

# **A Linguistic Approach to the Psychosis Continuum: (Dis)similarities and (Dis)continuities in How Clinical and Non-Clinical Voice-hearers Talk About Their Voices**

Introduction: 'Continuum' approaches to psychosis have generated reports of similarities and differences in voice-hearing in clinical and non-clinical populations. Previous research has reported similarities/differences at the cohort level, but not typically examined overlap or *degrees* of difference between groups.

Methods: We used a computer-aided linguistic approach to explore the language used to describe voice-hearing by a clinical group (Early Intervention in Psychosis service-users; N=40) and a non-clinical group (spiritualists; N=27). We identify semantic categories of terms statistically overused by one group compared with the other, and by each group compared to a control sample of non-voice-hearing interview data (log likelihood (LL) value 6.63+= $p < .01$ ; effect size measure: log ratio 1.0+). We examined relative frequencies of these terms at the individual level to consider if the dispersion of terms supports a continuum model.

Results: Notwithstanding significant cohort-level differences, there were varying degrees of overlap between the groups and considerable continuity in language use. Reports of negative affect were prominent in both groups ( $p < .01$ , log ratio: 1.12+). Challenges of cognitive control were also evident in both cohorts, with references to 'disengagement' accentuated in service-users ( $p < .01$ , log ratio: 1.14+).

Conclusion: A corpus linguistic approach to voice-hearing provides new evidence of differences between clinical and non-clinical groups. By considering variability at the individual level, we provide substantial evidence of continuity with implications for cognitive mechanisms underlying voice-hearing.

Keywords: psychosis; continuum; corpus linguistics; auditory verbal hallucinations; voice-hearing

## **Introduction**

In clinical settings, hearing voices (or auditory verbal hallucinations, AVH) is typically associated with a diagnosis of schizophrenia (Bauer et al., 2011; Sartorius et al., 1986), though hearing voices can also occur in other disorders such as bipolar disorder (Toh et al., 2016) and post-traumatic stress disorder (Brewin & Patel, 2010), where they can be distressing, debilitating, and often treatment-resistant experiences. Consistent with a 'continuum' approach to understanding psychotic experiences (Van Os et al., 2000), however, voices are also reported by individuals without a need for care, with a number of studies contrasting voice-hearing in clinical and non-clinical voice-hearers in terms of phenomenology (Powers et al., 2016; Sommer et al., 2010), cognition (Daalman et al., 2013), or brain activity (Diederer et al., 2012; Powers, Mathys & Corlett, 2017). The continuum approach suggests that such 'non-clinical voice-hearers' (NCVHs) are situated on a continuous dimension between non-voice-hearing individuals and voice-hearers with a psychiatric diagnosis (Baumeister et al., 2017). However, there have been concerns regarding the validity of the continuum approach (David, 2010), with suggestions that multiple continua might be involved (Johns et al., 2014) and that continuities in experience should not be taken to necessarily reflect continuity in terms of underlying neurocognitive mechanisms (Waters & Fernyhough, 2019).

A systematic review of studies of voice-hearing in NCVHs (Baumeister et al., 2017) demonstrated that the research cited in debates around a 'psychosis continuum' has typically relied on quantitative measures and the identification of significant differences (or not) between populations. For example, various studies have reported similarities between clinical voice-hearers (CVHs) and NCVHs in terms of the acoustic characteristics, loudness, location, number of voices heard, and personification of voices (De Boer et al., 2016; Powers et al., 2017). Nevertheless, the two groups have also been shown to differ on the interpretation of and beliefs about the experiences,

perceived ability to control them, and affective responses to them (Daalman et al., 2011; Johns et al., 2014; Powers et al., 2017; Woods & Wilkinson, 2017). Indeed, the most prominent reported differences between voice-hearing in CVHs and NCVHs relate to the emotional valence of the content, with the latter reporting more negative experiences, predicting the presence of a psychotic disorder in 88% of the participants (Daalman et al., 2011).

However, studies reporting group similarities and/or differences do not allow investigation of overlap between and within groups or degrees of difference on any given scale. This approach to 'difference' could be seen as actually reinforcing cohort boundaries, while shedding little light on what true continuity would look like for people who hear voices. David (2010) has called upon researchers interested in the 'psychosis continuum' to consider the variability of the phenomena within or between individuals, and to "define in advance what would constitute evidence of discontinuity" (p.1940). This demands an approach to measurement that allows researchers to evaluate discrepancies between (reported) experiences as incremental or discontinuous.

While linguistic choices in reports of voice-hearing may not straightforwardly reflect phenomenology, we suggest that the systematic analysis of similarities/differences and continuities/discontinuities in linguistic expression is an essential complement to other approaches to the psychosis continuum. Stanghellini, et al. (2012) argue that experiences of voice-hearing cannot be reliably assessed without some characterisation of the phenomenal *quality* of the experience, with one systematic review revealing a lack of qualitative studies of voice-hearing (Upthegrove et al., 2016). Those studies that do have a qualitative component typically involve semi-structured interviews and have demonstrated how some of the complexity of participants' sense-making and coping strategies can be captured in this way (Fenekou & Georgaca, 2010;

Milligan et al., 2013; Woods et al., 2015). The exploratory aspect of semi-structured interviews can "yield insights into what people who hear voices themselves regard as most important" (Woods et al., 2015, p33). This is consistent with a language-based approach to personal accounts, whereby the language choices made by participants are understood as reflecting the "relative importance of some narrative units as opposed to others" (Labov & Waletzky, 1997[1967], p32). Linguistics has been used recently to examine differing use of first- and second-person constructions in people with schizophrenia (Tovar et al., 2019), and to compare differences in the linguistic structure of voice "utterances" in clinical and non-clinical voice-hearing (De Boer et al., 2016). However, this work has largely focused on syntactic and grammatical differences that may reflect underlying deficit or disorder, rather than using linguistics to illustrate and elucidate experiential qualities of voice-hearing (but see Demjén, Semino, Marszalek, & Varese, 2019; Demjén, Marszalek, Semino, & Varese, 2020).

