

The Phenomenology of Voice-Hearing and Two Concepts of Voice

Sam Wilkinson, Department of Philosophy, University of Exeter

Joel Krueger, Department of Philosophy, University of Exeter

What has struck us both about the Voices in Psychosis (VIP) interview transcripts is that the experiences reported in them that are called ‘voices’ are, first, hugely varied, and, second, do not all happily fall under the label ‘auditory verbal hallucination’ (AVH).¹ Accordingly, a more phenomenologically sensitive engagement with detailed first-person reports (of which these transcripts are exemplary) is needed. In this context, phenomenology can be helpful in several ways. For one thing, it can help develop a more accurate description of the individual’s experience—i.e. one that preserves the integrity of both their experience *as lived* as well as their subsequent reports *about* that experience—which, in turn, can serve as a tool for cultivating a kind of empathy (Ratcliffe, 2012). Additionally, although phenomenology is generally not concerned with providing an account of underlying cognitive, neural, or biological processes, it can nevertheless assist with developing such accounts. For example, it can assist the development of neurocognitive and biosocial models of voice-hearing—which *are* primarily concerned with causal–mechanistic explanations—both by providing an explanatory target for such models and by assisting the evaluation of existing models against the experiences as described in first-person reports such as these (McCarthy-Jones et al., 2013).

Phenomenology starts with open-ended questions such as: ‘What is it like for an individual to be in a certain mental state, such as hearing a voice?’ and ‘What is the meaning of that experience for the individual?’ This phenomenological orientation is apparent in the consistent way the interviews in these transcripts begin. First, the interviewer acknowledges that voices ‘can mean lots of different things’ to different people, from something that ‘might involve sound’, to the feeling ‘that someone or something is communicating with you’, to something akin to ‘a form of telepathy or loud thoughts’. This is an acknowledgement to the individual being interviewed that voice-hearing is a rich and experientially diverse kind of experience, and that they should describe the experience as it happens *for them*. This experiential diversity is

¹ The label of AVH has been criticized by many (e.g. Slade and Bentall, 1988), including contributors to this volume.

also tacitly acknowledged by the open-endedness of the very first question—‘Could you try to describe to me some of the voices or voice-like experiences you’ve been having?’—along with an encouragement ‘to use your own terms, your own language’ in providing these descriptions.

Inspired by these transcripts—and as a way of gesturing towards a more nuanced phenomenological characterization of voice-hearing—we want to acknowledge the rich diversity of experiences found in these reports by isolating at least two things that ‘voice’ seems to mean in this context. We say ‘at least’ because we do not suggest that these two concepts are in any way exhaustive. However, notice that we do not need them to be. If the category of things called ‘voices’ in clinical contexts actually picks out a number of distinct phenomena (two, three, or four, etc.), then we should be mindful of that plurality in general and wary of using the term ‘auditory verbal hallucination’ in particular. A commitment to phenomenological sensitivity should breed this kind of caution.

Before we embark on this brief project, a word about method. We are an analytic philosopher and a phenomenologist, and so we will proceed by conceptual and phenomenological analysis. How do these things fit together? Talking about experience is a strange phenomenon, since it involves using a public tool to get at something very private. The public tool—language—is a rather blunt and coarse-grained instrument for capturing the precise nature of a given experiential episode, such as, for instance, the character and structure of hearing voices versus more general features of audition as a distinct perceptual modality. As long as it is not actively misleading, any given description of experience rules out a large number of possibilities about what that experience might be like. However, it does not rule out *all* of the possibilities; it does not narrow them down to one. In such a situation, the description is accurate, but inevitably too coarse-grained. In light of this relationship between the language used to talk about experience and the experience itself, we proceed in two steps:

- Step 1.** The conceptual analysis of ‘voice’ will pick out two different things that could be meant by ‘voice’. This is coarse-grained.
- Step 2.** The phenomenological analysis then elaborates on the precise possible nature of the experience in question.

In principle, then, the coarse-grained conceptual categories further subdivide phenomenologically. Little wonder thinking of voices as a unified phenomenon is misleading!

