

9. When Inner Speech Misleads

Sam Wilkinson and Charles Fernyhough

This chapter examines whether and when the experience of inner speech can be inaccurate and thereby mislead the subject. It presents a view about the representational content of speech experience generally and then applies it to inner speech in particular. On such a view, speech experience typically presents us with far more than simply the low-level acoustic properties of speech: it conveys the relevant mental states of the (actual or hypothetical) speaker. Similarly, inner speech presents inner speakers with their own mental states. In light of this, inner speech can mislead either by presenting the subject with mental states they do not in fact have, or by presenting these mental states as belonging to another agent. The chapter reflects on the sorts of contexts in which either of these could occur.

9.1. Introduction

Most philosophers think that at least some experiences have representational content: they represent the world as being a certain way.¹ Representational content dictates accuracy conditions, namely, what would need to be the case in order for the experience to be accurate. Inner speech, that “interior monologue” or familiar voice inside your head, is something that we experience, and that experience of inner speech seems to have representational content: it seems to “tell” the subject that something is going on in the world. Our central question is: What is it that the experience of inner speech is telling the subject is going on in the world, and could it, in some circumstances, be telling the subject something inaccurate? In other words: When, if ever, does the experience of inner speech mislead?

This may seem like a strange question to ask, and its importance may not be immediately obvious, but answering it has a number of significant implications. To start with, the

NLM Citation: Wilkinson S, Fernyhough C. When Inner Speech Misleads. In: Langland-Hassan P, Vicente A, editors. Inner Speech: New Voices [Select Chapters]. Oxford (UK): Oxford University Press; 2018. Chapter 9.

This chapter is open access under a CC-BY license.

Monographs, or book chapters, which are outputs of Wellcome Trust funding have been made freely available as part of the [Wellcome Trust's open access policy](#)

¹ We say “some experiences” because, although it is uncontroversial among representationalists (those philosophers who buy into the notion that experiences can have representational content) that, e.g., perceptual experiences have representational content, it is contentious whether other experiences that are less clearly about the world, e.g., pains, orgasms, etc. have such content.

question about whether the experience of inner speech can mislead requires us to answer a more basic question first: what sorts of things enter into the representational content of an experience of inner speech? This question is of tremendous importance since it tells us what the epistemic weight of an experience of inner speech is, namely, the content that it carries. In particular, if we view the experience of inner speech as important to self-knowledge, the content of the experience will tell us more precisely what the route to that self-knowledge is.

A more specific implication of an answer to this question is that there are unusual experiences (often in the context of psychiatric diagnoses), such as auditory verbal hallucinations (AVHs), which are taken by a number of theorists to involve inner speech (Frith 1992, Seal et al. 2004, Jones & Fernyhough 2007). If we think of AVHs as experiences of inner speech, we can usefully ask ourselves: is this experience of inner speech telling the subject something inaccurate? And if it is, what aspects of the world aren't, and which are, the way they are represented as being?

At this point it is important to clarify two things. First, there is the question of what exactly we mean by an "experience of inner speech". Some might want to say that inner speech simply is an experience. Others might want to say that inner speech is something that we do, and which we have an experience *of*. At this stage we remain neutral between these two, but it will become clear later on that our position is more in line with the latter. Second, it is important to clarify that we are talking about the *experiential content* of an experience of inner speech and not its *linguistic content*. We are not talking about utterances of inner speech linguistically expressing inaccuracies. Thus to draw an analogy with *outer* speech experience, if someone says "Madrid is the Capital of France", although they have said something inaccurate, my *experience* is accurate to the extent that it has accurately represented various features of the utterance, for example, the speech sounds produced, and perhaps more besides (a central part of this chapter is the controversy surrounding this). Now, the extent to which this analogy with outer speech holds is itself up for dispute and will depend upon how we think of inner speech.

We proceed as follows. We start by presenting an intuitively appealing view according to which an episode of inner speech is an imaginative episode, and therefore cannot mislead (at least not in the relevant sense). We criticize this view and reject it in favour of the view that inner speech is actually a kind of speech, rather than merely imagined speech. We then present a view about the representational content of speech experience generally, and then apply it to inner speech in particular. We end, in light of this, by presenting the different ways in which inner speech could potentially mislead.

9.2. Content without Commitment: Inner Speech as Imagination

It is important to distinguish representational content from psychological force. Perceiving and believing have representational content, but they also have a certain psychological force: they don't merely represent a content, they represent that content *as accurate*. In other words, they involve, by their very nature, a certain commitment to the

world being the way represented.² Other psychological states or events (such as suppositions or certain imaginings etc.) on the other hand may represent something but lack that kind of commitment to what is going on in the world. If you voluntarily imagine a pink unicorn, it cannot be regarded as an inaccurate experience just because there is no such thing in front of you (or no such thing in existence at all).³ The experience is not even in the running for accuracy. That said, the experience has representational content: it is of (or represents) a pink unicorn. What you have in this case of imagining is content (something is represented, it is about something) without commitment to accuracy. Another way of thinking about this lack of commitment to accuracy is that the imaginative episode is not presenting an aspect of the world over and above the experience itself (and so, trivially, it cannot do so inaccurately).

