A data-driven linguistic characterization of hallucinated voices in clinical and non-clinical voice-hearers
Corona Hernández, Hugo; Brederoo, Sanne; de Boer, Janna; Sommer, I. E. C.

Published in:
Schizophrenia Research

DOI:
10.1016/j.schres.2022.01.055

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
A data-driven linguistic characterization of hallucinated voices in clinical and non-clinical voice-hearers

H. Corona-Hernández a,*, S.G. Brederoo a, b, J.N. de Boer a, c, I.E.C. Sommer a, b

a Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, University of Groningen, the Netherlands
b Department of Psychiatry, University Medical Center Groningen, University of Groningen, the Netherlands
c Department of Biomedical Sciences of Cells, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Auditory verbal hallucinations
Clustering
Linguistics
Phenomenology
Schizophrenia
Voice-hearers

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.

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 (Laroi 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, epileptic subtype, and 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

* Corresponding author at: Department of Biomedical Sciences of Cells & Systems, University Medical Center Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, the Netherlands.
E-mail address: h.corona.hernandez@umcg.nl (H. Corona-Hernández).

https://doi.org/10.1016/j.schres.2022.01.055
Received 13 January 2021; Received in revised form 11 January 2022; Accepted 24 January 2022
Available online 9 February 2022
0920-9964/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
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; Stephe et al., 2003). Moreover, most studies on AVHs subtypes have focused exclusively on clinical AVHs (Chang et al., 2009; McCarthy-Jones et al., 2014b; Stephe 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 1 min, resulting in a total of at least 3 min 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 (1988) procedure 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

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Linguistic features for analysis in the AVHs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linguistic feature</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Attributive adjective</td>
<td>Word conveying properties of or adding features to a given noun, e.g., “pretty”</td>
</tr>
<tr>
<td>Definite article</td>
<td>Determiner for identifiable entities, e.g., “the”</td>
</tr>
<tr>
<td>Indefinite article</td>
<td>Determiner for non-identifiable entities, e.g., “a” or “an”</td>
</tr>
<tr>
<td>Locative adverb</td>
<td>Word for details of place or position, e.g., “homeward”</td>
</tr>
<tr>
<td>Mean length of utterance</td>
<td>The average number of words forming an utterance</td>
</tr>
<tr>
<td>Mean word length</td>
<td>The average number of letters forming a word</td>
</tr>
<tr>
<td>Moving-Average Type-Token Ratio</td>
<td>A ratio that expresses the number of different word forms relative to the total number of words</td>
</tr>
<tr>
<td>Nominal first-person singular pronoun</td>
<td>Word standing for the grammatical first person functioning as singular subject, i.e., “I”</td>
</tr>
<tr>
<td>Plural noun</td>
<td>Word standing for concrete or abstract entities, e.g., “tables”</td>
</tr>
<tr>
<td>Preposition</td>
<td>Word usually placed before a noun phrase and typically indicating spatial or temporal relations, e.g., “at” and “in”</td>
</tr>
<tr>
<td>Relative pronoun</td>
<td>Word that depends on an antecedent and that both introduces and plays a role in a new sentence, e.g., “who”</td>
</tr>
<tr>
<td>Simple future</td>
<td>Grammatical value of a later time, e.g., “will”</td>
</tr>
<tr>
<td>Simple past</td>
<td>Grammatical value of a previous time, e.g., “did”</td>
</tr>
<tr>
<td>Simple present</td>
<td>Grammatical value of a current time, e.g., “have”</td>
</tr>
<tr>
<td>Singular noun</td>
<td>Word standing for a concrete or abstract entity, e.g., “table”</td>
</tr>
<tr>
<td>Subordinating conjunction</td>
<td>Word linking an independent clause to a dependent clause, e.g., “because”</td>
</tr>
</tbody>
</table>
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).

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.

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).

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 Fig. 1.

Table 2

Demographic characteristics of participants.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Clinical voice-hearers (n = 21)</th>
<th>Non-clinical voice-hearers (n = 19)</th>
<th>Statistic value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>43.2 (11.38)</td>
<td>50.7 (16.19)</td>
<td>t(38) = 1.70</td>
<td>.09</td>
</tr>
<tr>
<td>Diagnosis, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>1 (5%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychosis not otherwise specified</td>
<td>5 (24%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>3 (14%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>12 (57%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females, n (%)</td>
<td>8 (38%)</td>
<td>8 (42%)</td>
<td>χ² (1) = 0.06</td>
<td>.79</td>
</tr>
<tr>
<td>Medication†, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>2 (11%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only antipsychotic</td>
<td>9 (50%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both antipsychotic and antidepressant</td>
<td>7 (39%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years hearing AVHs, mean (SD)</td>
<td>19.3 (15.2)</td>
<td>40.4 (18.56)</td>
<td>t(33) = 3.86</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

n = sample size, SD = standard deviation.

† Data available only for 18 clinical voice-hearers.

