Inner Speech: New Voices

Introduction

Peter Langland-Hassan & Agustín Vicente

In another possible world, not far from our own, inner speech occupies center stage in contemporary philosophical psychology. Researchers there see the "little voice in the head" as the ultimate theoretical Rubik's Cube: an introspectively familiar phenomenon—more common than a house sparrow—where independent puzzles intersect, and where answering one promises to disrupt solutions to the others.

Inner speech, they observe, involves language. And linguistic utterances are typically held to express thoughts. Yet inner speech—sometimes called "thinking in words"—is arguably a form of thought in its own right. So, then, are some uses of language cases of thinking? And are some cases of thinking also cases of expressing one's thoughts? Our grip on the distinction between thought and its expression begins to slip.

Inner speech is also a cognitive phenomenon, they note, as opposed to a perceptual one. Its occurrence is not dependent upon the ongoing uptake of an outside stimulus. And through its connection to language, inner speech seems suited to representing abstract states of affairs in the same manner as ordinary linguistic utterances. Like the sentences we utter or write on a page, it seems inner speech episodes can be true or false—not merely more or less accurate, in the manner of a picture. However, inner speech also seems to involve sensory imagery of a kind; what it is like to "hear" one's inner voice is, at least often, similar to what it is like to hear someone speak. And it has been argued—in this most proximate possible world—that states with sensory character occur in a format ill-suited to carrying the sort of abstract, truth-evaluable information represented by linguistic utterances. So our possible researchers are forced to confront the issue of how (or if) a single mental phenomenon can be both language-like and sensory-image-like.

The plot still thickens, they observe, when we recognize that inner speech is a paradigmatically *conscious* mental phenomenon, being the one obvious place where thought, language, and consciousness overlap. Our theories of what consciousness is, and of how some mental states, and not others, come to be conscious, will need to find a place for this language-like, image-like, expression-and-thought-like process. On the one hand, researchers' hopes are raised that inner speech can offer a needed gateway to consciousness for abstract, truth-evaluable thought contents; on the other, they are forced to weigh in on the vexed question of what it is to have a conscious thought, and of what *other* kinds of conscious thoughts we might possibly have. Soon they are knee-deep in broader debates about cognitive phenomenology, self-knowledge, and the possibility of having amodal, "unsymbolized" thoughts in one's stream of consciousness.

Thankfully, just as the air is getting too thin to breathe, our possible researchers are brought back to *terra firma* by inner speech's well-mapped neural geography. Unlike "rational thought" or "conscious reasoning" in general, inner speech's underlying brain regions are known with relative specificity, thanks to contemporary neuro-imaging and historical lesion studies of people with acquired language deficits. (For related work in *this* world, see Loevenbruck et al., and Geva, this volume). The fact that language production abilities can be lost or severely impaired while many other cognitive capacities remain intact—as evidenced in the aphasias— also speaks to the likely *modular* nature of inner speech. Researchers home in on inner speech in hopes they may gain theoretical traction on consciousness, language, and thought, without having to decode the entire human mind in the process. Yet there is again a corresponding cost: they are forced to enter entrenched debates concerning the nature and prevalence of mental modules within the human mind.

Whether or not language is rightly seen as modular, the more agreeable among them note, theories of inner speech can be superimposed on theories of language production and comprehension more generally. Thus, when trying to understand the nature of our inner voices, our possible researchers fuel their investigations with reams of developmental and behavioral research conducted in psycholingustics, in their uncannily familiar world, over the last seventy years. And recognizing that overt speech production is, in large part, a motoric action, they propose that inner speech can as well be viewed in those terms—as the internal preparation for

specific (linguistic) motor acts. This opens up inner speech up to investigation through yet another independent set of theories and concepts—those involved in explaining motor control. Opportunities emerge for new forms of explanation in the process. For instance, in one exciting development, the possible researchers have invoked theories of motor control to explain how a person's inner speech can, in cases of psychosis, seem to be under the control of an outside agent. (See Swiney (this volume) for related work by an actual world counterpart). Our possible researches do not shirk the responsibilities of explanatory coherence created by drawing on these disparate research programs: what they say about inner speech when viewing it as an aspect of motor control, they admit, must form a consistent picture with what they have to say about it when explaining its role in, say, self-knowledge. That's part of the fun of it.

In sum, for the theoretical psychologists and empirically-oriented philosophers of mind of this nearby world, there could hardly be a more tantalizing explanatory target than inner speech. It presents them with a set of seemingly intractable, intersecting "big ticket" questions about thought, language, and consciousness; and they have, close at hand, a variety means for making progress on them, by drawing on both ordinary introspection and existing empirical work concerning language, and the mechanisms underlying its production and perception. So it is plain to see why so much ink is split over inner speech in that possible world.

Of course, here in the actual world, philosophy and psychology have been consumed by somewhat different matters. This is despite the fact that things here are otherwise just as they are in the possible world just imagined. Inner speech *actually* involves the same confluence of theoretical puzzles, whose answers can be sought through use of the same empirical and philosophical tools. Nevertheless, until quite recently, inner speech has sat on the periphery of philosophical psychology. This is not to say that no one has made inner speech a focus of inquiry. Plato conceived of thought in terms of inner speech—as the soul in conversation with itself. Centuries later, the Russian psychologist Vygotsky painted a still-influential picture of inner speech's development, arguing that inner speech is the internalization of an ability that begins as outer speech; he thereby flipped the traditional view of outward speech as built on a prior ability to generate speech "in the head." And, within experimental psychology in the second half of the 20th Century, inner speech gained prominence within influential models of

working memory—under the guise of "activation in the phonological loop" (Baddeley, 1966, 1992, 2007).

Yet it was not until the 1990s that interest in inner speech began to percolate within analytic philosophy of mind, with a number of philosophers and philosophically-oriented psychologists assigning specific roles to inner speech within their theories of consciousness and self-awareness (Peter Carruthers, 1996; A. Clark, 1998; Dennett, 1991; Jackendoff, 1996). Since then, research on inner speech has steadily increased, with new voices commenting on and refining initial forays (Bermudez, 2003; Peter Carruthers, 2002; Langland-Hassan, 2014; Martinez-Manrique & Vicente, 2010). Novel links have since been drawn to new areas of study—including explanations of auditory verbal hallucination and "inserted thoughts" in schizophrenia (Fernyhough, 2004; Langland-Hassan, 2008; Swiney & Sousa, 2014), the question of how we gain knowledge of our own beliefs and desires (Byrne, 2011; P. Carruthers, 2011; Cassam, 2011; Knappik, 2017; Roessler, 2016), and the relation of natural language to conceptual thought (Gauker, 2011; Martínez-Manrique & Vicente, 2015).

Together with the twelve new essays published here, the increasing attention devoted to inner speech suggests we are moving closer to the possible world earlier described. Duly optimistic of inner speech's future as a central area of *actual* theoretical discussion, we have structured the balance of this introduction with new researchers in mind. The sections that follow correspond to the central questions fueling current debates in the philosophical psychology of inner speech. We explain within each section where chapters in this volume make contact with, and further expand upon, the controversies in question. In some cases, a particular chapter is relevant to more than one current controversy, and so is discussed under multiple headings. Thus, the commentaries provided in this introduction do not serve as chapter summaries, but rather as maps charting the chapters' relations to a common set of questions and concerns that extend beyond this volume. Readers simply interested in a brief synopsis of each paper may consult the abstracts provided at the beginning of each chapter.

This is an interdisciplinary volume—including chapters by philosophers, neuroscientists, and psychologists—aimed first at empirically oriented philosophers of mind, and, second, at cognitive scientists interested in foundational issues concerning the definition and methods for studying inner speech. Making progress on inner speech will require researchers in the allied

fields of cognitive science to be familiar with the theoretical proposals and methods characteristic of each. As in any context where different research programs are brought into conversation, there is a threat that researchers will talk past one another—or that they will, in fact, be talking about entirely different things. There is no avoiding that danger completely. It is endemic to the volatile areas of theoretical change that attract contemporary philosophers of mind, and theoretically-oriented psychologists. However, it is clear from the essays here that substantive interdisciplinary discussion is already underway. To take one example, no fewer than seven of the essays debate the theoretical construct of a "forward model"—a notion native to motor control theory—as it applies to inner speech. In other cases, the hard work of extracting the key points of agreement and dispute among different research programs remains to be done. It is work we hope readers will join us in.

We begin with what is seemingly the most straightforward query one could have about inner speech: what are is its parts?

What are the proper *parts* of inner speech, and how do they relate?

To immediate introspection, it might seem that inner speech is just one thing—a voice in the head—without any proper parts to speak of. But a moment's reflection suggests a more complex picture.   At least often, our inner speech episodes ("ISEs", for short) seem to have an auditory component. This is not to say that they literally make a sound. Rather, the experience of having inner speech resembles, in its phenomenology, the experience of hearing someone speak—more so than it resembles, say, seeing written words on a page. In more technical terms, we can say that ISEs have an associated auditory-phonological sensory character—or, again, they often *seem to*, to many. Yet this auditory-phonological component is not their only associated feature. Inner speech episodes also, like outer speech episodes, have an associated semantics. They appear to carry meanings corresponding to the natural language sentences we would use to describe them. Or so it seems, *prima facie*. Further, when we think of inner speech as a phenomenon of speech-production, a third component comes to light. Just as overt speech involves complex articulatory movements of the mouth, lips, tongue and vocal box, inner speech may likewise involve motoric imagery of those movements; or, if not imagery, it may involve

the generation of a motor plan needed for producing such an utterance. We can call this the articulatory component of inner speech.