Some might think that inner speech is like that. Inner speech, on this view, is like imagining yourself speaking (and hearing yourself speak). The experience does not inform you about something going on in the world, and, as such, it cannot be wrong since the world cannot act as a benchmark against which the experience can fall short. There simply is the experience, presenting itself pure and simple. At most, if this is true, an inner speech experience tells you the immediate and infallible fact that you are having that very experience. On such a view, inner speech may represent certain things, which would be reflected in the phenomenology of inner speech, much in the same way as imagining a pink unicorn represents certain things (like pinkness and unicorns), and this too is reflected in the phenomenology of the experience, but neither experience purports to tell you anything about the world beyond the experience itself. On such a view, inner speech, as a variety of imagination, cannot be inaccurate: it just is what it is.⁴

But is inner speech an instance of imagination? We think that the answer is no. A crucial step to seeing why this is the case involves an appreciation of the distinction between imagination and imagery. Imagination is a whole psychological event in its own right. *People* are engaged in acts of imagination. These acts of imagination enable them to appreciate, in potentially many different ways, non-actual scenarios, and, when they are engaged in such acts, they may be motivated to do so by a number of different things. They may be trying to judge whether they could have jumped over that river, reason about a social situation, or simply engage in imagination for the pleasure of it. Furthermore it is in the nature of imagination to have content without commitment (which is not to say

² In perception, it is a commitment that you can override: you don't have to take your perceptual experience at face value.

³ The term "imagination" gets used in lots of different ways for different purposes, for example, there are imagistic and propositional forms of imagining. What we mean by imagining is simply a mental state (or, better, episode) that represents something and has no commitment to its reality or actuality (it is hence to be contrasted with judgement and perception, which *do* have such commitments). Thus imagination may or may not recruit imagery, and is certainly not synonymous with imagery.

⁴ To put it another way there is no appearance/reality distinction. Since the phenomenon is an appearance, the appearance *is* the reality.

that it cannot serve, and fail to serve, a given function). These acts of imagination often will recruit or make use of imagery in many modalities, but there will also be aspects to the imaginative experience that aren't purely imagistic. Imagery, in contrast, is not *in itself* a complete psychological event. It features *as a component* of such events. Whereas people imagine things, people don't "imagine" or "do imagery". When people imagine things, imagery may be involved, but it is not all that is involved. And, crucially, imagery is also involved in many psychological events that aren't imaginings. For example, imagery may be involved in episodic recollections. It may even be involved in certain judgements (see, e.g. Langland-Hassan 2015). In other words it may be involved in psychological events that, unlike imagining (in the sense that we are using the term), have an inbuilt commitment to how things are (or *were*, in the case of memory) in the world.

In light of this, it is too quick to move from the (accurate) observation that inner speech involves *imagery* to the conclusion that an episode of inner speech is a case of *imagination*. And if it is not a case of imagination then it seems, at least in principle, that, as an experience, it can be committed to telling you something about the world.⁵

9.3. Inner Speech as Speech

If inner speech is not imagination, then what is it? In line with a number of other theorists (Vygotsky 1987/1934, Fernyhough 1996, Martinez-Manrique & Vicente 2010) our answer is: *it is speech*. It is speech in two important senses. First, it is a productive rather than recreative activity. Second, its primal use is in making *speech acts*: asserting, questioning, insulting etc. We take these points in turn.

9.3.1. Inner speech as productive rather than re-creative

To see the productive rather than re-creative nature of inner speech we need to ask ourselves not just, "What is inner speech?" or "What does it look like once developed?" but also: "How and why did it develop?" One attractive theory (which originates in Vygotsky 1987/1934), which carries both evolutionary and developmental plausibility, states that inner speech starts off as speech (namely, outer or "overt" speech). That is to say, whatever function inner speech plays, once it has developed, is played by outer speech in children who have not yet developed the capacity to engage in inner speech. This capacity to engage in inner speech is usually seen as partly constituted by the capacity to inhibit the overt production of speech.

According to this story, inner speech is the end product of a developmental trajectory that begins with private speech. "Private speech" refers to outer speech that is not produced for

⁵ One natural question at this point is whether inner speech can ever be an instance of imagining. This is a tricky question. In the first instance we want to say that paradigmatic inner speech isn't imagination. But the question of whether inner speech can sometimes be an instance of imagining seems to get things the wrong way round: imaginative episodes may be enabled by inner speech, but inner speech is not constructed out of imaginative episodes.

the benefit of anyone other than the speaker. Young children will first, under the guidance of a caregiver, learn to reason verbally, but out loud, for the benefit of guiding their thinking and attention. Over time, they learn to “internalize” this speech, to inhibit its overt production. However, as with many cases of motoric inhibition, vestiges of the motor processes remain. Evidence of motoric involvement in inner speech has been empirically supported by several electromyographical (EMG) studies, measuring muscular activity during inner speech, some of which date as far back as the early 1930s (e.g. Jacobsen 1931). In short, these discovered that, when you engage in inner speech, muscles in the face and throat, associated with speaking, are activated (see also Rapin et al. 2013).