In this paper, we used a computer-aided linguistic approach to explore similarities/differences and continuities/discontinuities in the language used in reports of experiences of voice-hearing by CVHs and NCVHs. We investigated the extent to which an analysis of the language used by different groups when describing voice-hearing experiences supports or challenges the notion of a psychosis continuum (or continua), and explored the degree of overlap between groups in terms of the language used. We applied well-established analytical techniques from Corpus Linguistics (McEnery & Hardie, 2012) to the transcripts of two sets of phenomenological interviews with a) a group of NCVHs: 'spiritualist' participants who reported regularly hearing voices ('clairaudience'); and b) 40 Early Intervention in Psychosis (EIP) service-users reporting regular AVHs. In prior research on non-clinical voice-hearing, opportunity sampling has very often involved recruiting from spiritualist churches and

similar spiritual communities, which will often identify individuals with benign experiences of voice-hearing from a young age (Alderson-Day et al., 2017; Powers, Kelley and Corlett, 2017; Peters et al., 2016; Sommer et al., 2010). Sampling largely from spiritual/spiritualist groups may run the risk of missing community-specific differences in unusual experiences (Luhmann, 2017; Woods & Wilkinson, 2018), but experimental evidence has highlighted overlapping cognitive and neural processes between these groups and clinical participants (Diederer et al., 2012; Powers, Mathys, and Corlett, 2017), lending some confidence to the comparison. Here – consistent with prior research – we use a sample of spiritualist voice-hearers as a proxy example of non-clinical voice-hearing more generally.

We employed an existing software tool for the semantic annotation of texts to analyse the frequencies of words pertaining to previously discussed aspects of voice-hearing (affect, control, meaning-making, and sensory characteristics), and to compare the groups both to each other and to a 'reference corpus' of more general interview talk. We further investigated the dispersion of references to each aspect at the individual level across and within groups, to explore the degree of overlap between the two groups.

## **Methods**

### ***Data***

We examined semi-structured interview data in which participants (N = 67) discussed their voice-hearing experiences with one of two interviewers trained in clinical and phenomenological interviewing. The samples – which had been collected previously as part of Durham University's *Hearing the Voice* project (Alderson-Day et al., 2020) –

consisted of 40 CVHs using Early Intervention for Psychosis services ( $N$  female = 17, mean age = 28.7,  $SD$  = 9.96, mean time in services = 114.1 days,  $SD$  = 64.8) and 27 NCVHs ( $N$  female = 19, mean age = 58.0,  $SD$  = 12.0): ‘spiritualist’ participants reporting regular ‘clairaudient’ voice hearing experiences, who had been recruited via the Spiritualist National Union. Developed from Woods et al. (2015), the interview builds on eight open-ended questions (with subsequent prompts for comparison and elaboration) used to elicit discussion (see *Supplementary Materials*). Interviews typically lasted one hour (range: 24–124 minutes) and were transcribed verbatim according to a standard protocol. Users of Early Intervention in Psychosis (EIP) services were invited to take part if they i) were aged 16–65<sup>1</sup>; ii) heard voices at least once a week for a month<sup>2</sup>; iii) were fluent English speakers; iv) had normal or corrected-to-normal vision; and, v) were in the first nine months of using EIP services. Exclusion criteria for service-users were: i) the presence of a suspected duration of untreated psychosis longer than five years; ii) any neurological diagnoses; or iii) having a hearing impairment that required the use of hearing aids. Spiritualists were invited to take part if they were i) aged 18–75; ii) reported hearing voices at least once per month for at least the last three months; and iii) had normal or corrected-to-normal vision. Exclusion criteria for the spiritualists were having had contact with services for mental health issues or a psychiatric or neurological diagnosis; a recent history of drug or alcohol

---

<sup>1</sup> This age range was determined by EIP services and restrictions for giving consent. The NCVH recruitment reflects the broadest range covered by standard research ethics procedures.

<sup>2</sup> This requirement is informed by previous research on frequency estimates for CVHs (e.g. Sommer et al., 2010; Kråkvik et al., 2014).

abuse (i.e. in the previous 3 months)<sup>3</sup>; or a diagnosed hearing impairment. All participants were asked the same interview questions (allowing for distinct elaborations relating to each participant's experiences). Participants provided written consent, including for the anonymised reproduction of direct quotes from their interviews. All procedures were approved by either a local NHS Research Ethics Committee or university research ethics committee.

### *Corpus linguistic methodology*

Corpus Linguistics refers to a range of approaches that use tailor-made software tools to study patterns of language use in collections of texts, or 'corpora' that are typically too large to be analysed manually (McEnery & Hardie, 2012). Here we conducted: (1) 'keyness' analyses, to identify statistically significant differences between the two sets of interviews, and between each set of interviews and a 'reference corpus' in terms of 'semantic domains' (areas of meaning); and (2) 'dispersion' analyses, by plotting the relative frequency values for key areas of meaning for each participant, to determine whether the distribution of these linguistic features across individuals supported the concept of a 'continuum'. Only the participants' contributions to the interviews were subject to linguistic analysis, but the interviewer's questions were considered to contextualize participants' answers as relevant.

