Computation of connectives-related and non-connectives cosine similarities

For each seven-words chunk having a connective, the connectives-related cosine similarity was operationalized as the cosine similarity between the embedding of the connective and the as-a-whole averaged embedding of the three previous and three following words. All segments of the transcripts no longer containing connectives (henceforth, free-of-connectives segments) were also split into chunks of seven words. For each free-of-connectives segment, the non-connectives similarity was operationalized as the cosine similarity between the embedding of the fourth word and the as-a-whole averaged embedding of the three previous and three following words. Thus, the same procedure was used to obtain the two different types of cosine similarities.

Connectives-related and non-connectives similarities were first calculated separately for all the chunks of words per participant. Then, six different measures of the cosine similarities were independently obtained per participant: maximum, mean, median, minimum, range, and variance. Finally, these six different measures of cosine similarity were averaged per group for each type of cosine similarity (see Figure 1). In all cases, cosine similarities could range from -1 to 1, with -1 standing for the lowest possible similarity, and 1 for the highest possible similarity.

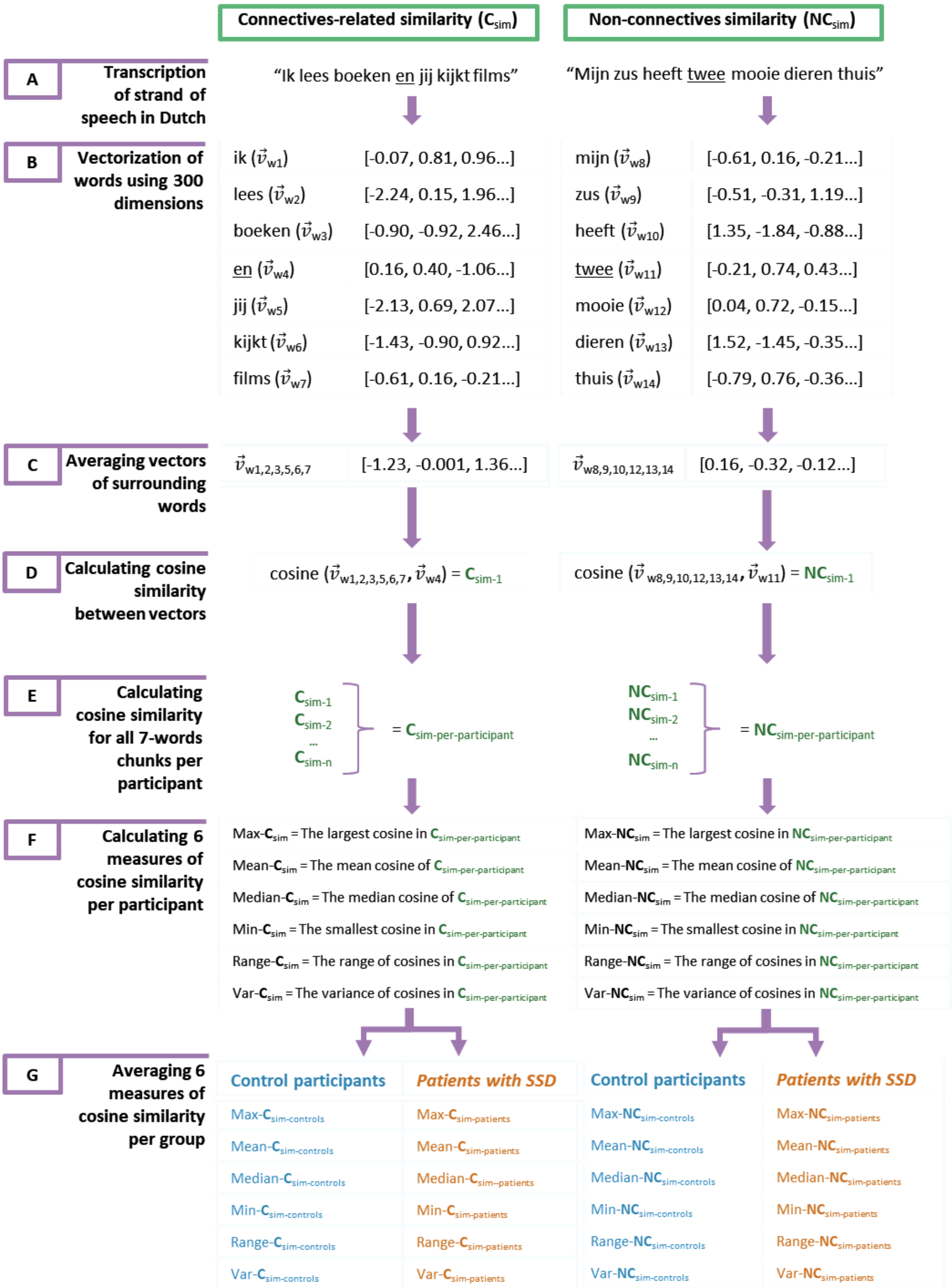


Figure 1. Steps (A-G) taken to calculate the connectives-related and the non-connectives cosine similarity. Note that the word "en" is a connective, while the word "twee" is not.

2.4. Analysis

2.4.1. Statistics

Groups were compared with regard to demographic continuous variables through independent one-way analysis of variance (ANOVA), and nominal variables through Chi-square tests without continuity correction.

To assess differences between groups in the proportion of types of connectives relative to all words used, we carried out generalized linear mixed-effects logistic regression models using the *glmer* function from the R "lme4" package (Bates et al., 2015), with proportion as dependent variable. Following Baayen (2008) and Winter (2020), group (patients with SSD *vs* control participants) and type of connective (comparison *vs* contingency *vs* expansion *vs* multiclass *vs* temporality) were considered as fixed-effects factors, and participants as random-effects factors (allowing by-participant varying intercepts and varying slopes). Implementing a forward-testing approach, we carried out stepwise model comparison between a series of independent regression models in order to arrive at the model that best fitted the data. Likelihood ratio tests were performed to assess whether there were significant differences between each pair of models being compared (Baayen, 2008).

To assess group differences for the connectives-related similarity, we used non-parametric multivariate analysis of variance (MANOVA), followed by post-hoc Wilcoxon rank sum tests with Holm correction.

For all correlations, we used the Spearman's rank non-parametric test, correcting for multiple comparisons when necessary. Statistical results with (adjusted) p-values < .05 were considered to be significant. All analyses were done in *RStudio* version 1.4.1103 (RStudio Team, 2019) running *R* version 4.1.0 (R Core Team, 2020).

2.4.2. Classification tasks

For reliable results, in all classification tasks the models were trained using 10-fold cross-validation, repeated ten times. This means that, for each iteration, the learning algorithm used nine-tenths partitions of the data for training, and one-tenth for testing. Accuracy, sensitivity, specificity, and area under the curve (AUC) for receiver operating characteristic (ROC) were obtained in order to assess the performance of each classifier. All tasks were conducted using the R "caret" package (Kuhn, 2021).

2.4.2.1. Connectives' vs non-connectives' features.

To assess whether connectives-related and non-connectives similarities might yield similar classification results, we first obtained and evaluated the performance of a control-classifier. Afterwards, we assessed how much improvement in performance this control-classifier could gain by independently adding either connectives' features or non-connectives similarity to it.

The control-classifier was built by training a random forest algorithm, using the mean, minimum and variance of general semantic similarity from sliding-windows between 5 and 10 words. This algorithm has been proven to be one of the best in performing binary classification (Fernández-Delgado et al., 2014). Mean, minimum, variance, and the 5-10 range of windows were chosen as parameters based on previous results showing that they were highly informative for classification (Voppel et al., 2021).

For the comparison between the connectives' features and the non-connectives similarity, number of features was controlled for (see Table 1). To rule out that amount of data influenced the results, a set of connectives-related similarity controlling for this was also calculated (see Table 1). These procedures to control for amount of data and number of features were exclusively done for this series of classification tasks.

2.4.2.2. Performance of connectives' features alone

Using connectives-related similarity per type of connective either alone or along with proportions of use per type of connective, random forest and support vector machine with polynomial kernel algorithms were trained to perform a binary classification between patients with SSD and control participants. For these classification tasks, the six measures of connectives-related similarity were calculated independently for each of the five types of connectives (see Table 1).

Table 1. Details of the different features used for the classification tasks.

Name of features used for classification	Use in classification task: Connectives' features *vs* non-connectives similarity	Use in classification task: Performance of connectives' features alone	Number of features	Description
Proportions per type of connective	✓	✓	5	Proportions of use of the comparison, contingency, expansion, multiclass, and temporality connectives.
Non-connectives similarity	✓		6	Maximum, mean, median, minimum, range and variance of all the free-of-connectives chunks.
General connectives-related similarity (random sub-sampling)	✓		6	Maximum, mean, median, minimum, range and variance of all connectives taken together (i.e., regardless of type of connective). These measures were calculated based on a random subsampling of connectives-chunks, which were matched to the number of free-of-connectives chunks per participant[1].
General connectives-related similarity (all chunks)	✓		6	Maximum, mean, median, minimum, range and variance of all connectives taken together (i.e., regardless of type of connective) calculated based on all connectives-chunks.
Per-type-connectives-related similarity with smallest adj. *p*-values	✓		6	Connectives-related similarity measures that, across the types of connectives, have the smallest adj. *p*-values as found through post-hoc analyses (see Table 7).
Proportions and per-type-connectives-related similarity significantly different between groups	✓		7	Connectives-related similarity measures and proportions per type of connective that were found to be significantly different between groups (see Table 5 and Table 7).
Sliding-window similarity	✓		18	Mean, minimum and variance of similarity from sliding-windows between 5 and 10 words. Partial set of the cosine similarities used in Voppel et al. (2021).
Connectives-related similarities per type of connective		✓	30	Maximum, mean, median, minimum, range and variance calculated for each of the 5 different types of connectives.

[1] Across all participants, there were more connectives-chunks than free-of-connectives chunks. This was due to partial overlap of some connectives-chunks when two connectives were used too close to each other. Accordingly, for each participant, a random subsampling of the connectives-chunks was carried out.

3. Results

3.1. Demographics and speech sample

The majority of patients in our sample had a diagnosis of psychosis not otherwise specified (46%), followed by schizophrenia (38%), schizoaffective (14%) and schizophreniform disorder (2%). There were no significant differences between groups in age (p = .98) and sex (p = .23). Years of education were significantly less for patients with SSD than for control participants (p = .001). In patients with SSD, the mean dose of antipsychotic medication as measured in chlorpromazine equivalence was 226.1 milligrams. Thirty-two patients used tight-binding antipsychotics, sixteen patients used loose-binding medication, and two patients were not receiving antipsychotic medication (see Table 2).

General characteristics of the participants' speech sample and use of connectives are presented in Table 3. Since patients with SSD had significantly less years of education than control participants, possible correlations between years of education and basic features of the speech sample were assessed (i.e., tokens and types). Neither number of running words (tokens) nor number of different word forms (types) correlated to years of education, (both *rho*(ρ) < 0.10, both p > .05).

Table 2. Demographic characteristics of participants.

Characteristic	Patients (n=50)	Controls (n=50)	Statistic value	*p*-value
Age in years, mean (SD)	31.6 (12.46)	31.5 (12.69)	F (1,98) = 0.001	.98
Females, n (%)	14 (28%)	9 (18%)	χ^2 (1) = 1.41	.23
Years of education, mean (SD) [a]	12.8 (2.22)	14.3 (2.39)	F (1,97) = 10.49	.001
Antipsychotic medication				
Chlorpromazine equivalence (mg), mean (SD) [b]	226.1 (142.81)			
Loose binding, n (%)	16 (32%)			
None	2 (4%)			
Tight binding	32 (64%)			
Diagnosis, n (%)				
Psychosis not otherwise specified	23 (46%)			
Schizophrenia	19 (38%)			
Schizoaffective	7 (14%)			
Schizophreniform	1 (2%)			
PANSS, mean (SD)				
Negative	13.4 (5.19)			
Positive	11.1 (4.41)			
General	26.5 (7.74)			
Total	51.2 (13.85)			

[a] Data available for all control participants, but only for 49 patients.
[b] Data available for 44 patients alone: two patients were not taken antipsychotic medication, and, for four patients, there was no information about dose equivalence.
n = sample size, SD = standard deviation.

Table 3. Characteristics of the speech produced by the participants per group.

Feature	Patients (n=50)	Controls (n=50)
Basic description of words		
Number of running words (*tokens*)		
Mean (SD)	1161.8 (753.27)	1819.4 (458.07)
Total	58,090	90,971
Number of different word forms (*types*)		
Mean (SD)	342.5 (152.46)	498.6 (89.02)
Total	17,125	24,930
Connectives		
Number of connectives of comparison		
Mean (SD)	39.6 (29.43)	62.4 (25.53)
Total	1982	3122
Number of connectives of contingency		
Mean (SD)	21.6 (19.40)	39.9 (18)
Total	1080	1997
Number of connectives of expansion		
Mean (SD)	46.9 (37.39)	79.9 (28.28)
Total	2348	3996
Number of connectives of multiclass		
Mean (SD)	28.6 (29.05)	58.4 (24.78)
Total	1430	2924
Number of connectives of temporality		
Mean (SD)	14.1 (11.51)	23.6 (10.40)
Total	709	1181

n = sample size, SD = standard deviation.

3.2. Proportion of connectives

Following stepwise model comparison with a forward-testing approach, the model that best fitted the data included an interaction between group and type of connective, as well as by-participant varying intercepts and varying slopes for type of connective per participant (see Table 4). Relative to connectives of comparison, patients with SSD had a lower probability of using connectives of contingency ($p < .001$), multiclass ($p < .001$), and temporality ($p < .001$), and a higher probability of using connectives of expansion ($p < .001$). Similarly, control participants had a lower probability of using connectives of contingency ($p < .001$) and temporality ($p < .001$), and a higher probability of using connectives of expansion ($p < .001$). When comparing the groups, patients with SSD had a lower probability of using connectives of contingency ($p = .008$) and multiclass ($p < .001$) than control participants (see Table 5 and Figure 2). The structure of the random-effects factors is shown in Table 6.

Considering patients with SSD had significantly less years of education, it was assessed whether this could have confounded the above-mentioned results. Independent two-tailed bivariate correlations were conducted, and false positive results were controlled using Holm correction. Results showed no significant correlations (all *rho*(ρ) < 0.21, all adj-*p* > .05).

Table 4. Stepwise procedure followed to obtain the logistic model that best fitted the data.

Factors in the model	Likelihood ratio test		
	Log-lik. increase	Statistic value	*p*-value
Participants as varying intercepts			
+Group	7	χ^2 (1) = 13.53	< .001
+Type of connective	3556062	χ^2 (4) = 7112125	< .001
+Varying slopes for type of connective per participant	1010919	χ^2 (14) = 2021838	< .001
+Interaction (Group*Type of connective)	9	χ^2 (4) = 17.19	0.0017

Note: the first row stands for the initial random model. Each subsequent row shows how the goodness of fit increased when the factor in the row was added to the model that included all preceding factors.

Table 5. Fixed-effects factors in the model that best fitted the data on proportions of connectives.

Model		Estimate	Standard Error	*z*-value	*p*-value
Control participants and connectives of comparison as baseline levels	(Intercept)	-3.24	0.04	-72.16	< .001
	Group				
	Patients	0.009	0.06	0.14	.88
	Type of connective				
	Contingency	-0.43	0.07	-6.11	< .001
	Expansion	0.34	0.05	6.03	< .001
	Multiclass	-0.07	0.06	-1.17	.23
	Temporality	-0.97	0.07	-12.57	< .001
	Interactions				
	Patients*Contingency	-0.26	0.10	-2.61	.008
	Patients*Expansion	-0.08	0.08	-0.98	.32
	Patients*Multiclass	-0.35	0.08	-4.16	< .001
	Patients*Temporality	-0.10	0.10	-0.99	.31
Patients and connectives of comparison as baseline levels	(Intercept)	-3.23	0.04	-71.94	< .001
	Group				
	Controls	-0.009	0.06	-0.14	.88
	Type of connective				
	Contingency	-0.70	0.07	-9.80	< .001
	Expansion	0.26	0.05	4.64	< .001
	Multiclass	-0.42	0.06	-7.06	< .001
	Temporality	-1.08	0.07	-13.98	< .001
	Interactions				
	Controls*Contingency	0.26	0.10	2.61	.008
	Controls*Expansion	0.08	0.08	0.98	.32
	Controls*Multiclass	0.35	0.08	4.16	< .001
	Controls*Temporality	0.10	0.10	0.99	.31

Note: estimates are "log odds" (i.e., logits). Positive estimates reflect an increase in probability and negative ones reflect a decrease. For a general guide to mixed-effects models in linguistics and their interpretation, see Baayen (2008) and Winter (2020).

Table 6. Random-effects parameters in the best model fitted to proportion of connectives.