Two Concepts of Voice

Concept 1

Let us begin with the concept that is closest to that of ‘auditory verbal hallucination’. This concept picks out a particular auditory quality of experience, namely, a

speech-like quality. This is the sense of ‘voice’ you might find in the utterance: ‘This synthesizer has a “voice” setting.’ Note how there are two further subtle ambiguities here. ‘Voice’, with this focus on auditory quality (pointing out that it *sounds* like a human voice, rather than, say, a violin or a klaxon), can refer to the subjective experience or the environmental stimulus (the speech sound, the pressure ripples in the air). It can also refer to someone’s capacity or disposition to produce a sound with a particular quality: the statement ‘She has a beautiful singing voice’ holds true even when the person is not singing. So, ‘voice’ in this sense, with its auditory qualitative focus, can take a subjective or objective flavour and can also pick out something actual or dispositional.

Many speakers in these transcripts seem to use this sense of voice. Alex, for instance, talks explicitly about the ‘pitch and tone’ of his voices, and says, for example, that the voices he hears are like ‘like someone’s whispering in your ear [. . .] Except it’s that, it’s not necessarily at a whisper’s volume.’ There is also explicit mention of the location of the sound and how that varies: ‘Sometimes it can be as if they’re sitting right next to you or standing alongside you. And then other times, I could be in another room and think it’s coming from a different room.’ What is interesting here is that, despite variations in volume, location, and felt proximity of the voices, Alex consistently experiences these voices *as voices*. In other words, he does not confuse them with other auditory stimuli, despite the fact, he says, that they are often hard to hear. Moreover, while Alex’s descriptions contain some mention of the thematic *content* of the voices—i.e. what they say to him (threats, insults, etc.)—the bulk of his report focuses on the speech-like qualities of these voices and how, despite a variety of ways these qualities are presented to him, he nevertheless experiences these qualities *as speech-like*. Often, he tells us, he is unable to understand the voices. However, he still recognizes the voices by their speech-like qualities: ‘It’s just slight distinctive differences in the voices, like in the pitch and the tone.’

Similarly, Bill consistently describes hearing voices that start like ‘a bark’ or ‘a shout or a noise, you know like someone that is more just ehm like a bang on the table or something like that, sort of . . . but more of a vocal version of that.’ However, Bill does not confuse these voices with a bang on a table or genuine barking. He tells us that ‘it’s similar to a dog bark but it’s not entirely a dog bark’—again, because the voices manifest experientially with a distinctive auditory quality that identifies them as voices. Like Alex, Bill spends some time describing what the voices say. However, most of his focus is on qualities of his experience of the voice, independent of this content. Again, these qualities are what appear to be most phenomenologically salient.

When we say that this concept of ‘voice’ picks out a particular ‘auditory quality’, there are two different ways in which this can be interpreted: as rich or sparse. This issue is related to an ongoing debate in philosophy of perception about whether the content of perceptual experience can ever be rich, namely, can ever go beyond the ‘low-level properties’ that are thought of as basic to that sensory modality. For vision, the question is whether all we ever really see are colour, shape, and shade, or whether

we literally see trees, or even, with relevant expertise, *oak* trees (Siegel, 2006). For audition, the question is whether what we experience goes beyond loudness, pitch, and timbre. Without going into the technicalities of this debate, we certainly intend a rich interpretation of ‘auditory quality’ here. We intend it to include properties that go well beyond the ‘low-level properties’ and include the experience of, for instance, personhood, femininity, aggression, and so on. One might think that allowing this richness risks collapsing Concept 1 into Concept 2, which we are about to introduce, but this would be to misunderstand the two concepts. However rich we take auditory experience to be, when we use ‘voice’ to express Concept 2, we are not picking out the auditory aspects of an experience at all (even though the experience may well be an auditory one).

