Accepted Version, Pre-Publication

WIREs Wiley Interdisciplinary Reviews

Article Title:
Inner Speech

Article Type:

- ○ OPINION
- ○ PRIMER
- ● OVERVIEW
- ○ ADVANCED REVIEW
- ○ FOCUS ARTICLE
- ○ SOFTWARE FOCUS

Authors:

First author
Peter Langland-Hassan*
https://orcid.org/0000-0002-9913-6471
University of Cincinnati
Langland-Hassan@uc.edu

Abstract

Inner speech travels under many aliases: the inner voice, verbal thought, thinking in words, internal verbalization, "talking in your head," the "little voice in the head," and so on. It is both a familiar element of first-person experience and a psychological phenomenon whose complex cognitive components and distributed neural bases are increasingly well understood. There is evidence that inner speech plays a variety of cognitive roles, from enabling abstract thought, to supporting metacognition, memory, and executive function. One active area of controversy concerns the relation of inner speech to auditory verbal hallucinations in schizophrenia, with a common proposal being that sufferers of AVH misidentify their own inner speech as being generated by someone else. Recently, researchers have used artificial intelligence to translate the neural and neuromuscular signatures of inner speech into corresponding outer speech signals, laying the groundwork for a variety of new applications and interventions.

1. INTRODUCTION

Inner speech lies at the busy crossroads of several active research programs in cognitive science. Thought and language appear together in a single phenomenon of "thinking in words" or "verbal thought"—one implicated in working memory, metacognition, abstract thought, executive control, auditory verbal hallucinations, and much more.

While inner speech has long attracted attention from philosophers and theoretically-minded psychologists (Bain, 1855; Broca, 1861; Lordat, 1843; Sokolov, 1972; Vygotsky, 2012/1962), advances in neuroimaging over the last thirty years have helped to make it a more tractable phenomenon for scientific study. The idea—promoted most prominently by Vygotsky—that inner speech is the controlled, internal *re*use of a capacity that first manifests in the form of *overt* speech remains a powerful current in some contemporary research. (For other early roots of this idea, see also Cooley (1992) and Piaget (1923)). However, much work on inner speech now extends well beyond that tradition into the more general territory of the relation of language to thought. For some, understanding inner speech amounts to grasping the deepest cognitive divide between humans and non-verbal animals (Frankish, 2018; Gauker, 2011; Paivio, 1990). Others highlight inner speech's close connection to perception and motor control as a means to grounding so-called "higher" or "abstract" forms of thought in perceptual and motor capacities found across the animal kingdom (Clark, 1998; Dove, 2014). Within cognitive neuropsychology, inner speech has played a central role in influential theories of auditory verbal hallucinations in schizophrenia (Frith, 1992; Jones & Fernyhough, 2007c; Langland-Hassan, 2008; Shergill, 2003; Wu, 2012).

What may appear, introspectively, to be a unitary phenomenon—the "voice in the head"—is now commonly seen as a multi-component process with dissociable stages or parts. Current research focuses on understanding the interrelation of those parts and their roles in supporting other cognitive processes and abilities, and in explaining psychiatric phenomena such as auditory verbal hallucinations. This Overview summarizes the central questions and emerging conclusions of that research. Other overviews of this landscape aimed at an interdisciplinary audience in cognitive science can be found in Alderson-Day and Fernyhough (2015), Perrone-Bertolotti et al. (2014) and in the interdisciplinary volume edited by Langland-Hassan and Vicente (2018).

2. WHAT IS MEANT BY 'INNER SPEECH'?

Despite the increasing amount of research conducted on inner speech, it can be difficult to characterize the phenomenon in a theory-neutral way. What is it that competing theories of inner speech all seek to explain or shed light upon? Unlike some notions in cognitive science—such as (perhaps) *working memory* (Baddeley, 1994)—the term *inner speech* lacks a commonly agreed operational definition. There is no outward behaviour, or pattern of stimuli and responses, that is simply equivalent to exercising inner speech. Nor is inner speech commonly viewed as a psychological "posit," put forward to explain specific behaviours or experimental results. In this way inner speech is unlike folk psychological states such as *belief*, *desire,* and *intention,* which, for many in cognitive science, are treated as theoretical posits—elements of a "theory of mind" whose ontological commitments are vindicated by the power of the theory they appear in to predict and explain human behaviour (Chihara & Fodor, 1965; Dennett, 1991; Fodor, 1987; Sellars, 1956).