---

<sup>3</sup> This was to prioritise participants whose experiences are clearly not related to drugs and alcohol (following e.g. Sommer et al., 2010), though does contrast with CVH populations, which have included participants who reported alcohol or drug use (e.g. Powers, Kelley and Corlett, 2017).

Keyness analysis is used to compare the relative frequencies of words, grammatical categories or semantic domains across two datasets (corpora) to determine the language features that occur with unusually high frequency ('key' or 'overused') in one dataset compared to another. 'Key' features are taken to be characteristic of that dataset. We employed the UCREL Semantic Analysis System (USAS) tagger in the online corpus software Wmatrix (Rayson 2008) to semantically annotate our data and carry out keyness analyses at the level of semantic domains. The USAS tagger automatically allocates the content of a corpus into a pre-determined set of categories that group together words with similar semantic meaning and this categorisation system has been applied across a range of studies (Rayson, 2008). It includes 21 general semantic domains (e.g. "Emotion") and 232 more specific sub-domains (e.g. "Sad" as a sub-domain of "Emotion") (see *Supplementary Materials*). The general categories are indicated by a letter and the sub-domains are numbered. Further subdomains indicate contrasting valence (e.g. "E4.1+ Happy" versus "E4.1- Sad") and degree, as with the category "X3.2+", which would contain the word 'loud', compared with "X3.2++", which would contain 'louder' (NB: In the Results and Discussion sections, we only include the alphanumeric codes when they are needed to distinguish between different subdomains). When applied to conversational data, the tool has a reported level of accuracy of approximately 91% (Rayson, 2008) and referring to the original interview context allows us to check for any tagging errors. As participants might use different words for similar concepts (e.g. 'frightened' or 'scared'), this approach allows us to identify recurrent themes in the data.

We relied upon two statistical measures to identify 'key' semantic domains:



- Log Likelihood (LL) – a measure of statistical significance sensitive to the size of the evidence that a difference exists (Dunning, 1993). Our findings meet a minimum confidence measure of 6.63, equivalent to  $p < .01$ .
- Log ratio – a measure of effect size, i.e., the binary log of the ratio of relative frequencies in the two datasets (<http://cass.lancs.ac.uk/log-ratio-an-informal-introduction/>). This indicates the degree of difference between relative frequency values. To focus on the semantic domains that demonstrated the most difference, we set a minimum threshold of 1.0, which indicates that the relative frequency for a dataset is at least twice that of the comparison data.<sup>4</sup>

We carried out two keyness comparisons:

- (1) a direct comparison of the interview responses of the two cohorts, identifying the semantic domains that are statistically overused by CVHs compared to NCVHs and vice versa. These key semantic domains reflect the largest *differences* in relative frequencies between the groups, reported in terms of effect size;
- (2) a comparison of each cohort's interviews against a corpus of oral history interviews taken from the British National Corpus and consisting of 777,132 words (Aston & Burnard, 1998). This reference corpus matched our own data in terms of genre and style, but was more general in terms of topics discussed.

Semantic domains that were overused by only one group – in comparison with the reference corpus – provided another indication of difference. Semantic domains that

---

<sup>4</sup> A binary log ratio value of 2 indicates four times as often; log ratio of 3 indicates eight times as often, etc.

were overused by *both* groups – in comparison with the reference corpus – indicate potential areas of *similarity* in the language use of NCVHs and CVHs. In such cases, examining effect size values makes it possible to consider *degrees* of similarity between the two groups, with respect to more general language use.

Semantic domains that emerged as key in these comparisons were grouped according to the aspects of the voice-hearing experience that they were used to describe (see Table 1). These groupings were arrived at inductively by exploring how the constituent terms in each domain were used in context, and bringing together domains that were related in their usage. For example, the following semantic domains were used by participants to express negative emotions in relation to voice-hearing experiences, and were therefore subsumed under the grouping *Negative affect* (e.g. 'And then I get very upset')

- E2- Dislike (e.g. including 'hate', 'hates')
- E4.1- Sad (e.g. 'upset', 'sad', 'grief')
- E5- Fear/shock (e.g. 'scared', 'scary', 'panic', 'fear')
- E6- Worry (e.g. 'anxiety', 'stress', 'distressing').

We focused our analysis on a selection of groupings that can be subsumed under four broad themes that have been discussed in previous research on the psychosis continuum (Baumeister et al., 2017): *Affect*, *Control*, *Meaning-Making* and *Sensory Input*. In our data, the groupings that are subsumed under each of the four themes relate to the following aspects of the experience of voice-hearing (see Table 1):

- *Affect*: experiencing positive or negative emotions, and negative self-evaluations;

- *Control*: attempting to disengage from the voices, exerting command over the voices, and developing skills in voice-hearing;
- *Meaning-Making*: understanding and interpreting what the voices say;
- *Sensory Input*: sensory and cognitive processes associated with the voices.

We then examined the dispersion of the semantic domains in each of these themes across all 67 interviews. This involved plotting the aggregated relative frequencies<sup>5</sup> of these domains as bar charts. Where interviews from the clinical and non-clinical cohorts alternate along the x-axis, there is potential evidence of continuity. Where there is a clear separation, either of individuals or of the cohorts along the x-axis, there is evidence of discontinuity.