Groups	Name	Variance	Standard deviation
Participant	(Intercept)	0.1012	0.3181
	Contingency	0.2587	0.5086
	Expansion	0.1672	0.4089
	Multiclass	0.1844	0.4294
	Temporality	0.3004	0.5481

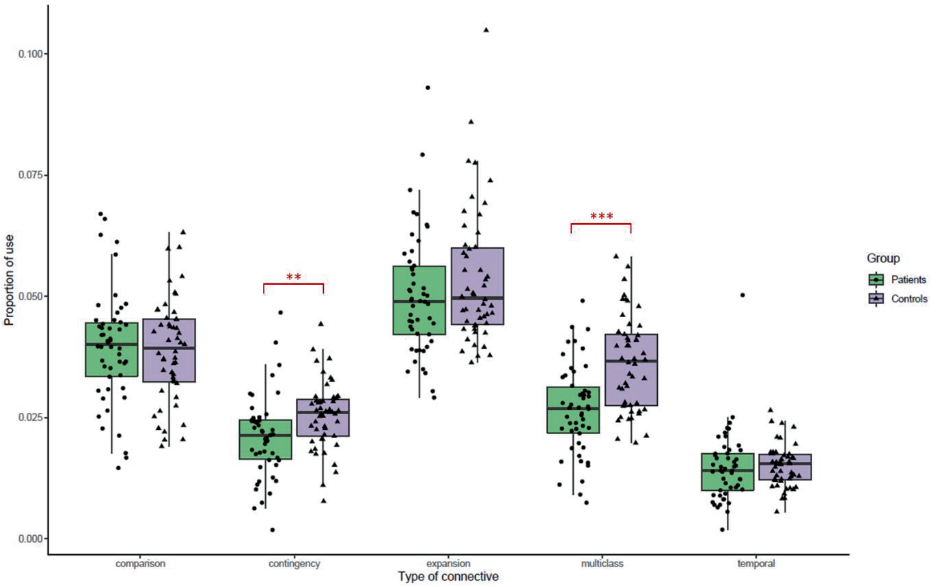


Figure 2. Mean proportion of use for each type of connective per group.

3.3. Connectives-related similarity

Multivariate analysis of variance showed that there were significant differences between groups in connectives-related similarity, F (10.3, 1018.3) = 5.2, $p < .001$. Post-hoc analyses showed that patients with SSD had higher minimum similarity of temporality connectives (adj-p < .001), as well as narrower range (adj-p = .002) and lower maximum similarity of expansion connectives (adj-p = .005) than control participants. Additionally, compared to controls, patients had narrower range (adj-p = .04) and higher minimum similarity of multiclass connectives (adj-p = .04) (see Table 7). Maximum similarity of expansion connectives positively correlated to years of education (*rho*(ρ) = 0.30, adj-p = .01), while the other four connectives-related similarities did not (all *rho*(ρ) ≤ 0.16, all adj-p > .05). In performing these analyses, a missing value in the variance of contingency connectives in one patient, and a missing value in the variance of temporality connectives in another patient, were substituted with zeros. Running the same analyses with the exclusion of these two patients did not change the results of these variables. For a full overview of the results with the exclusion of these two patients, see supplementary materials: 2. Additional analyses on connectives-related similarity.

Table 7. Differences in connectives-related similarity measures between groups.

Type of connective	Coherence measure	Patients (n=50), average (SD)	Controls (n=50), average (SD)	Statistic		r effect size
				W	Adj. *p*-value	
Comparison	Mean	0.51 (0.01)	0.50 (0.01)	881	.23	
	Median	0.51 (0.01)	0.51 (0.01)	920	.41	
	Maximum	0.62 (0.04)	0.63 (0.04)	1328	≈.99	
	Minimum	0.38 (0.04)	0.37 (0.02)	987	.91	
	Range	0.23 (0.05)	0.25 (0.05)	1421	≈.99	
	Variance	0.003 (0.001)	0.003 (0.001)	964	.68	
Contingency	Mean	0.48 (0.03)	0.47 (0.01)	827	.08	
	Median	0.49 (0.03)	0.48 (0.01)	814	.06	
	Maximum	0.59 (0.06)	0.59 (0.03)	1307	≈.99	
	Minimum	0.36 (0.06)	0.33 (0.03)	832	.09	
	Range	0.23 (0.09)	0.26 (0.06)	1565	.48	
	Variance	0.004 (0.003)	0.004 (0.001)	1129	≈.99	
Expansion	Mean	0.46 (0.01)	0.46 (0.01)	1190	≈.99	
	Median	0.45 (0.01)	0.45 (0.01)	1205	≈.99	
	Maximum	0.61 (0.05)	0.66 (0.06)	1792	.005	.37
	Minimum	0.33 (0.03)	0.31 (0.03)	937	.48	
	Range	0.28 (0.07)	0.34 (0.06)	1814	.002	.38
	Variance	0.005 (0.002)	0.005 (0.002)	1307	≈.99	
Multiclass	Mean	0.50 (0.02)	0.50 (0.01)	1235	≈.99	
	Median	0.49 (0.02)	0.49 (0.01)	1133	≈.99	
	Maximum	0.62 (0.06)	0.64 (0.03)	1592	.35	
	Minimum	0.38 (0.04)	0.35 (0.04)	797	.04	-.31
	Range	0.23 (0.08)	0.28 (0.05)	1704	.04	.31
	Variance	0.004 (0.002)	0.004 (0.001)	1239	≈.99	
Temporality	Mean	0.43 (0.02)	0.42 (0.01)	882	.23	
	Median	0.43 (0.03)	0.42 (0.01)	929	.46	
	Maximum	0.52 (0.06)	0.52 (0.05)	1342	≈.99	
	Minimum	0.35 (0.03)	0.31 (0.03)	603	.0002	-.44
	Range	0.16 (0.07)	0.20 (0.06)	1638	.16	
	Variance	0.003 (0.002)	0.003 (0.001)	1328	≈.99	

n = sample size, SD = standard deviation.

3.4. Classification tasks

3.4.1. Connectives' vs non-connectives' features

The control classifier (RF-c) yielded 83.5% accuracy. By adding non-connectives similarity to the classification (RF-non-conn), accuracy resulted in 84.9%. Matched in number of features and amount of data, the classifier that rather included connectives-related similarity (RF-conn-I) yielded 85% accuracy. The combination of the sliding-window measures and the connectives-related similarity matched in features and data (RF-conn-I) yielded the highest sensitivity (81.2%). The combination of the sliding-window measures and the connectives' features that were significantly different between groups (RF-conn-V) yielded the highest specificity (93.6%) (see Table 8).

3.4.2. Performance of connectives' features alone

Using connectives-related similarity per type of connective alone, the best classifier (RF-I) yielded 79.4% accuracy, 75% sensitivity and 83.8% specificity. Combining these features with the proportions of use per type of connective, the best classifier (SVM-II) yielded 85% accuracy, 83.8% sensitivity and 86.2% specificity (see Table 9).

Table 8. Comparison of classification performance between connectives' features and non-connectives similarity.

Name of classifier	Name of features included in classification task							Number of features included in the model	Accuracy	Sensitivity	Specificity	AUC-ROC
	Sliding-window similarity	Non-connectives similarity	General connectives-related similarity (random sub-sampling)	General connectives-related similarity (all chunks)	Per-type-connectives-related similarity with smallest adj. *p*-values	Proportions per type of connective	Proportions and per-type-connectives-related similarity significantly different between groups					
RF-c	✓							18	83.5%	79.4%	87.6%	0.89
RF-non-conn	✓	✓						24	84.9%	79.4%	90.4%	0.89
RF-conn-I	✓		✓					24	**85%**	**81.2%**	88.8%	0.88
RF-conn-II	✓			✓				24	84.7%	79.2%	90.2%	0.89
RF-conn-III	✓				✓			24	84.9%	80.4%	89.4%	0.90
RF-conn-IV	✓					✓		23	83.8%	80%	87.6%	0.88
RF-conn-V	✓						✓	25	84.4%	75.2%	**93.6%**	**0.90**

RF = Random Forest algorithm.

Table 9. Results of using connectives-related cosine similarities either alone or along with proportion of connectives in distinguishing patients from controls.

Algorithm	Name of classifier	Name of features included in classification task		Number of features included in the model	Accuracy	Sensitivity	Specificity	AUC-ROC
		Connectives-related similarities per type of connective	Proportions per type of connective					
RF	RF-I	✓		30	79.4%	75%	83.8%	0.84
	RF-II	✓	✓	35	80.1%	77.4%	82.8%	0.89
SVM-pk	SVM-I	✓		30	78.4%	77.6%	79.2%	0.85
	SVM-II	✓	✓	35	**85%**	**83.8%**	**86.2%**	**0.91**

RF = Random Forest; *SVM-pk* = Support Vector Machine with polynomial kernel.

4. Discussion

In this study, we analyzed linguistic coherence by comparing the relative use of different types of connectives and connectives-related similarity between patients with SSD and control participants. In parallel, we assessed how much connectives' features might improve a control-classifier, followed by an evaluation of the usefulness of connectives' features to achieve accurate results in automatically distinguishing patients with SSD from control participants.

Patients with SSD used significantly less contingency and multiclass connectives, while their use of the other types of connectives was not different from that of control participants. Although years of education differed between groups, it did not seem to affect these results. Regarding connectives-related similarity, patients with SSD had higher minimum similarity in both multiclass and temporality connectives, narrower range in both expansion and multiclass connectives, and lower maximum similarity in expansion connectives.

In the classification tasks comparing connectives' features and non-connectives similarity, both types of measures yielded similar overall performance in distinguishing patients with SSD from control participants. In the second series of classification tasks, combining connectives-related similarity per connective type with proportions of use per type of connective, the best classifier yielded 85% accuracy.

4.1. Proportion of connectives

We found significant differences between groups in their use of contingency connectives (subsuming cause, condition, and purpose connectives). This is partially in line with previous studies showing that connectives of cause are used relatively less by patients with SSD (Willits et al., 2018), but opposite results have also been found (Just et al., 2020). Aligning with previous research, we found no significant differences between patients with SSD and control participants in their use of expansion (also referred as additive connectives) and temporality connectives (Willits et al., 2018). Our results also showed no significant differences between groups in their use of comparison connectives (including concession, contrast, and similarity connectives). This is partially inconsistent with previous studies reporting that patients with SSD use less contrastive and differentiation connectives than control participants (Just et al., 2020; Willits et al., 2018).

No previous studies have paid specific attention to polysemic connectives, as the ones included in our multiclass category. In our study, the proportion of use of multiclass connectives was significantly smaller for patients with SSD than for control participants. Polysemic words are processed faster than unambiguous words (Eddington and Tokowicz, 2015; Klepousniotou and Baum, 2007). It is possible that this fast cognitive processing of polysemic connectives was different between patients with SSD and control participants in our sample, reflecting itself in the smaller proportion of polysemic connectives used by the patients. Yet, some variables that were not controlled for might have influenced our results. For instance, semantic activation of polysemic words is influenced by word frequency and context (Rice et al., 2019), and pragmatic inferences play a role in this as well (Carston, 2021). In parallel, it has been argued that connections between the word-forms and meaning of words in the mental lexicon are weaker in patients with SSD than in control participants (Kuperberg et al., 2019). Whether any of these factors relate to our results of the use of these polysemic connectives remains an open question.

Overall, our results suggest that speech of patients with SSD is characterized by a relative reduction in the use of contingency connectives (i.e., markers of cause, condition, and purpose) and multiclass connectives (i.e., markers that can establish more than one explicit type of semantic relation between clauses and/or sentences).

4.2. Connectives-related similarity

Previous research has shown that, compared to control participants, patients with SSD reach lower scores in semantic similarity (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2019). There is consistency between those findings and our results showing that patients with SSD had lower scores in three out of the five connectives-related similarity measures that significantly differed between groups.

In detail, our results showed that patients with SSD had narrower range in similarity of expansion connectives. Expansion connectives establish additive relations with either positive (e.g., books *and* notebooks) or negative polarity (e.g., books *or* notebooks) (Evers-Vermeul and Sanders, 2009; Sanders et al., 1992). The narrower range in similarity of expansion connectives might indicate that there is less semantic variation in the words, clauses and sentences added together by patients with SSD. Patients also had a lower maximum similarity of expansion connectives. This might mean that, compared to control participants, patients with SSD added together words, clauses and/or sentences that shared less semantic features. Of notice, maximum similarity of expansion connectives positively correlated to years of education. The results of maximum similarity of expansion connectives could therefore be a reflection of education level, rather than a difference between patients with SSD and control participants.

Intriguingly, in our study, patients with SSD showed higher minimum similarity of temporality connectives. With few exceptions (Panicheva and Litvinova, 2019), this is opposite to the majority of previous reports showing that patients with SSD often have lower semantic similarity scores than control participants (Elvevåg et al., 2007; Iter et al., 2018; Just et al., 2019). Temporality connectives establish ordered relations between series of events (Evers-Vermeul and Sanders, 2009; Sanders et al., 1992). Patients with SSD use temporality connectives as core linguistic devices to achieve coherence in narrative discourse (Saavedra, 2010). Interestingly, cognitively well-functioning patients with SSD can achieve temporal coherence similar to that of control participants (Holm et al., 2016). In our sample of patients with SSD, their total PANSS score (see Table 2) indicates that, cognitively, they were well-functioning (Leucht et al., 2005). Thus, our results show that patients with SSD can use temporality connectives as coherently as control participants during semi-structured interviews, suggesting that their use of temporality connectives might be related to cognitive well-functioning.

As well, patients with SSD had narrower range and higher minimum similarity of multiclass connectives. The narrower range might mean that patients had less semantic variation in the words conforming the previous and following context of multiclass connectives. Accordingly, the higher minimum similarity of multiclass connectives would reflect the low-end of such narrower range of semantic variation.

4.3. Classification tasks

4.3.1. Connectives' vs non-connectives' features

Among all classifiers, connectives' features and non-connectives cosine similarity yielded similar accuracies in distinguishing patients with SSD from control participants. When controlling for amount of data and number of predictors, general connectives-related similarity (RF-conn-I) seemed to increase sensitivity to classify patients with SSD. However, it is likely that this was due to the random sub-sampling procedure, because general connectives-related similarity (RF-conn-II) no longer increased sensitivity when the similarity was calculated based on all connectives-chunks. In contrast, the classifier using the connectives' features that were significantly different between groups (RF-conn-V) yielded 6% more specificity than the control classifier (RF-c), suggesting that connectives' features that were significantly different between groups are useful to correctly classify true negatives. Overall, our results suggest that connectives' features and non-connectives similarity can reach similar results in distinguishing patients with SSD from control participants.

4.3.2. Performance of connectives' features alone

SVM-II yielded 85% accuracy, 83.8% sensitivity and 86.2% specificity. These percentages are within the accuracy range reported in previous studies (for reviews, see Corcoran et al., 2020; Corcoran and Cecchi, 2020; Hitczenko et al., 2020). Accuracy of 85% had been previously obtained by our group using a full sliding-window general-semantic-similarity classifier (Voppel et al., 2021). SVM-II included connectives' features alone, and it was trained based on less speech input (i.e., only connectives and their surrounding words). This suggests that connectives and their surrounding words are linguistic loci that might concentrate important patterns to detect atypical coherence related to SSD.

4.4. Limitations and future directions

We acknowledge our study has limitations. Our proportion-of-connectives' results could not be straightforwardly compared to previous findings due to differences in the control group (healthy participants in our study and HIV+ in Willits et al., 2018) and in the analysis technique that was employed (mixed-effects logistic regression models in our study and principal component analysis in Mackinley et al., 2021). Also, the number of categories of connectives varied across studies (ranging from two to five), as well as the number of connectives per category and the number of types of connectives inside each main category. Similarly, there were inconsistencies in annotation schemes for connectives. For instance, in contrast to previous studies (Just et al., 2020; Mackinley et al., 2021; Willits et al., 2018), the annotation scheme that we used (Bourgonje et al., 2018; Stede et al., 2019) did not include a separate category for logical connectives, following the line of reasoning that there are no *logical connectives*, but rather abstract *logical operators* that then can have linguistic correlates in different types of connectives (Sanders et al., 1992). Furthermore, fillers and repetitions were not removed from the transcripts used for our analyses, even though they are known to influence cosine similarity calculations (Iter et al., 2018).

In replicating or expanding our results, future studies on the use of connectives in speech of patients with SSD should take into account a series of methodological challenges. The first is a consistent use of connectives categories across studies, which would aid knowledge accumulation. Secondly, our theory-based decision of using a seven-words window alone for our analyses decreased the Type I error rate of our findings. However, it remains to be determined whether this is the most appropriate window for the assessment of cosine similarity between connectives and their surrounding words. This was beyond the scope of the current study, but future research may specifically address the role of window size by directly comparing a range of different window sizes. In relation to this, it would be necessary to analyze what procedure to calculate the connectives-related similarity is the most reliable, valid and theoretically sounded. More recent computational semantic representations could be used to obtain the word embeddings for analyses, such as Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2019), possibly in combination with time-series analyses of semantic coherence (Xu et al., 2022). Also, mixed designs (e.g., Holm et al., 2016; Saavedra, 2010) might provide valuable details overlooked by quantitative approaches alone, increasing our comprehension of how thematic continuity and syntactic connectivity (in)dependently build up (in)coherence in patients with SSD.