Instead, inner speech first appears on the scene as an introspectively salient component of everyday experience. We gesture at it with phrases such as "the little voice in the head," or "talking to oneself silently." In this way, 'inner speech' is perhaps like 'mental imagery.' One can use the term to mark an introspectively salient phenomenon, while allowing for scientific disagreements about its ultimate nature (Block, 1981; Kosslyn, Thompson, & Ganis, 2006; Pylyshyn, 2002).

To linger a moment on this important point, let us look closely at an influential attempt at a more formal definition of inner speech—though one that still aims at theoretical neutrality. In an extensive review paper, Alderson-Day and Fernyhough (2015) propose that inner speech can be defined as "the subjective experience of language in the absence of overt and audible articulation" (2015, p. 1). Grandchamp *et al.* (2019, p. 2) endorse this definition as well, adjusting it slightly to: "the subjective experience of verbalization in the absence of overt articulation or sign." These definitions—with their appeal to "subjective experience"—highlight inner speech's close connection to introspection and first-person experience. Nevertheless, most in cognitive science—including Alderson-Day and Fernyhough, and Grandchamp *et al.*—do not assume that inner speech *must* occur consciously, nor that it must involve some attending "subjective experience." And, certainly, when it comes to characterizing the phenomenon in a theory-neutral way, it should not be assumed that inner speech must always have an attending "subjective experience."

Suppose, then, that we remove "subjective experience" from Alderson-Day and Fernyhough's definition. We are left with: *inner speech is language in the absence of overt and audible articulation*. Now, strictly speaking, what you are reading—these very words on the page—are instances of language in the absence of overt and audible articulation. So it seems we should specify that inner speech is *inner* language—in the sense of being *internal to the mind*—that occurs in the absence of overt articulation. That gives us: *inner speech is inner language in the absence of overt and audible articulation*. This is similar to the definition offered by Oppenheim and Dell, who propose that inner speech is "typically characterized as either the activation of abstract linguistic representations or a detailed articulatory simulation that lacks only the production of sound" (2010, p. 1147). And yet, we should leave open the possibility that inner speech can also occur simultaneously with overt articulation. For it may be that as we talk aloud we are also talking in the head; and, indeed, a number of current theories, which see inner speech as a component of outer-speech prediction and monitoring, hold this to be so (Perrone-Bertolotti et al., 2014; Pickering & Garrod, 2013; Postma, 2000). Adjusting our definitions again to remove the suggestion that inner speech cannot occur at the same time as outer speech, we are left with: *inner speech is inner language.* (Or, adjusting Oppenheim and Dell's definition: inner speech is the activation of abstract linguistic representations.) It is not clear that these characterizations move us forward from "the little voice in the head." Worse, they seem to suggest that 'inner speech' is simply a catchall phrase for any linguistic processing at all. We have taken a wrong turn—but where?

The problem is not with Alderson-Day and Fernyhough's or Oppenheim and Dell's characterizations of inner speech, but with the attempt to treat them as formal definitions or *theories* of inner speech. We should instead see them as rough-and-ready attempts to *point* at a phenomenon we would like to better understand, by noting some of its salient and common features: inner speech is a kind of mental use of language, or simulation of speaking, that, often enough, occurs consciously and in the absence of any overt articulation. Which, if any, of the features so-noted are essential to inner speech can be left open. We can imagine ourselves in the position of ornithologists who have on occasion caught sight of what looks to be a new species of bird. We can give the species a name and list the salient features we use to identify it, while leaving open to scientific investigation such questions as whether the birds so named really form a single species, whether each of its members has all (or only some) of the features we have used to identify them, and so on. Disagreements can be expected concerning borderline cases, as is common when rigorous theorizing and experimentation has begun.

Whether inner speech continues to be a useful construct to cognitive science will depend upon whether the subset of linguistic processing picked out by the term 'inner speech' constitutes an explanatorily useful kind. A possibility to bear in mind is that fragmentation will occur within that subset itself, such that theorists find it useful to posit distinct *kinds* of inner speech. At that point we again need to ask what it is that unites them all as kinds of inner speech, as opposed to language-related cognition generally. Nevertheless, despite this possibility of fragmentation, there are reasons (explored below) for thinking that inner speech's close relationship to first-person experience and introspection will in fact *increase* the chances that it remains an explanatorily useful kind. The fact that people can consciously control and report upon their own inner speech (in contrast to some other forms of linguistic processing) enables it to serve as a lever for therapeutic interventions, the development of new technologies, and further scientific theorizing.