## **Results**

Our keyness analyses generated a total of 148 overused semantic domains. The table provided in the *Supplementary Materials* contains the complete list, with effect size values (log ratio) indicating under which keyness comparison each semantic domain was found to be overused, and for which participant group. Our groupings were derived from this list.

Twelve of these groupings pertained to our four themes of *Affect*, *Control*, *Meaning-Making*, and *Sensory Input*, and will be the focus of our analysis (Table 1). These groupings accounted for 37 (25.0%) of the 148 key semantic domains. The remaining semantic domains were excluded from the current analysis because of one or

---

<sup>5</sup> Given that the interviews differ in length, we normalised raw frequencies to a percentage of the individual interview data.

more of the following reasons: they did not have a clear semantic link to other domains (e.g. "Non-resident", which includes 'homeless'); they were not relevant to our four overarching themes (e.g. "Personal Names"); they consisted of words that, in the context of our data, differed considerably in meaning (e.g. the semantic domain "E4.1+ Happy" included the word 'laughing', which referred to being amused but also ridiculed).

[Table 1. Themes and groupings of key semantic domains.]

### *Affect*

The *Affect* theme included three groupings: *Negative affect*, *Negative evaluations of the self*, and *Positive affect*. Two domains indicating *Negative affect* were overused by both participant groups (compared to the reference corpus): "Sad" (e.g. 'upset', 'grief') and "Fear/shock" (e.g. 'scared', 'panic') (see Table 1). In both cases, effect size values indicate a larger overuse by CVHs (1.95 and 2.52 for CVHs, compared with 1.12 and 1.42 for NCVHs, respectively). When comparing the two cohorts, both these domains were found to be overused by CVHs, alongside "Dislike" (e.g. 'hate', 'despise') and "Worry" (e.g. 'anxiety', 'stress'). All four semantic domains that make up *Negative evaluations of the self* were also overused by CVHs when compared to NCVHs: "Worthless", "Weak" (e.g. 'vulnerable', 'helpless'), "Foolish" (e.g. 'stupid', 'ridiculous') and "Inability/Unintelligence" (e.g. 'idiot', 'unable'). Correspondingly, two semantic domains that make up the *Positive affect* grouping were found to be overused by NCVHs compared to service-users: "Content" (e.g. 'proud', 'pleased') and "Confident" (e.g. 'reassurance', 'faith').

The distribution of relative frequencies across individuals for each of the three *Affect* groupings is displayed in Figure 1. The overlapping of values for NCVHs (white

bars) and CVHs (black bars) and incremental rise in values across participants suggests continuity. This is observed for each grouping. In the case of *Negative affect*, individual values for 44 of the 67 participants (65.7%) overlap (between 0.00-0.44). Similarly, 49 (73.1%) participant values for terms relating to *Negative evaluations of self* overlap. There is also overlap in the relative frequency values for *Positive affect* terms, although 16 CVHs (40%) do not use any terms in this grouping. In summary, while CVHs are more marked in their use of negative evaluative terms (for affect and the self), there is a notable overlap between the two cohorts.

[Figure 1. Individual relative frequency values for 'Affect' terms]

### ***Control***

The semantic domain "Inattentive" (e.g., 'ignore' and 'distract' in utterances such as 'I've been trying to distract myself') was overused by both CVHs and NCVHs compared to the reference corpus. Effect size values indicate a larger overuse in CVHs: 4.82 compared with 2.35 for NCVHs. The direct comparison showed that this semantic domain was also overused by CVHs when compared to NCVHs, along with two other domains relating to *Disengagement*: "Avoiding" (e.g. 'avoided', 'leave alone') and "Speech: Not communicating" (e.g. 'shut up', 'keep quiet') (see Table 1). In contrast, all three domains included in the grouping *Command over (the voices)* were overused by NCVHs compared to CVHs: "Using" (e.g. 'choos[ing] to use my mediumship'), "Open Finding; Showing" (e.g. 'open[ing] up and allowing [the voices] to come forward'), and "Giving" (e.g. 'getting rid' of the voices). This was also the case for *Developing understanding*, which includes "Education", "Learning" (e.g. 'take on board'), and "Ability and intelligence" (e.g. 'faculty', 'calibre').

The dispersion of values for the *Disengagement* grouping as a whole (Figure 2) shows an overlap in the relative values for NCVHs (0.00-0.06) and CVHs (0.00-0.40), suggesting continuity. However, nine NCVHs (33% of the group) did not use any terms for this grouping, alongside three CVHs (8%). There was more evidence for continuity between the two groups for *Command over (the voices)* and *Developing understanding* (Figure 2).

[Figure 2. Individual relative frequency values for 'Control' terms]

### ***Meaning-making***

The *Meaning-making* theme consists of a single grouping of five semantic domains, which were all overused by NCVHs compared to CVHs: "Suitable" (e.g. 'relevant', 'appropriate'), "Concrete/Abstract" (e.g. 'philosophical', 'practical'), "Important", "Evaluation: True" (e.g. 'evidence', 'prove'), and "Linguistic Actions, States and Processes; Communication" (e.g. 'message', 'means'). However, an examination of these terms in context shows that both groups reflected on the 'relevance' of their voice-hearing experiences to other aspects of their lives, evaluated the content of their voice-hearing experiences in terms of 'prominence', and described what the voices say as 'messages'.

The dispersion of values for this theme showed an overlap in the relative frequency values for NCVHs and CVHs up to a value of 0.35 (Figure 3), suggesting some degree of continuity. The exclusive representation of NCVHs above this value – consistent with the keyness analysis – suggests that *Meaning-making* terms are particularly characteristic of the spiritualist interviews.

