The linguistic signature of hallucinated voice talk in schizophrenia

Antonia Tovar [a,1], Paola Fuentes-Claramonte [b,c,1], Joan Soler-Vidal [d], Nuria Ramiro-Sousa [e],
Alfonso Rodriguez-Martinez [f], Carmen Sarri-Closa [d], Salvador Sarró [b,c], Jesús Larrubia [g],
Helena Andrés-Bergareche [g], Maria Carmen Miguel-Cesma [g], Pedro Pablo Padilla [g], Raymond Salvador [b,c],
Edith Pomarol-Clotet [b,c], Wolfram Hinzen [a,b,h,*]

[a] Department of Translation and Language Sciences, Universitat Pompeu Fabra, C/ Roc Boronat, 138, 08018 Barcelona, Spain
[b] FIDMAG Germanes Hospitalàries Research Foundation, C/ Dr. Antoni Pujadas, 38, 08830 Sant Boi de Llobregat, Barcelona, Spain
[c] CIBERSAM, Centro de Investigación Biomédica en Red de Salud Mental, Spain
[d] Benito Menni Complex Assistencial de Salut Mental, C/ Dr. Antoni Pujadas, 38, 08830 Sant Boi de Llobregat, Barcelona, Spain
[e] Hospital Sant Rafael, Passeig de la Vall d'Hebron, 107, 08035 Barcelona, Spain
[f] Parc de Salut Mar, Passeig Marítim 25-29, 08003 Barcelona, Spain
[g] Centro Neuropsiquiátrico N. S. del Carmen, Camino del Abejar, 100, 50190 Zaragoza, Spain
[h] Catalan Institute for Advanced Studies and Research, ICREA, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

ARTICLE INFO

ABSTRACT

Very few studies have investigated the formal linguistic aspects of auditory verbal hallucinations (AVHs), though speech is a defining aspect of AVHs. Hallucinated speech heard by 19 patients with schizophrenia and highly frequent voices was obtained online, as and when they spoke, and annotated for pre-selected linguistic variables. Results showed that, consistently across the sample, (i) the grammatical first Person was significantly less represented than both second and third person, and often absent altogether; (ii) overwhelmingly, isolated clauses with no grammatical connectivity (parataxis) were produced, as compared with subordinations, coordinations, and adjunctions; (iii) in all participants except one, virtually no noun phrases (NPs) were anaphoric ones, back-referring to previous NPs, illustrating again a lack of connectivity across utterances. (vi) Sentence-level content was largely personal rather than impersonal, and in impersonal utterances, it was generally vague. (v) Formal syntactic errors were consistently nearly absent, as were semantic level errors such as paraphasias. Voice talk was not generally stereotyped. These results indicate that, despite a certain amount of individual variation, there is a distinctive linguistic profile to voice speech, which constrains theories of AVHs and their neurocognitive basis.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Hallucinations are one of the core symptoms of schizophrenia and by far the most common form they take is hearing voices (auditory-verbal hallucinations, AVH) (Andreasen and Flaum, 1991; Bleuler, 1914; Slade and Bentall, 1998). In some ways, a considerable amount is known about the clinical features of AVH: in different patients (and sometimes in the same patient), the voices can be single or multiple, heard inside or outside the head; they may range in complexity from single words to sentences to conversations, refer to the patient in both

the 2nd or 3rd person, and they have themes that are often but not always being derogatory (Jones, 2010; McCarthy-Jones et al., 2014; Nayani and David, 1996). In one important way, however, knowledge about AVH is limited. Namely, while language is a defining feature of the notion of AVH, little is known about the linguistic features they show. One study (Stephane et al., 2003) used hierarchical clustering and multi-dimensional scaling to show that linguistic complexity (voices heard talking in words, sentences, or discourses), as assessed through semi-structured interviews, is one dimension along which the voice hearing experience can differ. One other study (de Boer et al., 2016) compared the linguistic complexity of voice talk in psychotic and non-psychotic voice hearers as based on direct transcriptions, and found that the former group had a lower mean length of utterance and verb complexity. More systematic linguistic profiling of voice talk based on more fine-grained linguistic measures could contribute important constraints on neurocognitive models of AVHs. Inner speech models, in particular, regard the voice-hearing experience as arising

from a failure of properly self-monitoring one's own inner speech (Frith, 1992; Frith and Done, 1989; Jones and Fernyhough, 2007). Leudar et al. (1997) already argued that the pragmatics of voice talk is consistent with this model. However, the degree of phenomenological fit between inner speech and AVH has remained a complex and open issue (Jones, 2010; Rosen et al., 2018). Insights at other levels of linguistic organization can provide further constraints to evaluate such models.

Examination of the linguistic features of voice talk is of interest also in the wider context of schizophrenia, where linguistic abnormalities have been documented to also occur in the patients' expressed speech and comprehension, particularly in patients with formal thought disorder (FTD, speech that is difficult to follow, sometimes to the point of complete incomprehensibility) (Chaika, 1974; Covington et al., 2005; Kuperberg, 2010; McKenna and Oh, 2005). A further well-replicated finding in thought-disordered schizophrenic speech is presence of unclear reference, particularly in the poorly specified use of personal pronouns such as *he, she, they*, etc., to refer listeners back to previous aspects of discourse (Barch and Berenbaum, 1996; Docherty et al., 2003; Rochester and Martin, 1979). There are certainly hints of some such linguistic abnormalities in one of the very few verbatim accounts of AVHs, from a patient reported by Kraepelin (1913). His voices consisted of a succession of short sentences, probably from different speakers, some of which were nonsensical ('We inhale you'; 'Because we have to fear your brain grease'). Linguistic abnormality has also been documented in spontaneous speech of patients with schizophrenia without FTD, where it takes the form of syntactic simplification and errors (Hoffman and Sledge, 1988; Morice and Ingram, 1982; Oh et al., 2002; Thomas et al., 1990).