3. METHODS FOR INVESTIGATING INNER SPEECH

There are many strategies for eliciting and measuring inner speech. Often two or more methods are used in collaboration. For instance, a task presumed to rely upon inner speech—such as silently judging whether two words rhyme—may be combined with simultaneous neuroimaging to locate the neural substrate of inner speech (Geva, Jones, et al., 2011b). With a theory of the neural bases for inner speech in hand, neuroimaging can then be conducted on a population of interest—such as those suffering from auditory verbal hallucinations—to see whether activations associated with inner speech occur when AVHs are reported (Di Biase et al., 2019; McGuire et al., 1995; Shergill, 2003; Simons et al., 2010). Such imagining can also be used in tandem with subjective questionnaires that query a variety of qualitative and phenomenological features of inner speech, to explore the neural variations corresponding to variations in how inner speech appears to subjects from the first person perspective (Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones, & Fernyhough, 2018; Grandchamp et al., 2019; Kühn, Fernyhough, Alderson-Day, & Hurlburt, 2014; McCarthy-Jones & Fernyhough, 2011).

Having noted the value of combining methods in this area, we can look in more detail at a variety of individual means for eliciting and studying inner speech.

3.1 Introspection: structured interviews, surveys, and self-reports

A person's introspective access to their own inner speech can be exploited as a means to gathering information about its nature and function (Alderson-Day & Fernyhough, 2014; Morin, 2018). One approach to doing so is the Descriptive Experience Sampling (DES) paradigm, developed by Hurlburt and colleagues (Heavey & Hurlburt, 2008; Hurlburt & Akhter, 2006). Participants in DES studies go about their daily activities while wearing a portable beeper that emits a beep at random (roughly 15 minute) intervals, after which they are to take notes on the thoughts, emotions, images, and other mental phenomenal that populated their conscious experience immediately prior to the beep. The paradigm's aim is to allow one to "sneak up on" oneself, so to speak, and, in so doing, decrease the likelihood that reports about the nature of one's conscious experience will be influenced by background theoretical beliefs and expectations.

In one application, the DES paradigm was used to assess the frequency of inner speech across a population, simply by analysing the percentage of "beep" reports where forms of inner speech were described. Data from one such DES study were used to arrive at a frequently-cited statistic that inner speech occurs, on average, during 25% of waking life—though it bears noting that individual reports vary widely, with some participants reporting that inner speech occurs nearly all the time and others rarely (Heavey & Hurlburt, 2008).

More controversially, DES has been used to taxonomize and discriminate different forms of inner speech. For instance, in addition to inner speaking, Hurlburt et al. cite results from DES studies in support of countenancing comparatively recondite phenomena such as "unworded inner speech"—i.e., the experience of inner speaking without the experience of words—and "unsymbolized thinking"—i.e., the experience of thinking without the subjective presence of any related sensory feature or symbolic structure. Further, they propose a strict distinction between inner speaking and the inner hearing of oneself speaking. In a development that could help to validate such distinctions, DES has been used in combination with neuroimaging (R.T. Hurlburt, B. Alderson-Day, Simone Kühn, & C. Fernyhough, 2016b; Kühn et al., 2014). In one study, experimenters found that "spontaneous" inner speech—as detected by DES—had significantly different neural correlates from the kinds of inner speech elicited by some common behavioural measures (Hurlburt et al., 2016b). In another study, neuroimaging differences were found corresponding to differences in DES reports of inner speaking versus inner hearing (Kühn et al., 2014).