Additionally, we need to understand how the use of connectives might be influenced by speech elicitation techniques and cognitive factors. For instance, speech elicitation techniques (e.g., re-telling a story or reading a text out loud) have been shown to influence linguistic outcomes (Kapantzoglou et al., 2017; Niebuhr and Michaud, 2015), and some types of connectives are more cognitively demanding to use than others (Evers-Vermeul and Sanders, 2009; Zufferey and Gygax, 2020). For these reasons, in future studies it would be informative to assess whether the use of connectives differs among speech elicitation techniques, and whether this might also depend on the varying cognitive demands of different connectives. This emphasizes the importance of exploring possible relations between patterns of connectives' use and cognitive outcomes in patients with SSD.

Also, future research should examine the generalizability of our findings on the use of connectives to other languages and across generations of speakers. There is evidence that patterns of word use are consistent across different linguistic families (Calude and Pagel, 2011). However, word order related to grammatical connectivity varies across languages (Lehmann, 2011), and word order changes throughout the history of languages (Gell-Mann and Ruhlen, 2011; Maurits and Griffiths, 2014). For instance, these days Subject-Object-Verb (SOV) is the canonical word order of a subordinated sentence introduced by a connective in Dutch (Jordens, 1988; Koster, 1975), while English has a fixed SVO structure (Comrie, 1981) and Spanish can have either of these (López Meirama, 2006). These syntactic structures might be different for future generations of Dutch, English or Spanish speakers. Thus, both cross-linguistic and historical-grammar factors await an exploration to further our understanding of the use of connectives as signifying patterns of speech (dis)organization in patients with SSD.

5. Conclusions

Connectives' features are informative and explainable variables that can be used to reliably assess disorganized speech in patients with SSD. The combination of this method with other linguistic components is a promising venue to further improve accuracy in categorizing individuals with SDD and control participants. Such fine-tuned automatic analyses of speech samples will help to reach the ultimate aim of advancing clinical practice.

Role of the funding source

The founding source(s) had no role neither in conducting this research nor in deciding to publish this manuscript.

CRediT author statement

H. Corona-Hernández: Conceptualization, Data curation, Formal analysis, Methodology, Writing-Original draft. **J.N. de Boer**: Conceptualization, Methodology, Writing-Review & Editing. **S.G. Brederoo**: Conceptualization, Methodology, Writing-Review & Editing. **A.E. Voppel**: Formal analysis, Writing-Review & Editing. **I.E.C. Sommer**: Supervision, Writing-Review & Editing.

Conflict of interest

All authors declare that they have no conflicts of interest.

Acknowledgement

HCH was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604).

Supplementary materials

List of connectives in Dutch

The original full list has 208 connectives (http://connective-lex.info/#{%22s%22:[%22discodict%22]}). After excluding one duplicated connective (i.e., “teneinde”), the list was reduced to 207, as reported in Bourgonje et al. (2018). Other connectives were also excluded: four complex connectives that allow interposition of lexical elements (e.g., “in de {n} plaats”), and fifteen connectives whose broad category was not clearly labeled in the web source. This resulted in this list of 188 connectives, which were used for our analyses.

Type of connective				
Comparison	**Contingency**	**Expansion**	**Multiclass**	**Temporality**
1. Aan de ene kant ... aan de andere kant	38. Aangezien	86. Afgezien van	130. Aanvankelijk	153. Aansluitend
2. Al	39. Als gevolg hiervan	87. Alleen	131. Als	154. Achteraf
3. Alhoewel	40. Als gevolg van	88. Alsmede	132. Dan	155. Allereerst
4. Alsof	41. Bijgevolg	89. Alsook	133. Dan ook	156. Alvorens
5. Daarentegen	42. Daar	90. Althans	134. Hierdoor	157. Daarna
6. Daartegenover staat dat	43. Daardoor	91. Anders gezegd	135. Hiervoor	158. Daarop
7. Desalniettemin	44. Daarmee	92. Behalve	136. Laat staan	159. Daarvoor
8. Desondanks	45. Daarom	93. Behalve dat	137. Laat staan dat	160. Eer
9. Doch	46. Daartoe	94. Bijvoorbeeld	138. Los daarvan	161. Eerder
10. Echter	47. Dankzij	95. Bovendien	139. Naarmate	162. Ervoor
11. Enerzijds ... anderzijds	48. Derhalve	96. Buiten	140. Namelijk	163. Hierop
12. Evenals	49. Dienovereenkomstig	97. Buiten dat	141. Om	164. Later
13. Eveneens	50. Dientengevolge	98. Daarbij	142. Om te beginnen	165. Naargelang
14. Evenmin	51. Door	99. Daarbovenop	143. Ook	166. Nadat
15. Evenwel	52. Doordat	100. Daarbuiten	144. Ook al	167. Naderhand
16. Evenzeer	53. Dus	101. Daarenboven	145. Tegelijk	168. Nadien
17. Evenzo	54. Gelet op	102. Daarnaast	146. Tegelijkertijd	169. Ondertussen
18. Hierentegen	55. Gezien	103. Dat wil zeggen	147. Tenslotte	170. Sedert
19. Hoewel	56. Hiertoe	104. Eerst	148. Terwijl	171. Sedertdien
20. Hoezeer	57. Immers	105. En	149. Tot slot	172. Sinds
21. In tegenstelling tot	58. In het kader van	106. Hetzij	150. Vervolgens	173. Sindsdien
22. Integendeel	59. In verband met	107. In plaats daarvan	151. Waarmee	174. Straks
23. Maar	60. Krachtens	108. In plaats van	152. Zoals	175. Tijdens
24. Net zomin	61. Met als gevolg dat	109. Kort gezegd		176. Toen
25. Niettegenstaande	62. Naar aanleiding daarvan	110. Kortom		177. Totdat
26. Niettemin	63. Naar aanleiding van	111. Met andere woorden		178. Vanaf
27. Nochtans	64. Omdat	112. Naast		179. Vanaf dat
28. Ofschoon	65. Omwille van	113. Naast dat		180. Voor
29. Ondanks	66. Op basis van	114. Noch		181. Vooraf
30. Ondanks dat	67. Op die manier	115. Of		182. Vooraleer
31. Ongeacht	68. Op grond daarvan	116. Oftewel		183. Voordat
32. Ongeacht of	69. Op grond van	117. Ofwel ... of(wel)		184. Waarna
	70. Opdat	118. Onder andere		185. Waarop
	71. Overeenkomstig	119. Onder meer		186. Wanneer
		120. Samenvattend		187. Zodra

33.	Toch	72.	Per slot van rekening	121.	Sterker nog	188.	Zolang
34.	Wel	73.	Ten behoeve van	122.	Te weten		
35.	Weliswaar … maar	74.	Ten gevolge van	123.	Tevens		
36.	Zonder	75.	Teneinde	124.	Uitgezonderd		
37.	Zonder dat	76.	Vandaar	125.	Verder		
		77.	Vandaar dat	126.	Voorts		
		78.	Vanwege	127.	Zij het		
		79.	Vastgesteld dat	128.	Zo		
		80.	Waardoor	129.	Zowel … als		
		81.	Waarom				
		82.	Want				
		83.	Wegens				
		84.	Zodat				
		85.	Zodoende				

Additional analyses on connectives-related similarity

One patient with SSD had a missing value in variance of contingency connectives, and a different patient had a missing value in variance of temporality connectives. Multivariate analysis of variance with the exclusion of these two participants (i.e., based on 48 patients with SSD and 50 control participants) again showed significant differences between groups in connectives-related similarity, F (10.4, 1002.5) = 4.99, $p < .001$. Results from these additional post-hoc analyses (see Table S1 below) largely corresponded with the results of the analyses on the sample at large (see Table 7). Two effects that were significant in the sample at large now failed to reach significance, namely *minimum* and *range* of multiclass connectives (both adj-p = .04 in the larger sample, and adj-p = .10 in this subsample). However, given the fact that the means and SDs did not change (*minimum*) or only marginally changed (*range*), this is likely the result of diminished power in the subsample. Importantly, the results of the analyses on the variables for which the excluded participants had missing values did not change (i.e., variance of contingency and temporality connectives). Independent non-parametric two-tailed Spearman bivariate correlations showed no significant correlations between the three significant variables and years of education (all *rho*(ρ) < |0.24|, all adj-p > .05).

Table S1. Differences in connectives-related similarity between groups, excluding the two patients whose cosine similarities had missing values.

Type of connective	Coherence measure	Patients (n=48), average (SD)	Controls (n=50), average (SD)	Statistic		*r* effect size
				W	Adj. *p*-value	
Comparison	Mean	0.51 (0.01)	0.50 (0.01)	808	.12	
	Median	0.51 (0.01)	0.51 (0.01)	846	.25	
	Maximum	0.63 (0.04)	0.63 (0.04)	1250	≈.99	
	Minimum	0.38 (0.04)	0.37 (0.02)	952	≈.99	
	Range	0.24 (0.05)	0.25 (0.05)	1344	≈.99	
	Variance	0.003 (0.001)	0.003 (0.001)	934	.88	
Contingency	Mean	0.48 (0.02)	0.47 (0.01)	777	.06	
	Median	0.49 (0.02)	0.48 (0.01)	764	.05	
	Maximum	0.59 (0.05)	0.59 (0.03)	1230	≈.99	
	Minimum	0.36 (0.06)	0.33 (0.03)	821	.15	
	Range	0.23 (0.09)	0.26 (0.06)	1466	.88	
	Variance	0.005 (0.003)	0.004 (0.001)	1030	≈.99	
Expansion	Mean	0.46 (0.01)	0.46 (0.01)	1178	≈.99	
	Median	0.45 (0.01)	0.45 (0.01)	1184	≈.99	
	Maximum	0.61 (0.05)	0.66 (0.06)	1755	.002	.39
	Minimum	0.33 (0.03)	0.31 (0.03)	927	.84	
	Range	0.28 (0.07)	0.34 (0.06)	1760	.002	.40
	Variance	0.005 (0.002)	0.005 (0.002)	1293	≈.99	
Multiclass	Mean	0.50 (0.02)	0.50 (0.01)	1179	≈.99	
	Median	0.49 (0.02)	0.49 (0.01)	1089	≈.99	
	Maximum	0.62 (0.06)	0.64 (0.03)	1492	.66	
	Minimum	0.38 (0.04)	0.35 (0.04)	796	.10	
	Range	0.24 (0.08)	0.28 (0.05)	1604	.10	
	Variance	0.004 (0.002)	0.004 (0.001)	1139	≈.99	
Temporality	Mean	0.43 (0.02)	0.42 (0.01)	874	.39	
	Median	0.43 (0.02)	0.42 (0.01)	906	.66	
	Maximum	0.52 (0.06)	0.52 (0.05)	1271	≈.99	
	Minimum	0.34 (0.03)	0.31 (0.03)	597	.0005	-.43
	Range	0.17 (0.06)	0.20 (0.06)	1543	.29	
	Variance	0.003 (0.002)	0.003 (0.001)	1253	≈.99	

n = sample size, SD = standard deviation.

References

American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed. American Psychiatric Association, Washington, D.C.

Andreasen, N.C., 1979. Thought, Language, and Communication Disorders: I. Clinical Assessment, Definition of Terms, and Evaluation of Their Reliability. Arch. Gen. Psychiatry 36, 1315–21.

Andreasen, N.C., Flaum, M., Arndt, S., 1992. The Comprehensive Assessment of Symptoms and History (CASH): An Instrument for Assessing Diagnosis and Psychopathology. Arch. Gen. Psychiatry 49, 615–623.

Baayen, R.H., 2008. Analyzing Linguistic Data: A Practical Introduction to Statistics using R. Cambridge University Press, New York, NY.

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67, 1–48.

Bedi, G., Carrillo, F., Cecchi, G.A., Slezak, D.F., Sigman, M., Mota, N.B., Ribeiro, S., Javitt, D.C., Copelli, M., Corcoran, C.M., 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. Npj Schizophr. 1, 1–7.

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., Matsuo, A., 2018. quanteda: An R package for the quantitative analysis of textual data. J. Open Source Softw. 3, 774.

Bourgonje, P., Hoek, J., Evers-Vermeul, J., Redeker, G., Sanders, T., Stede, M., 2018. Constructing a Lexicon of Dutch Discourse Connectives. Comput. Linguist. Neth. J. 8, 163–175.

Calude, A.S., Pagel, M., 2011. How do we use language? Shared patterns in the frequency of word use across 17 world languages. Philos. Trans. Biol. Sci. 366, 1101–1107.

Carston, R., 2021. Polysemy: Pragmatics and sense conventions. Mind Lang. 36, 108–133.

Comrie, B., 1981. Language universals and linguistic typology: syntax and morphology, 2nd ed. The University of Chicago Press, Chicago, Ill.

Corcoran, C.M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D.C., Bearden, C.E., Cecchi, G.A., 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry 17, 67–75.

Corcoran, C.M., Cecchi, G.A., 2020. Using Language Processing and Speech Analysis for the Identification of Psychosis and Other Disorders. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 5, 770–779.

Corcoran, C.M., Mittal, V.A., Bearden, C.E., E. Gur, R., Hitczenko, K., Bilgrami, Z., Savic, A., Cecchi, G.A., Wolff, P., 2020. Language as a biomarker for psychosis: A natural language processing approach. Schizophr. Res. 226, 158–166.

de Boer, J.N., Voppel, A.E., Begemann, M.J.H., Schnack, H.G., Wijnen, F., Sommer, I.E.C., 2018. Clinical use of semantic space models in psychiatry and neurology: A systematic review and meta-analysis. Neurosci. Biobehav. Rev. 93, 85–92.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Presented at the NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf, pp. 4171–4186.

Eddington, C.M., Tokowicz, N., 2015. How meaning similarity influences ambiguous word processing: the current state of the literature. Psychon. Bull. Rev. 22, 13–37.

Elvevåg, B., Foltz, P.W., Weinberger, D.R., Goldberg, T.E., 2007. Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. Schizophr. Res. 93, 304–316.

Evers-Vermeul, J., Sanders, T., 2009. The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition. J. Child Lang. 36, 829–854.

Fernández-Delgado, M., Cernadas, E., Barro, S., Amorim, D., 2014. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? J. Mach. Learn. Res. 15, 3133–3181.

Gell-Mann, M., Ruhlen, M., 2011. The origin and evolution of word order. Proc. Natl. Acad. Sci. 108, 17290–17295.

Givón, T., 2020. Coherence. John Benjamins Publishing Company, Amsterdam, The Netherlands.

Hitczenko, K., Mittal, V.A., Goldrick, M., 2020. Understanding Language Abnormalities and Associated Clinical Markers in Psychosis: The Promise of Computational Methods. Schizophr. Bull. sbaa141.

Holm, T., Thomsen, D.K., Bliksted, V., 2016. Life story chapters and narrative self-continuity in patients with schizophrenia. Conscious. Cogn. 45, 60–74.

Hong, K., Nenkova, A., March, M.E., Parker, A.P., Verma, R., Kohler, C.G., 2015. Lexical use in emotional autobiographical narratives of persons with schizophrenia and healthy controls. Psychiatry Res. 225, 40–49.

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

Jordens, P., 1988. The acquisition of word order in Dutch and German as L1 and L2. Second Lang. Res. 4, 41–65.

Just, S., Haegert, E., Kořánová, N., Bröcker, A.-L., Nenchev, I., Funcke, J., Montag, C., Stede, M., 2019. Coherence models in schizophrenia, in: Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology. Presented at the Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology, Association for Computational Linguistics, Minneapolis, Minnesota, pp. 126–136.

Just, S.A., Haegert, E., Kořánová, N., Bröcker, A.-L., Nenchev, I., Funcke, J., Heinz, A., Bermpohl, F., Stede, M., Montag, C., 2020. Modeling Incoherent Discourse in Non-Affective Psychosis. Front. Psychiatry 11. https://doi.org/10.3389/fpsyt.2020.00846

Kapantzoglou, M., Fergadiotis, G., Restrepo, M.A., 2017. Language Sample Analysis and Elicitation Technique Effects in Bilingual Children With and Without Language Impairment. J. Speech Lang. Hear. Res. 60, 2852–2864.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276.

Klepousniotou, E., Baum, S.R., 2007. Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. J. Neurolinguistics 20, 1–24.

Koster, J., 1975. Dutch as an SOV Language. Linguist. Anal. 1, 111–136.

Kuhn, M., 2021. caret: Classification and Regression Training.