[Figure 3. Individual relative frequency values for 'Meaning-making' terms]

### *Sensory Input*

Four semantic domains relating to sound were overused by both groups when compared with the reference corpus, reflecting the focus of some of the interview questions:

"Sensory: Sound" (e.g. 'hear', 'sounds'), "Sound: Quiet (X3.2-)" (e.g. 'quiet', 'muffled'), "Sound: Quiet (X3.2--)" (e.g. 'quieter') and "Sound: Loud" (e.g. 'noisy', 'audibly'). Effect size values show a similar degree of overuse for "Sound: Quiet (X3.2-)" and a greater overuse by CVHs for the other domains. The direct keyness comparison showed that "Sound: Loud" was overused by CVHs compared to NCVHs. In contrast, the semantic domains included under *Strength* were overused by NCVHs compared to CVHs, i.e.: "S1.2.5++ Tough/strong" (e.g. 'stronger') and "S1.2.5+++ Tough/strong" (e.g. 'strongest'). These terms tended to be used in reference to the intensity of the voices, as in 'So at the times of real severe emotional upset that external voice seems to come in stronger'.

The dispersion of values for *Loudness* (Figure 4) shows an overlap in the relative values for NCVHs (0.00–0.08) and CVHs (0.00–0.47), suggesting some degree of continuity. However, it should be noted that no terms for this grouping were found for five NCVHs (19%) and three CVHs (8%). Similarly, while NCVHs described their sensory experiences in terms of 'strength' more than CVHs, many participants in both groups did not use these terms at all.

[Figure 4. Individual relative frequency values for 'Loudness' and 'Strength' terms]

Participants were also asked about the phenomenological aspects of their voice-hearing beyond the perception of sound. Three semantic domains grouped under *Other senses* were found to be overused by CVHs: "Sensory: Touch", "Seen" ('noticed', 'notice'), and "Sensory: Smell". The latter was also overused in both cohorts when compared with the reference corpus, with a higher effect size value for CVHs.

## **Discussion**

Debates around the 'psychosis continuum' have highlighted areas of similarity and difference between clinical and non-clinical populations (Baumeister et al., 2017; Johns et al., 2014). In contrast with previous studies, our linguistic approach to the psychosis continuum makes it possible to investigate evidence of differences/similarities between individual CVHs and NCVHs (represented by Early Intervention in Psychosis service-users and spiritualists, respectively), as well as evidence of continuity/discontinuity within and across groups.

We showed that *Negative affect* terms were overused in both voice-hearing groups compared to more general interview talk (albeit to different extents), but also used much more frequently by CVHs when directly compared with NCVHs, along with terms related to *Negative evaluations of the self*. In contrast, NCVHs used more *Positive affect* terms, consistent with previous research reporting the differences in emotional valence in the responses of clinical and non-clinical groups (Daalman et al., 2011). However, our approach allowed us to extend this comparison to demonstrate that, despite these overall differences, there is also overlap and therefore some degree of continuity between the cohorts.

Likewise, consistent with previous research, our findings with respect to *Control* showed that terms related to *Disengagement* were overused by CVHs, while NCVHs



more often referred to having *Command over* their voices and *Developing understanding*. This is consistent with research that has found that CVHs are more likely to try to actively ignore their voices (Kråkvik et al., 2015) and that there are differences between CVHs and NCVHs in their perceived ability to control voices (Powers et al., 2017). However, again, we observed overlap between the two cohorts in these dimensions, suggesting that control of voices may not be an 'all-or-nothing' component of voice-hearing, and instead may be dimensional, with differences between CVHs and NCVHs being matters of degree rather than dichotomy. Considering that inhibitory control is one cognitive mechanism that has been associated with hallucination-proneness in both clinical and non-clinical groups (Waters et al., 2003; Paulik et al., 2007; Alderson-Day et al., 2019), these findings suggest a scenario where the spontaneity and intrusiveness of voices may naturally vary across individuals, or be flexibly tempered with training and practice (Luhmann et al., 2018).

Similarly, previous research has also highlighted differences in the way that CVHs and NCVHs interpret their voice-hearing (Powers et al., 2017) and we found that five semantic domains relating to *Meaning-making* were significantly overused by NCVHs, compared with CVHs. This supports (now mainstream) cognitive models of voice-hearing where meta-cognitive appraisals – rather than underlying sensory atypicalities – mark out the main differences between help-seeking and healthy individuals (Peters et al., 2017).

Language about sensory components highlighted some of the subtler differences between our two cohorts. Both NCVHs and CVHs overused terms relating to the audible aspects of their voice-hearing in comparison with general interview talk, but *Loudness* terms were used much more frequently by CVHs than NCVHs, while the opposite pattern was observed with regard to use of *Strength* terms. This provides one

example of the care which must be taken when assessing the phenomenology of voice-hearing experiences across such different groups, with spiritualist participants sometimes using more metaphorical than literal language to describe the sensory experience (see Luhrmann, 2017). Although both CVHs and NCVHs may describe "auditory" phenomena, this can mean many different things to different individuals (Woods et al., 2015).