When breaking inner speech into parts, it is possible to draw far subtler distinctions than these three—as evidenced throughout this volume. But they will suffice to show where the main disputes lie. A first question we can ask is how each of the associated components relates to inner speech itself. Is each a *necessary* component of inner speech, such that an ISE simply cannot occur without having associated articulatory, auditory-phonological, and semantic components? Or, if not, is any *one* of them necessary? And is any one *sufficient*, by itself, for an episode of inner speech? Such questions aim at distinguishing inner speech from other psycholinguistic phenomena only loosely associated with it. Of course, whether it is wise to pursue an essentialist picture of this sort—as opposed to viewing inner speech as a complex cognitive syndrome with no single essential feature—is one of the questions at issue.

Christopher Gauker (Chapter X), defends a strikingly original perspective on the question of inner speech's components, arguing that inner speech *itself* lacks any auditory-phonological component. For Gauker, auditory verbal imagery is a kind of mental representation by which we become aware of our own inner speech; but inner speech itself is non-sensory in nature and consists in the internal tokening of words and sentences of a natural language. While auditory verbal imagery is therefore closely *associated with* inner speech, for Gauker, it is not a proper part of it. He offers the following analogy: just as we distinguish between the speech of others—consisting, as it does, in certain sound waves in the air—and our perceptual means for becoming aware of that speech, so too should we distinguish between our own inner speech and the auditory imagery by which we become aware of it. Unlike outer speech, however, inner speech "does not consist in sounds, but rather, in events of some kind in the brain of the speaker." Gauker puts this distinction to use in developing a theory where ISEs are not episodes of "revealing one's thoughts" to oneself, but rather are a form of "internalized conversation" by which we engage in problem solving.

Peter Langland-Hassan (Chapter X) presents a dramatically different picture, on which the auditory-phonological imagery associated with inner speech is in fact essential to it, seeking to put this intuitive view on a firmer theoretical foundation than it is normally thought to have. His argument begins with the introspective observation that inner speech episodes always

appear, to the people having them, to occur in a specific natural language—typically, though not always, the native language of the speaker. Langland-Hassan then argues that, when we consider the different possibilities for what it could be that makes our inner speech appear to occur in a particular natural language, its having an auditory-phonological component emerges as the only plausible candidate. Viewing the auditory-phonological component as an ISEs essential feature, he maintains, serves to ensure that subsequent empirical investigations of inner speech's nature will remain properly tethered to the introspective means by which inner speech is identified in the first place.

Two chapters by teams of psychologists explore and further elaborate upon the different components of inner speech, endorsing views that allow for a great deal of heterogeneity in what may be considered inner speech. In a rich critical review of recent empirical and theoretical research on inner speech, **Helene Loevenbruck et al.** (Chapter X) argue that inner speech should be seen as "both motor and sensory," having articulatory and sensory components essentially. However, they specify that the sensory components need not be auditory in nature, but may at times be visual or even tactual. Yet Loevenbruck and colleagues limit these conclusions to what they term "expanded" inner speech and do not foreclose the possibility that inner speech also has "abstract, amodal" semantic components.

Meanwhile, **Hurlburt & Heavey** (Chapter X), draw upon their Descriptive Experience Sampling paradigm[1] to argue that inner speech takes a great many forms that are often overlooked, or even dismissed as impossible. For instance, some of their participants report having inner speech that is "unworded", in the sense that they experience inner speaking in the absence of specific inner words. Such inner speech, they maintain—in contrast to Langland-Hassan's thesis—"is not at all phonetic." Nor, they argue, is all inner speech meaningful. In support, they cite participants who report experiencing inner speech while having no subjective awareness of their utterances' meanings.

Sharon Geva (Chapter X), a neuroscientist, reveals how different means for empirically investigating inner speech are forced to make different assumptions about its nature and

---

[1] In Descriptive Experience Sampling ("DES"), participants go about their ordinary activities while wearing a device that emits a beep at random intervals (roughly every half hour). They are instructed to immediately write notes on the nature of their subjective experience just prior to the beep. Later, in an interview with an experimenter, their notes are developed into detailed reflections on the nature of their conscious experiences.

components. Her discussion offers historical context on the neuroscientific investigation of inner speech, while explaining the delicate experimental manipulations currently used in investigations of inner speech's neural underpinnings. To take one example, Geva notes that some imaging studies have explored the neural correlates of inner speech by recording activation patterns when a participant is asked to repeat a certain word silently, in the head. Yet many other studies have incorporated a variety of silent rhyme and homophone judgment tasks, where participants are asked to silently judge whether two presented words rhyme, or are homophones. The latter sort of task assumes that inner speech carries auditory-verbal phonological information; whereas, the task that merely asks participants to repeat a word in the head has no such explicit assumption— nor do tasks that measure inner speech by measuring neural activation during silent reading. If Gauker is correct that auditory verbal imagery is not a proper part of inner speech itself, then the (putative) inner speech assessments that require participants to make rhyme and homophone judgments are not assessments of inner speech after all. Such tasks could be agreed, by both sides, to at least assess auditory verbal imagery; yet they would likely remain at odds over whether there was a further phenomenon of *inner speech* in need of separate assessment. In addition, if, as both Loevenbruck and colleagues and Hurlburt & Heavey propose, there are non-phonetic, "condensed" versions of inner speech, then an inability of a participant to silently judge rhymes and phonemes will not itself be evidence of a lack of inner speech (in the "condensed" sense). The overall picture Geva paints is one of a tumultuous but rapidly-developing field of empirical inquiry, where theoretical assumptions and newly available neuroscientific methods— such as transcranial magnetic stimulation—are in continual interplay. While progress has undoubtedly been made in understanding inner speech's neural underpinnings, Geva's chapter should encourage reflection on the extent to which that knowledge relies upon debatable assumptions concerning inner speech's nature.

What is the relation of inner speech's components to each other?

A second issue animating research on the components of inner speech concerns the relation of the components *to each other*: are they each distinct mental states that at times co-occur? Or are they more intimately linked, being fused, somehow, into a mental whole?

Relatedly, is each component well-conceived as a mental representation? Or are some of the components non-representational?

In an earlier paper, Langland-Hassan (2014) argues that there is a tension in holding that an inner speech episode is a single mental state with *both* auditory-phonological and semantic features. Rather, he argued, if inner speech is to have both components, it must be understood as consisting in the co-occurrence of (at least) two ontologically distinct mental representations—one with semantic content, and the other with auditory-phonological content. This, he proposed, creates difficulties for views that assign cognitive functions to ISEs on the basis that they have both sensory and semantic features simultaneously. **José Luis Bermúdez** (Chapter X) considers Langland-Hassan's argument and finds it lacking. According to Bermúdez, Langland-Hassan "slides illegitimately from the harmless truth that inner speech episodes have auditory-phonological properties to the false claim that inner speech episodes represent those auditory-phonological properties." Bermúdez proposes a view where inner speech episodes have auditory-phonological character, but where their having this character is not a matter of their representing any phonetic features of words. Instead, their having such sensory character is an instance of a more general phenomenon where "perceptual experiences can have properties that are not part of their representational content." This allows Bermúdez to hold that inner speech episodes are discrete mental states that have only one form of representational content—viz., semantic ("propositional") content—while also maintaining that they have the kinds of auditory-phonological character commonly associated with them.

Peter Carruthers (Chapter X) also argues that ISEs are unitary (not separately occurring) mental states with both semantic and auditory-phonological features. Yet, unlike Bermúdez, he aims to make this view consistent with the idea that both semantic and auditory-phonological contents are *represented by* inner speech episodes. To that end, he introduces the notion of an *event file,* which is a kind of cognitive repository for different forms of information one might have about a particular speech event, modelled on the notion of an *object file* familiar from work on feature-binding in visual perception. "Speech is segmented into distinct events (generally sentences)," he proposes, "with multiple properties drawn from many different levels of processing bound into each event-file." Such files serve as summations of all this information and can contain "both conceptual and nonconceptual information." This may include

information about the sound of a voice, the semantic content of what it is saying, and the intent (e.g., sincere or sarcastic) with which it is said. Insofar as everything in the event-file is globally broadcast, and—according to Carruthers—thereby made conscious together, one can hold that the disparate informational states comprising the event file constitute a unified conscious mental episode—a single ISE.

Wilkinson & Fernyhough (Chapter X) also discuss the relationship of the semantic and auditory-phonological components of inner speech, arguing that there is an important difference in how they should be conceived. While an inner speech episode will represent both the sound of an utterance *and* its associated semantic content, they propose, the episode will only be assessable for accuracy when considering its semantic content. On their view, whether an inner speech episode is to be considered "correct" or "incorrect" hangs entirely on whether the semantic content it represents is had by another propositional mental state (such as a belief or desire) had by the subject. "Saying 'I'm such an idiot!' in inner speech" they remark, "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." By contrast, the auditory-phonological representations that partly comprise inner speech are, they claim, not assessable for accuracy and are instead cases of "content without commitment."