The aim of the present study was to provide a more systematic linguistic profiling of voice speech in patients with a diagnosis of schizophrenia. Our primary research question was which linguistic variables would characterize this profile and which would not. We also attempted to corroborate the existence of some of the features of AVH suggested by the existing clinical and research literature. Specifically, given that voices are commonly heard so as to speak in the 2nd (2P, e.g. 'you' or 2P agreement on the verb), and 3rd Person (3P, e.g. 'he/it/ the God/a man'), we examined the distribution of grammatical person. Since voices are also commonly reported to take the form of brief phrases and this was a feature in the example of Kraepelin as well as the study of de Boer et al. (2016), we also examined grammatical connectivity between clauses. This specifically included the proportion of clauses that were embedded within other clauses, and how they were embedded, namely as adjuncts (e.g. *You are bad because you killed your mother*), as clausal arguments (e.g. *I think you are bad*), as coordinated (*You are bad and you killed your mother*), or without any grammatical connectivity at all (parataxis). The different types of noun phrases (NPs) produced, i.e. configurations like *the devil, your body, she*, etc. can also reflect degrees of connectivity between subsequent utterances: e.g., correct use of *anaphoric* NPs such as *the man* or *he* depends on the referent of these NPs having been mentioned previously, thus reflecting a narrative connection. Voices are also usually reported as being personal, i.e. directed at the patient, often in a negative way, so we also examined the frequency with which they were personal as opposed to being impersonal in content, in the sense that they reflect states of the world that obtain irrespective of the speaker's relation to them and his or her mental states.

Our specific hypotheses based on clinical impression and previous phenomenological reports were that (i) use of the 1st Person (1P) singular and plural (e.g. Spanish equivalents of 'I', 'my', or 1P agreement on the verb) would be uncommon, (ii) that clauses would exhibit little grammatical connectivity, with a preponderance of parataxis, (iii) that the proportion of anaphoric NPs would be low, and (iv) that sentence-level content would be largely personal.

Finally, in a more exploratory fashion, we investigated the pattern of formal syntactic errors in voice speech, since normal expressive language in schizophrenia has been found to feature an increased number of such errors in production (Marini et al., 2008; Morice and Ingram, 1982; Morice and McNicol, 1986; Tavano et al., 2008; Thomas et al., 1990), while in language perception, patients with schizophrenia show less sensitivity to syntactic errors as compared with neurotypical controls (Moro et al., 2015). Moreover, since single-sentence semantic-level anomalies characterized the language of formal thought disorder in a previous study (Oh et al., 2002), we explored whether such anomalies, along with other features of thought disorder such as paraphasias and neologisms, would also be found in voice speech.

2. Methods

2.1. Participants

The patient sample consisted of 19 patients with a diagnosis of schizophrenia or schizoaffective disorder, recruited from five psychiatric hospitals in Barcelona (Benito Menni CASM, Hospital Sagrat Cor de Martorell, Hospital Sant Rafael, Parc de Salut Mar) and Zaragoza (Centro Neuropsiquiátrico N.S. del Carmen), Spain. Results are reported from 18 patients, since inspection of one transcript revealed that only 9 words were produced by this patient's voice, grouped into three sentences separated by long pauses to which our analysis scheme could not be applied, after which the patient was removed prior to analysis. All participants met DSM-IV-TR criteria for schizophrenia or schizoaffective disorder, based on review of their clinical history by the patient's psychiatrist and a member of the research team. They were excluded if they (a) were younger than 18 or older than 65, (b) had a history of alcohol or substance abuse/dependence in the last year, (c) had a history of head injury, neurological disorder or medical disorders affecting cognition, (d) had hearing loss, and (f) had had treatment with electroconvulsive therapy in the last six months. All patients were also required to have a current IQ in the normal range (i.e. >70). Current IQ was prorated from 4 subtests from the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997): Vocabulary, Similarities, Matrix reasoning and Block design. The patients were all on antipsychotic treatment (typical [$n =$ 1], atypical [$n = 7$], both kinds [$n = 10$], missing data for one patient). All participants gave written informed consent. All the study procedures were approved by the local research ethics committee and adhered to the Declaration of Helsinki.

2.2. Clinical assessment

The patients were administered the PANSS (Kay et al., 1987) to assess positive and negative psychotic symptoms. AVHs were rated with the Psychotic Symptom Rating Scales (PSYRATS, Haddock et al., 1999), auditory hallucinations subscale (AHS). This subscale consists of a semi-structured interview with 11 items referring to frequency, duration, controllability, loudness, location; severity and intensity of distress; amount and degree of negative content; beliefs about the origin of voices; and disruption caused by the AVHs (see Supplementary Table 1 for a list of the qualitative aspects of AVH in the present sample as obtained from the PSYRATS). To be included in the study, patients were required to score at least 'once an hour' in the 'frequency' item of the PSYRATS-AHS. This was done to ensure that the patients would experience voice hearing during the time of the assessment. Voice frequency was further examined by asking the patients to signal (tap on the table) each time they experienced an instance of AVHs during a period of 5 min in a quiet environment.