Kuperberg, G.R., Weber, K., Delaney-Busch, N., Ustine, C., Stillerman, B., Hämäläinen, M., Lau, E., 2019. Multimodal neuroimaging evidence for looser lexico-semantic networks in schizophrenia: Evidence from masked indirect semantic priming. Neuropsychologia 124, 337–349.

Landauer, T.K., Foltz, P.W., Laham, D., 1998. An introduction to latent semantic analysis. Discourse Process. 25, 259–284.

Lehmann, C., 2011. Towards a typology of clause linkage, in: Haiman, J., Thompson, S.A. (Eds.), Clause Combining in Grammar and Discourse. John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 181–226.

Leucht, S., Kane, J.M., Kissling, W., Hamann, J., Etschel, E., Engel, R.R., 2005. What does the PANSS mean? Schizophr. Res. 79, 231–238.

López Meirama, B., 2006. Semantic and Discourse-Pragmatic Factors in Spanish Word Order, in: Clements, J.C., Yoon, J. (Eds.), Functional Approaches to Spanish Syntax: Lexical Semantics, Discourse and Transitivity. Palgrave Macmillan, London, pp. 7–52.
Maat, H.P., Sanders, T., 2006. Connectives in Text, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 33–41.
Mackinley, M., Chan, J., Ke, H., Dempster, K., Palaniyappan, L., 2021. Linguistic determinants of formal thought disorder in first episode psychosis. Early Interv. Psychiatry 15, 344–351.
MacWhinney, B., 2000. The CHILDES Project: Tools for Analyzing Talk, 3rd Edition. ed. Lawrence Erlbaum Associates, Mahwah, NJ.
Maher, B., 1972. The Language oi Schizophrenia: A Review and Interpretation. Br. J. Psychiatry 120, 3–17.
Maurits, L., Griffiths, T.L., 2014. Tracing the roots of syntax with Bayesian phylogenetics. Proc. Natl. Acad. Sci. 111, 13576–13581.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013a. Efficient Estimation of Word Representations in Vector Space, in: ICLR Workshop.
Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013b. Distributed Representations of Words and Phrases and their Compositionality, in: Proc. of the 26th International Conference on Neural Information Processing Systems. pp. 3111–3119.
Niebuhr, O., Michaud, A., 2015. Speech data acquisition: the underestimated challenge. Kiel. Arb. Linguist. Phon. KALIPHO 3, 1–42.
Panicheva, P., Litvinova, T., 2019. Semantic Coherence in Schizophrenia in Russian Written Texts. Presented at the 25th Conference of Open Innovations Association (FRUCT), pp. 241–249. https://doi.org/10.23919/FRUCT48121.2019.8981535
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rezaii, N., Walker, E., Wolff, P., 2019. A machine learning approach to predicting psychosis using semantic density and latent content analysis. Npj Schizophr. 5, 9. https://doi.org/10.1038/s41537-019-0077-9
Rice, C.A., Tokowicz, N., Fraundorf, S.H., Liburd, T.L., 2019. The complex interactions of context availability, polysemy, word frequency, and orthographic variables during lexical processing. Mem. Cognit. 47, 1297–1313.
RStudio Team, 2019. RStudio: Integrated Development Environment for R. RStudio, Inc., Boston, MA.
Saavedra, J., 2010. Quantitative Criteria of Narrative Coherence and Complexity in Persons With Paranoid Schizophrenia. J. Nerv. Ment. Dis. 198, 349–355.
Sanders, T., Maat, H.P., 2006. Cohesion and Coherence: Linguistic Approaches, in: Brown, K., Anderson, A.H., Bauer, L., Berns, M., Hirst, G., Miller, J. (Eds.), The Encyclopedia of Language & Linguistics. Elsevier Publishing, pp. 591–595.
Sanders, T.J.M., Spooren, W.P.M., Noordman, L.G.M., 1992. Toward a taxonomy of coherence relations. Discourse Process. 15, 1–35.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59, quiz 34-57.
Stede, M., Scheffler, T., Mendes, A., 2019. Connective-Lex: A Web-Based Multilingual Lexical Resource for Connectives. Discours 24.
Tang, S.X., Kriz, R., Cho, S., Park, S.J., Harowitz, J., Gur, R.E., Bhati, M.T., Wolf, D.H., Sedoc, J., Liberman, M.Y., 2021. Natural language processing methods are sensitive to

sub-clinical linguistic differences in schizophrenia spectrum disorders. Npj Schizophr. 7.

van der Vliet, N., Redeker, G., 2014. Explicit and implicit coherence relations in Dutch texts, in: Gruber, H., Redeker, G. (Eds.), The Pragmatics of Discourse Coherence: Theories and Applications. John Benjamins Publishing Company, Amsterdam, The Netherlands, pp. 23–52.

van Eerten, L., 2007. Over het Corpus Gesproken Nederlands. Ned. Taalkunde 12, 194–2015.

Voppel, A.E., de Boer, J.N., Brederoo, S.G., Schnack, H.G., Sommer, I.E.C., 2021. Quantified language connectedness in schizophrenia-spectrum disorders. Psychiatry Res. 304, 114130. https://doi.org/10.1016/j.psychres.2021.114130

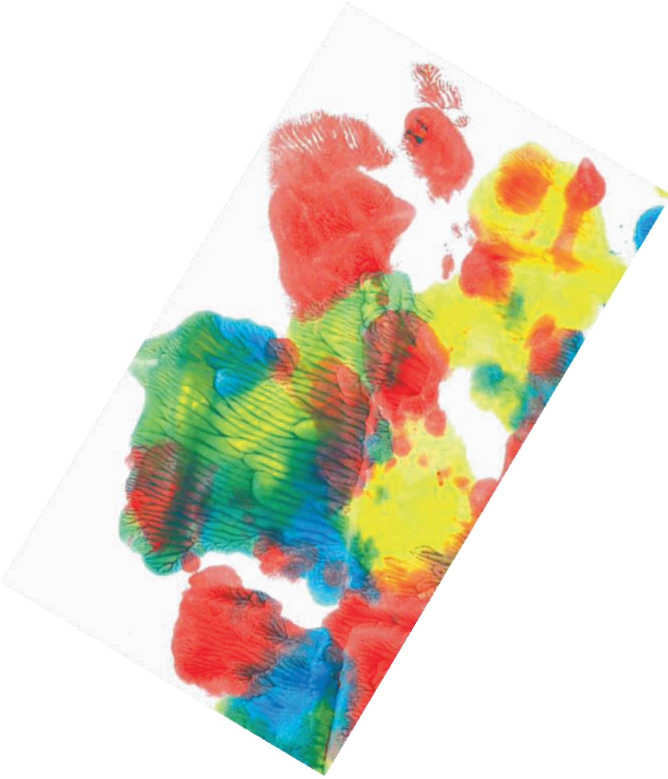
Wijffels, J., 2020. word2vec: Distributed Representations of Words.

Willits, J.A., Rubin, T., Jones, M.N., Minor, K.S., Lysaker, P.H., 2018. Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. Schizophr. Res. 197, 365–369.

Winter, B., 2020. Statistics for Linguists: An Introduction Using R. Routledge, New York, NY.

Xu, W., Wang, W., Portanova, J., Chander, A., Campbell, A., Pakhomov, S., Ben-Zeev, D., Cohen, T., 2022. Fully automated detection of formal thought disorder with Time-series Augmented Representations for Detection of Incoherent Speech (TARDIS). J. Biomed. Inform. 126, 103998.

Zufferey, S., Gygax, P., 2020. Do teenagers know how to use connectives from the written mode? Lingua 234, 102779.



CHAPTER 5

Natural language processing markers for psychosis and other psychiatric disorders: emerging themes and research agenda from a cross-linguistic workshop

Hugo Corona Hernández[1], Cheryl Corcoran[2], Amélie M Achim[3], Janna N de Boer[1,4], Tessel Boerma[5], Sanne G Brederoo[1,6], Guillermo A Cecchi[7], Silvia Ciampelli[1], Brita Elvevåg[8], Riccardo Fusaroli[9-11], Silvia Giordano[12], Mathias Hauglid[13], Arjan van Hessen[5,14], Wolfram Hinzen[15], Philipp Homan[16], Sybren F de Kloet[17], Gina R Kuperberg[18-20], Kritika Maheshwari[21,22], Natalia B Mota[23,24], Alberto Parola[9,10], Roberta Rocca[10], Sanne Schuite-Koops[1], Iris E C Sommer[1,6], Khiet Truong[14], Alban E Voppel[1], Marieke van Vugt[25], Frank Wijnen[5], and Lena Palaniyappan[26,27]

[1]University of Groningen, University Medical Center Groningen, Department of Biomedical Sciences of Cells & Systems, Groningen, the Netherlands; [2]Icahn School of Medicine at Mount Sinai, Department of Psychiatry, New York, United States of America; [3]Université Laval, Département de Psychiatrie et Neurosciences, VITAM Centre de Recherche en Santé Durable, Cervo Brain Research Centre, Québec, Canada; [4]Utrecht University, University Medical Center Utrecht, Department of Psychiatry, Utrecht, the Netherlands; [5]Utrecht University, Department of Languages, Literature and Communication, Institute for Language Sciences, Utrecht, the Netherlands; [6]University Medical Center Groningen, University Center of Psychiatry, Groningen, the Netherlands; [7]IBM T.J. Watson Research Center, Yorktown Heights, New York, United States of America; [8]University of Tromsø - the Arctic University of Norway, Department of Clinical Medicine, Tromsø, Norway; [9]Aarhus University, Department of Linguistics, Cognitive Science and Semiotics, School of Communication and Culture, Aarhus, Denmark; [10]Aarhus University, Department of Culture, Interacting Minds Center, Cognition and Computation Communication, School of Culture and Society, Aarhus, Denmark; [11]University of Pennsylvania, Linguistic Data Consortium, Pennsylvania, United States of America; [12]University of Applied Sciences and Arts of Southern Switzerland, Department of Innovative Technologies, Lugano, Switzerland; [13]University of Tromsø - the Arctic University of Norway, Faculty of Law, Tromsø, Norway; [14]University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science, Twente, the Netherlands; [15]Institució Catalana de Recerca i Estudis Avançats and Universitat Pompeu Fabra, Barcelona, Spain; [16]Psychiatric Hospital of the University of Zurich, Department of Psychiatry, Psychotherapy, and Psychosomatics, Zurich, Switzerland; [17]Neurocast BV, Amsterdam, the Netherlands; [18]Tufts University, Department of Psychology, Medford, Massachusetts, United States of America; [19]Harvard Medical School, Department of Psychiatry, Boston, Massachusetts, United States of America; [20]Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, United States of America; [21]University Medical Centre Groningen, Department of Genetics, Groningen, the Netherlands; [22]Delft University of Technology, Ethics & Philosophy of Technology Section, Faculty of Technology, Policy, and Management, Delft, the Netherlands; [23]Federal University of Rio de Janeiro, Institute of Psychiatry, Rio de Janeiro, Brazil; [24]Research department at Motrix Lab – Motrix, Rio de Janeiro, Brazil; [25]University of Groningen, Bernoulli Institute of Mathematics, Computer Science & Artificial Intelligence, Groningen, the Netherlands; [26]McGill University, Douglas Mental Health University Institute, Department of Psychiatry, Montreal, Quebec, Canada; [27]Western University, Robarts Research Institute, Schulich School of Medicine and Dentistry, London, Ontario, Canada.

Schizophrenia bulletin. 2023; 49(2): S86–S92.

Introduction

A multidisciplinary workshop entitled "Crosslinguistic speech patterns: biosocial markers of psychiatric disorders" was held with the support of a Distinguished Lorentz Fellowship granted to Iris Sommer, in conjunction with the DISCOURSE in Psychosis Consortium (October 31st-November 4th 2022, Leiden University, the Netherlands). We (the attendees) included clinical practitioners and experts in diverse scientific disciplines, such as artificial intelligence (AI), clinical psychology, cognitive neurosciences, computational sciences, ethics, law, linguistics, psychiatry, and technology industry. A main aim of the workshop was to deliberate on potential challenges with respect to the discovery, characterization, validation, and potential utilization of natural language processing (NLP) markers for psychosis and other psychiatric disorders using computational technologies, with the ultimate goal of implementing them ethically in clinical settings. Related to this, we discussed who the main stakeholders key to this enterprise are, including individuals with lived experience, their families, the clinicians who serve them, research scientists with diverse areas of expertise, and ethicists. Ethical issues were discussed in detail, emphasizing their relationship to regulatory concerns that may differ by country and by stakeholder status.

NLP markers for psychiatric disorders

Definition and potential roles

Aligning with a broad characterization of markers in digital medicine (Vasudevan et al., 2022), we agreed that an NLP marker is a digitally acquired, computationally derived, quantifiable measure or set of measures of human language production reflecting the state of biological, neurocognitive, and social processes that contribute to it. While acknowledging the breadth of oral and sign language-related processes (i.e., production and comprehension of spoken/sign/written language), we mostly focused on speech production for a few key reasons. In psychiatric practice, spoken language is considered to be indicative of mental states, which are reflected in its meaning (i.e., semantic content), form (i.e., grammar), and acoustic features. Metrics of spoken language can easily be derived from audio recordings obtained during routine clinical practice in psychiatry, as well as in naturalistic, ecologically valid contexts (e.g., at home). While many developing markers are obtained using NLP techniques (e.g., cosine semantic similarity metrics), markers derived using other computational approaches focused on human communication processes (e.g., acoustics of speech signal and non-verbal behaviors such as facial expression) are also included in the broad definition of NLP markers.

We recognized that NLP markers might have a *descriptive* role useful for screening, stratification in trials, and as a marker of outcome (e.g., prediction of relapse). In parallel, NLP markers might also have a *mechanistic* role, making them indicative of underlying pathological mechanisms at cellular, physiological and/or circuit-based levels, which could lead to target engagement for the development of new therapeutics, and plausibly improve prediction accuracy, stratification and monitoring of treatment response.

NLP markers for clinical actions

A set of potential clinical actions and goals were nominated for the use of NLP markers in psychiatry (see Table 1), based on discussions of examples and existing avenues of research. These comprise mostly *descriptive* NLP markers that as yet are limited in accuracy, carrying the risk of both false positives and false negatives. It was agreed that much work needs to be done before any of these use cases could be implemented in the clinic, and that ethical issues, commensurate with other fields of *neurotechnology* that prioritize people's neurorights (Goering et al., 2021; Yuste et al., 2017), are paramount in developing NLP markers for psychiatric disorders.

The group agreed that the field as yet lacks comprehensive large-scale "candidate-selection" studies for several clinical decisions (e.g., treatment response monitoring and prediction of aggression/violence). We reviewed the promising proof-of-concept studies that support the construct validity of candidate NLP markers that correlate with standard clinical ratings (e.g., associations between cosine similarity metrics and tangentiality (Bilgrami et al., 2022) in individuals at clinical risk for psychosis) and that are predictive of some outcomes of interest, such as transition to psychosis from risk states (Corcoran et al., 2018). Robust external replications (Corcoran et al., 2018), prospective validations, cross-linguistic comparisons, and reliability estimates on assay performance are also needed, and clinical trials on integrating NLP markers with routine practice are yet to begin.

The measurement and evaluation of NLP markers for specific clinical actions (Holmlund et al., 2022) can be guided by a principled approach with three steps (Foltz et al., 2022). First, current clinical knowledge, prior research results, and data-driven approaches should guide the selection of promising features to validate NLP markers for specific clinical actions. Second, optimal procedures for measuring those features should be defined. Third, arguments both in favor and against making changes in current clinical practice related to the employment of NLP markers should be thoroughly examined, addressing issues of validity, reliability, utility, acceptability, and costs.

Understanding the constraints of NLP markers on generalization (e.g., heterogeneity and inherent volunteer bias in training data) is crucial, requiring debiasing strategies during acquisition, training and validation stages and safeguards during implementation. There was general agreement on the need to collect large diverse samples to determine how NLP markers generalize over populations varying in age, sex, ethnicity, and education, for instance. Constraints on implementation of NLP markers must be considered right from the start in developing predictive models for clinical use. Data-sharing obstacles should be tackled (Palaniyappan et al., 2022) so that interested parties can collaborate inter-institutionally (Kairouz et al., 2021) to advance the field.

Table 1. Cases in point and scientific questions relevant to the validation and potential use of NLP markers in psychiatry.