The disparate and often conflicting views put forward in this volume concerning the components of inner speech and their relation are evidence that these questions will remain active areas of research. Theorists are pulled in two directions simultaneously: towards properly describing the nature and cognitive role of the familiar "voice in the head"; and towards breaking that phenomenon into parts that will enable us to better understand, and scientifically explore, its nature. It is unclear to what extent the two impulses can be reconciled. In pulling inner speech apart, will we better understand the means by which it carries out important cognitive functions? Or will we discover that there is no *it* there to carry out such functions, but rather a panoply of discrete states that tend to occur together, each of which has a somewhat different nature and cognitive role? These still unsettled questions loom large in the background of the next central controversy concerning inner speech: is inner speech thought itself, or the expression of thought?

Is inner speech the expression of thought, or thought itself?

The words and sentences we say aloud are not usually held to be thoughts. The predominant view, both today and historically, is that our public utterances serve to communicate or *express* thoughts, where thoughts themselves are something distinct from those utterances. Thoughts are typically held to be private mental events of a kind, which we can choose to express with an overt utterance, or not.

By its very nature, inner speech seems to straddle the fence between thought and expression. On the one hand, it seems we are able perform different types of speech acts in inner speech: commands, assertions, questions, and so on. Most of the things we say in inner speech could be said aloud, and *vice versa*. So if overt utterances serve to express thoughts, it may seem that ISEs do as well. This idea—that inner speech episodes are genuine speech acts on a par with overt utterances—is an important claim within both **Machery's** and **Wilkinson and Fernyhough's** contributions to this volume. On the other hand, ISEs have features paradigmatic of thought itself: they are private mental events we can choose to express aloud with an overt utterance, or not. Where, then, should we place inner speech? Are ISEs thoughts that we at times express with overt speech? Or are they the (inner) expression of thoughts—thoughts that themselves occur in some other mental medium? Or is this a false dichotomy? Are ISEs somehow both thoughts *and* expressions of thoughts simultaneously?

As one might expect, answers here will be influenced by one's broader commitments about the nature of language and cognition. If we use the term 'thought' in the broadest sense possible—to mark any sort of mental episode whatsoever—it should be uncontroversial that ISEs are thoughts of a kind. The more difficult, historically more contentious, question is whether ISEs are thoughts in some more robust sense. Are they, for instance, necessary for *conceptual,* or *propositional* thought? (As proposed, e.g., by Gauker (2011)). Do they constitute steps in reflective reasoning processes available only to humans (Frankish, 2004)? Do they play an essential role in focusing concentration on our own reasoning (Bermudez, this volume)? One way to collapse these questions is to ask whether ISEs have cognitive roles that extend beyond the preparation for and recollection of overt speech. As we will see, the contributors to this volume defend quite different—and often conflicting—views on this question. Before

contrasting some of those views, it will help to fill in the historical background to these debates, in order to appreciate the larger motivation for the views on offer.

       The idea that inner speech is itself a form of thinking has a long history. Plato, in *Thaethetus,* holds that "the soul when thinking appears to me to be just talking—asking questions of herself and answering them, affirming and denying... I say, then, that to form an opinion is to speak, and opinion is a word spoken,—I mean, to oneself and in silence, not aloud or to another." In the mid-20th Century, philosophers such as Ryle and Wittgenstein, and psychologists including Watson, Vygotsky, and Sokolov, put forward similar views. Ryle (1972; 133), for instance, held that "[t]hinking, then, can be saying-things-tentatively-to-oneself with the specific heuristic intention of trying, by saying them, to open one's own eyes, to consolidate one's own grasp, or to get oneself out of a rut, etc."[2] Nevertheless, despite a tendency in some areas of psychology to describe inner speech as "verbal thought" or "thinking in words," the predominant trend within both the philosophy of mind and cognitive science in the late 20th Century was to assign inner speech very little in the way of robust, extra-communicative cognitive functions. Chomsky's attack on behaviorism, together with the rise of modern computers, created fertile ground for a view that strongly analogizes human cognition to digital computation. Thinking came to be seen as the manipulation of language-like symbols according to syntactic rules, where these symbols are realized by physical states of the brain. This view found its most influential statement in Jerry Fodor's *The Language of Thought* (1975).

       According to Fodor's Language of Thought hypothesis, our core thought-processes take place in a language-like representational medium distinct from any natural language. This "Language of Thought" (LOT), sometimes called *Mentalese,* is said to have its own representational units and syntax, enabling the generation of complex representations out of simple parts. The LOT hypothesis has been held to explain a number of features of cognition, including its systematicity and productivity (Fodor & Pylyshyn, 1988), the ability of rational thought processes to preserve truth, and the means by which words are learned (Fodor, 1975).

---

[2] Similarly, Wittgenstein (1953; paragraph 332): "While we sometimes call it 'thinking' to accompany a sentence by a mental process, that accompaniment is not what we mean by a 'thought.'——Say a sentence and think it; say it with understanding.— And now do not say it, and just do what you accompanied it with when you said it with understanding!—". On Vygotsky and Watson, see below.

The LOT hypothesis also gave flesh to what has been called the "Gricean program" in the philosophy of language, which seeks to explain the representational character of linguistic utterances in terms of the representational character of thought. According to the Gricean program, we mentally represent the world as being in a certain way, and linguistic utterances simply try to convey to other people how we represent the world. Levelt's (1989) influential theory of speech production, developed around the same time, seemed to fit the Fodorian and Gricean mold as well: the speaker begins with a "message" to be conveyed—where the message is realized in a non-linguistic conceptual structure of some kind—and then carries out successive stages of linguistic processing in order to generate natural-language words suitable for expressing the message aloud.

Ironically, at the very moment many decided that thought must take place in a language-like medium, it was determined, with equal vehemence, that the language in question could not be one we speak. Plato's intuitive claim that we think in a natural language—that inner speech simply is "thinking in words"—came under heavy criticism, for a number of reasons (see, e.g., Pinker, 1994). Perhaps most influential was what we can call the *argument from explicitness,* which purported to show that a natural language cannot *possibly* be a vehicle for the sort of thoughts we in fact have. Versions of this argument can be found in Pinker (1994), Fodor (2001), Carston (2002) and Martínez Manrique and Vicente (2005, 2008). The key idea behind the argument is that natural language sentences are semantically ambiguous in ways our thoughts are not. For instance, a sentence such as 'John's car is empty' may say something about the car that John owns, or the car that John usually drives, or the car that John likes, or the car that John bet on, and so on. And even what it says about such car is open to multiple interpretations: the car can be empty because it lacks passengers and driver, because everything that is not steering wheel, seats, etc., has been taken away, or because it does not even have steering wheel, seats, and so on. Nor does a simple sentence such as 'the leaves are green' determine a particular set of truth-conditions on its own: the leaves can be taken to be green if they are naturally green, even though they have been painted in another color; alternatively, they can be taken to be green if they have been painted green, even if they are naturally red (Travis, 1996). One can try to resolve a particular ambiguity by being more precise; but, arguably, the more precise version will still admit of ambiguities that do not occur at the level of thought (Fodor, 2001, Carston, 2002, Recanati, 2004). If we nevertheless wish to hold that our linguistic utterances have precise,

unambiguous meanings, then, it seems, our thoughts themselves must be semantically precise; our utterances could then inherit their precision from the thoughts they express. But, in that case, our thoughts cannot occur through the use of the same ambiguous sets of symbols. Or, so says the argument from explicitness.

Not everyone accepts this line of reasoning (see, e.g., Gauker (2011, p. XX), this volume); there are possible explanations for how linguistic utterances gain their semantic determinacy other than by their connection to semantically explicit thoughts, such as through their context of utterance. But, for the many who are moved by the argument from explicitness, the question arises as to what, exactly, we are up to with our frequent episodes of inner speech, if we are not using them to have conceptual thoughts. This question sets the background for many discussions of inner speech in recent analytic philosophy. A popular response has been that ISEs are (internal) expressions of thoughts, and that these expressive episodes have special properties unlike those of the (amodal, semantically determinate) thoughts they serve to express. Just what those properties are, and what roles they enable inner speech to play, are matters of controversy, even among those who agree in seeing inner speech as an expressive phenomenon. The most common proposal within philosophy is that ISEs provide a means by which can become aware of our own (conceptual, propositional) thoughts—thoughts that are themselves amodal and not otherwise accessible to the introspecting subject. Several versions of this view are on offer in this volume (see the entries by **Carruthers**, **Machery**, **Wilkinson & Fernyhough**, and **Bermudez**). They all work from the idea that inner speech must have some relatively *narrow* cognitive role—e.g. in making us aware of our thoughts—given that our propositional thoughts themselves do not occur in a natural language. Details of these views will be discussed in the section concerning inner speech's connection to self-knowledge and self-awareness, below. Our focus in the balance of this section will be on chapters that buck the trend of seeing ISEs as primarily expressive of thought, and which hold, instead, that ISEs constitute some of our core reasoning processes themselves.