Concept 2

The second concept that ‘voice’ can express picks out, not so much an auditory quality of a sound (however richly this is conceived), but a particular agent: what Felicity Deamer and colleagues (Deamer and Hayward, 2018; Deamer and Wilkinson, 2015) have called the ‘speaker behind the voice’. This concept is intrinsically related to the experience of something with its own perspective and agency, and goes beyond the superficially auditory. (Indeed, not only does this build onto the auditory experience, it also sometimes circumvents it altogether, as in the case of soundless voices (Wilkinson, 2019).) Often, it involves the binding of this agentic experience into a singular agent representation persisting over time (see Wilkinson and Bell, 2016). This is the sense of ‘voice’ that is very commonly used in the context of hearing voices. Indeed, it is implicit in a question that is so often asked of voice-hearers: ‘How many voices do you hear?’ This question is asking the voice-hearer *how many re-identifiable agents are represented over time* across several experiential episodes. The question makes no sense otherwise. It is not a question about the qualitative variety of some relatively two-dimensional auditory experience. Indeed, it is not even a question about the variety of what is said (i.e. its thematic or semantic content), or even about the qualitative personality of the represented agents. The personality of a ‘voice’, in this sense that picks out the identity of an agent, can change over time. Voices can shift from being nice to nasty. Personality traits are features of these voices, but they are not definitive of them. The question ‘How many voices do you hear?’ is one about specific individual identities. In many of the VIP transcripts, the voices have personality traits, but these do not appear fully bound to clearly re-identifiable individuals. The question of how many voices are heard by these voice-hearers is therefore not answerable, since although they hear voices in both the speech–sound and agentic sense, the voice-as-agent is transient and untrackable. In some of the transcripts, however, the agents can be tracked, and hence counted.

We find an example of this second concept of voice in Dan’s transcript, when he says that ‘over the last couple of months I’ve sort of been hearing I think up to

seven different voices'. Of phenomenological interest here is the fact that Dan experiences these voices as individuated by their spatial continuity (i.e. where each of these voices originate from, respectively, within his inner psychoacoustic landscape). Sometimes, he says, 'when it's inside the head, [it's] very, very loud and it's like it consumes all the space around you, and you feel like you're kind of suffocating in it'. At other times, the voices are externally located, but very close by (e.g. 'it's like it's pounding right in your ear, it's like it's, they're talking right into your ear'), or even more spatially remote ('sometimes it sounds like it's somebody maybe within the same room as me, or sat next to me'). Other voice-hearers in these transcripts seem to individuate their voices by personality trait (e.g. Sean individuates one of his voices as an angry person).

The take-home message from this is that the grounds on which is built the singular representation that constitutes the voice (Concept 2) as a persisting entity over time can vary enormously, as can the 'depth' of agency represented in the experience (Wilkinson and Bell, 2016). These grounds, however, cannot constitute the singular representation itself. For example, one of Dan's voices, though originally bound to a singular representation on the basis of location, could, in principle, move location. Similarly, Sean's angry voice might cheer up. Note that there must be this singular representation, however sparse, in order to enable enough continuity over time to allow the voice-hearer to judge that (or even wonder whether) *this* voice, heard now, is the same as *that* voice, heard yesterday.

A Paradox in Voice-Hearing

There is one final phenomenological feature of these transcripts that we would like to mention—a feature that further highlights the experiential complexity of voice-hearing, while also reaffirming the need to remain committed to preserving this complexity in our descriptions and analyses. This complexity can be seen by highlighting what appears to be a kind of tension, or contradiction, running through many of the transcripts. In short, the experience of hearing voices is described as being both like *and* unlike voices heard in everyday life (often simultaneously).

On the one hand, the voices are regularly described as having an experiential profile like voices heard in everyday life. They have auditory qualities and speech-like properties that everyday voices have: they have thematic or expressive content that is often directed towards the hearer (sometimes positive and affirming, although more often negative and distressing), and they are bound to distinct owners in that they are tied to re-identifiable agents over time. On the other hand, even though voices share these qualities with everyday voices, they also present other qualities that further complicate their phenomenological profile, and are thus often experienced as radically *unlike* everyday voices in some important ways. This can be seen in the way that some of the participants describe the experience of voice-hearing as being clearer, or somehow more pronounced, than the experience of perceiving a voice in everyday