Table 3

Details of the AVHs’ shadows per group of voice-hearers.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Voice-hearers</th>
<th>Clinical (n = 21)</th>
<th>Non-clinical (n = 19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of AVHs, mean (SD)†</td>
<td>Voices occur at least once an hour</td>
<td>3.2 (0.89)</td>
<td>2.4 (1.28)</td>
</tr>
<tr>
<td>Number of different word forms (types)</td>
<td>Voices occur at least once a day</td>
<td>77.4 (45.21)</td>
<td>97.1 (53.56)</td>
</tr>
<tr>
<td>Mean (SD) Total</td>
<td></td>
<td>690</td>
<td>795</td>
</tr>
<tr>
<td>Mean (SD) Total</td>
<td></td>
<td>176.2 (122.96)</td>
<td>199.6 (139.14)</td>
</tr>
<tr>
<td>Recording time Mean</td>
<td></td>
<td>3701</td>
<td>3794</td>
</tr>
<tr>
<td>Recording time Total</td>
<td></td>
<td>5 m 47 s (2 m 54 s)</td>
<td>3 m 19 s (1 m 38 s)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 h 1 m 35 s</td>
<td>1 h 3 m 4 s</td>
</tr>
</tbody>
</table>

† This information was available only for 35 out of 40 participants.

According to the Psychotic Symptom Rating Scales (PSYRATS) (Haddock et al., 1999).
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 $\rho(\rho < 0.4$, all $p > .05$).

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 4

Characteristics of the participants per cluster of AVHs.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cluster with n = 18</th>
<th>Cluster with n = 22</th>
<th>Statistic value</th>
<th>p-Value</th>
<th>Cramer’s V effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years</td>
<td>47.2 (14.70)</td>
<td>46.5 (14.14)</td>
<td>t (38) = 0.14</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>Clinical voice-hearers</td>
<td>4 (22.22%)</td>
<td>17 (77.27%)</td>
<td>$\chi^2 (1) = 12.03$</td>
<td>&lt;.001</td>
<td>0.54</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bipolar disorder</td>
<td>1 (4.54%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Psychosis not otherwise specified</td>
<td>1 (5.5%)</td>
<td>4 (18.18%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Schizoaffective disorder</td>
<td>1 (5.5%)</td>
<td>2 (9.09%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Schizophrenia</td>
<td>2 (11.11%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medication$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>1 (5.5%)</td>
<td></td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Only antipsychotic</td>
<td>3 (16.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Both antipsychotic and antidepressant</td>
<td>7 (31.81%)</td>
<td>10 (45.45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Years hearing AVHs$^b$</td>
<td>33 (18.5)</td>
<td>26.3 (19.64)</td>
<td>t (33) = 1.02</td>
<td>.31</td>
</tr>
</tbody>
</table>

n = sample size, SD = standard deviation.

$^a$ Data available only for 18 clinical voice-hearers.

$^b$ Data available only for 18 clinical and 17 non-clinical voice-hearers.

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 indicate 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 (Klein, 1998), and the proportion of this type of constructions was smaller in compact-AVHs (see Supplementary Table SS). 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; Svenoniis, 2007). Importantly, prepositional constructions often specify information in connection with verbs (Svenoniis, 2007; Talmy, 2000). For instance, someone can either “live” in a house, “eat” at lunchtime, or “walk” toward a door. Indeed, a post-hoc exploratory evaluation of our data showed

### Table 5

Linguistic variables across the two clusters of AVHs.

<table>
<thead>
<tr>
<th>Linguistic variable</th>
<th>Expanded-AVHs cluster (n = 18), mean (SD)</th>
<th>Compact-AVHs cluster (n = 22), mean (SD)</th>
<th>Statistic value</th>
<th>p-Value</th>
<th>r effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns</td>
<td>165.1 (151.56)</td>
<td>156.3 (146.54)</td>
<td>0.03</td>
<td>.98</td>
<td></td>
</tr>
<tr>
<td>Simple present</td>
<td>96.9 (210.28)</td>
<td>143.6 (246.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple past</td>
<td>1623.5 (487.25)</td>
<td>1770.6 (447.15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple present</td>
<td>1305.3 (391.56)</td>
<td>1027 (204.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple past</td>
<td>1601.49 (236.70)</td>
<td>(272.15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simple present</td>
<td>397.5 (233.62)</td>
<td>290.5 (246.76)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n = sample size, SD = standard deviation.
Table 6
Phenomenological features of the two clusters of AVHs.