2.3. Procedure

Patients were placed in a quiet environment and asked to repeat verbatim everything their voices said over a 5 to 25-minute period (mean 11 min and 45 s, one patient reported to stop hearing voices after 2 min and therefore was only recorded for that period). Patients were

recorded with the built-in microphone in a Dell laptop using in-house developed software written in TCL.

2.4. Transcription and annotation

Transcriptions were strictly literal and hence included all repetitions, unintelligible speech (which was marked as such), and indications of speech pauses, which were indicated without time specifications as these had no impact in the linguistic variables. Uncertain words and other non-linguistic sounds and screams were also reported, but not analyzed. Relevant non-verbal aspects of communication were specified, such as tone. Other significant data were reported, such as when the patient started to sing or laughed. Linguistic variables were manually annotated in CLAN (MacWhinney, 2000) by a linguist rater (AT), who

Table 1
Linguistic variables, definitions, and examples.

Variable	Definition	Example
1st Person NP	NP in grammatical 1st Person	*I* don't know.
2nd Person NP	NP in grammatical 2nd Person	*You* will fail.
3rd Person NP	NP in grammatical 3rd Person	*He* didn't say *that*.
Subordination	One clause is embedded in another and dependent on it	You want to hurt people *that are around you*.
Coordination	Two clauses are coordinated when both are at same level.	Why are you talking in Spanish *and* why are you saying this?
Adjuncts	Dependent clauses specifying further information that is not grammatically required.	This has happened *because you are a junkie*.
Parataxis	Utterances not connected grammatically to others.	Ceremony. Don't be afraid.
Anaphoric NP	NPs picking up the referent of a previous NP.	*He* has not repeated *this*.
Personal content	The content of the utterances is directly referring to speakers/hearers or objects/events present in the speech context.	*I* love *you*.
Impersonal content	Sentences stating facts of the world without reference to speaker/hearers.	Spain is a democracy.
Vagueness of content (only annotated in impersonal contents)	The fact referenced lacks specificity or is overly generic.	It is a mortal sin. Respiration is the method.
Impersonal content non vague	A specific fact of the world is referenced informatively.	Here one can speak.
Formal grammatical error	Violations of grammatical well-formedness conditions.	And you pay her in this way [laísmo]. (Y la pagaste de esa manera.)
Paraphasia	Word choice for a given referent where another word is expected.	Red Riding Hood walked into the *park* [instead of forest].
Violations of semantic selectional restrictions	Conceptually impossible combination of lexical features	The *pond fell* in the front doorway.
Neologisms	Creation of a new word	*Noises* [non-existent word in Spanish in a context where Moses, the biblical figure, appears to have been intended].
Clanging	Association of words based on similar sounds or inside the same lexical field	Tú eres un culturista. El niño no tiene cultura [lit.: You are a bodybuilder. The child has no culture]

had previously not been involved in the study, including in the transcriptions (performed by PF-C).

In the linguistic analysis we annotated, for every single NP, which grammatical Person (1P, 2P or 3P, subdivided into singular and plural) was grammatically specified. In the domain of syntactic complexity, four modes of grammatically connecting clauses were distinguished: subordination (embedding), coordination, adjuncts and parataxis, where the last of these reflects the absence of a grammatical connection. NPs were further annotated for whether they were anaphoric or non-anaphoric, where anaphoricity means that the NPs refer to an object previously mentioned by another NP. In terms of the sentence-level content of utterances, we distinguished between personal and impersonal content, where personal means that speech participants or other objects of the immediate context are the subject of the utterance, while impersonal content concerns facts about the world relatively independent of the speech context. Formal syntactic errors, i.e. violations of grammatical well-formedness conditions, were also counted. Finally, semantic-level anomalies were annotated according to four variables, capturing paraphasias, violations of semantic selectional restrictions, neologisms and clanging. Definitions and examples (in English translation) are presented in Table 1.

2.5. Statistical analysis

All absolute counts of grammatical Persons in singular and plural were normalized by the total quantity of noun phrases (NPs) produced. This was done both for total occurrences of each of the three grammatical Persons, and separately for their respective occurrences (i) in isolation and (ii) as occurring jointly with other grammatical Persons within the same sentence (e.g., 'I love you', where 1P co-occurs with 2P, or 'I like him', where 1P co-occurs with 3P). We proceeded in the analogous way for the four different modes of clausal connectivity, which we normalized relative to the total utterances produced; and with the relative proportions of anaphoric vs. non-anaphoric NPs, of impersonal vs. personal sentence-level content, of utterances with formal syntactic errors, and with semantic-level anomalies. Results are stated descriptively as averaged percentages. In addition, and wherever appropriate, related-samples analyses of variance (Friedman's related samples two-way analysis of variance by ranks) were carried out to test for significant deviance from equal distributions of instances of the same variable under different conditions, specifically the three grammatical Persons in the case of grammatical Person, and the four modes of clausal connectivity. The α-level was set at 0.05. Multiple comparisons were corrected for by multiplying the observed *p*-value from the significance tests by the number of tests, *k*. Then if any *k* P is <0.05, the test is significant at the 0.05 level. Only significant comparisons are reported.

3. Results

3.1. Demographics and clinical information

Demographic and clinical information for the participants is shown in Table 2. Subjective AVH frequency, as indicated by asking the patients to tap each time they experienced a voice during a 5-minute period, ranged from 1 to 360. The mean was 51.27 (SD = 98.39), median = 13.50 (Interquartile range, IQR = 33.5).