There have been brain-imaging studies (fMRI) presenting results that are very much in keeping with the distinction between a productive phenomenon, namely, inner speech proper, and a re-creative imaginative phenomenon, imagined speech. In particular, Tian & Poeppel (2012) and Tian, Zarate, and Poeppel (2016) have shown that there are two very different ways of generating auditory-verbal imagery, namely, of activating relevant areas of auditory sensory cortices in the absence of external sensory stimulation. One, which corresponds to inner speech (which they call “articulation imagery”) is induced through “motor simulation”, i.e., is initiated “top-down” by activation in areas of prefrontal and motor cortex associated with speaking. The other, which corresponds to inner hearing/imagined speech, is induced, in line with more standard accounts of imagery (including in other modalities, such as vision), via a memory-based mechanism (e.g. Kosslyn 1994), i.e., by the re-creation of a sensory event (derived, to some extent, from past sensory events). While the former mechanism involves trying to produce something directly (and its inhibition results in imagery being activated as part of the sensory predictions of the completed action), the latter involves trying to re-create the sensory effects of a past or constructed scenario. There is a sense in which imagining hearing something entirely new (i.e., not previously experienced) is “producing something”, but not in the same sense that inner speaking is productive. Unlike the latter, it involves the recreation of the sensory effect of an event, in this case an event that has never happened.

This distinction between a productive and re-creative phenomenon may map onto a phenomenological distinction between two different forms that auditory-verbal imagery can take. Using descriptive experience sampling (DES), Hurlburt and colleagues (Hurlburt, Heavey, and Kelsey 2013) isolated two differently reported phenomena: “inner speaking” on the one hand, and “inner hearing” on the other. The former may correspond to the top-down mechanism of generating imagery that Tian and colleagues isolated; the latter, to the more bottom-up mechanism. Nevertheless, equating Hurlburt’s “inner speaking” with “inner speech” does not suffice to show that “inner speech” is not a case of imagination. The reason for this is that it seems plausible that inner speaking can take part in imaginative episodes as well as in more authentic or ecologically valid instances of inner speech. If you imagine yourself going up to someone and speaking to them, nothing prevents this from engaging the sort of top-down imagery that Tian and colleagues isolate, or in having phenomenological features more akin to inner speaking than to inner hearing. What we actually need is three-way distinction among the phenomena that make

use of auditory-verbal imagery: (i) a genuinely productive phenomenon (which we are about to introduce, and which constitutes ecologically valid inner speech); (ii) a re-creative productive phenomenon (like the case of imagining yourself speak to someone, which involves inner speaking); and (iii) a re-creative sensory phenomenon (inner hearing). Whereas (ii) involves the same (or much of the same) apparatus as (i), it is used in a different context and for a different purpose (i.e., for the recreation of a counterfactual scenario). On the other hand, (iii) recruits sensory imagery for a similar re-creative activity as (ii). The genuinely productive phenomenon, namely, (i), is what we examine now.

9.3.2. Inner speech acts as the main form of inner speech

Following Roessler (2016) we can distinguish between a “mere act of inner speech” and an “inner speech act”, in a way that perfectly mirrors the distinction between a “mere act of speech” and a “speech act”. Although there are different accounts of speech acts (see Austin 1962, Searle 1969, Bach & Harnish 1979 for some classic formulations) everyone agrees that speech acts are closely tied to the speaker’s mental state in a way that mere acts of speech are not. If you change the mental state in relevant ways, then you change the speech act in relevant ways. Indeed, if you remove the mental state, then you thereby remove the speech act altogether. Examples will make things clearer. Reciting a poem, or repeating an address so as to remember it, is an act of speech, but it is not a speech act. This is, in part, because the speaker, in reciting, or repeating, does not mean what is being said, and any potential variations in the subject’s mental states are compatible with the same act being performed (and variations in what is repeated or recited do not thereby signal similar variations in the subject’s mental states). In stark contrast, sincerely asserting, requesting, demanding, questioning are speech acts. These require the person performing them to be in certain states of mind. For example, an assertion (if sincere) requires the asserter to believe what they are asserting, a question (if sincere) requires the questioner to have the desire to know the answer to the question, and so on.

This fact adds further weight to the point that inner speech is not imagined speech, but rather is speech. Consider the following:

1. Jane asserted that p
2. Jane imagined asserting that p
3. Jane asserted in inner speech that p

Whereas 3 implies 1, 2 does not. In fact, if anything, 2 implies that 1 is false: merely imagining asserting rules out actually asserting (just like imagining raising your right hand rules out you actually doing so). On the other hand, an assertion in inner speech is a perfectly good instance of assertion.⁶ And insofar as 1 and 3 are both assertions, they both, if sincere, require that Jane be in a certain mental state (i.e., believing that p). In a

⁶ Things are somewhat complicated by the fact that some theorists (e.g. Searle 1969) make it a requirement that an assertion have an interlocutor. It seems to us that we regularly make private

related manner, assertions that *p* are treated as evidence for the attribution of the mental states that they (if sincere) require (or express), in this case, believing that *p*. Thus if someone asserts, "Paris is the capital of France", you will (other things being equal) think that they believe that Paris is the capital of France. The same applies to other kinds of speech acts, and other kinds of speech acts are intimately tied to other kinds of mental state. Orders and requests are tied to goals, questions are tied to desires to know, compliments are tied to positive evaluations, insults to negative evaluations, etc. And when people request, question, compliment, or insult, if we take them to be sincere, we thereby take them to be in those mental states.