<table>
<thead>
<tr>
<th>PSYRATS item</th>
<th>Phenomenological feature of the AVHs</th>
<th>Expanded-AVHs cluster, description of median [median]</th>
<th>Compact-AVHs cluster, description of median [median]</th>
<th>Expanded-AVHs (n = 18), mean (SD) (^a)</th>
<th>Compact-AVHs (n = 22), mean (SD) (^b)</th>
<th>Statistic W</th>
<th>Adj. p-Value</th>
<th>r effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Frequency</td>
<td>Voices occur at least once an hour ([3])</td>
<td>Voices occur continuously or almost continuously ([4])</td>
<td>2.5 (1.23)</td>
<td>3.1 (1.04)</td>
<td>111.5</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Duration</td>
<td>Voices last for several minutes ([2])</td>
<td>Voices last for at least 1 h ([3])</td>
<td>2.2 (1.03)</td>
<td>2.8 (1.20)</td>
<td>110</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Location</td>
<td>Voices outside the head, but close to ears or head ([2])</td>
<td>Voices outside the head, but close to ears or head ([2])</td>
<td>1.9 (1.03)</td>
<td>2.1 (1.04)</td>
<td>134</td>
<td>.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Loudness</td>
<td>About same loudness as own voice ([2])</td>
<td>About same loudness as own voice ([2])</td>
<td>1.6 (0.70)</td>
<td>1.9 (0.87)</td>
<td>125</td>
<td>.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Origin</td>
<td>More than 50% but less than 100% of voices originate from external causes ([3])</td>
<td>Less than 50% of voices originate from external causes ([2])</td>
<td>2.7 (1.25)</td>
<td>2.6 (1.29)</td>
<td>162.5</td>
<td>.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Amount of negative content</td>
<td>No unpleasant content ([0])</td>
<td>More than 50% of voice content is unpleasant or negative ([3])</td>
<td>0.8 (1.36)</td>
<td>2.5 (1.47)</td>
<td>62</td>
<td>.01</td>
<td>−.52</td>
<td></td>
</tr>
<tr>
<td>7 Degree of negative content</td>
<td>Not unpleasant or negative ([0])</td>
<td>Personal verbal abuse relating to self-concept ([3])</td>
<td>0.8 (1.36)</td>
<td>2.6 (1.20)</td>
<td>55</td>
<td>.009</td>
<td>−.56</td>
<td></td>
</tr>
<tr>
<td>8 Amount of distress</td>
<td>Voices not distressing at all ([0])</td>
<td>More than 50% but less than 100% of distressing voices ([3])</td>
<td>1.1 (1.69)</td>
<td>2.8 (1.47)</td>
<td>73</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Intensity of distress</td>
<td>Voices not distressing at all ([0])</td>
<td>Voices are distressing to a moderate degree ([2])</td>
<td>0.9 (1.34)</td>
<td>2 (1.19)</td>
<td>80.5</td>
<td>.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Disruption to life</td>
<td>No disruption to life ([0])</td>
<td>Voices cause moderate amount of disruption to life ([2])</td>
<td>0.7 (1.25)</td>
<td>2 (1.28)</td>
<td>75.5</td>
<td>.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Control</td>
<td>Some control most of the time ([1])</td>
<td>Only occasional control ([3])</td>
<td>1.7 (1.72)</td>
<td>2.8 (1.34)</td>
<td>96</td>
<td>.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) p-Values were adjusted using Holm correction.
\(^b\) This information was available only for 17 of 18 participants.
\(^c\) This information was available only for 18 of 22 participants.

Fig. 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.

Examples:

**Example 1**

- Dutch: Ja het wordt toch niks.
- Dutch: Then you no longer have to look like that old woman who laid in that nursing home at that time.
- Translation: "Yes, it won’t work".

**Example 2**

- Dutch: Allemaal negatieve dingen.
- Dutch: "I should not say anything, but that child of theirs surely has a considerable influence on their relationship".
- Translation: "All negative things".

**Example 3**

- Dutch: Dan hoef je er niet meer uit te zien als die oude vrouw die toen in dat verpleeghuis lag.
- Dutch: “Then you no longer have to look like that old woman who laid in that nursing home at that time”.
- Translation: "I should not say anything, but that child of theirs surely has a considerable influence on their relationship".

**Example 4**

- Dutch: "Then you no longer have to look like that old woman who laid in that nursing home at that time".
- Dutch: "Then you no longer have to look like that old woman who laid in that nursing home at that time".
- Translation: "Then you no longer have to look like that old woman who laid in that nursing home at that time".
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; Laroi 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 dialogue (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 features like these, it can be hypothesized that they are triggered by 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; Laroi et al., 2019; Sommer and Diederan, 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). Cokal 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 (Cokal et al., 2018). Our results are surprising in this respect, as both definite and indefinite articles were less frequent in compact-AVHs 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 (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 (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 by a unique neurocognitive model, or whether different models are needed to explain 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., Constens 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., 2018).
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., Pilaga, 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 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 anti-psychotic medication can influence language production (de Boer et al., 2020; Salome 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 funding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CRediT authorship contribution statement

H. Corona-Hernandez: 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.

Declaration of competing interest

All authors declare that they have no conflicts of interest.

Acknowledgement

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).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2022.01.055.

References

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.