life. In other words, the voice is experientially manifest with an intensity that can be overwhelming. Of course, similar experiences can happen when hearing voices in everyday life, such as when the ambient vocal noise at a pub or party becomes too much and one flees outside for an auditory break. However, unlike the pub and party cases, these participants are describing experiences of voices that seem to emanate from *within*; their overwhelming character comes not from a relentless penetration of sounds coming in through one's ears (although some have this quality, which further complicates things), but rather from the way in which they bubble up from within one's inner psychoacoustic landscape. Accordingly, these experience can seem 'more real than reality' (Karlsson, 2008) because the voice is not felt to have an external origin. Whereas voices from the external world are mediated by a variety of factors (the speaker's body, occlusion by other ambient noise, etc.) that specify their felt character, many of the voices in these reports lack this public character.

The phenomenological lesson, then, is that many of the voice-hearing experiences described in these transcripts seem to have a peculiar form of directedness, or what phenomenologists call 'intentionality' (Ratcliffe, 2017, p. 91). The voices are experienced, on the one hand, as somehow less real because they are not rooted in the public world; yet, on the other hand, they are also felt to be more real *for precisely this reason*. The hearer cannot distance themselves from the voice, which is why they are often perceived as having extraordinary power over the hearer (Chadwick and Birchwood, 1994) and compel hearers to feel as though they lack the agency to reinterpret or repudiate their negative appraisals.

There is much more that could be said about these transcripts, of course. But the takeaway message is simple: voice-hearing in psychosis is complex, both conceptually and phenomenologically, in ways that far exceed the standard technical term 'auditory verbal hallucination.' Acknowledging this complexity can help us refine our descriptions, deepen our understanding, and, ideally, develop a more empathetic stance towards voice-hearers and their rich variety of voice-hearing experiences.

References

- Chadwick, P., and Birchwood, M. J. (1994). The omnipotence of voices: a cognitive approach to auditory hallucinations. *British Journal of Psychiatry*, 164, 190–201.
- Deamer, F., and Hayward, M. (2018). Relating to the speaker behind the voice: what is changing? *Frontiers in Psychology*, 9.
- Deamer, F., and Wilkinson, S. (2015). The speaker behind the voice: therapeutic practice from the perspective of pragmatic theory. *Frontiers in Psychology*, 6.
- Karlsson, L.-B. (2008). 'More real than reality': a study of voice hearing. *International Journal of Social Welfare*, 17(4), 365–73.
- McCarthy-Jones, S., Krueger, J., Larøi, F., Broome, M., and Fernyhough, C. (2013). Stop, look, listen: the need for philosophical phenomenological perspectives on auditory verbal hallucinations. *Frontiers in Human Neuroscience*, 7(127), 1–9.
- Ratcliffe, M. (2012). Phenomenology as a form of empathy. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 55(5), 473–95.

- Ratcliffe, M. (2017). *Real Hallucinations: Psychiatric Illness, Intentionality, and the Interpersonal World*. Cambridge, MA: MIT Press.
- Siegel, S. (2006). Subject and object in the contents of visual experience. *Philosophical Review*, 115(3), 355–88.
- Slade, P. D., and Bentall, R. P. (1988). *Sensory Deception: A Scientific Analysis of Hallucination*. London: Croom Helm.
- Wilkinson, S. (2019). Hearing soundless voices. *Philosophy, Psychiatry and Psychology*, 26(3), 27–34.
- Wilkinson, S., and Bell, V. (2016). The representation of agents in auditory verbal hallucinations. *Mind and Language*, 31(1), 104–26.

About the Authors

Sam Wilkinson is a Senior Lecturer in Philosophy at the University of Exeter. As a Postdoctoral Research Fellow on Hearing the Voice, he was the project's resident conceptual hygienist but also worked on predictive processing and the experience of agency.

Joel Krueger is an Associate Professor in Philosophy at the University of Exeter and a former Postdoctoral Research Fellow on Hearing the Voice. He works on various issues in phenomenology, philosophy of mind, and cognitive science, especially on topics in embodied and extended cognition, emotions, and psychopathology.