AI task	Variable	Clinical goal	Example of candidate NLP marker	Research questions
Detection	•Diagnosis	•Establish a categorical diagnosis (despite questionable validity).	Emotion-related acoustic features in speech differentiate unipolar depression and bipolar disorder (Huang et al., 2020).	What are the likely pathognomonic NLP markers for the different psychiatric disorders?
	•Symptoms	•Improve detection and quantification of symptoms to more efficiently provide patients with Measurement-Based Care (Lewis et al., 2019).	In CHR youths, pause length and percentage of pauses positively correlated with total severity of negative symptoms (Stanislawski et al., 2021).	With what periodicity should the assessment of symptoms occur (e.g., once or twice per day) and for how long (e.g., one vs three months) to obtain reliable estimates?
	•Warning signs	•Identify CHR individuals timely. •Study pre-symptomatic phases of mental disorders.	Prior to the first psychiatric hospitalization of patients with SSD, a relative increase in the use of swearwords and words related to perceptual processes and negative emotions (Birnbaum et al., 2020).	Are there transdiagnostic and pathognomonic early-warning NLP markers? Do early-warning NLP markers manifest similarly across the lifespan?
Monitoring	•Treatment effects	•Monitor response to treatment actively, including side effects. •Minimize adverse effects from medication and increase adherence to optimal treatment.	In adults with major depression, pause behavior and mean fundamental frequency (pitch) differentiated treatment response (Mundt et al., 2012).	Can NLP markers that vary with a treatment effect provide sufficient information to make decisions regarding changing, augmenting, or discontinuing treatments?
Prediction	•Aggression/violence	•Reduce the number of injuries, amount of harm and damage resulting from aggression or violence. •Reduce the use of coercive measures against aggressive or violent individuals.	In youth referred for psychiatric risk assessment, features such as words related to violence and temporal phrases related to the frequency of violent thoughts or acts were significantly associated with the risk of school violence (Ni et al., 2020).	Might NLP markers be predictive of types of aggression/violence (e.g., verbal vs physical) and who the target is?
	•Psychosis onset	•Stratification of CHR individuals for targeted preventive interventions.	Prior to initial psychosis onset in CHR, decrease in semantic cosine similarity, greater variance in similarity, and less usage of possessive pronouns (Corcoran et al., 2018).	How early should NLP markers be measured in order to predict onset reliably? What predictive value will NLP markers have in non-help seeking samples?
	•Prognosis	•Estimate the course of a patient's psychiatric disorder and/or the probability of recovery.	In first episode psychosis, a drop in syntactic complexity over 6 months indicated a later diagnosis of schizophrenia (Silva et al., 2022).	Which NLP markers best predict outcomes such as social functioning, symptoms' remission, or vocational recovery?
	•Relapse	•Estimate relapse to improve preventive care.	For patients with psychosis, in the month preceding relapse there was a relative increase in the use of words related to swearing, anger, death, and a decrease in words related to work, friends, and health along with more first and second person pronouns (Birnbaum et al., 2019).	What is the best and actionable time frame for gathering relapse-prediction NLP markers data (e.g., every 2-4 weeks)?
	•Suicidality	•Improve assessment of suicidal ideation. •Prevent suicidal acts.	Among USA veterans, a combined set of acoustic and linguistic features improved detection of suicidal ideation (Belouali et al., 2021).	Can NLP markers predict suicidal ideation and behavior with greater accuracy than existing risk calculators? Can NLP markers accurately distinguish between suicidal ideation and non-suicidal, negative thoughts?
Selection	•Optimal treatment	•Select an optimal treatment to increase the probability of recovery.	In individuals with depression, scores of words with emotional content were predictive of treatment success with psilocybin (Carrillo et al., 2018).	Can NLP markers assist in the identification of the optimal treatment for a given patient?

AI: artificial intelligence; **CHR**: clinical high risk; **NLP**: natural language processing; **SSD**: schizophrenia-spectrum disorders; **USA**: United States of America.

NLP markers and mechanistic research

Significant progress has been made in understanding the neural basis of language processing (Hagoort & Beckmann, 2019) and its interaction with neurocognitive processes such as attention (Goller et al., 2020) or memory (Shain et al., 2022). Spoken language conveys information about impairments in thought and cognition in psychiatric disorders (de Boer et al., 2020). Thus, the mechanisms that underlie NLP markers might be in close proximity to the etiology of psychosis and other psychiatric disorders (Uher & Zwicker, 2017). To test this, there is a need for carefully designed hypothesis-driven experiments in clinical samples. By developing causal-mechanistic explanations for promising NLP markers (R. Fusaroli et al., 2022; Parola et al., 2020) (i.e., delineating the neural mechanisms that account for their characteristics), in the near future NLP markers could be used as proxy outcomes reflecting whether clinical interventions exert an effect on the underlying mechanisms of a given disorder.

Attendees highlighted that language production is the result of genetic (Mekki et al., 2022) and developmental (Rudolph & Leonard, 2016) processes. Further, while an individual's anatomical (Dediu et al., 2019) and cognitive (Shafto & Tyler, 2014) characteristics constrain its features, language production is influenced by pharmacological (M. Fusaroli et al., 2022), contextual (Nölle et al., 2020), and socio-demographic (Palaniyappan, 2021) factors. Therefore, we considered that, with respect to mechanistic investigations of candidate NLP markers, we must improve the consistency of how we acquire, preprocess, and analyze speech data, how we parse effect(s) of potential confounders on the characteristics of candidate NLP markers, and how we interpret candidate NLP markers to ensure robust replications. We acknowledged that candidate NLP markers could map onto multilevel biosocial causal frameworks, and group-aggregated results of NLP markers might be used as priors to inform any personalized care (Barron et al., 2021). Rigorous and large-scale clinical studies evaluating predictive models alongside experimental mechanistic studies should allow us to identify explainable candidate NLP markers.

Imagining a clinical decision support system incorporating NLP markers

Discussions of a putative clinical decision support system (CDSS) incorporating NLP markers highlighted that candidate markers must be validated with "ground truth" clinical rating scales, and evidence that they have real-life functional correlates should be provided. We also agreed that NLP markers must be integrated with other sources of clinical information (Barron et al., 2021), and that training related to their acquisition and interpretation should have minimal burden on clinicians. Further, along with accessibility to and acceptance of candidate markers by clinicians and patients (Brederoo et al., 2021), any CDSS incorporating NLP markers should achieve expected standards of transparency, trust, and efficient and safe functioning (Sutton et al., 2020) for regulatory approvals before widespread clinical use (Koutsouleris et al., 2022). In the absence of a formal CDSS, clinical settings can implement NLP markers in pilot testing using human-in-the-loop iterative methodologies (Chandler et al., 2022) to begin to flesh out these issues.

Ethical challenges

We anticipate the implementation of any CDSS incorporating NLP markers to face a series of ethical challenges (many of which have been debated for decades). Spoken language reflects psychological states and is considered to be personal data, raising nuanced concerns about data protection and privacy legislation (United Nations Conference on Trade and Development, 2021). The use of audio and video recordings require us to adhere to a set of ethical principles to “preserve people’s privacy, identity, agency and equality” (Yuste et al., 2017). Likewise, (inter)national AI-laws (Hauglid, 2022) should regulate the process of scaling up any putative CDSS incorporating NLP markers for routine use. Moreover, broader concerns over AI explainability, clinical reasoning, and patients’ autonomy also persist (Keeling & Nyrup, 2021). Specifically, unease about misuse (e.g., discrimination) or potential harms (e.g., missing a relapse event) arising from mistakes in utilizing NLP markers is widespread. In this context, NLP markers must also be first validated and assessed for accuracy, reliability, acceptability, scalability, utility and cost before any consideration can be made for making them an integral part of clinical care. All these ethical issues must be addressed in an explicit and transparent manner. Importantly, previous efforts have suggested that these challenges are surmountable (e.g., the European MONARCA project (Puiatti et al., 2011)), but call for an interdisciplinary action plan.

Conclusions and future directions

Psychiatric practice is deeply rooted in human language and the communicative interchanges it allows. With unprecedented developments in digital health technology and NLP, we are now at the cusp of systematically building on language-related data to derive clinical benefits. Our consortium will work to build an alliance of lived-experience experts, clinicians, and caregivers in further collaborative work. Constructing benchmark transdiagnostic datasets requires sustained global multicenter collaborations. Researchers in the language sciences could inform the development of cross-linguistic NLP markers that incorporate phenomena of linguistic variation, thus increasing generalizability and avoiding the bias of underrepresenting certain languages or communities of speakers. Empirical cognitive neuroscience and psycholinguistic studies investigating the mechanistic basis of NLP markers can enhance their use in experimental medicine and treatment discoveries. The results could inspire novel linguistic remediations and speech and language therapy approach in psychiatry. A partnership of computational and data scientists with end-users (i.e., clinicians and patients) will enable the implementation of informed modelling pipelines fitting the needs of clinical use. Along with stakeholders in the health technology industry, we will work to improve the accessibility to and acceptability of acquisition and analytics procedures. The success of a safe and responsible use of any CDSS incorporating NLP markers requires support and guidance from ethicists, policy and legal experts, and regulatory bodies. With a commitment to act on these points, a diverse, inclusive, interdisciplinary and global collective for mental-health NLP markers can create the conditions to optimize health care with readily accessible and widely acceptable technology.

Funding

This article was enabled by a Distinguished Lorentz Fellowship granted to Iris Sommer in 2022 for Computational Linguistics to aid Diagnosis and Treatment Monitoring in Psychiatry. HCH was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604). LP received research support from the Tanna Schulich Chair of Neuroscience and Mental Health (Schulich School of Medicine, Western University: 2019-2022); Monique H. Bourgeois Chair in Developmental Disorders and Graham Boeckh Foundation (Douglas Research Centre, McGill University) and salary award from the Fonds de Recherche du Québec-Santé (FRQS). Part of the networking efforts of the DISCOURSE in Psychosis consortium is funded by a grant from the Tannenbaum Open Science Initiative at the Neuro, McGill University. AP is supported by a Marie Skłodowska-Curie Actions (MSCA) – Individuals Fellowship H2020-MSCA-IF-2018 (Grant agreement ID: 832518, Project: MOVES). AMA received support from the Fonds de Recherche de Québec-Santé (FRQS). CC was supported by two grants of the National Institute of Mental Health: Using the RDoC Approach to Understand Thought Disorder: A Linguistic Corpus-Based Approach (5R01MH115332) and Thought Disorder and Social Cognition in Clinical Risk States for Schizophrenia (5R01MH107558). KM received funding from the NWO ELSA AI Lab Northern Netherlands.

Authorship contribution

HCH, IECS, and LP wrote the first draft of the manuscript. HCH, CC, BE, IECS, and LP developed the structure and arguments of the manuscript based on the contributions made by all of the online- and onsite-attendees of the workshop. All authors read, critically revised, and approved the final manuscript. AMA, GAC and GRK were not able to attend the workshop.

Disclosures

LP reports personal fees from the Canadian Medical Association Journals for serving as chief editor, speaker/consultant fee from Janssen Canada and Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. RF reports past consultant fees from F. Hoffmann-La Roche. NBM works at Motrix, an EduTech startup, and has been a consultant to Boehringer Ingelheim.

References

Barron, D. S., Baker, J. T., Budde, K. S., Bzdok, D., Eickhoff, S. B., Friston, K. J., Fox, P. T., Geha, P., Heisig, S., Holmes, A., Onnela, J.-P., Powers, A., Silbersweig, D., & Krystal, J. H. (2021). Decision Models and Technology Can Help Psychiatry Develop Biomarkers. *Frontiers in Psychiatry*, *12*, 706655. https://doi.org/10.3389/fpsyt.2021.706655

Belouali, A., Gupta, S., Sourirajan, V., Yu, J., Allen, N., Alaoui, A., Dutton, M. A., & Reinhard, M. J. (2021). Acoustic and language analysis of speech for suicidal ideation among US veterans. *BioData Mining*, *14*. https://doi.org/10.1186/s13040-021-00245-y

Bilgrami, Z. R., Sarac, C., Srivastava, A., Herrera, S. N., Azis, M., Haas, S. S., Shaik, R. B., Parvaz, M. A., Mittal, V. A., Cecchi, G., & Corcoran, C. M. (2022). Construct validity for computational linguistic metrics in individuals at clinical risk for psychosis: Associations with clinical ratings. *Schizophrenia Research*, *245*, 90–96. https://doi.org/10.1016/j.schres.2022.01.019

Birnbaum, M. L., Ernala, S. K., Rizvi, A. F., Arenare, E., R. Van Meter, A., De Choudhury, M., & Kane, J. M. (2019). Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *Npj Schizophrenia*, *5*, Article 1. https://doi.org/10.1038/s41537-019-0085-9

Birnbaum, M. L., Norel, R., Van Meter, A., Ali, A. F., Arenare, E., Eyigoz, E., Agurto, C., Germano, N., Kane, J. M., & Cecchi, G. A. (2020). Identifying signals associated with psychiatric illness utilizing language and images posted to Facebook. *Npj Schizophrenia*, *6*, Article 1. https://doi.org/10.1038/s41537-020-00125-0

Brederoo, S. G., Nadema, F. G., Goedhart, F. G., Voppel, A. E., De Boer, J. N., Wouts, J., Koops, S., & Sommer, I. E. C. (2021). Implementation of automatic speech analysis for early detection of psychiatric symptoms: What do patients want? *Journal of Psychiatric Research*, *142*, 299–301. https://doi.org/10.1016/j.jpsychires.2021.08.019

Carrillo, F., Sigman, M., Fernández Slezak, D., Ashton, P., Fitzgerald, L., Stroud, J., Nutt, D. J., & Carhart-Harris, R. L. (2018). Natural speech algorithm applied to baseline interview data can predict which patients will respond to psilocybin for treatment-resistant depression. *Journal of Affective Disorders*, *230*, 84–86. https://doi.org/10.1016/j.jad.2018.01.006

Chandler, C., Foltz, P. W., & Elvevåg, B. (2022). Improving the Applicability of AI for Psychiatric Applications through Human-in-the-loop Methodologies. *Schizophrenia Bulletin*, *48*(5), 949–957. https://doi.org/10.1093/schbul/sbac038

Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., Bearden, C. E., & Cecchi, G. A. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*, *17*(1), 67–75.

de Boer, J. N., Brederoo, S. G., Voppel, A. E., & Sommer, I. E. C. (2020). Anomalies in language as a biomarker for schizophrenia: *Current Opinion in Psychiatry*, *33*(3), 212–218.

Dediu, D., Janssen, R., & Moisik, S. R. (2019). Weak biases emerging from vocal tract anatomy shape the repeated transmission of vowels. *Nature Human Behaviour*, *3*(10), 1107–1115. https://doi.org/10.1038/s41562-019-0663-x

Foltz, P. W., Chandler, C., Diaz-Asper, C., Cohen, A. S., Rodriguez, Z., Holmlund, T. B., & Elvevåg, B. (2022). Reflections on the nature of measurement in language-based automated assessments of patients' mental state and cognitive function. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2022.07.011

Fusaroli, M., Simonsen, A., Borrie, S. A., Low, D. M., Parola, A., Raschi, E., Poluzzi, E., & Fusaroli, R. (2022). *Identifying medications underlying communication atypicalities in psychotic and affective disorders: A pharmacosurveillance study within the FDA Adverse Event Reporting System* (p. 2022.09.05.22279609). medRxiv. https://doi.org/10.1101/2022.09.05.22279609

Fusaroli, R., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., & Weed, E. (2022). Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Research: Official Journal of the International Society for Autism Research*, *15*(4), 653–664. https://doi.org/10.1002/aur.2661

Goering, S., Klein, E., Specker Sullivan, L., Wexler, A., Agüera y Arcas, B., Bi, G., Carmena, J. M., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Kellmeyer, P., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., Teicher, M., … Yuste, R. (2021). Recommendations for Responsible Development and Application of Neurotechnologies. *Neuroethics*, *14*, 365–386. https://doi.org/10.1007/s12152-021-09468-6

Goller, F., Choi, S., Hong, U., & Ansorge, U. (2020). Whereof one cannot speak: How language and capture of visual attention interact. *Cognition*, *194*, 104023. https://doi.org/10.1016/j.cognition.2019.104023

Hagoort, P., & Beckmann, C. F. (2019). Key issues and future directions: The neural architecture for language. In P. Hagoort (Ed.), *Human language: From genes and brains to behavior* (pp. 527–532). MIT Press.