For all of the aspects of the voice-hearing experience we considered, the dispersion analyses at the level of individual participants show an overlap in the relative frequencies of relevant terms for members of the two groups, supporting the notion of a continuum even in the context of a group-level difference. This offers a response to David's (2010) call to consider the variability within or between individuals. Lawrie (2016) has argued that "[t]o overturn current practice would require convincing proof or at least some persuasive evidence that the psychosis continuum approach adds something in clinical settings" (p. 126), and so we must also consider what we learn from a continuum model. In elaborating on the similarity/difference approach to the psychosis continuum, we can begin to identify aspects of the experience where the difference between CVHs and NCVHs would appear to be a matter of degree. For example, *Meaning-making* terms are used in similar ways by members of both groups, e.g. 'the voice kind of was relevant to that' (CVH) and 'But they are relevant to the person' (NCVH). However, they occur much more frequently in the interviews with NCVHs. Similarly, we have shown that there is a qualitative difference in how participants discuss the intensity of their voice-hearing (i.e. with CVHs describing 'loudness' and NCVHs more likely to describe 'strength') but that overall cohort differences may actually only reflect a few key individuals. In these instances, we should be wary of overstating the relationship between such a feature and clinical status,

and consider whether this indicates something of the phenomenology of the individuals' experience, or – as researchers have observed (Luhmann, 2017) – is more reflective of a particular vocabulary, for example, that associated with mediumship. Either way, these findings suggest that seemingly simple features of voices (e.g., loudness) are complex, and may not exist on a single dimension.

It could be considered a limitation of our work that our NCVH participants are all self-identified spiritualists, in that they may not be representative of the breadth of NCVH experiences. However, this allows us to contextualise our work among other studies of NCVHs, in which NCVH populations often comprise a high proportion of spiritualists (e.g. Sommer et al., 2010; Peters et al., 2016; Powers, Kelley and Corlett, 2017). Spiritualism and non-clinical status are also likely to be related – for example, due to positive appraisals of voices and expectations regarding controllability of the experience that form part of the spiritualist worldview (Peters et al., 2017) – so it is difficult to characterise a participant as either one or the other. This highlights that more work is needed to understand the interaction between spiritualism and clinical status (e.g. looking at both clinical and non-clinical spiritualist participants) and how this compares to wider varieties of non-clinical voice-hearing (Luhmann et al., 2019). In that sense, our work provides some valuable insights into the NCVH spiritualist experience, and a basis for further exploration of language use among 'healthy' voice-hearers.

Since the recruitment of CVHs targeted those newly engaging with services, there are also likely to be discrepancies between the two groups regarding the length of time they have been hearing voices (and subsequently, the development of coping strategies, for example). Similarly, the mean ages of the NCVHs (58.0) and CVHs (28.7) could have implications for how voice-hearing is experienced and reported by

participants. The participant groups – on average – belong to the same 'life stage' (Labov, 2001) and so we would expect that individual differences would have more of an influence on language use than age. Differences according to such participant demographics (such as age, gender, region) are likely to be minimised when the reports are aggregated and subject to statistical tests. Furthermore, our semantic domain approach allows us to group words referring to related concepts (Rayson, 2008), even if there are differences in the specific lexis according to gender, class or individual preference. Nevertheless, participants could be grouped according to more discrete age and/or duration of voice-hearing categories to investigate any potential effects on language use and analyses of individual (language) differences would need to be contextualised within individual case studies.

In addition to investigating differences in clinical and non-clinical spiritualist voice-hearers, further research should also consider the extent to which self-reports are shaped by such interpretive frameworks. While our analysis focused on semantic domains that we could relate to four key themes in the discussion of a psychosis 'continuum', some of the remaining key semantic domains (see table in *Supplementary Materials*) might be relevant to exploring other aspects pertinent to the potential need for care. For example, terms relating to alcohol and recreational drugs were overused by CVHs, while terms related to the sense of professional identity that is associated with voice-hearing were overused by NCVHs.

## **Conclusion**

In sum, we have shown that Corpus Linguistics offers ways of discerning patterns in the self-report of voice-hearers that can highlight (dis)similarities and (dis)continuities within and across clinical boundaries. In combining keyness analysis with an

exploration of dispersion at the individual level, we have been able to report differences and some similarities between voice-hearing groups, as well as presenting measures that identify and document aspects of the voice-hearing experience that are variously 'continuous' across our clinical and non-clinical groups.

## Acknowledgments

We would like to thank all of our participants for their contributions.

This research was supported by the Wellcome Trust grant WT108720 and the ESRC grant number ES/R008906/1 for the Centre for Corpus Approaches to Social Science.

Conflict of Interest: None.

## References

- Alderson-Day, B., Lima, C. F., Evans, S., Krishnan, S., Shanmugalingam, P., Fernyhough, C. & Scott, S. K. (2017). Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain*, *140*, 2475–2489. doi:10.1093/brain/awx206
- Alderson-Day, B., Smailes, D., Moffatt, J., Mitrenga, K., Moseley, P., & Fernyhough, C. (2019). Intentional inhibition but not source memory is related to hallucination-proneness and intrusive thoughts in a university sample. *Cortex*, *113*, 267-278.
- Alderson-Day, B., Woods, A., Moseley, P., Common, S., Deamer, F., Dodgson, G., & Fernyhough, C. (2020, May 15). Voice-hearing and personification: Characterising social qualities of auditory verbal hallucinations in early psychosis. Retrieved from <https://psyarxiv.com/rvfty>
- Aston, G. & Burnard, L. (1998). *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh; Edinburgh University Press. Available from: <http://corpus.leeds.ac.uk/teaching/aston-burnard-bnc.pdf>.
- Bauer, S. M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., Chaudry, H. R., Idemudia, S. E., Gscheider, S., Ritter, K. and Stompe, T. (2011). Culture and the prevalence of hallucinations in schizophrenia. *Comprehensive Psychiatry*, *52*(3), 319–325. doi:10.1016/j.comppsy.2010.06.008