Of course, there is one rather perplexing feature of inner speech, construed as an inner speech act, which is: why do we engage in it at all? Usually when we assert, question, or insult in outer speech, we have an addressee. We are speaking our minds to someone else. When we assert, question or insult in inner speech, who are we doing it for? Who are we speaking our minds *to*? The answer is: ourselves.

Organisms that live in groups, that cooperate and communicate, can do so very successfully without inner speech, and also without the need to directly introspect. They simply need to express themselves to their conspecifics. These communicative acts do not require the organism to have reflected on, or even have prior access to, its own mental state: the expression can be spontaneous and unreflective. However, once produced, these communicative acts can be perceived and interpreted by the agent who produced them. But of course, this cannot be regularly used as a way of accessing your mental states, since that would involve making your beliefs, desires, plans, and evaluations entirely public. That would often be, at best, socially unacceptable, and, at worst, downright dangerous. Inner speech can be understood in part as a solution to this problem of indiscretion: it is a way of expressing, and hence accessing and reflecting upon, your own state of mind without thereby having to risk giving that information away to others.⁷

There are many theorists who would be in general agreement with this picture (e.g. Jackendoff 1996, Clark 1996, Carruthers 2011). One interesting feature of positing this role for inner speech is that it suggests that we (at least sometimes, perhaps always) lack other more direct means of reflecting on our mental states. Our view is that inner speech certainly helps a great deal with reflection on our minds, but there are certainly ways of so reflecting that don't make use of inner speech.

assertions (and that these carry the same features as "normal" assertions, e.g., have the same sincerity conditions). This can either be accommodated by (contra Searle) removing the dialogic requirement, or by claiming that human inner speech is in some important sense dialogic. As will become clearer, we would opt for the latter.

⁷ We say "in part" because although inner speech, like outer speech, offers us improved self-knowledge, it doesn't always, or even often, serve that purpose. Much of the time it regulates our behaviour and focuses our attention.

9.4. The Experiential Content of Speech Experience

If inner speech is, in an important sense, speech, it is reasonable to assume that we may learn about its content by examining the content of speech experience *per se*. As it happens, there is currently a lively philosophical debate about the content of the auditory experience of speech (see O’Callaghan 2011 and Brogaard forthcoming). This debate is a specific version of a more general debate about the content of perceptual experience generally. There are those, sometimes called “liberals”, who want to allow that “high-level properties” can enter into the contents of perceptual experience (e.g. Siegel 2006, Bayne 2009), and there are those, sometimes called “conservatives” (e.g. Dretske 1995, Tye 1995), who claim that only “low-level properties” can.

A concrete case will be helpful here. Suppose you are looking at a green apple, and suppose that you know that it’s a Granny Smith apple, and, furthermore, that it was grown in Chile. The apple in question has:

- (i). a certain shape and colour
- (ii). the property of being an apple
- (iii). the property of being a Granny Smith apple
- (iv). the property of having been grown in Chile

The further down this list you go, in terms of accepting that it could enter into the content of perceptual experience, the more “liberal” you are about the admissible content of perceptual experience. Even the most liberal of liberals will admit that (iv) just isn’t the right kind of property for your perceptual experience to convey. You may come to know that the apple was grown in Chile, but you can’t have known that solely on the basis of your perceptual experience. Liberals, however, may claim that (iii) can enter into the content of perceptual experience for, say, someone familiar with Granny Smiths. And they will certainly say that (ii) enters into the content of perceptual experience for those of us familiar with apples. The conservative, on the other hand, wants to say that only (i) is the purview of perceptual experience: (ii) and (iii) go beyond what perceptual experience can represent.

This debate came to prominence in the light of a classic argument in favour of liberalism that proceeded by presenting what might be called “contrast cases” (see Siegel 2006 for perhaps the classic example of a contrast case). In contrast cases, you compare two cases where the “low-level” properties represented (e.g. colour and shape) remain constant, but the “high-level” properties represented are different because, in one of the cases, the high-level properties cannot be represented due to lack of knowledge or expertise. For example, looking at one and the same oak tree will be phenomenologically different depending on whether you know nothing of tree species, or whether you are an expert. The idea is that the two cases differ in specifically *perceptual* phenomenology, and that this should be attributed to the representation of high-level properties in perceptual experience. The expert, in automatically recognizing the oak, has represented *in her perceptual experience* the property of being an oak, whereas the novice hasn’t.

When applied to speech perception, the very same phenomenal contrast arguments can be used, and are perhaps even more convincing since language is an area where expertise has especially powerful effects on experience. If you think about the phenomenological difference between hearing a language that you understand and one that you don't, it seems plausible that understanding a spoken language makes it *sound* different (see, e.g. Strawson 1994). This has led some people to attribute this to the representation of meaning in auditory speech experience. Thus, you don't merely get loudness, pitch, and timbre represented: you also get "high-level" properties like meaning (in a way that is akin to how you get the high-level property "oak tree" in the visual case).

O'Callaghan (2011) has recently criticized the view that meanings are represented in the auditory experience of speech. He does, however, accept that there is a phenomenological contrast between hearing speech when you understand the language and when you don't, and he accepts that the contrast is one of perceptual (rather than emotional or cognitive) phenomenology. What he thinks explains the difference is the representation, in one instance, of, not the standard low-level properties of loudness, pitch, and timbre, but properties a bit "higher" (we might call them mid-level properties), namely language-specific phonological properties ("language-specific" in the sense of specific to, say, French as opposed to German).