Christopher Gauker (Chapter X) defends an overtly anti-Gricean view of language and its relation to thought, arguing that inner speech is indeed our only means for conceptual thought. Echoing the quote from Plato, above, Gauker maintains that all propositional thought (conscious or not) takes the form of a conversation the thinker has with herself, a conversation that can

become "audible," during overt speech, or not. The Gricean view he rejects has it that in conversation we express propositional thoughts that have already occurred in some other format. Gauker argues, to the contrary, that conversation is a way of cooperating and joint problem-solving that involves articulating new propositional thoughts and confronting perspectives. When we converse with ourselves we do not inform ourselves about what we were already thinking, but rather articulate our thinking in a way that enhances problem-solving abilities. This general position is reminiscent of Vygotsky's work—discussed below—and has close allies in contemporary Vygotskyans such as Fernyhough (2004, 2017).

Gauker addresses the argument from explicitness by proposing that "elements of thought that do not themselves bear propositional contents may play a role in determining the propositional content of an incomplete thought." A mental image of a measuring cup half filled, for instance, may serve resolve the ambiguity of the reference of 'that' when one says, in inner speech, "that should be enough" (Gauker, this volume).   In other cases, superficial ambiguities may be resolved by the observable circumstances—an utterance of "I need to fill the tank," said while driving, will refer to the car's gas tank, and not a military vehicle. In short, Gauker argues, a piece of inner speech can be ambiguous in itself as long as it is produced against a cognitive and environmental background that resolves its ambiguities. Gauker then attempts to shift the burden back in the direction of Gricean and Fodorian views, arguing that they lack genuine empirical backing, and that such approaches bring with them puzzles as deep as the ones they propose to answer.

Keith Frankish (this volume) also endorses the idea that ISEs are themselves episodes of propositional thought, although he restricts the use of inner speech to *conscious* thinking, and focuses on episodes of reasoning. Frankish is a well-known defender of a dual-process approach to reasoning (Evans & Frankish, 2009; Frankish, 2004). Dual-processes accounts posit two general types of reasoning processes in the human mind: a fast, intuitive, unconscious kind of reasoning (Type-1 thinking), and a slow, serial, conscious form of reasoning linked to language (Type-2 thinking). For Frankish, the main difference between Type-1 and Type-2 thinking lies in that Type-1 processes are autonomous and non-intentional. In contrast, Type-2 processes involve intentional action. An example of a Type-2 processes is a mathematical calculation intentionally broken in steps, whose execution is under the control of the subject. Using inner

speech in problem solving is also an example of a Type-2 process—one that, according to Frankish, implements a cyclical structure under partial intentional control. Frankish holds that this kind of reasoning, in which we iterate cycles of inner speech in problem solving situations, is plausibly an internalization of linguistic exchanges in problem-solving social settings.

Expanding upon and refining earlier work (Morin, 2005), **Alain Morin** (Chapter X) argues that inner speech is linked to a host of high-level cognitive phenomena related to the self, including mental time travel, self-concept formation, self-evaluation, self-esteem, self-knowledge, and self-description. For Morin, there are certain kinds of self-reflective thoughts we can only have through the use of inner speech. For that reason, we place him among those who see inner speech as constituting "thought itself," and not merely as an internal medium for expressing one's thoughts. One of Morin's central claims is that having the self as an object of our thinking is not possible without inner speech. In this he shares affinities with Bermúdez and other authors, such as Clark (1998), who hold that inner speech is necessary for higher forms of self-reflection. Morin's hypothesis has far broader scope, however, as it is not limited to knowledge of one's current thoughts and propositional attitudes. The evidence Morin presents is rich and multi-faceted, including neurophysiological data, introspective reports of stroke victims, structured questionnaires, and experimental results spanning developmental, social, and personality psychology. His theoretical claims gain additional traction in cohering with influential theories in social psychology that connect self-identity and self-awareness to "narrative" views of the self. Morin endorses the Vygotskyan idea that, in inner speech, we not only internalize speech but social exchanges themselves, including conversations that force us to see ourselves and what we do from different perspectives. He further argues that (inner) verbalization helps to create the required distance between the self that experiences and the self that observes, and presents evidence that inner speech crystalizes awareness of our own traits and inner lives through the verbal labelling of private sensations and emotional responses.

For each of Gauker, Frankish, and Morin, inner speech plays substantive roles in cognition over and above facilitating reflection on our own current mental states. In the next section, we look at views that focus more narrowly on the latter.

In what ways does inner speech facilitate self-knowledge?

Consider your awareness of your own thoughts at this moment. How might it differ if you were no longer able to say anything silently in your head?[3] Would it present an obstacle to your knowing your own judgments, decisions, beliefs, desires, and so on? Or would you have just the same knowledge of those other mental states—just an inability to put that knowledge into words? Over the last two decades, some of the most exciting work in philosophical psychology has centered on these questions, examining the ways in which inner speech may—or may not—facilitate different forms of self-knowledge. 'Self-knowledge,' in this context, refers to knowledge of one's own current mental features—one's beliefs, desires, thoughts, intentions, personality traits, and cognitive abilities.[4]

In the late 1990s, Ray Jackendoff (1996) and Andy Clark (1998) were some of the first to publish papers arguing that inner speech enables a special kind of self-reflective thought about one's own thoughts—what Clark termed "second-order cognitive dynamics." Clark's idea was that natural language sentences have certain "context resistant" and "modality transcending" features that make them suitable objects of reflection when we are tasked with assessing the validity of our own reasoning. José Bermúdez (2003) later endorsed and expanded upon Clark's line of thought, arguing that awareness of inner speech is the primary means by which humans gain awareness of their own propositional thought processes, which are otherwise amodal and not available to introspection. In a similar vein, Peter Carruthers (2011) has defended a sophisticated view on which inner speech is a central resource exploited in generating knowledge of one's own propositional attitude states (such as beliefs and desires), while Alex Byrne (2011) has argued that inner speech utterances offer the most reliable route to knowing the contents of our ongoing thoughts.

A number of interesting new proposals concerning the relation of inner speech to self-knowledge are made in this volume—some by familiar voices in these debates, others by new

---

[3] Loss of inner speech appears to occur, to differing degrees, in people with aphasia, who have corresponding outer speech deficits (Geva, Bennett, Warburton, & Patterson, 2011; Langland-Hassan, Faries, Richardson, & Dietz, 2015)).

[4] This definition is somewhat broader than some philosophical treatments of self-knowledge, which focus on propositional attitude states—such as beliefs and desires—to the exclusion of cognitive abilities and personality features.

participants. **José Luis Bermúdez** (Chapter X) offers a robust defense of his earlier view against objections that have subsequently appeared in the literature, including, in particular, those of Martinez-Manrique & Vicente (2010) and Langland-Hassan (2014). Of special note is Bermúdez's nuanced discussion of Martínez-Manrique and Vicente's argument from explicitness. Against the claim the inner speech utterance serve to make us aware of our propositional thoughts, Martínez-Manrique & Vicente (2010) argue that there are "a myriad of examples that show that our usual utterances are semantically underdetermined at one level or another." (See the discussion of the argument from explicitness above, under the question, "Is inner speech the expression of thought, or thought itself?"). If extracting the explicit meaning of a sentence requires disambiguation, it is unclear how awareness of internal natural language sentences could suffice for awareness of one's own—presumably explicit and unambiguous—propositional thoughts. Bermúdez confronts this challenge by proposing three different ways of resisting the conclusion that the semantic indeterminacy of linguistic utterances prevents it from making us aware of our explicit propositional thoughts. One response argues that the sort of indeterminacy in question goes "all the way down," and is therefore present in our propositional thoughts themselves; the other two responses accept that propositional thoughts may be semantically determinate in ways inner speech is not, while still finding a suitable role for inner speech in making us aware of those thoughts.

In keeping with earlier work, **Peter Carruthers** (Chapter X) proposes that our inner speech episodes come to have semantic contents associated with them through use of the same cognitive mechanisms that processes and interpret outer speech. This leads him to consider a new puzzle: Given that the same processes underlie the comprehension of both inner and outer speech, why is it that we do not ever ask ourselves "What did I just think?" in the way that we often have to ask others, "What did you just say?" In answering, Carruthers draws on theories of speech interpretation that place heavy reliance on the *accessibility* of recently activated mental representations (e.g. Sperber & Wilson, 1995, 2002). In another interesting new proposal, Carruthers argues that not only the sensory and semantic content of an inner speech utterance is made conscious, but so is its associated attitude: we can "hear" ourselves *as judging,* or *as deciding* something. Nevertheless, Carruthers thinks we are far more reliable in assigning proper semantic contents to our inner speech episodes than we are in determining the attitudes they express—even if both processes involve some degree of interpretation. Yet, for reasons he

explores here for the first time, this comparative lack of reliability does nothing to undermine the certainty we feel when considering whether we know the nature of the attitudes we express in inner speech. Along the way he offers new explanations for why we at times interpret ourselves as having mere "passing thoughts" that don't express *any* determinate attitude, when we do not tend to interpret others' statements as expressing no attitude.