- Baumeister, D., Sedgwick, O., Howes, O. and Peters, E. (2017). Auditory verbal hallucinations and continuum models of psychosis: A systematic review of the healthy voice-hearer literature. *Clinical Psychology Review*, 51, 125–141. doi:10.1016/j.cpr.2016.10.010
- Brewin, C. R. & Patel, T. (2010). Auditory pseudohallucinations in United Kingdom war veterans and civilians with posttraumatic stress disorder. *The Journal of Clinical Psychiatry*, 71(4), 419–425. doi:10.4088/JCP.09m05469blu
- Daalman, K., Boks, M. P. M., Diederer, K. M. J., De Weijer, A. D., Blom, J. D., Kahn, R. S. & Sommer, I. E. C. (2011). The Same or Different? A Phenomenological Comparison of Auditory Verbal Hallucinations in Healthy and Psychotic Individuals. *The Journal of Clinical Psychiatry*, 72(3), 320–325. doi:10.4088/JCP.09m05797yel
- Daalman, K., Sommer, I. E., Derks, E. M. & Peters, E. R. (2013). Cognitive biases and auditory verbal hallucinations in healthy and clinical individuals. *Psychological Medicine*, 43(11), 2339–2347. doi:10.1017/S0033291713000275
- David, A. S. (2010). Why we need more debate on whether psychotic symptoms lie on a continuum with normality. *Psychological Medicine*, 40, 1935–1942. doi:10.1017/s0033291710000188
- De Boer, J. N., Heringa, S. M., Van Dellen, E., Wijnen, F. N. K. & Sommer, I. E. C. (2016). A linguistic comparison between auditory verbal hallucinations in patients with a psychotic disorder and in nonpsychotic individuals: Not just what the voices say, but how they say it. *Brain and Language*, 162, 10–18. doi:10.1016/j.bandl.2016.07.011
- Demjén, Z., Semino, E., Marszalek, A., & Varese, F. (2019). Metaphor framing and distress in lived-experience accounts of voice-hearing. *Psychosis*, 11(1), 16–27. doi:10.1080/17522439.2018.1563626
- Demjén, Z., Marszalek, A., Semino, E., & Varese, F. (2020). “One gives bad compliments about me, and the other one is telling me to do things” – Power, (Im)Politeness and Agency in Lived-experience Accounts of Voice-hearing. In Demjén, Z. (ed.) *Applying Linguistics in Illness and Healthcare Contexts*. pp. 17–43. London: Bloomsbury.
- Diederer, K. M., Daalman, K., de Weijer, A. D., Neggers, S. F., van Gastel, W., Blom, J. D., Kahn, R. S. & Sommer, I. E. (2012). Auditory hallucinations elicit similar brain activation in psychotic and nonpsychotic individuals. *Schizophrenia Bulletin*, 38(5), 1074–1082. doi:10.1093/schbul/sbr033
- Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, 19(1), 61–74. Retrieved from <https://www.aclweb.org/anthology/J93-1003.pdf>
- Fenekou, V and Georgaca, E. (2010). Exploring the experience of hearing voices: A qualitative study. *Psychosis: Psychological, Social and Integrative Approaches*, 2(2), 134–143. doi:10.1080/17522430903191783

- Johns, L., Kompus, K., Connell, M., Humpston, C., Lincoln, T., Longden, E., Preti, A., Alderson-Day, B., Badcock, J. C., Cella, M., Fernyhough, C., McCarthy-Jones, S., Peters, E., Raballo, A., Scott, J., Siddi, S., Sommer, I. E. & Larøi, F. (2014). Auditory verbal hallucinations in persons with and without a need for care. *Schizophrenia Bulletin*, 40, S255–S264. doi:10.1093/schbul/sbu005
- Kråkvik, B., Larøi, F., Kalhovde, A. M., Hugdahl, K., Kompus, K., Salvesen, Ø., Stiles, T. C., & Vedul-Kjelsås, E. (2015). Prevalence of auditory verbal hallucinations in a general population: a group comparison study. *Scandinavian Journal of Psychology* 56(5), 508–515. doi:10.1111/sjop.12236
- Labov, W. (2001). *Principles of Linguistic Change. Volume 2: Social Factors*. Oxford: Blackwell.
- Labov, W. & Waletzky, J. (1997[1967]). Narrative analysis: Oral versions of personal experience. *Journal of Narrative and Life History*, 7(1/4), 3–38. Retrieved from <https://www.jbe-platform.com/content/journals/10.1075/jnlh.7.02nar>
- Lawrie, S. (2016). Whether "psychosis" is best conceptualized as a continuum or in categories is an empirical, practical and political question. *World Psychiatry*, 15(2), 125–126. doi:10.1002/wps.20325
- Linden, D. E. J., Thornton, K., Kuswanto, C. N., Johnston, S. J., Van den Ven, V. & Jackson, M. C. (2011). The brain's voices: comparing nonclinical auditory hallucinations and imagery. *Cerebral Cortex*, 21(2), 330–337. doi:10.1093/cercor/bhq097.
- Luhrmann, T. M. (2017). Diversity within the psychotic continuum. *Schizophrenia Bulletin*, 43(1), 27–31. doi:10.1093/schbul/sbw137
- Luhrmann, T. M., Alderson-Day, B., Bell, V., Bless, J. J., Corlett, P., Hugdahl, K., Jones, N., Larøi, F., Moseley, P., Padmavati, R., Peters, E., Powers, A. R. & Peters, E. (2019). Beyond trauma: a multiple pathways approach to auditory hallucinations in clinical and nonclinical populations. *Schizophrenia bulletin*, 45(Supplement\_1), S24–S31. doi:10.1093/schbul/sby110
- McEnery, T. & Hardie, A. (2012). *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
- Milligan, D., McCarthy-Jones, S., Winthrop, A. & Dudley, R. (2013). Time changes everything? A qualitative investigation of the experience of auditory verbal hallucinations over time. *Psychosis: Psychological, Social and Integrative Approaches*, 5(2), 107–118. doi:10.1080/17522439.2012.667438
- Paulik, G., Badcock, J. C., & Maybery, M. T. (2007). Poor intentional inhibition in individuals predisposed to hallucinations. *Cognitive Neuropsychiatry*, 12(5), 457–470. doi:10.1080/13546800701394329
- Peters, E., Ward, T., Jackson, M., Morgan, C., Charalambides, M., McGuire, P., Woodruff, P., Jacobsen, P., Chadwick, P. Garety, P. A. (2016). Clinical, socio-demographic and psychological characteristics in individuals with persistent