O'Callaghan's reasons for adopting such a view stem from another contrast case that compares homophones. He claims that there is no phenomenological difference between hearing homophones, even if we perceive them as having different meanings. So, to take an example, if we hear an utterance of "bank" (the financial institution) and "bank" (the edge of a river), they sound the same. As a result, O'Callaghan claims that it isn't meaning that explains the phenomenological difference, since here we have different meanings but the same phenomenology. Rather, what better explains the difference between hearing languages you do and don't understand is familiarity or expertise with the phonology of the known language, which affects the temporal and qualitative features the relevant speech sounds are experienced as having.

As Brogaard (forthcoming) rightly points out, this argument from homophones has the weakness that it arguably isn't words that are the relevant vehicles of meaning, but entire utterances, namely, sentences used in context. We would go a step further and say that, whatever "meaning" is taken to be (it refers to different abstract entities for different purposes) the relevant sense in which meaning is represented in speech experience is in the sense of "speaker meaning", namely, the underlying mental state of the speaking agent that is expressed by the speech act. What makes it the case that two assertions of "I'm going down to the bank" are experienced differently based on attributing different meanings to the word "bank" is that in one case you take the speaker to be expressing (their belief in) their intention to go to a financial institution, while in the other you take the speaker to be expressing (their belief in) their intention to go to the edge of the river. That said, the phenomenological difference between the two uses of "I'm going down to the bank" is very subtle, and some people may deny its existence. Clearer examples are cases of syntactic ambiguity ("I'm glad I'm a man and so is Lola"), or cases of sarcasm (e.g.

saying “Well done” in a berating, rather than congratulatory manner). Of course, in such cases (especially sarcasm) the acoustic properties of the utterance are often altered by the speaker in order to promote one interpretation over the other. This, however, doesn’t mean that two identical speech sounds won’t be experienced as phenomenologically different if interpreted differently.

However, the conservative can say that, although there is a phenomenological difference, it is attributable to phenomenological differences associated with *judgements* about the speaker’s mental state, rather than experiences of these mental states. Thus when I hear “I’m glad I’m a man and so is Lola”, it is phenomenologically different to *judge* that the speaker is expressing gladness that he and Lola are both men, than to judge that the speaker is expressing that both Lola and he are glad that he is a man. In other words, the phenomenology is different, but it is not *experiential* phenomenology. One problem with this suggestion is that the relevant phenomenology remains even when we know that the speaker doesn’t have that mental state (e.g., is acting on stage), or when the speaker is just a vague, hypothetical construct (e.g., as when abstractly considering different utterances, or hearing announcements at the train station). It doesn’t therefore seem that it can be something to do with judgement. Granted, it could be a phenomenology associated with something less committal than judgement, but that remains non-experiential. Whatever this state may be, the phenomenology seems stimulus-bound, bound to the experience, so why not view it as part of the experience?⁸

The other thing that the conservative might say, which is very much in line with what O’Callaghan says, is that judging, or even merely hypothesizing that a speech sound expresses a certain (even hypothetical) mental state has a top-down influence on how the low-level stuff is experienced. That doesn’t mean that *perceptual* experience represents anything over and above those low-level properties. The difference in phenomenology is indeed a difference in properly perceptual content, but this difference just is a difference in those low-level properties. In other words, a premise of the liberal’s contrast case doesn’t hold, since the low-level properties aren’t being kept constant after all.

This seems like a plausible response, but then the debate becomes one of conceptual cartography. What do you mean by “*perceptual* experience”? In particular, the liberal could just say that these and similar top-down influences are so rife in even the most basic forms of perception, that that just is what perceptual experience is. If even the experience of low-level properties is enabled by top-down influences, where do you draw the line? One might say that you draw the line at that which remains the same when the sensory inputs are kept the same. But arguably, even at the very front line of sensation, top-down effects have influence (see e.g. Lee 2002 for vision, Davis and Johnsruide 2007 for

⁸ When you hear the Kinks song, Lola, you don’t literally attribute different mental states to Ray Davies depending on how you disambiguate “I’m glad I’m a man and so is Lola”. But you do experience it differently depending on how you disambiguate it, and this comes down to hypothetical mental state attribution (namely, the mental state that you would attribute *if* you took it seriously).

audition). And if top-down influences enable us to hear sounds a certain way, why not allow that top-down influences enable us to experience meanings? Granted, such a response on the part of the liberal raises problems about sensory modality. If the speaker meaning (the mental state, real or hypothetical) behind an utterance is represented in perceptual experience, then surely it must be represented in a sensory modality, namely, audition? But isn't it implausible to say that mental states are *auditorily* represented? I surely cannot literally *hear* your beliefs.

It doesn't matter for our purposes whether the phenomenological changes in the experience are down to properly "perceptual" features of the experience, or to some kind of non-perceptual experiential accompaniment (e.g., some kind of stimulus-bound cognitive or affective phenomenology). What matters to us here is that the overall experience of speech is extremely representationally rich, regardless of whether all of the features can be thought of as perceptual or non-perceptual. In particular, we think that, as well as the low-level features of loudness, pitch, and timbre, the states of mind underlying utterances can at times also be part of what is represented in the experience of those utterances.