Hauglid, M. K. (2022). What's That Noise? Interpreting Algorithmic Interpretation of Human Speech as a Legal and Ethical Challenge. *Schizophrenia Bulletin*, *48*(5), 960–962. https://doi.org/10.1093/schbul/sbac008

Holmlund, T. B., Chandler, C., Foltz, P. W., Diaz-Asper, C., Cohen, A. S., Rodriguez, Z., & Elvevåg, B. (2022). Towards a temporospatial framework for measurements of disorganization in speech using semantic vectors. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2022.09.020

Huang, K.-Y., Wu, C.-H., Su, M.-H., & Kuo, Y.-T. (2020). Detecting Unipolar and Bipolar Depressive Disorders from Elicited Speech Responses Using Latent Affective Structure Model. *IEEE Transactions on Affective Computing*, *11*(3), 393–404. https://doi.org/10.1109/TAFFC.2018.2803178

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., & Cummings, R. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, *14*(1–2), 1–210. http://dx.doi.org/10.1561/2200000083

Keeling, G., & Nyrup, R. (2021). Explainable Machine Learning, Patient Autonomy, and Clinical Reasoning. In C. Véliz (Ed.), *The Oxford handbook of digital ethics* (online). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780198857815.013.27

Koutsouleris, N., Hauser, T. U., Skvortsova, V., & Choudhury, M. D. (2022). From promise to practice: Towards the realisation of AI-informed mental health care. *The Lancet Digital Health*, *4*(11), e829–e840. https://doi.org/10.1016/S2589-7500(22)00153-4

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry*, *76*(3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Mekki, Y., Guillemot, V., Lemaître, H., Carrión-Castillo, A., Forkel, S., Frouin, V., & Philippe, C. (2022). The genetic architecture of language functional connectivity. *NeuroImage*, *249*. https://doi.org/10.1016/j.neuroimage.2021.118795

Mundt, J. C., Vogel, A. P., Feltner, D. E., & Lenderking, W. R. (2012). Vocal Acoustic Biomarkers of Depression Severity and Treatment Response. *Biological Psychiatry*, *72*(7), 580–587. https://doi.org/10.1016/j.biopsych.2012.03.015

Ni, Y., Barzman, D., Bachtel, A., Griffey, M., Osborn, A., & Sorter, M. (2020). Finding warning markers: Leveraging natural language processing and machine learning technologies to detect risk of school violence. *International Journal of Medical Informatics*, *139*, 104137.

Nölle, J., Fusaroli, R., Mills, G. J., & Tylén, K. (2020). Language as shaped by the environment: Linguistic construal in a collaborative spatial task. *Palgrave Communications*, *6*, Article 1. https://doi.org/10.1057/s41599-020-0404-9

Palaniyappan, L. (2021). More than a biomarker: Could language be a biosocial marker of psychosis? *Npj Schizophrenia*, *7*, Article 1. https://doi.org/10.1038/s41537-021-00172-1

Palaniyappan, L., Alonso-Sanchez, M. F., & MacWhinney, B. (2022). Is Collaborative Open Science Possible With Speech Data in Psychiatric Disorders? *Schizophrenia Bulletin*, *48*(5), 963–966. https://doi.org/10.1093/schbul/sbac058

Parola, A., Simonsen, A., Bliksted, V., & Fusaroli, R. (2020). Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophrenia Research*, *216*, 24–40. https://doi.org/10.1016/j.schres.2019.11.031

Puiatti, A., Mudda, S., Giordano, S., & Mayora, O. (2011). Smartphone-centred wearable sensors network for monitoring patients with bipolar disorder. *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3644–3647. https://doi.org/10.1109/IEMBS.2011.6090613

Rudolph, J. M., & Leonard, L. B. (2016). Early Language Milestones and Specific Language Impairment. *Journal of Early Intervention*, *38*(1), 41–58. https://doi.org/10.1177/1053815116633861

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*, *346*(6209), 583–587.

Shain, C., Blank, I. A., Fedorenko, E., Gibson, E., & Schuler, W. (2022). Robust Effects of Working Memory Demand during Naturalistic Language Comprehension in Language-Selective Cortex. *Journal of Neuroscience*, *42*(39), 7412–7430. https://doi.org/10.1523/JNEUROSCI.1894-21.2022

Silva, A. M., Limongi, R., MacKinley, M., Ford, S. D., Alonso-Sánchez, M. F., & Palaniyappan, L. (2022). Syntactic complexity of spoken language in the diagnosis of schizophrenia: A probabilistic Bayes network model. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2022.06.011

Stanislawski, E. R., Bilgrami, Z. R., Sarac, C., Garg, S., Heisig, S., Cecchi, G. A., Agurto, C., & Corcoran, C. M. (2021). Negative symptoms and speech pauses in youths at clinical high risk for psychosis. *Npj Schizophrenia*, *7*(3), Article 1. https://doi.org/10.1038/s41537-020-00132-1

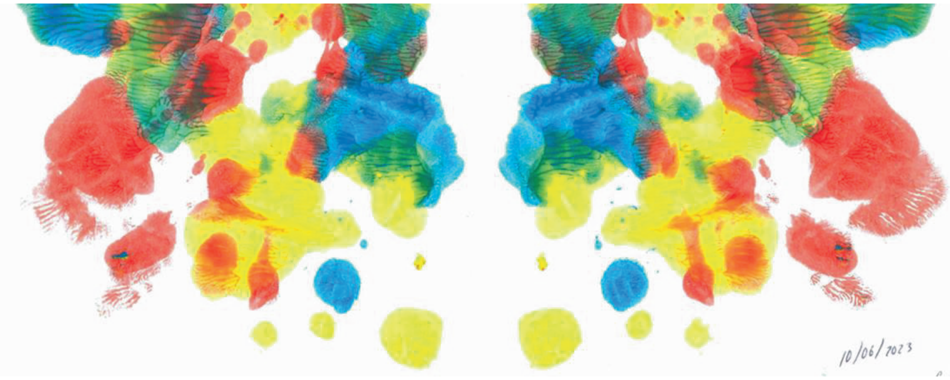
Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *Npj Digital Medicine*, *3*, Article 1. https://doi.org/10.1038/s41746-020-0221-y

Uher, R., & Zwicker, A. (2017). Etiology in psychiatry: Embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry*, *16*(2), 121–129. https://doi.org/10.1002/wps.20436

United Nations Conference on Trade and Development. (2021, December 14). *Data Protection and Privacy Legislation Worldwide*. https://unctad.org/page/data-protection-and-privacy-legislation-worldwide

Vasudevan, S., Saha, A., Tarver, M. E., & Patel, B. (2022). Digital biomarkers: Convergence of digital health technologies and biomarkers. *Npj Digital Medicine*, *5*. https://doi.org/10.1038/s41746-022-00583-z

Yuste, R., Goering, S., Arcas, B. A. Y., Bi, G., Carmena, J. M., Carter, A., Fins, J. J., Friesen, P., Gallant, J., Huggins, J. E., Illes, J., Kellmeyer, P., Klein, E., Marblestone, A., Mitchell, C., Parens, E., Pham, M., Rubel, A., Sadato, N., … Wolpaw, J. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*, *551*(7679), 159–163. https://doi.org/10.1038/551159a


10/06/2023

CHAPTER 6

General discussion

In this thesis, transcripts of auditory verbal hallucinations (AVHs) and speech were analyzed by means of natural language processing (NLP) techniques and artificial intelligence (AI) algorithms. The overarching aim was to assess the usefulness of linguistic features and computationally-derived measures in reliably, validly, and semi-automatically performing psychiatric clinical-like tasks, i.e., computational procedures mimicking clinicians' medical work (e.g., distinguishing speech from individuals with SSD from that of controls). More broadly, the potential descriptive and mechanistic roles of NLP-based markers in studying and providing care to patients with psychosis and/or other psychiatric disorders were outlined too. It was also emphasized that both long-ago recognized and new, emerging multidimensional challenges must be faced and overcome before implementing NLP-based markers in clinical decision support systems (CDSS) for routine clinical practice in psychiatry (e.g., helping with early detection of warning signs and/or differential diagnosis, selecting optimal treatment, and/or predicting relapse after remission).

Summary of main findings

In **chapter 2**, the main aim was to explore whether putative linguistic subtypes of AVHs could be identified by means of a cluster analysis (i.e., a computationally unsupervised learning technique) when applied to transcripts of AVHs without a priori grouping of the AVHs. The main finding of this study was the identification of two data-driven linguistic subtypes of AVHs in Dutch: compact-AVHs and expanded-AVHs. Compared to expanded-AVHs, compact-AVHs had fewer determiners and prepositions, and were shorter too, having a rounded averaged length of 5 words. Phenomenologically, compact-AVHs were characterized by a larger amount of negative content and a higher degree of negativity. Of note, compact-AVHs were mainly experienced by clinical voice-hearers. It was noted that the linguistic characteristics of these two data-driven subtypes of AVHs showed both consistencies and inconsistencies with previous findings in which an a priori group dichotomization of the AVHs was done. Further, it was discussed that distinct neurocognitive mechanisms might underlie these linguistic subtypes of AVHs, and that choosing a psychological treatment for AVHs might be informed by this linguistic subtyping procedure.

Chapter 3 showed how sentiment analysis, i.e., a computational technique allowing to automatically label and score words on negative, neutral, and/or positive valence, can address the problem of reliably quantifying the negative content of voices. This chapter further explored whether the averaged general valence (i.e., the "mean objective sentiment") of the voices related to the perceived (i.e., phenomenological) negativity of the voices. The results showed that, proportionally, clinical voice-hearers heard more negative-valence voices, and that this linguistic negative-valence of the voices associates with the perceived negativity, the amount of distress and the disruption of life related to the voices. Importantly, it was stressed that, while current psychological treatments for voices with negative content often focus on changing the beliefs about the voices, these treatments should also intend to modify and/or handle the negative content itself.

In **chapter 4**, an exploratory study tested whether explicit linguistic markers (i.e., connectives) that establish different relations of coherence in discourse (e.g., a comparison between entities or a temporal sequence of events) can be the operational anchor of the assessment of disorganized speech (DS) in patients with schizophrenia-spectrum disorders (SSD). It was found that, in proportion, connectives with a contingent meaning (e.g., "because") and those with polysemic meaning (e.g., "as") were less used by patients with SSD than by control participants. Also, results showed that, compared to control participants, patients with SSD had instances of both higher and lower coherent use of three different connective types. A full set of 35 connectives-related features was further shown to lead a support vector machine learning algorithm to achieve a cross-validated accuracy of 85% in distinguishing between groups.

Chapter 5 broadened the overview of what type of NLP-based digital markers for psychosis and other psychiatric disorders are currently under scientific and clinical scrutiny, highlighting that those markers could mainly fulfill descriptive (e.g., a mark for stratification in trials) and/or mechanistic (e.g., a mark indicating a disruption in neurocognitive mechanisms) roles. Further, this chapter outlined ten clinical actions that these NLP-based digital markers can be used for, while acknowledging a series of open questions for each of them. As a core message, this chapter stressed that ethical issues permeate deeply into multifaceted challenges that currently prevent any NLP-based digital marker to be implemented in CDSS for actual delivery of care. Yet, it also remarked that, through long-lasting and strictly-regulated international multidisciplinary collaboration, it will be possible to attain enough valid and reliable scientific knowledge, societal, legal, and clinical trust, and acceptable and accessible technology to start incorporating NLP-based digital markers in routine clinical psychiatric care.

Interpretation of main findings

Even if decades ago studies on voices/AVHs and psychiatric disorders already used linguistic labels while exemplifying (e.g., Campbell, 1930) or characterizing the voices (e.g., Mott et al., 1965), theory-based linguistic approaches date from less than 30 years ago (de Boer et al., 2016;

Hoffman et al., 1994; Leudar et al., 1997; Tovar et al., 2019). Even more, NLP-based approaches applied on actual transcripts of what the voices say to the voice-hearers started to emerged just recently (e.g., Turkington et al., 2019), and they remain scarce.

Subtyping voices/AVHs uniquely by their linguistic characteristics (**chapter 2**) and analyzing their negative-valence content by means of computational sentiment analysis (**chapter 3**) present themselves then as still-unconventional but promising approaches in the study of voices/AVHs. For the studies described in **chapters 2** and **3,** it was assumed that the hallucinated "voice-speech", i.e., the speech-like words, phrases, sentences and/or dialogues that were heard by the voice-hearers, presumably instantiated linguistic characteristics resembling those of actual human speech. Thus, by implementing linguistic theory, AI algorithms, and NLP techniques for preprocessing, quantification, analysis, and interpretation of the voices/AVHs data, the findings offered in this thesis extend our understanding about the linguistic variation of the voices/AVHs (**chapter 2**) and the distribution of their affect-related negative content (**chapter 3**) among clinical and non-clinical voice-hearers.

The need for questioning whether different neurocognitive models of voices/AVHs might account for these linguistic subtypes of voices/AVHs and for the negative-valence content that those voices might contain can be seen as an original contribution of this thesis. Several of these models exist (Brown & Kuperberg, 2015; Frith & Done, 1988; Grandchamp et al., 2019; Jones & Fernyhough, 2007; McGuire et al., 1995; Rollins et al., 2019; Sommer & Diederen, 2009; Waters et al., 2006). Yet, none of them offers a unifying framework that outlines the necessary and sufficient components and processes to parallelly posit hypotheses about how different subtypes of linguistic voices/AVHs can originate, and how it might be that negative-valence content can vary in presence across each linguistic subtype of voices/AVHs. For instance, the hierarchical-generative model (Brown & Kuperberg, 2015) and recent versions of inner-speech and self-monitoring models (Grandchamp et al., 2019; Jones & Fernyhough, 2007) offer detailed descriptions of the involvement of language-related processes in the generation of voices/AVHs. Hypotheses could be made about the generation of either compact-AVHs or expanded-AVHs based on those descriptions. Yet, none of these "language-focused" models addresses how negative-valence content might become part of the voices/AVHs. As another example, the non-dominant hemisphere model (Sommer & Diederen, 2009) posits that the right homologue of Broca's area might be responsible for the negative-valence content often found in relatively short-length voices/AVHs (Sommer & Diederen, 2009). However, this model leaves unexamined how it might be that different subtypes of voices/AVHs can originate, and how it might be that negative-valence content is found across linguistic subtypes of voices/AVHs.

Beyond research settings, subtyping voices/AVHs based on linguistic features, quantifying their negative-valence content, and elaborating neurocognitive mechanistic models to test their origin and characteristics gain further relevance in light of the possibility of using this knowledge to alleviate suffering linked to voice-hearing (Cancel et al., 2018; McCarthy-Jones et al., 2014). For instance, even though they might be effective, pharmacological interventions for voices/AVHs are often discontinued by patients due to common aversive effects resulting from antipsychotic medication use (Horowitz et al., 2022). Likewise, brain-stimulation treatments have been shown to be non-efficient to treat medication-resistant voices/AVHs (Guttesen et al., 2021). Since subtyping voices/AVHs based on linguistic characteristics has not been systematically done in previous research on treatments for voices/AVHs, it might be speculated that the presence of different linguistic subtypes of voices/AVHs built part of the heterogeneity in the intervention-groups. This in turn might have played a confounding role in those studies, partly accounting for the found inefficiency of the pharmacological or brain-stimulation interventions for the voices/AVHs.

Considering the plurality of brain regions involved in the origin of voices/AVHs across disorders (Rollins et al., 2019), it might be that different underlying mechanisms correspond to different linguistic subtypes of voices/AVHs. Thus, by successfully identifying the linguistic subtype(s) of voices/AVHs affecting a voice-hearer, pharmacological, psychological, and/or brain-stimulation interventions might become more effective, assuming those linguistic subtypes of voices/AVHs would respond better to interventions targeting those mechanisms. Furthermore, in the case of voices/AVHs that are even resistant to those treatments or for those voice-hearers who prefer to benefit from psychological interventions only, management of the negative-valence content would be then required (see hypothetical examples offered in **chapter 2** and **3**).

Overall, the points made so far about the NLP-based approach towards voices/AVHs can be interpreted under the view that NLP digital markers' roles can be either descriptive or mechanistic, as argued in **chapter 5**. For instance, the idea of subtyping the voices/AVHs of a given individual for the purpose of determining their group in a research study on treatment response would instantiate a descriptive role. In contrast, in actual clinical practice, taking a given linguistic subtype of voices/AVHs as a mark of a (set of) disruption(s) in specific underlying mechanisms that can be targeted by means of a given treatment would represent a mechanistic role. It is relevant to insist, however, that either with a descriptive or a mechanistic role, NLP digital markers can be used for other clinical actions too (see **chapter 5**), and the study presented in **chapter 4** shows in fact how in a diagnosis-like binary classification task an NLP connectives-related semantic marker could be used to assess DS in patients with SSD.

In **chapter 4**, it was shown that, both in their proportion of use and in their contextual utilization, some types of connectives (e.g., polysemic connectives) are particularly susceptible to reflect patterns that associate with SSD. Thus, in a fictitious scenario, a patient could receive a diagnosis of SSD relying on evidence that the proportion and incoherent use of certain connectives (i.e., behavior) reflects altered language processes (cognition) as a result of an SSD (i.e., biology). A contentious thought is that, for the medical decision taken in this fictitious scenario, attaining linguistic explainability would then be more crucial than addressing mathematical or computational accounts of the algorithm, considering that current psychiatric diagnostic classifications heavily rely on behavioral, cognitive, and biological constructs (Owen, 2014). Explainability remains an advocated and warranted characteristic of AI in medicine (Reddy, 2022), even though it has been argued that there is "a false hope for explainable AI" (Ghassemi et al., 2021). To a small but an illustrative extent, it can be argued that the findings presented in **chapter 4** highlight that linguistic explainability might become a cornerstone of digital markers for psychiatric disorders, leading clinicians to hold positive attitudes towards the possible future use of these AI-based technologies in actual care-delivery settings (Doraiswamy et al., 2020; Terra et al., 2023; Young et al., 2021).