3.2. Linguistic variables

3.2.1. Differences in the use of grammatical persons

Use of 1P in isolation, in both singular (S) and plural (P), was rare across the group and completely absent in 12 (singular) and 13 (plural) out of 18 cases (Table 3). A Friedman test (related samples two-way analysis of variance by ranks) confirmed statistically significant differences between total uses of the three Persons ($\chi^2 = 19.681$, $p < .001$). Post hoc pairwise comparisons corrected for multiple comparisons

Table 2
Demographic, clinical and neuropsychological variables.

	Mean (SD)
Sex (M/F)	15/4
Age (years)	43.89 (7.71)
Illness duration (years)	19.58 (9.81)
PANSS-total	77.39 (16.09)
PANSS-positive	20.00 (5.13)
PANSS-negative	22.22 (6.72)
PANSS-general	35.33 (10.41)
PSYRATS-AHS	27.83 (8.04)
Estimated IQ	93.59 (11.50)
Current antipsychotic dose (CPZ equivalent, mg)	751.99 (526.35)

Table 4
Median percentages of uses four grammatical connectivity types of clauses.

Clausal connectivity	Subordination	Adjuncts	Coordination	Parataxis
%	4,0	5,5	3,9	86,7
Total	47	57	37	566
Median	1	1	1	18
Mean	2,47	3,00	1,95	29,79
IQR	3	3	2,5	25,5
Min.	0	0	0	0
Max.	12	15	12	108
% subjects with 0 instances	36,84	36,84	42,11	5,26

showed significant differences between total use of 1P and both 2P ($p < .001$), and 3P ($p = .001$). The same test confirmed that distributions in the combinations of different Persons was not equal either ($\chi^2 = 11.019$, $p < .004$). Pairwise comparisons revealed significant differences between combinations of 1w2 (1P with 2P) and of 2w3 (2P with 3P) ($\chi^2 = 11.019$, $p < .014$), with the latter significantly more frequent than the former (all significance values Bonferroni corrected). For examples see Supplementary materials.

3.2.2. Modes of clausal connectivity

Median percentages of uses of grammatical connectivity types summarized in Table 4 show that parataxis was by far the most frequent connection type (for examples see Supplementary material). A Friedman test confirmed significant differences between distributions of different types of connectivity ($\chi^2 = 38.92$, $p < .001$). Post hoc pairwise comparisons revealed significant differences in the comparison of ratios of parataxis with coordination ($p < .03$), with subordination ($p < .000$), and with adjuncts ($p < .001$).

3.2.3. Ratio of anaphoric vs. non-anaphoric NPs

Inspections of raw counts of anaphoric NPs revealed one outlier who produced 26 out of a total of 32 instances of anaphoric NPs occurring in this corpus. Of the remaining 17 participants, the voices of 13 had no anaphoric NPs, 2 participants had 1, and 2 had 2. The median percentage of anaphoric NPs of our total NPs was 0% (mean 0,6%).

3.2.4. Impersonal vs. personal content

Median percentages of Personal and Impersonal content out of total utterances were 92,3% and 7,7%, respectively. A Wilcoxon signed-rank test showed a statistically significant difference between utterances with personal vs. non-personal utterance content, $z = 3.44$, $p < .001$. Impersonal content turned out to be often vague upon inspection (e.g. *la respiración es el método* [the respiration is the method]), and only 5 utterances in the corpus (out of a total 72 impersonal utterances) were classed as impersonal-non-vague (see Supplementary materials).

3.2.5. Formal syntactic errors

An extremely low number of utterances with formal syntactic errors was found, corresponding to a raw total count of 6 in the entire corpus.

3.2.6. Single-sentence semantic-level anomalies

Very few instances of such anomalies were found: 0 cases of paraphasias, 1 semantic selectional restrictions (*este niño ni lleva la cultura* [this child does not even carry the culture]), 1 neologism (*Moises*) and 1 possible instance of clanging (*sabes sabueso*).

4. Discussion

Results confirmed our hypothesis that 1P would be significantly under-represented in comparison to the other two grammatical persons, which between themselves did not significantly differ. This difference is seen (i) when total uses of the three Persons are compared, (ii) in the case of combinations of 1P and 2P, which were significantly less frequent than combinations of 2P and 3P, and (iii) in the case of 1P singular and plural as used in isolation, which were absent altogether in a majority of participants. In such isolated uses, the NPs in question are used in simple sentences in which only this single NP occurs, which allows assessing uses of grammatical Persons controlling for the effects of the co-presence of other Persons. This is also relevant since use of 1P in combination, where it occurred, often did so in such a way that it was still directly related to the voice hearer as addressed in 2P, as e.g. in *te quiero* (I love you), without much additional personal information provided about the 1P referent, i.e. the voice itself. In short, 1P can be 'carried along' by a discourse largely centered around 2P and could in this sense be in part a by-product of this fact and the use of a verb with two nominal arguments.

On the other hand, 2P plural and 3P plural in isolation were also very rare or absent, raising the question of why the absence of 1P in isolation should be special or be of any theoretical interest. In the case of 2P plural, a natural explanation is the nature of the speech situation itself: there is only *one* voice hearer to whom the voice speech is directed, and while, in theory, 2P plural could also be used to address several other voices, this apparently simply never happened in sentences figuring a single NP. Such a natural explanation is not available, on the other hand, for the case of the absence of 1P in isolation, since there always *is* at least one voice talking. Why then should it be that this voice, which is normally heard as personified, almost never says 'I' in simple sentences with only one NP, and does so much more rarely also when NPs occur in combination with others?