9.5. The Experiential Content of Inner Speech

Although we think that the lessons are transferrable, we cannot assume that the content of your experience of someone else's outer speech is similar to the content of your own inner speech. To argue for this step by step, it is helpful to move from experience of someone else's speech to a qualitatively intermediate stage on the way to inner speech: the experience of your own speech spoken out loud.

So, what is the difference between hearing someone else's speech and experiencing your own speech? First of all, you don't only experience your own speech by hearing it. You are proprioceptively and sensorially aware of your speech production apparatus. But that's not the only thing: you have a sense of agency, in both the sense that you tend to be aware that it is *you* who is speaking, and also in the more specific sense that *what* you say tends not to come as a surprise. In spite of this difference, you also tend to know who is speaking when you hear others speak, and what comes out of the mouths of people you know really well tends not to come as a surprise either (and conversely you can sometimes surprise yourself). Another difference between your speech (and your action generally) and the speech of others, is that it is embedded in a rich and pervasive context that you (normally) have unparalleled access to (not least because you are always with yourself). You tend to speak as part of an overall complex serial process (namely, your life) in the service of your plans, goals, habits, machinations. And you are there to witness it all, effortlessly taking in the past and projecting into the future.

We agree that there are major asymmetries between the epistemology of your own speech and the speech of others. These are asymmetries that parallel the difference between experiencing yourself act, and perceiving others act. That said, these are epistemological differences, rather than differences in experiential content. Your access to the experiential

content of your own speech may be more direct, more secure, and aided by a pervasive context, but that doesn't mean that it doesn't have a similar kind of content to that of someone else's speech. Thus, we suggest, the experience of your own outer speech, like the experience of someone else's speech, doesn't only represent the low-level acoustic properties of your speech (as well as low-level features that are lacking in the experience of someone else speaking, such as tactile and proprioceptive information about your speech production) but also mental state information.

It is a small step from the experience of our own outer speech to the experience of our inner speech. Of course, the experiences are qualitatively rather different, but they achieve, or at least can achieve, the same thing. For example, you can (in situations when you are alone, or social norms allow it), replace your inner speech with outer speech with very much the same effect. Encouraging yourself with "Come on!" during a game of tennis, or asking yourself "What did I come upstairs for?" can be done in either inner or outer speech with similar effect (although there may be an added motivational effect to saying the former out loud).

In outer speech, there is lots of fine-grained auditory information in the content of the experience. If you were to mishear the pitch at which you were speaking, that would be relatively unimportant in most cases, but it would still be an inaccurate experience. You can imagine someone who hears pitch distortions, but still manages to pick up on the content and nuances of utterances. Another interesting case to reflect on in this instance is when congenitally deaf people learn to speak. In these cases they are producing speech sounds that they themselves cannot hear, for the benefit of a hearing interlocutor. But in what sense is their own experience, with its proprioceptive and tactile feedback elements, failing to adequately represent their speech? Sure, it is not representing the sounds they are producing, but does that mean that it is not still representing what is by far the most valuable aspect of the speech, namely, what is being conveyed? Clearly the deaf speaker has an experiential appreciation of what they are saying in the absence of hearing what they are saying. In short, even in cases where sounds are being produced, what is more significant are the mental states—the speaker meanings—expressed in speech.

How does all of this apply to inner speech? Given that, in inner speech, there is no real-world auditory information to accurately represent, the information about mental states seems to be even more at the heart of what is carried by the vehicle of inner speech. An experience of inner speech, insofar as it is of an inner speech act, typically represents the state of mind expressed by that speech act, and, however minimally, the individual whose state of mind it is, namely, *you*. Thus when you assert something in inner speech, the conscious experience of that represents your belief in that which you have asserted, and, somewhat trivially, represents it as belonging to you. This much can also be said about hearing someone (yourself or someone else) sincerely assert something in outer speech. However, in contrast to this, it is hard to see how things like loudness, pitch, and timbre (or even phonology) can be represented *in a relevantly committal way* in inner speech. They may be (and often probably are) represented insofar as they contribute to the phenomenology of the experience (just like an imagining of a pink unicorn represents a

pink unicorn in a way that is reflected in the episode's phenomenology), but it seems that there is no feature of the world that would make that aspect of the experience accurate or not.⁹ To make matters clearer, consider the fact that in hearing your own outer speech, you might mishear, e.g., the pitch at which you spoke, and this is determined by an objective feature of the world (namely, the actual pitch of the sounds you produced). It is not clear how something like this would work for *inner* speech. Although you could imagine someone complaining to a doctor and saying "I'm hearing my own voice 2 tones lower than it actually is", such a complaint would make no sense in inner speech. There is no epistemic distance between how inner speech really is in terms of pitch and how it is experienced as being (just like there is no epistemic distance between my imagining of a pink unicorn and how it is experienced as being: how it is experienced as being constitutes the imaginative episode). However, in contrast, the mental and agentive aspects are the same in both inner and outer speech. And whether your experience of those aspects is accurate is determined by objective features of the world, namely, your actual state of mind. Such objective features are precisely what enables us to draw the boundary between sincerity and insincerity. "Great drawing!" you might say, in response to your friend's woeful attempt at a sketch, with a feigned air of sincerity in your voice so as not to offend them. The actual insincerity of that speech act is an objective fact about the world, largely determined by the fact that you don't *really* positively evaluate their drawing. Similarly, saying, "I'm such an idiot!" in inner speech is accurate to the extent that you are genuinely reprimanding yourself, which, like any speech act, requires you to be in a very particular mental state.