psychotic experiences with and without a “need for care”. *World Psychiatry*, 15(1), 41–52. doi:10.1002/wps.20301

Peters, E., Ward, T., Jackson, M., Woodruff, P., Morgan, C., McGuire, P., & Garety, P. A. (2017). Clinical relevance of appraisals of persistent psychotic experiences in people with and without a need for care: an experimental study. *The Lancet Psychiatry*, 4(12), 927–936. doi:10.1016/S2215-0366(17)30409-1

Powers, A. R., Kelley, M. S. & Corlett, P. R. (2017). Varieties of voice-hearing: psychics and the psychosis continuum. *Schizophrenia Bulletin*, 43, 84–98. doi:10.1093/schbul/sbw133

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science*, 357(6351), 596–600. doi:10.1126/science.aan3458

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519–549. doi:10.1075/ijcl.13.4.06ray

Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J. E., & Day, R. (1986). Early manifestations and first-contact incidence of schizophrenia in different cultures. A preliminary report on the initial evaluation phase of the WHO Collaborative Study on determinants of outcome of severe mental disorders. *Psychological Medicine*, 16(4), 909–928. Retrieved from <https://www.cambridge.org/core/journals/psychological-medicine/article/early-manifestations-and-firstcontact-incidence-of-schizophrenia-in-different-cultures-a-preliminary-report-on-the-initial-evaluation-phase-of-the-who-collaborative-study-on-determinants-of-outcome-of-severe-mental-disorders/3C7A646CA9EC43AC75057AA05F25FB29>

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

Stanghellini, G., Langer, Á. I., Ambrosini, A. & Cangas, A. J. (2012). Quality of hallucinatory experiences: differences between a clinical and non-clinical sample. *World Psychiatry*, 11(2), 110–113. doi:10.1016/j.wpsyc.2012.05.007

Toh, W. L., Castle, D. J., Thomas, N., Badcock, J. C. & Rossell, S. L. (2016). Auditory verbal hallucinations (AVHs) and related psychotic phenomena in mood disorders: analysis of the 2010 Survey of High Impact Psychosis (SHIP) data. *Psychiatry Research*, 243, 238–245. doi:10.1016/j.psychres.2016.06.035

Tovar, A., Fuentes-Claramonte, P., Soler-Vidal, J., Ramiro-Sousa, N., Rodriguez-Martinez, A., Sarri-Closa, C., Sarró, S., Larrubia, J., Andrés-Bergareche, H., Miguel-Cesma, M. C., Padilla, P. P., Salvador, R., Pomarol-Clotet, E., & Hinzen, W. (2019). The linguistic signature of hallucinated voice talk in schizophrenia. *Schizophrenia Research*, 206, 111–117. doi:10.1016/j.schres.2018.12.004



- Upthegrove, R., Broome, M. R., Caldwell, K., Ives, J., Oyeboode, F. & Wood, S. J. (2016). Understanding auditory verbal hallucinations: a systematic review of current evidence. *Acta Psychiatrica Scandinavica*, 133, 352–367. doi:10.1111/acps.12531
- Van Os, J., Hanssen, M., Bijl, R. V. & Ravelli, A. (2000). Strauss (1969) revisited: a psychosis continuum in the general population? *Schizophrenia Research*, 45(1-2), 11–20. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0920996499002248?via%3Dihub>
- Waters, F. A., Badcock, J. C., Maybery, M. T., & Michie, P. T. (2003). Inhibition in schizophrenia: association with auditory hallucinations. *Schizophrenia research*, 62(3), 275–280. doi:10.1016/s0920-9964(02)00358-4
- Waters, F., & Fernyhough, C. (2019). Auditory hallucinations: Does a continuum of severity entail continuity in mechanism? *Schizophrenia Bulletin*, 45(4), 717–719. doi:10.1093/schbul/sbz002
- Woods, A., Jones, N., Alderson-Day, B., Callard, F. & Fernyhough, C. (2015). Experiences of hearing voices: analysis of a novel phenomenological survey. *Lancet Psychiatry*, 2(4), 323–331. doi:10.1016/S2215-0366(15)00006-1
- Woods, A. & Wilkinson, S. (2017). Appraising appraisals: role of belief in psychotic experiences. *Lancet Psychiatry*, 4(12), P891–892. doi:10.1016/S2215-0366(17)30434