Table 3
Percentages (Medians) of grammatical Persons (1P/2P/3P: totals; 1/2/3 PS: 1/2/3 Person Singular in isolation; 1/2/3PP: 1/2/3 Person Plural in isolation; w: one Person in combination with another in the same utterance).

Person	1P	2P	3P	1PS	2PS	3PS	1PP	2PP	3PP	1w2	1w3	2w3
%	12,6	48,4	44,3	0,0	19,4	8,8	0,0	0,0	0,5	0,0	5,9	16,1
Total	110	381	391	16	174	107	8	2	31	19	67	186
Median	2	13	12	0	7	3	0	0	0	0	1	5
Mean	5,79	20,05	20,58	0,84	9,16	5,63	0,42	0,11	1,63	1,00	3,53	9,79
IQR	5	23,5	20	1	13,5	4,5	0,5	0	1,5	1	3	11
Min.	0	2	2	0	0	0	0	0	0	0	0	0
Max.	45	72	99	10	30	24	4	2	13	5	31	48
% subjects with 0 instances	36,84	0	0	68,42	5,26	15,79	73,68	94,74	52,63	57,89	47,37	15,79

The answer would be immediate if voice talk was largely the impersonal one of a news reader on TV, say, who will barely if ever use 1P except in direct quotations. The observation is startling in the present case, however, where speech content precisely is largely personal. Voice speech is not merely largely personal, moreover, but often takes place in a quasi-conversational setting, in which the voice hearer is, for the most part, directly addressed in 2P by the voice, or else is the 3P topic of a conversation of one voice with another. There seems to be no reason at all, in such a setting, why 2P and 3P should be nearly four times as frequent in sentences with two nominal arguments and be used near-exclusively in simple sentences. A tendency to under-use 1P by a speaker in a personal conversational setting would mean that this speaker does not tend to talk about himself (engages in self-reference): he is not *present* in the conversation as a first-person self. Insofar as the voice hearer also does not reply, there is thus little or no 1P reference in this conversational setting at all. This is even more remarkable in light of the fact that numerous elements of such speech, such as swearing, are considered 'expressive' and expressives are inherently related to a speaker's self (Potts, 2007). What transpires, then, is that selfhood at this deeper, emotional level dissociates in AVH from the kind of 'conversational' or thinking self that is referenced in English with the 1P pronoun.[2]

As for the rarity of 3P plural in isolation, this suggests that in sentences featuring a single NP, no pluralities of things or persons were ever referred to: the voices stuck to single, individual referents. While this may reflect a form of 'concretism' in voice speech that is worth noting, our finding about 1P again does not have such a straightforward answer. As noted, the absence or rarity of 1P in a personal, conversational setting could not simply be an accident, and a form of concretism would certainly not predict it either. One could further speculate that this finding might be due to the tendency of the voices to often engage in insults, threats, or praise (which itself, in order to be explanatory as a hypothesis, would then have to have some other explanation, unrelated to the poverty of the language and the distribution of grammatical persons used). However, while voice talk is indeed frequently marked by such speech act types, it is by no means the case that this type of speech necessarily entails the absence of 1P: *I hate you*, *I will kill you*, *Don't piss me off*, *I love you*, *Listen to me*, *Don't talk to me*, *You are worse than my dog*, *I am the Master*, *I am fed up*, etc., are all sentences in the insulting/praising/commanding mode, and they all contain 1P. Moreover, while this argument might make us expect more 2P use than 1P use, it would not explain why 1P in total was underused compared with both 2P and 3P, and that this also applied to combinations of 1P with 2P relative to combinations of 2P and 3P.

While we thus have no convincing explanation to offer for our finding, it is worth highlighting it *as* an explanandum for any theory of AVH, which we would characterize descriptively as follows. Person distinctions are *deictic* distinctions: they regulate whether a person referenced as taking part in some event is identical with a speech participant or not: with the speaker in the case of 1P, and with the hearer in the case of 2P, or with neither of them in the case of 3P. The normal conversational speech situation is that there is a speaker identifying himself in 1P, i.e. the 'center' or 'origo' (Buehler, 1934) of the deictic space in which speech takes place. This person speaks to a second person (a 'you', the 2P) about the world (the 'it/he/she', i.e. the 3P). Voice talk in this sense thus effectively lacks a deictic center – it is deictically decentered. Theories of AVHs need to explain this deictic shift, which is in line with theoretical approaches pointing to a disruption of deictic anchoring of thought in schizophrenia potentially not specific to AVH (Crow, 2010; Hinzen et al., 2016). The result is also significant with regards to studies of voice hearing in non-clinical populations. It is noteworthy that all examples given in de Boer et al. (2016) of voice speech in which 1P occurred, were from the non-clinical sample.

Results on grammatical connectivity of clauses also confirmed our predictions. Clauses are the minimal structural configurations that are meaningful at a propositional level. Connections between clauses in normal speech are reflected in the kind of grammatical devices we distinguished here: subordinations, coordinations, and adjuncts, all of which are crucial to narrative, dialogue, and discourse. A restricted kind of lack of connectivity occurs in specific rhetorical contexts (e.g. *he came*, *he saw*, *he won*), or when a thematic connection makes connectors superfluous (e.g. *Inspector Clouseau arrived in the evening. The house was dark. He rang the bell.*). But cases annotated as 'parataxis' in our study were not of these kinds: they either exhibit no grammatical connectivity at all or else exhibit only semantic connectivity in the sense of being broadly subordinated under a given broad lexical-thematic field (e.g. *killing*, see example 5). Effectively, then, voice speech in our sample is language with a strong tendency to be reduced to the single-sentence level, lacking connectivity and embedding. This was further confirmed by the virtual absence of anaphoric NPs – and ipso facto the absence of referential connections across utterances that this entails.