9.6. The Ways in Which Inner Speech Can (and Can't) Mislead

As we've already mentioned, it is not clear that auditory or even phonological properties which may well be represented in inner speech (as reflected in the phenomenology of some inner speech) are subject to inaccuracy. They are cases of "content without commitment", since it really isn't clear what objective feature of the world, over and above the experience of those auditory or phonological properties themselves, might act as a benchmark against which they could fall short. In contrast, the agent and their mental state is precisely such an objective part of the world, and is one that, crucially, an inner speech act may well, in principle, misrepresent. Now let's examine ways in which this can be misrepresented.

In producing an inner speech act and becoming aware of it, you become aware, in the good case, that (i) the mental state expressed is such and such (i.e., the speech act has a certain meaning), and (ii) you are the agent of the speech act (i.e., it is you, and not someone else, who has the mental state the speech act expresses). As a result, it seems to us that you can in principle have two errors, which might sometimes occur together.

⁹ Note that this will apply even in the case of having a dialogue with another person in inner speech. You might be talking to your mother in your head, for example, and be getting her voice all wrong, but you would still be talking to your mother.

- A. Misrepresenting the state of mind you actually have.
- B. Misrepresenting the agent of the speech act (namely, whose speech act it is).

We take it that, on reflection, A is relatively common. Both culpable insincerities and innocent inaccuracies regularly creep into our inner speech, and do so with negative impact upon our self-knowledge (although they may have positive impact in non-epistemic ways, e.g., on psychological comfort or well-being).

B, on the other hand, seems much less common. However, if the model according to which (at least some) auditory verbal hallucinations (AVHs) are misattributed episodes of inner speech is correct, then it might be something that one sees in those instances. And if that is the case, one interesting question, which to our knowledge has never been addressed in this way, is whether in these cases you get misrepresentations of just B, or of both A and B? In other words, does the speech act the voice-hearer experiences in some cases express a mental state that she herself “really has deep down”? In which case, is it something that is protectively disowned in a failure to recognize it as her own mental state? Some cases of AVH in the context of very strong feelings of shame and self-loathing may be like this.¹⁰ Or is it that the voice-hearer has produced an episode of inner speech that is somehow expressively inaccurate, namely, doesn't express a mental state that the voice-hearer actually has, and so is attributed to another agent (as in, for example, Stephens and Graham's (2000) model of ego-dystonic thoughts being misattributed as alien “voices”). This latter option in a sense does not involve the same degree of lack of self-knowledge. Although there is failure to detect self-production, the voice-hearer has in fact *accurately* detected that she doesn't have the relevant mental state. This then raises a number of interesting further questions. For example, if the episode of inner speech (the inner speech act) that constitutes the AVH is not an expression of the voice-hearer's own mental state, then *whose* mental state is it? Where does the voice (in the sense of the agent producing the speech act) *come from*? This may then lead us to hypothesize that voice hearers countenance rich and relatively autonomous representations of communicative agents (see Deamer & Wilkinson 2015, Wilkinson & Bell 2016). Then, of course, the question arises as to where this agent representation comes from. Why is it voice-hearers who have a propensity to represent agents in this way? Perhaps at this point it makes sense to suggest that it isn't only voice-hearers who have this propensity. Indeed it can be argued (see McCarthy-Jones & Fernyhough 2011) that normal inner speech can be dialogic and is shot through with representations of agents other than ourselves making speech acts. For example, reasonably large proportions of respondents endorse statements about hearing the voices of other people in inner speech (McCarthy-Jones & Fernyhough 2011, Alderson-Day et al. 2014). Hence our inner speech sometimes expresses mental states that we don't in fact have but that we hypothesize someone else might have. This would clearly

¹⁰ See Woods et al. (2015) for a recent phenomenological survey exploring, among other things, the varied emotional states that surround the experience of hearing voices (depression is reported in 29 per cent and shame in 14 per cent of their participants).

be helpful, and may play a crucial role in underpinning social, and perhaps even normative, reasoning.

If the inner speech model of AVHs is accurate, these reflections on inner speech acts add a dimension of complexity to the phenomenon in question. These experiences are not just straightforward hallucinations of sounds that aren't there (although in many cases they are partly this). They are experiences of mental states, had by people with mental states, and so admit of the different dimensions of inaccuracy explained above.

9.7. Conclusion

In this chapter, we have combined a particular view about the nature of inner speech with a liberal view about the representational content of speech experience. Our view of inner speech thinks of inner speech as speech in a number of important ways. It is a productive rather than re-creative phenomenon, and, as with speech, its normal ecological use is in the performing of speech acts. Our liberal view about the content of speech experience allows that, *in addition to* the speech sounds that you hear when someone speaks to you, what their speech means, where this is seen as speaker meaning (e.g. their communicative intentions), also enters into the content of the experience. Applying this to inner speech, although there appears to be no constraint of accuracy on the experience of speech sounds (since there are no objective speech sounds produced that could be accurately or inaccurately represented), there is a constraint for the agentive elements of the experience. The mental state that you happen to be in when you engage in inner speech is an objective fact, and your episode of inner speech could mislead you about it. Within this framework there is a great deal of work that could be done in ascertaining when and how people mislead themselves in inner speech, and as a result develop flaws in their self-knowledge. In more extreme cases, this approach could be used to explore cases of AVH.