Methodological considerations

In the empirical studies presented in **chapters 2-4**, the samples of the targeted populations were clearly specified and chosen to address the aim(s) of each study, and, to the largest possible extent, sociodemographic and clinical factors that are known to influence the outcomes that were of interest were controlled for. The instruments, procedures, and statistical, NLP, and AI techniques used to carry out the analyses were chosen both largely on the ground of previous research and according to the needs of the innovative approach taken to carry out each study. When necessary, supplementary analyses were carried out as well to test the consistency and/or interpretability of the results. The experts-based article presented in **chapter 5** was written as a report of a multidisciplinary workshop that took place largely following a nominal group technique (i.e., a consensus development method). Thus, overall, the studies presented in **chapters 2-5** followed rigorous methodological standards.

Yet, a series of methodological considerations constrain the extent of the reliability, validity, and generalizability of the results presented in **chapters 2-4**. One limitation is that, in these three studies, the sample sizes were relatively small (N = 40 in **chapters 2** and **3** independently, and N = 100 in **chapter 4**). It is widely known that a small sample size can make the replicability of studies' results harder to achieve, and it increases the chance of obtaining false negative results too (Althubaiti, 2023). Evidently, this might have affected the studies in **chapters 2** and **3** to a larger extent, although the findings of **chapter 4** are not exempted from these possibilities.

A second limitation relates to the elicitation methods. As acknowledge in **chapters 2** and **3**, either because of emotional (e.g., shame), cognitive (e.g., verbal repetition skills) or situational factors (e.g., distractors), the shadow procedure used to obtain the verbatim repetitions of the voices/AVHs might not have allowed to precisely reflect what the voice-hearers experienced as voices/AVHs. Similarly, considering that speech task might change the outcomes related to DS (Parola et al., 2022), it might be that the elicitation technique used to collect the speech samples for the study presented in **chapter 4** (i.e., open-ended questions) was not the best suited to assess DS based on connectives-related features. For instance, rather than open-ended questions, a task could have been developed to lead participants to give a verbal response with a type of discourse (e.g., explanations vs narrations) along with specific grammatical complexity (e.g., syntactic coordination vs syntactic subordination), allowing to refine the analysis and understanding of how these factors influence the (in)coherent use of connectives.

A third limitation is that antipsychotic medication might have an impact on speech characteristics (de Boer et al., 2020; Sinha et al., 2015). Thus, for all the patients who participated in these studies, either their verbatim repetitions of the voices/AVHs or their speech produced during the PRAAT interview could have been further influenced by the antipsychotic medication they were taking. Specifically, whether the shadowing procedure is equally reliable between periods with or without antipsychotic medication to obtain verbatim repetitions of voices/AVHs is unknown, and a similar question can be asked for the use of open-ended questions to elicit speech samples.

The set of features used for each independent analysis carried out in **chapters 2-4** places constraints on the findings reported there too. Even if for actual speech the set and number of features needed to account for linguistic variation has been extensively studied based on a multidimensional approach (Biber, 1995, 2012; Biber & Conrad, 2009), the same cannot be said for voices/AVHs. Thus, even if the set and number of linguistic features used for the analyses presented in **chapters 2** and **3** were determined relying on previous research, our results can only represent an exploration of what linguistic variation of voices/AVHs might be.

Likewise, a specific set and number of connective types were used to carry out the analyses presented in **chapter 4**, but evidence still needs to accumulate to determine what set and number of connectives allows to most reliably and validly assess DS in individuals with SSD.

Another limitation derives from the particular type and number of AI algorithms and/or NLP techniques that were used to carry out the analyses in each study. For instance, in **chapter 2**, the clustering analysis was done using Canberra distance and Ward's method alone, despite the existence of multiple distances and methods for the clustering (Moisl, 2015) and no current guidance to determine the most optimal combination of these to analyze voices/AVHs. In **chapter 3**, it was acknowledged that swear words were not considered for analysis due to her absence in the Pattern sentiment-analysis tool (De Smedt & Daelemans, 2012), despite the fact that swear words are an exemplar case of linguistic units that can largely convey negative valence (Stapleton et al., 2022). In **chapter 4**, only Word2vec (Mikolov et al., 2013) was used for creating the word embeddings that were used for the cosine similarity analysis, even though it has been repeatedly pointed out that it cannot account for contextualized embeddings (Foltz et al., 2022), which would have improved the analysis of the polysemic connectives, for instance. On a daily basis, AI and NLP advances expand rapidly (Sawicki et al., 2023; Zhang & Lu, 2021), making very difficult to determine what the "state-of-the-art" in AI and NLP techniques is. Parallelly, scientific mastery of AI and NLP (i.e., AI literacy) is a laborious, time-consuming task that comprises the acquisition of computational, mathematical, and linguistic knowledge and skills, among others (Ng et al., 2021). Thus, regarding the use of AI algorithms and NLP techniques, the results presented in **chapters 2-4** should only be seen, inevitably, as "state-*out*-of-the-art".

Finally, as extensively addressed in **chapter 5** and largely discussed by other researchers, NLP measures and computational language-based models suffer from "human-like", "stereotyped" biases, such as unfair socio-demographic representativeness and gender inequality (Bailey et al., 2022; Caliskan et al., 2017; Hovy & Prabhumoye, 2021; Straw & Callison-Burch, 2020), raising widespread ethical, legal, cultural-societal, and clinical concerns (Hauglid, 2022; Hauglid & Mahler, 2023). Acknowledgedly, the studies presented in **chapters 2-4** were unable to avoid these "human-like" biases. For instance, the corpus used to create the computational semantic model of the word embeddings in **chapter 4**, i.e., *Het Corpus Gesproken Nederlands* (van Eerten, 2007), arguably represents only a proportion of the larger socio- and ethnolinguistic variation of the Dutch language that is spoken in the Netherlands. In parallel with this, only native Dutch speakers took part of the studies presented in this thesis, despite the fact that positive psychotic symptoms are experienced by adult immigrants living in the Netherlands too (Stouten et al., 2019; Vanheusden et al., 2008). Also, individuals belonging to sexual minority groups who live in the Netherlands have a higher cumulative incidence of psychotic symptoms as compared to heterosexual individuals (Gevonden et al., 2014), but none of the studies presented in **chapters 2-4** focused on these populations. In sum, these constraints, even if thoughtfully chosen on the basis of practical methodological considerations or limitations, should not be overlooked in interpreting the findings of the **chapters 2-4**.

Future directions

Towards a linguistic theory of and methodology for voices/AVHs

In the studies presented in **chapters 2** and **3**, linguistic theory and methods born from studies on spoken language were used to analyze linguistic characteristics in verbatim repetitions of voices/AVHs. As yet, however, it can only be speculated that any theoretical "speech-based" linguistic framework actually accommodates the "speech-like" characteristics of voices/AVHs. In other words, it is still unknown to what extent voices/AVHs' linguistic characteristics might be accounted for by linguistic theory and methods that were developed to study actual articulated speech.

Currently, a large corpus of voices/AVHs is missing, which hampers drawing estimations about distributional and frequency properties that are present in the linguistic features of voices/AVHs at a large scale. For instance, in written and spoken language, determiners and prepositions have a relatively large frequency of use (Bentz et al., 2017; van Heuven et al., 2014), but whether this is the case in voices/AVHs is undetermined. The compact-AVHs had fewer determiners and prepositions than the expanded-AVHs, but, in the absence of knowledge on distributional and frequency properties of these word categories in voices/AVHs at a large scale, it is impossible to know which of these subtypes of voices/AVHs might instantiate a more common pattern of "voices/AVHs-based" linguistic variation.

To expand the current knowledge on linguistic variation in voices/AVHs, verbatim repetitions of voices/AVHs should be collected across clinical and non-clinical populations of voice-hearers in multiple languages. If these data were collected longitudinally, the extent to which the linguistic characteristics of voices/AVHs (including their negative-valence content) vary over time could be examined, and whether only one or several subtypes of voices/AVHs are experienced by each voice-hearer could be assessed. The (cross-linguistic) construct validity of the compact-AVHs and the expanded-AVHs could be tested, and the opportunity to identify other linguistic subtypes of voices/AVHs would arise. Of note, it should be examined whether existing methods used to study linguistic variation in actual speech, e.g., multidimensional analysis (Biber, 2012; Biber & Conrad, 2009; Eckert, 2012), have similar reliability in analyzing voices/AVHs. Parallelly, it should be considered whether new methodologies are needed, e.g., one combining NLP with micro-phenomenology (Petitmengin, 2006), maybe allowing to integrate the linguistic and the experiential dimension of voice-hearing equally in detail.

Understanding linguistic organization in voices/AVHs and actual speech

Both experiencing voices/AVHs (Ćurčić-Blake et al., 2017; Rollins et al., 2019) and producing actual speech (Giglio et al., 2022; Rolls et al., 2022) rely on neurocognitive mechanisms related to human language processing. In the case of actual speech production, these mechanisms underpin psycholinguistic and neurolinguistic processes that, based on elementary linguistic units (e.g., lexical items) and operations (e.g., grammatical construction of sentences) (Hagoort, 2019), allow linguistic organization to arise. When a given strand of actual spoken language entails linguistic organization attaining hierarchical continuity and connectivity across multiples levels (e.g., syntax, semantics, and pragmatics), it is considered to be coherent (Givón, 2020). While how actual speech can become disorganized has been largely described (Andreasen, 1979; Covington et al., 2005; Hinzen & Rosselló, 2015), little is still known about the extent to which voices/AVHs show "disorganized-like" linguistic characteristics, although grammatical violations have been reported in verbatim instances of voices/AVHs (de Boer et al., 2016; Tovar et al., 2019). It has been posited that abnormalities in language processes underly both disorganized speech and voices/AVHs (Brown & Kuperberg, 2015). Supporting this, there is evidence of shared brain mechanisms related to voices/AVHs and DS (Chang et al., 2022). Yet, a comparison of linguistic disorganization between voices/AVHs and actual speech of the voice-hearer is currently non-existent. Exploring this is paramount for at least two reasons. On the one hand, it would show whether linguistic disorganization affects voices/AVHs and actual speech similarly or not, suggesting shared underlying mechanisms. On the other hand, on top of the possibility to predict whether voices/AVHs or DS or both would emerge (Brown & Kuperberg, 2015), hypotheses could be drawn about the specific type of language-related processes abnormalities that either allow or prevent linguistic disorganization to occur in both voices/AVHs and actual speech.

Linguistic concerns related to NLP-based digital markers for psychiatric disorders

Besides the multidisciplinary challenges raised in **chapter 5**, linguistic limitations remain related to the development and validation of NLP-based digital markers for psychiatric disorders. To exemplify, consider the approach followed in **chapter 4** to assess DS in SSD. In carrying out this study, the fact that syntax, semantics and pragmatics interact (Belloro, 2019) was not taken into account. Assessing connectives related to DS should comprise and try to account for all these levels of information, not only for the semantic one. To attempt this, different self-contained meaningful units (e.g., phrases, [in]dependent sentences, full answers, etc.) should be used to assess the use of the connectives at different scales, which was completely overlooked in this study too. It is evident then that the NLP connectives-related semantic marker used in **chapter 4** to mimic a binary diagnosis task holds an undermined linguistic interpretability, further questioning what its descriptive role really depends on. In a broader sense, thus, guaranteeing the "linguistic design" to test and validate any NLP-based digital marker must be a priority, not only for its theoretical accountability, but for clinical ease of interpretability too.

General conclusions

This thesis showed that identifying linguistic subtypes of voices/AVHs and objectively quantifying negative-valence content in voices/AVHs' samples obtained from both clinical and non-clinical voice-hearers can be attained combining linguistics, artificial intelligence algorithms, and natural language processing techniques. Similarly, the assessment of disorganized speech in patients with schizophrenia-spectrum disorders on the basis of their use of linguistic connectives was possible by exploiting a combination of these theories and methods. More broadly, characteristics, roles, uses, and limitations of existing NLP-based digital markers for psychiatric disorders were summarized, and some venues to overcome the challenges that currently prevent their implementation in clinical practice were suggested too. Overall, despite their susceptibility to socio-cultural, contextual, and neurocognitive factors, hallucinated and spoken linguistic patterns hold potential to become clinically relevant markers of psychiatric disorders. These disorders pervade all human groups, so these linguistic markers of psychiatric disorders should be studied in all human groups. Since human groups, languages, brains and minds change, I wonder whether all linguistic markers of psychiatric disorders change too. Or maybe some. Or maybe none. Hallucinated and spoken linguistic patterns relate to processes, so whether NLP-based digital markers for psychiatric disorders retain (part of) this dynamic nature is a key pending problem that must be disentangled to refine our understanding both of these markers and these disorders. I truly hope that, by combining our current knowledge with upcoming findings about NLP-based digital markers, many individuals affected by psychiatric disorders might experience the alleviation of their suffering.

References

Althubaiti, A. (2023). Sample size determination: A practical guide for health researchers. *Journal of General and Family Medicine*, *24*(2), 72–78. https://doi.org/10.1002/jgf2.600

Andreasen, N. C. (1979). Thought, Language, and Communication Disorders: I. Clinical Assessment, Definition of Terms, and Evaluation of Their Reliability. *Archives Of General Psychiatry*, *36*(12), 1315–1321.

Bailey, A. H., Williams, A., & Cimpian, A. (2022). Based on billions of words on the internet, people = men. *Science Advances*, *8*(13), eabm2463. https://doi.org/10.1126/sciadv.abm2463

Belloro, V. A. (Ed.). (2019). *La Interfaz Sintaxis-Pragmática. Estudios teóricos, descriptivos y experimentales*. De Gruyter.

Bentz, C., Alikaniotis, D., Samardžić, T., & Buttery, P. (2017). Variation in Word Frequency Distributions: Definitions, Measures and Implications for a Corpus-Based Language Typology. *Journal of Quantitative Linguistics*, *24*(2–3), 128–162. https://doi.org/10.1080/09296174.2016.1265792

Biber, D. (1995). *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge University Press.

Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, *8*(1), 9–37. https://doi.org/10.1515/cllt-2012-0002

Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge University Press. https://doi.org/10.1017/CBO9780511814358

Brown, M., & Kuperberg, G. R. (2015). A Hierarchical Generative Framework of Language Processing: Linking Language Perception, Interpretation, and Production Abnormalities in Schizophrenia. *Frontiers in Human Neuroscience*, *9*, 1–23.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science (New York, N.Y.)*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Campbell, C. M. (1930). Hallucinations; their nature and significance. *American Journal of Psychiatry*, *86*(4), 607–618. https://doi.org/10.1176/ajp.86.4.607

Cancel, A., Rousseau, P.-F., & Jardri, R. (2018). Peut-on guider le traitement des hallucinations sur la base de la sémiologie et de la phénoménologie ? *Annales Médico-Psychologiques, Revue Psychiatrique*, *176*(8), 819–823. https://doi.org/10.1016/j.amp.2018.08.008

Chang, X., Zhao, W., Kang, J., Xiang, S., Xie, C., Corona-Hernández, H., Palaniyappan, L., & Feng, J. (2022). Language abnormalities in schizophrenia: Binding core symptoms through contemporary empirical evidence. *Schizophrenia,* *8*(1), 95. https://doi.org/10.1038/s41537-022-00308-x

Covington, M. A., He, C., Brown, C., Naçi, L., McClain, J. T., Fjordbak, B. S., Semple, J., & Brown, J. (2005). Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research*, *77*(1), 85–98. https://doi.org/10.1016/j.schres.2005.01.016

Ćurčić-Blake, B., Ford, J. M., Hubl, D., Orlov, N. D., Sommer, I. E., Waters, F., Allen, P., Jardri, R., Woodruff, P. W., David, O., Mulert, C., Woodward, T. S., & Aleman, A. (2017). Interaction of language, auditory and memory brain networks in auditory verbal hallucinations. *Progress in Neurobiology*, *148*, 1–20.

de Boer, J. N., Heringa, S. M., van Dellen, E., Wijnen, F. N. K., & Sommer, I. E. C. (2016). A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it. *Brain and Language*, *162*, 10–18.

de Boer, J. N., Voppel, A. E., Brederoo, S. G., Wijnen, F. N. K., & Sommer, I. E. C. (2020). Language disturbances in schizophrenia: The relation with antipsychotic medication. *NPJ Schizophrenia*, *6*(1), 24. https://doi.org/10.1038/s41537-020-00114-3

De Smedt, T., & Daelemans, W. (2012). Pattern for Python. *Journal of Machine Learning Research*, *13*, 2031–2035.