In line with results from phenomenological studies, linguistic annotations based on analyzing referencing revealed that a median 92,3% of utterances were personal-level utterances, and that the vast majority of the remaining non-personal utterances were defective in the sense that their content was not clearly identifiable. This shows that where voice speech deviated from its pattern of operating at the personal level, it was rarely the case that language strayed away from the speech context and lived up to its normal function and potential, namely capturing informative content about the world in objective, impersonal terms. It is inherent to language in its normal use in mental health that it can convey thoughts which, while expressing the mental states (beliefs, desires) of the speaker, capture states of the world that hold independently of that speaker (e.g. *Gold is yellow*, *John is married*).

Finally, with regards to syntactic and semantic errors, our hypothesis concerning syntactic errors was not confirmed. Voice speech was virtually error-free in either the syntactic or the semantic terms we distinguished, demonstrating that, within schizophrenia symptomatology, it is its own register that has a distinctive profile. We speculate that the low grammatical complexity of such speech may act as a kind of protector of its syntax, since if speech lacks complexity, syntactic errors are less likely to occur. As for semantic-level distinctions, lexical and phrasal selection are clearly not what makes such speech deviant.

Summarizing, what emerged in this study as distinctive of voice speech are: (i) the shift in the deictic space in which speech takes place, away from 1P to non-1P, while at the time being largely personal; (ii) the effective lack of grammatical connectivity, with speech reducing to the single-clause level without embedding, and (iii) reduction to the personal level language which does not reach informative levels depending on the presence of objective (impersonal) meaning.

These results constrain theories of AVHs. In particular, on the still widely maintained 'inner speech' theory of AVHs going back to Frith (1992), voice talk is inner speech misattributed to an external source (Jones and Fernyhough, 2007). As noted the degree of 'fit' of this model with phenomenological features of AVHs has remained complex and contentious (Jones, 2010; Rosen et al., 2018). But it seems difficult to fit this model to almost any of the linguistic features identified here. It is important in this regard to distinguish inner speech from a different linguistic genre, namely self-talk (Holmberg, 2011), which tends to involve a person referring to herself in 2P (e.g. *You idiot!*, *Now pull yourself together!*, *You messed this up again!*). By contrast, inner speech or ruminations will often, if not mostly, be 1P (*I really don't like this*, *I still need to do the shopping*, *Damn, my secretary is still waiting*, *Why does he stare at me like this*, etc.), though they can also be impersonal (*They'll never repair this bridge*, *Joe has really gotten old*, etc.). Voice speech in our sample did not have the former feature, nor were impersonal utterances qualitatively of the same kind. The 'perceptual' theory (Mørch-Johnsen et al., 2017), according to which AVHs arise from abnormal activations of

---

[2] We owe this point to conversations with Martina Wiltschko.

temporal cortex involved in language perception, predicts the specificity of the patterns we have found even less.

These failures could be rescued by more specific such models. Thus, one particular conception of inner speech regards it as based on internalized interpersonal dialogue as a raw material (Fernyhough, 2004). This would still not predict the low proportions of grammatical 1P, whether referring to the voice hearer or to other persons or voices featuring in such dialogue. This conception also proposes that external dialogical speech, when internalized, undergoes important transformations leading to syntactic and semantic changes, particularly a form of 'syntactic abbreviation' involving the dropping of subjects in favor of predicates, and to fragments of verbal images rather than fully formed utterances, until inner speech ultimately 'loses most of the acoustic and structural qualities of external speech' (Fernyhough, 2004: p.55). Yet again, none of these features were observed in the AVH studied here, where utterances were mostly fully formed, very little ungrammaticality was found, and virtually no anaphoric NPs were present (which dropped subjects in normal speech would typically be, corresponding to 'old' information); moreover, despite voices present identified as people talking to or about the voice hearer continuously, the dialogical or conversational involvement of the voice hearer was very rare in our sample; and it is doubtful whether a voice's running commentary on what the voice hearer is thinking, or a series of instructions or commands that the voice hearer is following, can count as 'dialogical'. Ultimately, the question of whether inner speech exhibits a similar linguistic profile as hallucinated voice talk, whether with regard to grammatical Person or grammatical connectivity, could only be answered by a direct comparison using the same method of annotation in order to obtain such a profile. But this is hardly possible since, by definition, inner speech is not hallucinated voice talk, i.e. AVH, and hence is not presented and transcribable in the same way, though other methods are available (Alderson-Day and Fernyhough, 2015; Hurlburt et al., 2017).

This illustrates that and how linguistic analysis of AVH can provide an objective basis against which theories of AVHs can be evaluated and neurolinguistic correlates of such distorted speech can be explored. Specific linguistic features of voice talk documented here may also provide important clues for their neural correlates. This in particular applies to the processing of deictic distinctions (grammatical Persons), which has been explored with fMRI in autism spectrum conditions as compared with neurotypical individuals (Mizuno et al., 2011), but has not yet been a focus of research in AVH.