References

- Alderson-Day, B., McCarthy-Jones, S., Bedford, S., Collins, H., Dunne, H., Rooke, C., & Fernyhough, C. (2014). Shot through with voices: Dissociation mediates the relationship between varieties of inner speech and auditory hallucination proneness. *Consciousness and Cognition* 27: 288–96. PubMed PMID: 24980910.
- Austin, J. L. (1962). *How to Do Things with Words*. Clarendon Press.
- Bach, K. & Harnish, R. (1979). *Linguistic Communication and Speech Acts*. MIT Press.
- Bayne, T. (2009). Perception and the reach of phenomenal content. *Philosophical Quarterly* 59 (236): 385–404.
- Brogaard, B. (forthcoming). In Defense of Hearing Meanings. *Synthese*: 1–17.
- Davis, M. H. & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing Research* 229: 132–47. PubMed PMID: 17317056.

- Deamer, F. & Wilkinson, S. (2015). The speaker behind the voice: therapeutic practice from the perspective of pragmatic theory. *Frontiers in Psychology* 6: 817. PubMed PMID: 26124738.
- Carruthers, P. (2011). *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford University Press.
- Clark, A. (1996). Linguistic anchors in the sea of thought? *Pragmatics and Cognition* 4 (1): 93–103.
- Fernyhough, C. (1996). The dialogic mind: A dialogic approach to the higher mental functions. *New Ideas in Psychology* 14: 47–62.
- Frith, C. (1992). *The cognitive neuropsychology of schizophrenia*. Hove: Lawrence Erlbaum.
- Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition* 22: 1477–94. PubMed PMID: 24184987.
- Jackendoff, R. S. (1996). How language helps us think. *Pragmatics and Cognition* 4 (1): 1–34.
- Jacobsen, E. (1931). Electrical measurements of neuromuscular states during mental activities, VII: Imagination, recollection, and abstract thinking involving the speech musculature. *American Journal of Physiology* 97: 200–9.
- Jones, S. R. & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition* 16 (2): 391–9. PubMed PMID: 16464616.
- Kosslyn, S. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Langland-Hassan, Peter (2015). Imaginative Attitudes. *Philosophy and Phenomenological Research* 90 (3): 664–86.
- Lee, T.S. (2002). Top-down influence in early visual processing: A Bayesian perspective. *Behaviors and Physiology* 77(4–5): 645–50.
- Macpherson, F. (ed.) (2011). *The Senses: Classic and Contemporary Philosophical Perspectives*. New York: Oxford University Press.
- Martínez-Manrique, F. & Vicente, A. (2010). What the ...! The role of inner speech in conscious thought. *Journal of Consciousness Studies* 17 (9–10): 141–67.
- McCarthy-Jones, S. & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and Cognition* 20: 1586–93. PubMed PMID: 21880511.
- O’Callaghan, C. (2011). Against hearing meanings. *Philosophical Quarterly* 61 (245): 783–807.
- Rapin, L., Dohen, M., Polosan, M., Perrier, P., & Loevenbruck, H. (2013). An EMG study of the lip muscles during covert auditory verbal hallucinations in schizophrenia. *Journal of Speech, Language and Hearing Research* 56: S1882–S1893.
- Roessler, J. (2016). Thinking, Inner Speech, and Self-Awareness. *Review of Philosophy and Psychology* 7 (3): 541–57.

- Seal, M. L., Aleman, A., & McGuire, P. K. (2004). Compelling imagery, unanticipated speech and deceptive memory: Neurocognitive models of auditory verbal hallucinations in schizophrenia. *Cognitive Neuropsychiatry* 9(1–2): 43–72. PubMed PMID: 16571574.
- Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Siegel, S. (2006). Which properties are represented in perception? In Tamar S. Gendler & John Hawthorne (eds.), *Perceptual Experience* (pp. 481–503). Oxford University Press.
- Stephens, G. L. & Graham, G. (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts*. Cambridge, MA: MIT Press.
- Strawson, G. (1994). *Mental Reality*. Cambridge, MA: MIT Press.
- Tian, X. & Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience* 6: 314. PubMed PMID: 23226121.
- Tian, X., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex*, 77, 1–12. PubMed PMID: 26889603.
- Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1987). Thinking and speech. In R.W. Rieber & A.S. Carton (eds.), *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology* (pp. 39–285). New York: Plenum Press. (Original work published 1934.)
- Wilkinson, S. & Bell, V. (2016). The Representation of Agents in Auditory Verbal Hallucinations. *Mind and Language* 31 (1): 104–26. PubMed PMID: 26900201.
- Woods, A., Jones, N., Alderson-Day, B., Callard, F., & Fernyhough, C. (2015). Experiences of hearing voices: analysis of a novel phenomenological survey. *The Lancet Psychiatry* 2 (4): 323–31. PubMed PMID: 26360085.