Edouard Machery (Chapter X) concurs with Carruthers and Bermúdez that inner speech plays an important role in enabling us to have knowledge of our own mental states. But, he argues, the kind of knowledge that inner speech enables of our beliefs is much different than that which it provides of our desires. And, further, this tells us something interesting about the difference between beliefs and desires themselves. "For some beliefs," Machery observes, "the listener needs nothing more than the speech act to be justified in believing that the speaker believes so and so." Because of this, such beliefs "are transparent." On the other hand, desires, Machery argues, are never transparently communicated by assertions: coming to know someone's desire on the basis of her utterance always requires an inferential step not necessary for belief. An important result is that one can come to know one's own beliefs "just by talking to [oneself]," while this is never the case for one's desires. When it comes to knowing the latter, Machery argues, something beyond awareness of one's own inner speech is needed. Machery traces this difference to the fact that beliefs are essentially linked to our capacity for linguistic communication and intra-species cooperation. Our ability to transparently know a person's beliefs on the basis of her assertions traces to the specific social role beliefs play. Self-knowledge of beliefs is then a mere "by-product or a spandrel of the evolved social nature of beliefs and of the internalization of communication."

Wilkinson & Fernyhough (Chapter X) agree with Machery that inner speech involves the generation of *speech acts*: assertings, questionings, insultings, and so on. They also advance the popular view that inner speech 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." Yet, like Carruthers, they do not see inner speech an infallible source of information about one's state of mind. Both the agent of the inner speech act, and the mental state expressed by that speech act, are elements that might be misrepresented by an episode of inner speech. The latter happens when, for instance, we "lie to ourselves," or simply fail to consciously register our

true thoughts and motives. Erring in assessing the agent of and inner speech act occurs more rarely, but, they argue, may contribute to the experience of auditory verbal hallucinations (AVHs) in schizophrenia. While Wilkinson & Fernyhough's chapter deals only briefly with the question of inner speech's relation to AVHs, both have written extensively on the topic elsewhere (Alderson-Day & Fernyhough, 2014; Jones & Fernyhough, 2007; Wilkinson, 2014; Wilkinson & Bell, 2016).   Their work on the topic is part of a larger trend—the subject of our next section—that seeks to understand core symptoms of psychosis in terms of deficits or irregularities in the cognitive architecture underlying inner speech.

What role can inner speech play in explanations of auditory verbal hallucinations and "inserted thoughts"?

When people report "hearing voices" in psychosis, they seem to mistake an internal voice of their own creation for someone else's. But how could that be? How could someone fail to recognize their inner speech as being, in fact, their own? To sharpen the question: in what sense do we normally feel *in control* of our own inner speech, such that losing that sense of control would lead us to associate the inner speech with that of an outside agent? These remain some of the most hotly debated questions in the philosophical psychology of inner speech.

An attractively simple answer is that the voices being "heard" not so much cases of misidentified inner speech as they are auditory-verbal *hallucinations*—perceptual experiences that occur in the absence of appropriate sensory input. If, phenomenologically, such episodes are just like hearing someone else speak, this might explain why patients think that such voices derive from another's agency (Wu, 2012; Wu & Cho, 2013). Perhaps surprisingly, this style of explanation has had relatively few adherents. One reason lies in patient reports themselves: it is common in surveys concerning the characteristics of AVHs for patients to report hearing voices that are not subjectively loud or rich in sensory features (Graham & Stephens, 2000; Hoffman, Varanko, Gilmore, & Mishara, 2008; Laroi et al., 2012; Nayani & David, 1996). Often the voices heard are described by patients as being more like their own verbal thought (Hoffman et al., 2008) and are at times even called "soundless" (Junginger & Frame, 1985). Further, in addition to "hearing voices," some patients with schizophrenia report having another person's *thoughts* inserted into their minds—this being the symptom of known as thought insertion. In

such reports there is no suggestion that the thoughts in question are at all similar to cases of hearing someone speak. Yet the door remains open to their being unusual cases of inner speech, as inner speech is itself often described in commonsense terms that equate it with thought—e.g., as "thinking in words" or "verbal thought." Thus, unusual inner speech could conceivably come to be described as an "inserted" thought, while still lacking strong sensory features (Langland-Hassan, 2008).

A second reason more nuanced explanations have been sought is that people with schizophrenia exhibit a range of passivity symptoms suggestive of a broader disruption in their sense of agency (Blakemore, Smith, Steel, Johnstone, & Frith, 2000; Frith, 1992; Spence et al., 1997). By "sense of agency" here we mean the feeling each of us has of being in control of our own actions—whether these are bodily actions, like grasping a cup of coffee, or (putative) mental actions, like conducting mental arithmetic. When we carry out an action, we normally have a sense of being in control of the action—we do not feel as though it is something happening *to* us—and in this way have a sense of agency concerning the action. Yet people with schizophrenia suffering passivity symptoms—also known as delusions of control—often appear to lack this sense of agency, reporting that their thoughts and actions are under control of an external agent. If we can understand the mental states and cognitive mechanisms by which we come to feel in control of our actions in the normal case, we may better grasp how that sense of agency could become disrupted in schizophrenia. Supposing that speaking in the head is as much an action as speaking aloud—or, indeed, as reaching for a coffee cup—then, arguably, one and the same style of explanation may ultimately apply to both the experience of "hearing voices" and bodily delusions of control in schizophrenia.

Perhaps the most influential proposal along these lines is Christopher Frith's (1992) comparator hypothesis. (N.B., Frith himself has subsequently abandoned some aspects of that hypothesis (Frith, 2012)). Working within a framework originally developed to explain bodily control—and the swift correction of errors in bodily movement (Miall, Weir, Wolpert, & Stein, 1993; Wolpert, Miall, & Kawato, 1998)—Frith proposed that the experience of AVHs and thought insertion in schizophrenia may be due to impairments in mechanisms underlying motoric self-monitoring. This self-monitoring was thought to occur through processes of prediction and comparison, with actions being accompanied by predictions of the likely sensory consequences

of successfully carrying out the actions. The mechanism responsible for generating the prediction, on the basis of a "copy" of the relevant motor command, has been called a "forward model." When the prediction of the sensory consequences of an act do not match actual sensory input—as judged by a "comparator" mechanism—the subject is alerted to correct its movement via an error signal. Connecting this signal back to the "sense of agency," feeling in control of one's actions, on this view, becomes a matter of predictions from the forward model matching sensory input at the comparator, while feeling out of control—or, indeed, as though some other agency has intervened on one's action—is associated with mismatches at the comparator, and subsequent error signals. This is the comparator hypothesis in its barest outlines.

Whatever one makes of the comparator hypothesis as an explanation of *bodily* delusions of control, significant questions can be raised as to whether it can properly be applied to a mental activity such as inner speech—and, indeed, whether it even makes sense to speak of *mental* actions, conceiving of some thought episodes in motoric terms. Multiple briefs have been filed on both sides of this debate.[5] Some argue that the comparator-style explanation, as applied to AVHs and "inserted" thoughts, is misguided in fundamental respects (Synofzik, Vosgerau, & Newen, 2008; Vicente, 2014; Vosgerau & Newen, 2007; Wilkinson, 2014; Wu, 2012). While others have responded that, properly understood—and suitably amended—something like the comparator hypothesis can indeed be extended to explain the disruptions in the sense of agency that lead to reports of AVHs and inserted thoughts (Campbell, 1999; Ford & Mathalon, 2005; Jones & Fernyhough, 2007; Knappik, 2017; Langland-Hassan, 2008, 2015; Stephan, Friston, & Frith, 2009; Swiney & Sousa, 2014). A challenge faced by such proposals is to place the ordinary experience of inner speech within such models in a way that is consistent with their application to AVHs and inserted thoughts. If, for instance, we conceive of ordinary ISEs as predictions of auditory-verbal input that never occurs—because the overt speech act is aborted before any speech sounds are registered through the ears—then all inner speech should generate a "mismatch" at the comparator, triggering an error signal as a result. Perhaps this in fact occurs, as its doing so might explain why ISEs become conscious and attended by the subject. Yet mismatches at the comparator could not then be called on to explain the missing sense of agency

---

[5] Additionally, one can question the idea that AVHs and thought insertion are misattributed regular inner speech given that, e.g., a male subject may experience a female voice, and the voice may address him in the second person (while he usually talks to himself in the first person) (Langdon, et al., 2009). Gregory (2016) proposes that AVHs result from the misattribution of *imagined* speech.

that purportedly gives rise to reports of AVHs and inserted thoughts. For ordinary inner speech would be missing that sense of agency as well.

Two papers in this volume confront these issues head-on, while adding new levels of specificity to the processes and mechanisms that may play a role in the generation and self-monitoring of inner speech episodes. **Loevenbruck *et al.*** (Chapter X) posit three separate stages of prediction and comparison that occur during speech production. A first comparison takes place between actual sensory feedback and the desired sensory state. This is the most familiar sort of prediction and comparison posited by such approaches, and one that faces the question just discussed. But, for Loevenbruck and colleagues, this is not the level at which the sense of agency relevant to feeling in control of one's inner speech occurs. Instead, it is a separate prediction and comparison between "desired and predicted states" that, they argue, may lie at the root of AVHs and inserted thoughts. This stage of prediction and comparison normally serves to "tune" the system's ability to issue proper motor commands on the basis of its desired states. In essence, the language-production system has in place mechanisms for checking whether the motor command it is issuing is of the right sort to bring about a desired end state. Building on a related proposal by Swiney & Sousa (2014), they argue that mismatches at this stage will result in a failure to attenuate an inner speech episode, with this attenuation failure in turn leading to the ISE's having an unusual "alien" phenomenology. They then review a wealth of evidence relating the key components of this multi-layered prediction and control architecture to specific areas of the brain, developing what is likely the most detailed brain-to-cognitive-architecture mapping of inner speech to date.