Doraiswamy, P. M., Blease, C., & Bodner, K. (2020). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artificial Intelligence in Medicine*, *102*, 101753. https://doi.org/10.1016/j.artmed.2019.101753

Eckert, P. (2012). Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology*, *41*(1), 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Foltz, P. W., Chandler, C., Diaz-Asper, C., Cohen, A. S., Rodriguez, Z., Holmlund, T. B., & Elvevåg, B. (2022). Reflections on the nature of measurement in language-based automated assessments of patients' mental state and cognitive function. *Schizophrenia Research*. https://doi.org/10.1016/j.schres.2022.07.011

Frith, C. D., & Done, D. J. (1988). Towards a Neuropsychology of Schizophrenia. *British Journal of Psychiatry*, *153*(4), 437–443.

Gevonden, M. J., Selten, J. P., Myin-Germeys, I., de Graaf, R., ten Have, M., van Dorsselaer, S., van Os, J., & Veling, W. (2014). Sexual minority status and psychotic symptoms: Findings from the Netherlands Mental Health Survey and Incidence Studies (NEMESIS). *Psychological Medicine*, *44*(2), 421–433. https://doi.org/10.1017/S0033291713000718

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet Digital Health*, *3*(11), e745–e750. https://doi.org/10.1016/S2589-7500(21)00208-9

Giglio, L., Ostarek, M., Weber, K., & Hagoort, P. (2022). Commonalities and Asymmetries in the Neurobiological Infrastructure for Language Production and Comprehension. *Cerebral Cortex*, *32*(7), 1405–1418. https://doi.org/10.1093/cercor/bhab287

Givón, T. (2020). *Coherence*. John Benjamins Publishing Company.

Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., Lachaux, J.-P., Dohen, M., Perrier, P., Garnier, M., Baciu, M., & Lœvenbruck, H. (2019). The ConDialInt Model: Condensation, Dialogality, and Intentionality Dimensions of Inner Speech Within a Hierarchical Predictive Control Framework. *Frontiers in Psychology*, *10*, 1–30.

Guttesen, L. L., Albert, N., Nordentoft, M., & Hjorthøj, C. (2021). Repetitive transcranial magnetic stimulation and transcranial direct current stimulation for auditory hallucinations in schizophrenia: Systematic review and meta-analysis. *Journal of Psychiatric Research*, *143*, 163–175. https://doi.org/10.1016/j.jpsychires.2021.09.001

Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science (New York, N.Y.)*, *366*(6461), 55–58. https://doi.org/10.1126/science.aax0289

Hauglid, M. K. (2022). What's That Noise? Interpreting Algorithmic Interpretation of Human Speech as a Legal and Ethical Challenge. *Schizophrenia Bulletin*, *48*(5), 960–962. https://doi.org/10.1093/schbul/sbac008

Hauglid, M. K., & Mahler, T. (2023). Doctor Chatbot: The EU's Regulatory Prescription for Generative Medical AI. *Oslo Law Review*, *10*(1), 1–23. https://doi.org/10.18261/olr.10.1.1

Hinzen, W., & Rosselló, J. (2015). The linguistics of schizophrenia: Thought disturbance as language pathology across positive symptoms. *Frontiers in Psychology*, *6*, 1–17.

Hoffman, R. E., Oates, E., Hafner, J., Hustig, H. H., & McGlashan, T. H. (1994). Semantic organization of hallucinated ‘‘voices” in schizophrenia. *The American Journal of Psychiatry*, *151*(8), 1229–1230.

Horowitz, M. A., Moncrieff, J., Haan, L. de, Bogers, J. P. A. M., Gangadin, S. S., Kikkert, M., Veling, W., & Sommer, I. E. C. (2022). Tapering antipsychotic medication: Practical considerations. *Psychological Medicine*, *52*(1), 32–35. https://doi.org/10.1017/S0033291721003299

Hovy, D., & Prabhumoye, S. (2021). Five sources of bias in natural language processing. *Language and Linguistics Compass*, *15*(8), e12432. https://doi.org/10.1111/lnc3.12432

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*, *16*(2), 391–399.

Leudar, I., Thomas, P., McNally, D., & Glinski, A. (1997). What voices can do with words: Pragmatics of verbal hallucinations. *Psychological Medicine*, *27*(4), 885–898. https://doi.org/10.1017/s0033291797005138

McCarthy-Jones, S., Thomas, N., Strauss, C., Dodgson, G., Jones, N., Woods, A., Brewin, C. R., Hayward, M., Stephane, M., Barton, J., Kingdon, D., & Sommer, I. E. (2014). Better Than Mermaids and Stray Dogs? Subtyping Auditory Verbal Hallucinations and Its Implications for Research and Practice. *Schizophrenia Bulletin*, *40*(Suppl 4), S275–S284.

McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., David, A. S., Frackowiak, R. S., & Frith, C. D. (1995). Abnormal monitoring of inner speech: A physiological basis for auditory hallucinations. *Lancet*, *346*(8975), 596–600. https://doi.org/10.1016/s0140-6736(95)91435-8

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. *Proc. of the 26th International Conference on Neural Information Processing Systems*, 3111–3119.

Moisl, H. (2015). *Cluster Analysis for Corpus Linguistics*. De Gruyter Mouton.

Mott, R. H., Small, I. F., & Anderson, J. M. (1965). Comparative Study Of Hallucinations. *Archives of General Psychiatry*, *12*(6), 595–601. https://doi.org/10.1001/archpsyc.1965.01720360067011

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*. https://doi.org/10.1016/j.caeai.2021.100041

Owen, M. J. (2014). New approaches to psychiatric diagnostic classification. *Neuron*, *84*(3), 564–571. https://doi.org/10.1016/j.neuron.2014.10.028

Parola, A., Lin, J. M., Simonsen, A., Bliksted, V., Zhou, Y., Wang, H., Inoue, L., Koelkebeck, K., & Fusaroli, R. (2022). Speech disturbances in schizophrenia: Assessing cross-linguistic generalizability of NLP automated measures of coherence. *Schizophrenia Research*, S0920-9964(22)00274-2. https://doi.org/10.1016/j.schres.2022.07.002

Petitmengin, C. (2006). Describing one’s subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, *5*(3), 229–269. https://doi.org/10.1007/s11097-006-9022-2

Reddy, S. (2022). Explainability and artificial intelligence in medicine. *The Lancet Digital Health*, *4*(4), e214–e215. https://doi.org/10.1016/S2589-7500(22)00029-2

Rollins, C. P. E., Garrison, J. R., Simons, J. S., Rowe, J. B., O’Callaghan, C., Murray, G. K., & Suckling, J. (2019). Meta-analytic Evidence for the Plurality of Mechanisms in Transdiagnostic Structural MRI Studies of Hallucination Status. *EClinicalMedicine*, *8*, 57–71.

Rolls, E. T., Deco, G., Huang, C.-C., & Feng, J. (2022). The human language effective connectome. *NeuroImage*, *258*, 119352. https://doi.org/10.1016/j.neuroimage.2022.119352

Sawicki, J., Ganzha, M., & Paprzycki, M. (2023). The state of the art of Natural Language Processing—A systematic automated review of NLP literature using NLP techniques. *Data Intelligence*, 1–47. https://doi.org/10.1162/dint_a_00213

Sinha, P., Vandana, V. P., Lewis, N. V., Jayaram, M., & Enderby, P. (2015). Evaluating the effect of risperidone on speech: A cross-sectional study. *Asian Journal of Psychiatry*, *15*, 51–55. https://doi.org/10.1016/j.ajp.2015.05.005

Sommer, I. E., & Diederen, K. M. (2009). Language production in the non-dominant hemisphere as a potential source of auditory verbal hallucinations. *Brain*, *132*(10), e124.

Stapleton, Karyn., Beers Fägersten, Kristy., Stephens, Richard., & Loveday, Catherine. (2022). The power of swearing: What we know and what we don't. *Lingua*, *277*, 103406. https://doi.org/10.1016/j.lingua.2022.103406

Stouten, L. H., Veling, W., Laan, W., & Van der Gaag, M. (2019). Psychopathology, cognition and outcome in Dutch and immigrant first-episode psychosis patients. *Early Intervention in Psychiatry*, *13*(3), 646–656. https://doi.org/10.1111/eip.12561

Straw, I., & Callison-Burch, C. (2020). Artificial Intelligence in mental health and the biases of language based models. *PloS One*, *15*(12), e0240376. https://doi.org/10.1371/journal.pone.0240376

Terra, M., Baklola, M., Ali, S., & El-Bastawisy, K. (2023). Opportunities, applications, challenges and ethical implications of artificial intelligence in psychiatry: A narrative review. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery*, *59*(1), 80. https://doi.org/10.1186/s41983-023-00681-z

Tovar, A., Fuentes-Claramonte, P., Soler-Vidal, J., Ramiro-Sousa, N., Rodriguez-Martinez, A., Sarri-Closa, C., Sarró, S., Larrubia, J., Andrés-Bergareche, H., Miguel-Cesma, M. C., Padilla, P. P., Salvador, R., Pomarol-Clotet, E., & Hinzen, W. (2019). The linguistic signature of hallucinated voice talk in schizophrenia. *Schizophrenia Research*, *206*, 111–117.

Turkington, G. D., Lawrence, T., & Stuart, W. (2019). Linguistic analysis of the valence, arousal and dominance of auditory hallucinations and internal thoughts in schizophrenia: Implications for psychoeducation and CBT. *Cogent Psychology*, *6*, 1–5.

van Eerten, L. (2007). Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde*, *12*(3), 194–2015.

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vanheusden, K., Mulder, C. L., van der Ende, J., Selten, J.-P., van Lenthe, F. J., Verhulst, F. C., & Mackenbach, J. P. (2008). Associations between ethnicity and self-reported hallucinations in a population sample of young adults in The Netherlands. *Psychological Medicine*, *38*(8), 1095–1102. https://doi.org/10.1017/S0033291707002401

Waters, F., Badcock, J., Michie, P., & Maybery, M. (2006). Auditory hallucinations in schizophrenia: Intrusive thoughts and forgotten memories. *Cognitive Neuropsychiatry*, *11*(1), 65–83.

Young, A. T., Amara, D., Bhattacharya, A., & Wei, M. L. (2021). Patient and general public attitudes towards clinical artificial intelligence: A mixed methods systematic review. *The Lancet. Digital Health*, *3*(9), e599–e611. https://doi.org/10.1016/S2589-7500(21)00132-1

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, 100224. https://doi.org/10.1016/j.jii.2021.100224

Nederlandse samenvatting

In dit proefschrift werden auditieve verbale hallucinaties (AVH's) en ongeorganiseerde spraak onderzocht.

In **Hoofdstuk 2** presenteren we een datagestuurde en natuurlijke taalverwerking (NLP in Engels) aanpak voor het bestuderen van transcripties van woordelijk verzamelde AVH's. De transcripties van de AVH's, verkregen van zowel klinische als niet-klinische stemhoorders, werden onderworpen aan automatische taalkundige classificatie en aan automatische berekening van kwantitatieve taalkundige maten van complexiteit, resulterend in een reeks van 16 kenmerken. Niet uitgaande van een *a priori* verschil op groepsniveau, werd een niet-gesuperviseerde clusteranalyse uitgevoerd met behulp van deze kenmerken om clusters van AVH's te onderscheiden. Er werden twee verschillende AVH-clusters gevonden, die verband hielden met de klinische status. We interpreteren de AVH-clusters volgens verschillende neurocognitieve modellen, en veronderstellen dat hun taalkundige kenmerken artsen kunnen helpen beslissingen te nemen met betrekking tot psychologische interventies voor AVH’s.

Hoofdstuk 3 laat zien hoe sentimentanalyse (d.w.z. de bepaling van positieve, neutrale en negatieve taalkundige valentie) op basis van NLP-technieken kan worden gebruikt om het gemiddelde van deze ‘sentimenten’ tussen transcripties van AVH’s van klinische en niet-klinische stemhoorders te vergelijken. Vergelijkingen tussen automatisch gelabelde ‘sentimenten’ van de AVH-uitingen tussen de twee groepen lieten zien dat AVH’s uit de klinische groep een groter aandeel negatieve stemmen bevatten dan de AVH’s van de groep niet-klinische stemmenhoorders. Het gemiddelde sentiment (d.w.z. de valentie) van de uitingen bleek statistisch significant gecorreleerd te zijn met de ervaren negativiteit, hoeveelheid leed en ontwrichting van het leven. We bespreken deze resultaten in het licht van de huidige discussies over de negatieve inhoud van AVH's, en we benadrukken het belang van hoe deze bevindingen psychologische interventies voor AVH's kunnen verbeteren.

Hoofdstuk 4 gaat in op de uitdaging met computermodellen de (in)coherentie in spraak van individuen met schizofrenie-spectrumstoornissen (SSD) en controledeelnemers te beoordelen door connectieve (in)coherentie te ontwarren van (in)coherentie zonder connecties. Het aandeel van verschillende soorten verbindingen en hun gebruikswijze werden tussen groepen beoordeeld. Er werden verschillen gevonden in de verhouding tussen het gebruik van twee soorten verbindingen tussen de groepen, en vijf van NLP afgeleide maatstaven met betrekking tot de wijze van gebruik van de soorten verbindingen verschilden ook aanzienlijk tussen de groepen. Verder werd een algemene nauwkeurigheid van 85% verkregen bij het maken van onderscheid tussen patiënten met SSD en controledeelnemers op basis van slechts 35 connectiviteitsgerelateerde kenmerken. We betogen dat connectiviteitsgerelateerde kenmerken kunnen worden gebruikt om de (in)coherentie in SSD betrouwbaar te beoordelen, terwijl we erkennen dat er nog verschillende methodologische obstakels moeten worden overwonnen.

Hoofdstuk 5 gaat verder dan alleen (in)coherentie en de beoordeling ervan voor (differentiële) diagnosedoeleinden en verbreedt de discussie over op NLP of gesproken taalverwerking (SLP in Engels) gebaseerde digitale markers van psychose en andere psychiatrische stoornissen. Dit hoofdstuk definieert tien klinische prioriteiten waarvoor NLP/SLP-gebaseerde digitale markers momenteel rigoureus en intensief worden ontwikkeld, waarbij tegelijkertijd wordt benadrukt dat bij het streven om dergelijke digitale markers te creëren multidimensionale uitdagingen komen kijken die interinstitutionele en internationale samenwerking vereisen om te worden opgelost.

Hoofdstuk 6 begint met het samenvatten van de belangrijkste bevindingen van de onderzoeken gepresenteerd in de hoofdstukken 2, 3, 4 en 5, gevolgd door een identificatie van de sterke punten en beperkingen die de onderzoeken gemeen hebben. In het laatste deel van hoofdstuk 6 gaan we verder dan de gemeenschappelijke basis die door de gepresenteerde onderzoeken wordt gedeeld, en doen we een poging om aspecten die elk onderzoeksonderwerp kenmerken samen te smelten of met elkaar te verweven, d.w.z. een poging om delen van schijnbaar op zichzelf staande onderwerpen samen te brengen. Mijn hoop is dat dit het mogelijk zal maken om af te bakenen of te onthullen wat nog steeds “onbekend” is en de moeite waard is om te onderzoeken met het doel te begrijpen wat AVH’s en spraak met elkaar verbindt en van elkaar onderscheidt bij patiënten met psychose.

Acknowledgments

With few but meaningful words I want to express my gratitude to a number of persons whose guidance, companionship and support were present during my PhD trajectory, leading to its successful end.

Prof. Iris Sommer, towards you I feel respect and thankfulness in depth. Having accepted me as PhD student in the Netherlands altered the course of my life and that of my family. I need more than written lines to manifest how transcendental this is, but I wish that through these words the bare minimum of these feelings can be expressed.

Sanne Brederoo and Janna de Boer, I feel fortunate that, throughout my PhD trajectory, you both nurtured interdisciplinarity at the core of my learning process. Together, the three of us looked at the different angles of language and speech in the context of psychiatric disorders, and, since I know that you know how happy this made me feel, I just want to add this: such happiness is among the happiest happiness I have ever experienced.

Emile d'Angremont and Silvia Ciampelli, my friends, I appreciate you both so much. Appreciation born from joy and delight in the academic talks we had and from the kindness and relaxation I have felt many times while sharing food or personal stories together.

A mi esposa Andrea y a mi hija Elisa: ¡qué reciprocidad de amor y de alegría! De su cariño y bienestar los míos se regocijan. Andrea, en especial, gracias por tu apoyo, por tu tiempo, por tus ideas, por tu creatividad, por tus cuidados y por tu compañía. Mi tiempo en Países Bajos y durante mi doctorado son sólo una parte de nuestro tiempo. ¡Qué dicho tiempo siga y siga!

A mi familia en México: lejos están mis progenitores y mi hermana. Lejos están los progenitores de mis progenitores. Cerca viven y los veo, empero, en mi vida y en mis obras.

To my research group PhD colleagues, fellow PhD students from the BCN, and all the professors who instructed me during my PhD: quietly, from you I learned about statistics, neuroscience, psychiatry and many more disciplines. I may have openly thanked you before. And here so do I once more.

To myself: thank you, Hugo, for the PhD experience as a whole. May you/I be able to continue joyfully enjoying the gift of knowledge, doubts, and discoveries.