Our results suggest further studies integrating language as a dimension of AVH that is not captured by traditional phenomenological descriptions and methods, and exploring links between linguistic analysis and phenomenological data as obtained from interviews. de Boer et al. (2016) already explored these links but did not find a correlation between a sum score of syntactic complexity in AVH and the amount of negative emotional content. Nor is it clear conceptually why such a link should obtain, i.e. which implications such content would have for structural linguistic complexity; or why it should obtain between other features of voices, such as whether they are heard inside or outside of the head, whether their loudness is like that of the own voice, or what beliefs are held about them. In line with this, post hoc inspection of qualitative aspects of voices as captured by the PSYRATS (Supplementary Table 1), did not suggest any tendencies or correlations, though small numbers prevent formal correlational analyses. This may indicate that the linguistics of the voices is a dimension of AVH at least partially independent of their non-linguistic phenomenological dimensions.

5. Conclusion

This study reveals that AVHs have a less explored linguistic dimension that is subject to formal analysis and shows a distinctive profile, which may illuminate the nature of voice hearing as such and informs

neurocognitive models. Specifically, while voice talk was generally neither stereotyped nor ungrammatical, grammatical complexity was seen to virtually reduce to clauses and NPs with no grammatical connectivity to other clauses or NPs, respectively. The use of grammatical Person showed a striking pattern, in which the grammatical first Person was significantly less represented than both the second and third persons and often absent altogether, whether referring to the voice hearer or another voice heard. As Person distinctions are deictic distinctions, this may suggest a disruption in the deictic anchoring that is a necessary requirement whenever thought or speech is generated, linking thought content to a person thinking it.

CRediT authorship contribution statement

Antonia Tovar: Conceptualization, Formal analysis, Writing - original draft. **Paola Fuentes-Claramonte:** Conceptualization, Data curation, Formal analysis, Writing - original draft. **Joan Soler-Vidal:** Data curation. **Nuria Ramiro-Sousa:** . **Alfonso Rodriguez-Martinez:** Data curation. **Carmen Sarri-Closa:** Data curation. **Salvador Sarró:** Data curation. **Jesús Larrubia:** Data curation. **Helena Andrés-Bergareche:** . **Maria Carmen Miguel-Cesma:** Data curation. **Pedro Pablo Padilla:** Data curation. **Raymond Salvador:** Formal analysis, Writing - review & editing. **Edith Pomarol-Clotet:** Conceptualization, Data curation, Funding acquisition, Supervision. **Wolfram Hinzen:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing - original draft.

Role of the funding source

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

Conflict of interest

All authors declare that they have no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2018.12.004.

References

Alderson-Day, B., Fernyhough, C., 2015. Relations among questionnaire and experience sampling measures of inner speech: a smartphone app study. Front. Psychol. 6, 1–8. https://doi.org/10.3389/fpsyg.2015.00517.
Andreasen, N.C., Flaum, M., 1991. Schizophrenia: the characteristic symptoms. Schizophr. Bull. 17, 27–49. https://doi.org/10.1093/schbul/17.1.27.
Barch, D.M., Berenbaum, H., 1996. Language production and thought disorder in schizophrenia. J. Abnorm. Psychol. 105, 81–88. https://doi.org/10.1037/0021-843X.105.1.81.
Bleuler, E., 1914. Die Kritiken der Schizophrenien. Z. Ges. Neurol. Psychiatr. 22, 19–44.
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. https://doi.org/10.1016/j.bandl.2016.07.011.
Buehler, K., 1934. Sprachtheorie. Die Darstellungsfunktion der Sprache. Verlag von Gustav Fischer, Jena.
Chaika, E., 1974. A linguist looks at "schizophrenic" language. Brain Lang. 1, 257–276.
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. Schizophr. Res. 77, 85–98. https://doi.org/10.1016/j.schres.2005.01.016.
Crow, T.J., 2010. The nuclear symptoms of schizophrenia reveal the four quadrant structure of language and its deictic frame. J. Neurolinguistics 23, 1–9. https://doi.org/10.1016/j.jneuroling.2009.08.005.

Docherty, N.M., Cohen, A.S., Nienow, T.M., Dinzeo, T.J., Dangelmaier, R.E., 2003. Stability of formal thought disorder and referential communication disturbances in schizophrenia. J. Abnorm. Psychol. 112, 469–475. https://doi.org/10.1037/0021-843X.112.3.469.

Fernyhough, C., 2004. Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. New Ideas Psychol. 22, 49–68. https://doi.org/10.1016/j.newideapsych.2004.09.001.

Frith, C.D., 1992. The Cognitive Neuropsychology of Schizophrenia. Lawrence Erlbaum Associates, Hove.

Frith, C.D., Done, D.J., 1989. Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. Psychol. Med. https://doi.org/10.1017/S003329170001240X.

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. https://doi.org/10.1017/S0033291799008661.

Hinzen, W., Rosselló, J., McKenna, P., 2016. Can delusions be understood linguistically? Cogn. Neuropsychiatry 21, 281–299. https://doi.org/10.1080/13546805.2016.1190703.

Hoffman, R.E., Sledge, W., 1988. An analysis of grammatical deviance occurring in spontaneous schizophrenic speech. J. Neurolinguistics https://doi.org/10.1016/0911-6044(88)90008-5.

Holmberg, A., de Vries, M., 2011. Referring to yourself in self-talk. In: Zwart, J. (Ed.), Structure Preserved: Studies in Syntax for Jan Koster. John Benjamins, Amsterdam, pp. 185–192.

Hurlburt, R.T., Alderson-Day, B., Fernyhough, C., Kühn, S., 2017. Can inner experience be apprehended in high fidelity? Examining brain activation and experience from multiple perspectives. Front. Psychol. 8, 1–8. https://doi.org/10.3389/fpsyg.2017.00043.