In her own chapter for this volume, **Lauren Swiney** (Chapter X) begins with a helpful discussion of the history and theoretical rationale for comparator approaches as they have been applied to inner speech and the explanation of passivity phenomena. After contrasting different versions of the comparator account in their abilities to explain patients' phenomenological reports, Swiney considers such views in the light of independently developed theories of speech production (Oppenheim, 2013; Oppenheim & Dell, 2008; Pickering & Garrod, 2013). This is a welcome step forward in the dialectic, as psycholinguistic theories concerning speech production have not always been linked in clear ways to the stages and mechanisms of prediction and comparison that are posited within models of motor control. Swiney then goes on to consider the

relation of traditional comparator accounts to the new wave of *predictive processing* theories, which have sought to apply a Bayesian predictive processing framework to the explanation of almost all areas of perception and cognition (Andy Clark, 2013; Friston, 2009; Hohwy, 2013). After showing how some predictive processing theories can be seen as elaborations of the familiar comparator architecture, she then considers an alternative "active inference account" of predictive processing, on which all action involves the registering of at least some error, on the basis of which subsequent predictions are continually refined.

There are a number of delusions of control that active inference accounts are well placed to explain, Swiney argues, yet, like earlier comparator approaches, they face questions when applied to AVHs and inserted thoughts; for they also seem to require that inner speech is, in the normal case, accompanied by actual sensory perception of speech, of which it is a prediction. This leads Swiney to consider a quite different interpretation of inner speech's place within a cognitive architecture centered on predictive processing, which she calls the 'reality monitoring approach'. This view no longer sees inner speech as a process that normally occurs as a prediction of overt speech, but rather as *the result of* a prior prediction that occurs in the absence of any bottom-up sensory information. On this view, the experience of inner speech is not itself a prediction, but is rather the result of a top-down speech-production signal meeting with no bottom up information with which it might be compared and adjusted. Swiney proposes several ways in which inner speech, so-understood, might acquire an unusual phenomenology of a sort that could explain AVHs and inserted thoughts.

In a third chapter of this volume touching on the relation of inner speech to AVHs and inserted thoughts, **Langland-Hassan** (Chapter X), proposes a new way of thinking about the relation of the two symptoms. AVHs, it is widely agreed, have sensory character related to the auditory perception of words. If, as Langland-Hassan argues, inner speech episodes have an auditory verbal component, it makes sense to offer explanations of AVHs that appeal to this sensory component. In particular, the kinds of views discussed by Swiney and Loevenbruck et al., on which inner speech is construed as a kind of sensorimotor prediction or perceptual simulation, will seem apt. On the other hand, as earlier discussed, when patients report experiencing "inserted" thoughts, they often emphasize the non-sensory, even "soundless" nature of the experience. This has led some to conclude that explanations of thought insertion that

connect the symptom to sensorimotor mechanisms underlying inner speech are misguided (Vosgerau & Newen, 2007). Langland-Hassan proposes a new way of arbitrating this debate: if the mental episodes reported as "inserted" thoughts by a patient seem, to that patient, to occur in a natural language, then we should infer that they do indeed have auditory-phonological characteristics—and so can, at least potentially, be explained by appeal to deficits in sensorimotor prediction mechanisms after all. He argues this conclusion follows from the main argument of his paper, which is that the best explanation for why a mental episode seems to occur in one natural language, as opposed to another, is that it has auditory phonological features distinctive of that language. If that argument succeeds, then, in cases where an inserted thought seems to a patient to occur in a particular language, the best explanation for its seeming that way will be that it has auditory-phonological features—even if the patient is not inclined to describe it in that way. On the other hand, if patients do not report that their "inserted" thoughts seem to occur in a natural language, then critics may be correct that sensorimotor prediction and comparison accounts are ill-suited to explain their occurrence. Langland-Hassan concludes that progress can be made on this question if diagnostic patient interviews are amended to include questions concerning the language (if any) in which an AVH or inserted thought seems, to the patient, to occur.

Vygotsky's complicated legacy

The Russian/Soviet psychologist Lev Vygotsky is one of the most influential historical authors on the topic of inner speech. Although many of contemporary debates surrounding inner speech do not substantively connect with the tradition he inaugurated, Vygotsky's ideas are still widely discussed. His influence can be seen throughout this volume. Vygotsky (as well as other mid-to-late 20th Century Russian psychologists, including Luria, Sokolov, and Bakhtin) gave inner speech a prominent role not only in cognition but also in the development of the individual. Vygotsky's most famous thesis is that inner speech is social speech internalized. That is, inner speech is the internalization of an activity that begins as an external, social phenomenon. This thesis has implications regarding the probable age at which children start using inner speech, as well as the functions that inner speech may serve.

Vygotsky and current Vygotskyans focus mainly on what may be termed the *regulatory* functions of inner speech. Many of the linguistic exchanges between children and caregivers, for instance, can be seen as regulatory in character: the child is engaged in a task (e.g. solving a jigsaw puzzle) and his/her caregivers may help by providing advice, instructions, corrections, commands, warnings, and motivating speech. Vygotsky's central idea was that this regulatory give-and-take between agents becomes internalized in two steps. First, the child reproduces these exchanges privately but aloud, in what Piaget called "ego-centric" speech; this involves the child's transferring her social behavior patterns her own "sphere of intra-personal psychic functions" (p.35). Later on, she reproduces the pattern of social interactions silently, developing her inner speech in the process. This entails, on the hand, that the internalization of speech happens relatively late in development, and, on the other, that the main role of inner speech is to serve as a culturally embedded form of self-regulation. However, as Geva points out in her chapter, even in Vygotsky's time, other researches posited that inner speech has prior functions related to overt speech acquisition and production, which would entail that infants are able to speak to themselves much earlier than Vygotsky proposed, and that the functions of inner speech are not purely regulative. Many contemporary authors concur in viewing inner speech primarily as a part of a monitoring mechanism related to speech production, which is already in active use during the early stages of the child's speech acquisition and production (e.g., Oppenheim and Dell, 2008, Pickering and Garrod, 2013, Carruthers, this volume, Loevenbruck et al., this volume, Swiney, this volume). In addition, there is some evidence that, even early in development, children make use of linguistic labeling, which is arguably a form of inner speech (Khan, 2013). A significant challenge for the Vygotskian tradition is to reconcile the developmental trajectory and cognitive functions it assigns to inner speech with the roles inner speech plays within these other research programs.

According to Vygotsky, inner speech appears in a variety of semantic and syntactic forms, some of them already visible in private or egocentric speech (Vygotsky, 1986; Berk, 1992, Winsler, et al. 2003, Frawley, 1997). Most commonly, for Vygotsky, inner speech is an abbreviated or condensed form of speech. Subjects are typically dropped out, such that the format of our inner speech consists only of predicates: "predication is the natural form of inner speech; psychologically it consists of predicates only" (1986: 243). Vygotsky thought this condensation can even reach a point where words are no longer used and the subject speaks in

"pure meanings" (1986: 247). Semantically, inner speech was held by Vygotsky to be idiosyncratic, with swift transitions from (conventional) meaning to (idiosyncratic) sense, such that a word said in inner speech will convey a host of connotations and associations known only by the subject. In all, our inner speech differs so much from our overt speech that people would not be able to understand the way we talk to ourselves if transcribed. Despite these difference, Vygotsky held that there is one central feature of overt speech that inner speech retains, which is its dialogical character. Our inner speech, being an internalization of social linguistic exchanges, takes the form of an internal conversation with ourselves. Related to this, inner speech retains the regulatory functions of overt speech mentioned above.

In what follows, we will comment on the extent to which Vygotsky's ideas are alive in the current literature on inner speech, and within this volume itself. At least three of Vygotsky's ideas still enjoy popularity, namely: that inner speech is for self-regulation; that inner speech is the internalization of conversation; and that inner speech is condensed and idiosyncratic.

Inner speech for self-regulation

Despite its centrality to Vygotsky's work, the claim that inner speech is put to use in self-regulation is not an idea that, by itself, could qualify one as importantly Vygotskyan. For instance, while Vygotskyan psychologists have studied the relationship between private speaking (speaking aloud but for oneself) and succeeding in task-switching (Emerson and Miyake, 2003) or in Tower of Hanoi-style tasks, that involve planning and executive control, the finding that (inner) speech may be necessary for efficient planning is compatible with a number of different approaches towards the genealogy, nature and format of inner speech. Inner speech could be used in these tasks simply because it has a role in focusing our attention, for instance—a role that derives from experiences with overt speech in general, and not regulatory exchanges in particular.

Moreover, taken on its own, the notion of self-regulation is quite vague. Understood in a broad sense, it includes most if not all cognitive functions. Short-term planning and executive control, reasoning and motivating functions, as well as meta-cognitive functions including self-reflection and self-evaluation can all be seen as "self-regulative" in some sense. And practically all authors who have written on the functions of inner speech would agree that inner speech has at least one of these functions. Yet only some among them self-identify as Vygotskians, with

others being avowedly non-Vygotskian. Clark (1998), for instance, advances the hypothesis that inner speech is necessary for a metacognitive processes he calls "second-order dynamics." According to Clark, "[t]he coding system of public language is thus especially apt to be co-opted for more private purposes of inner display, self-inspection and self-criticism, exactly as predicted by the Vygotskian treatments…" (1998; 174). Yet, roughly the same idea is defended by both Jackendoff (1997) and Bermúdez (2003) on different grounds, with both being avowedly non-Vygotskian. Thus, while a theorist may be Vygotskian, in a weak sense, simply by proposing a self-regulative function for inner speech, it must be borne in mind that such a view will not by itself reflect any deep debt to Vygotsky, or influence by his views.

Inner speech as internalization of conversations

Another way of being Vygotskyan is to endorse the explanation that Vygotsky gave of how inner speech comes to have the self-regulatory function that it has. According to Vygotsky, inner speech is used in self-regulation because it is an internalization of self-regulatory conversations, such that inner speech is, in a sense, a conversation that we have with ourselves, where we display different perspectives on the world and on ourselves. Some contemporary authors have defended positions close to this one, but embedded in a very different framework. Daniel Dennett (1991), for instance, has suggested that inner speech may be thought as the way the different subsystems in the brain "talk" to each other (see also Clowes (2007), and Frankish (this volume)). This kind of view is also endorsed by defenders of massive modularity who see in language (and so in inner speech) a lingua franca that can receive information from, and send information to, many other modules (Carruthers, 2006). The contemporary author who pushes this Vygotskyan line most strongly is the psychologist Charles Fernyhough (2004, 2009, 2017; but see also Wertsch, 1985). Fernyhough's idea is that inner speech is essentially dialogic in that it involves the simultaneous accommodation of different perspectives that are flexibly coordinated and present the typical triadic relations of external dialogues (between themselves, the other perspectives and objects in the world). The inner dialogue that takes place between these different perspectives is, as are all dialogues, a process that is open-ended and self-regulating (Fernyhough, 2009).

Frankish, **Morin**, and **Gauker** all defend some version of the conversational hypothesis in their chapters for this volume. Frankish, in accounting for the genealogy of Type-2 reasoning,

wherein we typically iterate cycles of inner speech, appeals to the idea that we internalize linguistic exchanges in problem-solving social settings. Gauker takes a stronger position, based on the idea that conversation, internal or external, is a way of cooperating and joint problem-solving that involves articulating propositional thoughts and confronting perspectives. Morin, in turn, appeals to the idea that inner speech involves perspective-switching to explain how inner speech can have the role it has in self-reflection. Among the conversations that we internalize and then run in our heads are conversations that force us to see ourselves and what we do from different perspectives, which is essential to solve the epistemic problem that we are for ourselves. The strong connection each makes to the notion of conversation—where conversational skills begin as *social* skills—qualifies each as Vygotskian in a more substantive sense than if they merely posited a self-regulatory function for inner speech.

Inner speech as condensed and idiosyncratic

As earlier mentioned, Vygotsky also held that the way in which we speak to ourselves differs greatly from the way we speak to other people. Inner speech is fragmentary or condensed, lacks subjects, and has idiosyncratic meanings. Few contemporary authors endorse this last Vygotskyan claim. (Though see Fernyhough (2004), who maintains that inner speech is typically condensed or abbreviated, while allowing that it can take more elaborate formats). **Hurlburt & Heavey** (this volume) report finding no trace, in their studies using Descriptive Experience Sampling, of the peculiarities of inner speech predicted by Vygotsky: "We find that inner speech usually has vocalization aspects quite similar to that of external speech," they note, adding,

> We find agglutination in the Vygotskian sense extremely rarely. We find no more idiomatic sense aspects in inner speech than we do in external speech. In short, we find nothing in our DES studies that supports the Vygotsky claims about inner speech, nothing that suggests any intimate-speaker-and-listener-have-shared-access characteristic that Vygotsky describes.

As earlier noted, Vygotsky also held that inner speech could evolve into a "meanings only" kind of speech, where both linguistic subjects and predicates are dropped. Vygotsky predicted that this form of inner speech/thinking should be found in later stages of development, when self-talk has become so intimate that the subject does not need to produce phonological forms to

reproduce linguistic exchanges. If such a phenomenon existed, it would perhaps be experienced as a kind of "pure" language-like thinking, but without specific associated words. The possibility that there would be this kind of conscious inner speech clashes with the claims of contemporary theorists who emphasize the importance of inner speech's sensory-phonological characteristics to the functions it serves (see, inter alia, Jackendoff, 1997, 2012, Prinz, 2011, Bermúdez, 2003, this volume, Carruthers, 2015, this volume, Frankish, this volume). According to these theorists, the reason that inner speech is involved in conscious thinking is that such thought requires a sensory vehicle; we cannot, they claim, consciously experience non-perceptual representations. At least on its face, Vygotsky's proposal conflicts these contemporary approaches, insofar as inner speech, for him, can play its regulatory roles without having associated sensory character.

What evidence do we have for Vygotsky's meanings-only, "ultra-condensed" inner speech? Curiously, some of the best support may come from Hurlburt and colleagues' work, which calls into question other of Vygotsky's claims. A common experience reported during DES is what Hurlburt and colleagues call "unsymbolized thinking," where these are episodes of thinking that do not include the experience of any conventional symbols, such as phoneme strings or graphic imagery, nor any image-like states, such as visual imagery. The reality of the phenomenon is contested. It has been suggested that reports of unsymbolized thinking stem from confabulation (Engelbert and Carruthers, 2011), that there may be the effect of masking due to the beep used in the experiments (Tye and Wright, 2011), or that unsymbolized thinking is not, after all, totally devoid of sensory/perceptual representations (Byrne, 2011). Hurlburt has responded to these claims (Hurlburt, 2011, Hurlburt and Heavey, this volume), and could be seen to gain added support from current philosophical debates concerning cognitive phenomenology, wherein a number of philosophers—including Siewert (1998) and Kriegel (2015) —argue that we experience this kind of bare conscious thinking. The dispute appears far from settled. An issue likely to attract further attention is whether reports of unsymbolized thinking can be explained in Vygotskian terms, as a kind of rarified inner speech that is aborted before phonological encoding. (See, e.g., Vicente & Jorba (2017) for a proposal along these lines).

Conclusion

Inner speech offers a rare example of a mental process that is both introspectively salient and empirically tractable, where nebulous constructs such as consciousness, language, thought,

concepts, self-awareness, and agency can be investigated in a concrete manifestation. The relative historical neglect of inner speech as a cognitive scientific and philosophical area of study comes with the benefit that exciting discoveries remain to be made, and new explanatory connections unearthed. This anthology has been edited with those future discoveries in mind. Readers will find the tools they need to understand the leading research programs investigating inner speech across the cognitive sciences, and to grasp the central theories and disputes animating current work on inner speech. Our aim is to have opened the door a bit wider to the formulation of innovative, empirically informed, theoretical hypotheses about the nature and functions of our inner voices.

References

Alderson-Day, B., & Fernyhough, C. (2014). More than one voice: investigating the phenomenological properties of inner speech requires a variety of methods. *Conscious Cogn, 24*, 113-114. doi:10.1016/j.concog.2013.12.012

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quaerterly Journal of Experimental Psychology, 18*, 362-365.

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559. doi:10.1126/science.1736359

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford University Press.

Berk, L. (1992). Children's private speech: An overview of theory and the status of research. In R. M. Diaz & L. E. Berk (eds.), *Private speech: From social interaction to self-regulation* (pp. 17–53). Hillsdale, NJ: Erlbaum, Inc.

Bermudez, J. L. (2003). *Thinking without Words*. Oxford: Oxford University Press.

Blakemore, S.-J., Smith, J., Steel, R., Johnstone, E., & Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychological Medicine, 30*(5), 1131-1139.

Byrne, A. (2011). Knowing that I am Thinking. In A. Hatzimoysis (Ed.), *Self-Knowledge* (pp. 105-124). Oxford: Oxford University Press.

Campbell, J. (1999). Schizophrenia, the space of reasons, and thinking as a motor process. *The Monist, 82*(4), 609-626.

Carruthers, P. (1996). *Language, thought, and consciousness : an essay in philosophical psychology*. Cambridge England ; New York, NY, USA: Cambridge University Press.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences, 25*, 657-726.

Carruthers, P. (2006). *The Architecture of Mind*. Oxford: Oxford Univeristy Press.

Carruthers, P. (2011). *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.

Carruthers, P. (2015). The Centered Mind. New York: Oxford University Press.

Carston, R. (2002) *Thoughts and Utterances*. London: Blackwell.

Cassam, Q. (2011). Knowing What You Believe. *Proceedings of the Aristotelian Society, 111*(1pt1), 1-23.

Clark, A. (1998). Magic words: how language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and Thought: Interdisciplinary Themes* (pp. 162-183). Cambridge: Cambridge University Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

Clowes, R. (2007). A Self-Regulation Model of Inner Speech and Its Role in the Organisation Of Human Conscious Experience. *Journal of Consciousness Studies* 14: 59–71.

Dennett, D. C. (1991). *Consciousness Explained*. New York: LIttle, Brown & Company.