Jones, S.R., 2010. Do we need multiple models of auditory verbal hallucinations? Examining the phenomenological fit of cognitive and neurological models. Schizophr. Bull. https://doi.org/10.1093/schbul/sbn129.

Jones, S.R., Fernyhough, C., 2007. Neural correlates of inner speech and auditory verbal hallucinations: a critical review and theoretical integration. Clin. Psychol. Rev. 27, 140–154. https://doi.org/10.1016/j.cpr.2006.10.001.

Kay, S.R., Flszbein, A., Opfer, L.A., 1987. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr. Bull. 13 (2), 261–276.

Kraepelin, E., 1913. Dementia Praecox and Paraphrenia. Livingstone, Edinburgh.

Kuperberg, G.R., 2010. Language in schizophrenia part 1: an introduction. Lang. Linguist. Compass 4, 576–5898. https://doi.org/10.1111/j.1749-818X.2010.00217.x.Language.

Leudar, I., Thomas, P., McNally, D., Glinski, A., 1997. What voices can do with words: the pragmatic of verbal hallucinations. Psychol. Med. 27, 885.

MacWhinney, B., 2000. The CHILDES Project: Tools for Analyzing Talk. 3rd edition. Lawrence Erlbaum Associates, Mahwah, NJ.

Marini, A., Spoletini, I., Rubino, I.A., Ciuffa, M., Bria, P., Martinotti, G., Banfi, G., Boccascino, R., Strom, P., Siracusano, A., Caltagirone, C., Spalletta, G., 2008. The language of schizophrenia: an analysis of micro and macrolinguistic abilities and their neuropsychological correlates. Schizophr. Res. 105, 144–155. https://doi.org/10.1016/j.schres.2008.07.011.

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. Schizophr. Bull. https://doi.org/10.1093/schbul/sbs156.

McKenna, P.J., Oh, T.M., 2005. Schizophrenic Speech: Making Sense of Bathroots and Ponds That Fall in Doorways. Cambridge University Press, New York.

Mizuno, A., Liu, Y., Williams, D.L., Keller, T.A., Minshew, N.J., Just, M.A., 2011. The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. Brain 134, 2422–2435. https://doi.org/10.1093/brain/awr151.

Mørch-Johnsen, L., Nesvåg, R., Jørgensen, K.N., Lange, E.H., Hartberg, C.B., Haukvik, U.K., Kompus, K., Westerhausen, R., Osnes, K., Andreassen, O.A., Melle, I., Hugdahl, K., Agartz, I., 2017. Auditory cortex characteristics in schizophrenia: associations with auditory hallucinations. Schizophr. Bull. 43, 75–83. https://doi.org/10.1093/schbul/sbw130.

Morice, R.D., Ingram, J.C.L., 1982. Language analysis in schizophrenia: diagnostic implications. Aust. N. Z. J. Psychiatry 16, 11–21. https://doi.org/10.3109/00048678209161186.

Morice, R., McNicol, D., 1986. Language changes in schizophrenia: a limited replication. Schizophr. Bull. 12, 239–251. https://doi.org/10.1093/schbul/12.2.239.

Moro, A., Bambini, V., Bosia, M., Anselmetti, S., Riccaboni, R., Cappa, S.F., Smeraldi, E., Cavallaro, R., 2015. Detecting syntactic and semantic anomalies in schizophrenia. Neuropsychologia https://doi.org/10.1016/j.neuropsychologia.2015.10.030.

Nayani, T.H., David, A.S., 1996. The auditory hallucination: a phenomenological survey. Psychol. Med. https://doi.org/10.1017/S003329170003381X.

Oh, T.M., McCarthy, R.A., McKenna, P.J., 2002. Is there a schizophasia?: a study applying the single case approach to formal thought disorder in schizophrenia. Neurocase https://doi.org/10.1093/neucas/8.3.233.

Potts, C., 2007. The expressive dimension. Theor. Linguist. https://doi.org/10.1515/TL.2007.011.

Rochester, S.R., Martin, J.R., 1979. Crazy Talk: A Study of the Discourse of Schizophrenic Speakers. Plenum, New York.

Rosen, C., McCarthy-Jones, S., Chase, K.A., Humpston, C.S., Melbourne, J.K., Kling, L., Sharma, R.P., 2018. The tangled roots of inner speech, voices and delusions. Psychiatry Res. 264, 281–289. https://doi.org/10.1016/j.psychres.2018.04.022.

Slade, P.D., Bentall, R.P., 1998. The Johns Hopkins series in contemporary medicine and public health. Sensory Deception: A Scientific Analysis of Hallucination. Johns Hopkins University Press, Baltimore, MD, US.

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. https://doi.org/10.1016/S0920-9964(03)00013-6.

Tavano, A., Sponda, S., Fabbro, F., Perlini, C., Rambaldelli, G., Ferro, A., Cerruti, S., Tansella, M., Brambilla, P., 2008. Specific linguistic and pragmatic deficits in Italian patients with schizophrenia. Schizophr. Res. https://doi.org/10.1016/j.schres.2008.02.008.

Thomas, P., King, K., Fraser, W.I., Kendell, R.E., 1990. Linguistic performance in schizophrenia: a comparison of acute and chronic patients. Br. J. Psychiatry 156, 204–210. https://doi.org/10.1192/bjp.156.2.204.

Wechsler, D., 1997. Wechsler Adult Intelligence Scale. 3rd edition. .