Emerson, M. J. & Miyake, A. (2003). The Role of Inner Speech in Task Switching: A Dual-Task Investigation. *Journal of Memory and Language* 48: 148–68.

Engelbert, M. & Carruthers, P. (2011). Descriptive Experience Sampling: What is it good for? *Journal of Consciousness Studies* 18:130-149.

Evans, J., & Frankish, K. (2009). The duality of mind: An historical perspective. In J. Evans & K. Frankish (Eds.), *In Two Minds: Dual Processes and Beyond* (pp. 1-32). Oxford: Oxford University Press.

Fernyhough, C. (2004). Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. *New ideas in Psychology, 22*(1), 49-68.

Fernyhough, C. (2009). Dialogic thinking. In A. Winsler, C. Fernyhough, & I. Montero (eds.), *Private speech, executive functioning, and the development of verbal self-regulation*. Cambridge, UK: Cambridge University Press, 42-52.

Fernyhough, C. (2017) *The Voices Within: the History and Science of How We Talk to Ourselves*. Profile Books.

Fodor, J. (1975) *The Language of Thought*. Cambridge, MA: Harvard University Press

Fodor, J. (2001). Language, thought and compositionality. *Mind and Language*, 16: 1–15.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1), 3-71.

Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations. *International Journal of Psychophysiology, 58*(2-3), 179-189.

Frankish, K. (2004). *Mind and Supermind*. Cambridge: Cambridge University Press.

Frawley, W. (1997) *Vygotsky and Cognitive Science*. Cambridge, MA: Harvard University Press.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences, 13*(7), 293-301.

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hove, UK: Lawrence Erlbaum.

Frith, C. D. (2012). Explaining delusions of control: the comparator model 20 years on. *Consciousness and Cognition, 21*, 52-54.

Gauker, C. (2011). *Words and Images: An Essay on the Origin of Ideas*. Oxford: Oxford University Press.

Geva, S., Bennett, S., Warburton, E. A., & Patterson, K. (2011). Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology, 25*(3), 323-343. doi:10.1080/02687038.2010.511236

Graham, G., & Stephens, G. L. (2000). *When Self-Consciousness Breaks*. Cambridge, MA: MIT Press.

Gregory, D. (2016) Inner Speech, Imagined Speech, and Auditory Verbal Hallucinations. *Review of Philosophy and Psychology* 7: 653-673.

Hoffman, R., Varanko, M., Gilmore, J., & Mishara, A. L. (2008). Experiential features used by patients with schizophrenia to differentiate 'voices' from ordinary verbal thought. *Psychological Medicine, 38*, 1167-1176.

Hohwy, J. (2013). The Predictive Mind. Oxford: Oxford University Press.

Hurlburt, R.T. (2011). *Investigating Inner Experience: Moments of Truth*. Cambridge, MA: Cambridge University Press.

Jackendoff, R. (1996). How language helps us think. *Pragmatics and Cognition, 4*(1), 1-34.

Jackendoff, R. (2012) *A user's guide to thought and meaning*. New York: Oxford University Press.

Jones, S. R., & Fernyhough, C. (2007). Thought as action: inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition, 16*(2), 391-399. doi:10.1016/j.concog.2005.12.003

Junginger, J., & Frame, C. L. (1985). Self-Report of the Frequency and Phenomenology of Verbal Hallucinations. *The Journal of Nervous and Mental Disease, 173*(3), 149-155.

Khan, M. (2013). *Thinking in words: Implicit verbal activation in children and adults* (Ph.D. thesis). Harvard University, Cambridge, MA

Knappik, F. (2017). Bayes and the first person: consciousness of thoughts, inner speech and probabilistic inference. *Synthese*, 1-28.

Kriegel, U. (2015). *The Varieties of Consciousness*. Oxford: Oxford University Press.

Langland-Hassan, P. (2008). Fractured Phenomenologies: Thought Insertion, Inner Speech, and the Puzzle of Extraneity. *Mind & language, 23*(4), 369-401.

Langland-Hassan, P. (2014). Inner Speech and Metacognition: In Search of a Connection. *Mind and Language, 29*(5), 511-533.

Langland-Hassan, P. (2015). Hearing a Voice as one's own: Two Views of Inner Speech Self-Monitoring Deficits in Schizophrenia. *Review of Philosophy and Psychology*, 1-25. doi:10.1007/s13164-015-0250-7

Langland-Hassan, P., Faries, F., Richardson, M., & Dietz, A. (2015). Inner Speech Deficits in People with Aphasia. *Frontiers in Psychology, 6*, 528.

Langdon, R., Jones, S. R., Connaughton, E. & Fernyhough, C. (2009). The Phenomenology of Inner Speech: Comparison of Schizophrenia Patients With Auditory Verbal Hallucinations and Healthy Controls. *Psychological Medicine* 39: 655–63.

Laroi, F., Sommer, I. E., Blom, J. D., Fernyhough, C., Ffytche, D. H., Hugdahl, K., . . . Waters, F. (2012). The characteristic features of auditory verbal hallucinations in clinical and nonclinical groups: state-of-the-art overview and future directions. *Schizophr Bull, 38*(4), 724-733. doi:10.1093/schbul/sbs061

Levelt, W. J. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Martinez-Manrique, F., & Vicente, A. (2010). 'What the...!' The role of inner speech in conscious thought. *Journal of Consciousness Studies, 17*(9-10), 141-167.

Martínez-Manrique, F., & Vicente, A. (2015). The activity view of inner speech. *Frontiers in Psychology, 6*(232). doi:10.3389/fpsyg.2015.00232

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, R. C. (1993). Is the cerebellum a Smith Predictor? *Journal of Motor Behavior, 25*, 203-216.

Nayani, T. H., & David, A. (1996). The auditory hallucination: a phenomenological survey. *Psychological Medicine, 26*, 177-189.

Oppenheim, G. (2013). Inner speech as a forward model? *Behavioral and Brain Sciences, 36*(04), 369-370. doi:doi:10.1017/S0140525X12002798

Oppenheim, G., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*, 528-537.

Pinker, S. (1994). *The Language Instinct*, William and Morrow Co.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(04), 329-347. doi:doi:10.1017/S0140525X12001495

Prinz, J. (2011). The Sensory Basis of Cognitive Phenomenology, in Bayne, T. & Montague, M. (eds.). (2011). *Cognitive Phenomenology*. Oxford: Oxford University Press, 174-196.

Recanati, F. (2004). *Literal Meaning*. Cambridge: Cambridge University Press.

Roessler, J. (2016). Thinking, Inner Speech, and Self-Awareness. *Review of Philosophy and Psychology, 7*(3), 541-557.

Ryle, G. (1972). Thinking and Saying. *Rice University Studies*, 123-134.

Siewert, Ch. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.

Spence, S. A., Brooks, D. J., Hirsch, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain: a journal of neurology, 120*(11), 1997-2011.

Sperber, D. & Wilson, D. (1995). *Relevance*. Second Edition. London: Blackwell.

Sperber, D. & Wilson, D. (2002). Pragmatics, modularity, and mindreading. *Mind and Language*, 17: 3-23.

Stephan, K. E., Friston, K. J., & Frith, C. D. (2009). Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia Bulletin, 35*(3), 509-527.

Swiney, L., & Sousa, P. (2014). A new comparator account of auditory verbal hallucinations: how motor prediction can plausibly contribute to the sense of agency for inner speech. *Front Hum Neurosci, 8*.

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition, 17*(1), 219-239. doi:http://dx.doi.org/10.1016/j.concog.2007.03.010

Travis, C. (1996). Meaning's Role in Truth, *Mind*, 105: 451-466.

Vicente, A. (2014). The comparator account on thought insertion: some open questions. *Phenomenology and the Cognitive Sciences, 13*, 335-353.

Vicente, A. & Jorba, M. (forth.) The Linguistic Determination of Conscious Thought Contents. *Nous*

Vicente, A. and Martínez Manrique, F. (2005). Semantic Underdetermination and the Cognitive Uses of Language. *Mind and Language* 20: 537–58.

Vicente, A., & Martínez-Manrique, F. (2008). Thought, Language, and the Argument from Explicitness. *Metaphilosophy*, 39: 381–401.

Vosgerau, G., & Newen, A. (2007). Thoughts, motor actions, and the self. *Mind and Language, 22*, 22-43.

Wilkinson, S. (2014). Accounting for the phenomenology and varieties of auditory verbal hallucination within a predictive processing framework. *Consciousness and Cognition, 30*(0), 142-155. doi:http://dx.doi.org/10.1016/j.concog.2014.09.002

Wertsch, J. V. (1985) *Vygotsky and the Social Formation of Mind*. Cambridge, MA: Harvard University Press.

Wilkinson, S., & Bell, V. (2016). The representation of agents in auditory verbal hallucinations. *Mind & language, 31*(1), 104-126.

Winsler, A., De León, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language*, 30, 583–608

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal Models in the cerebellum. *TRENDS in Cognitive Science, 2*, 338-347.

Wu, W. (2012). Explaining Schizophrenia: Auditory Verbal Hallucination and Self-Monitoring. *Mind and Language, 27*(1), 86-107.

Wu, W., & Cho, R. (2013). Mechanisms of auditory verbal hallucination in schizophrenia. *Frontiers in Schizophrenia, 4*.