Other broadly introspective means for assessing inner speech involve the use of structured questionnaires, such as the Self Talk Inventory (STI; Burnett, 1996), the Self-Talk Use Questionnarie (STUQ; Hardy, Hall & Hardy, 2005), and the Self Talk Scale (STS; Brinthaupt, Hein, & Kramer, 2009). More recently, the Varieties of Inner Speech Questionnaire (VISQ) (McCarthy-Jones & Fernyhough, 2011) and the VISQ-R (Alderson-Day et al., 2018) have taken specific aim at distinguishing different forms of inner speech. Inspired by Vygotsky's (1934/1987) developmental theory of inner speech, the VISQ asks participants to rate the phenomenological properties of their inner speech according to four factors: dialogicality (the extent to which inner speech occurs as a back-and-forth conversation), the evaluative or motivational force of the inner speech, whether the "voices" of others occur in one's inner speech, and the degree of "condensation" of inner speech (i.e. the extent to which sentences are abbreviated in inner speech, while their meaning appears retained) (McCarthy-Jones & Fernyhough, 2011). Several studies have used the VISQ and VISQ-R to statistically link reported features of inner speech to psychiatric conditions, including anxiety, depression, hallucination-proneness, and self-esteem (Alderson-Day et al., 2014; de Sousa, Sellwood, Spray, Fernyhough, & Bentall, 2016).

Other recently developed subjective assessment tools include the Nevada Inner Experience Questionnaire (Heavey et al., 2019) and the Internal Representations Questionnaire (IRQ) (Roebuck & Lupyan, 2020). Many of these tools have distinct aims and foci, making their results difficult to compare. For instance, whereas the questions on the VISQ-R aim to uncover different *forms* of inner speech, the IRQ seeks to shed light on the *propensity* of individuals to use inner speech in problem solving not related to interpersonal communication. Roebuck and Lupyan (2020) found correlations between the propensity to generate internal verbalizations (as measured by the IRQ) and performance on objective matching tasks that probed the influence of semantic and phonological features on response times.

3.2 Behavioural Measures

Even if there are no outward behaviours that can be taken as strictly equivalent to engaging in inner speech, there are many that are considered strong evidence for the active use of inner speech. For instance, the ability to silently judge whether two written words rhyme, or whether two words are homophones, is standardly taken as good evidence for intact inner speech (Geva, Bennett, Warburton, & Patterson, 2011; Langland-Hassan, Faries, Richardson, & Dietz, 2015). Such studies intermix word pairs that rhyme despite having dissimilar spellings (e.g., 'eye' and 'tie') and that have similar spellings while not rhyming (e.g. 'wood' and 'blood'), to assure that participants cannot show competence through a visual strategy. A related measure is to ask participants to silently count the number of syllables in a word (Levine, Calvanio, & Popovics, 1982), though this is potentially confounded by the general difficulty many have in counting syllables correctly.

Abilities in these tasks are standardly taken as indirect measures of general inner speech capacities. For instance, when people with aphasia are shown to have difficulty judging rhymes silently, this can be taken as evidence that their outer speech deficits are mirrored by inner speech deficits (Geva, Bennett, et al., 2011; Langland-Hassan et al., 2015; Stark, Geva, & Warburton, 2017). Note, however, that some studies have shown an imperfect correlation between inner and outer speech abilities in people with aphasia (Geva, Jones, et al., 2011a; Levine et al., 1982). The precise relationship between the two—and the extent to which one may be better preserved than the other in people with aphasia—remains a matter of ongoing investigation (Fama, Hayward, Snider, Friedman, & Turkeltaub, 2017; Fama et al., 2019; Stark et al., 2017).

It is worth reflecting on potential confounds in using silent rhyming tasks as assessments of inner speech abilities, as doing so raises interesting questions about what we should take inner speech to be. First, it is possible that one might generate words internally but simply have difficulties in judging whether they rhyme. Arguably, this would not be an impairment in inner speech *per se*, but rather in judging rhymes. Steps should therefore be taken to ensure that a deficit in silently rhyming is not simply a deficit in judging rhymes generally by, for instance, asking participants to judge rhymes presented to them aloud (Langland-Hassan et al., 2015). A problem, however, is that even listening to words—one after the other—so as to judge whether they rhyme may require quickly repeating the words back to oneself; such self-generated repetition could itself be considered a form of inner speech. In that case, it becomes difficult to distinguish the ability to judge rhymes from the ability to generate inner speech. A possible response here is that holding an aurally perceived word or phrase in working memory is not equivalent to generating that word or phrase in inner speech—perhaps on the grounds that a mental episode is only inner speech when it results from one's own articulatory intentions. As a related matter, we should consider the possibility that a population with impaired working memory might be capable of generating inner speech, yet incapable of "holding" such episodes in mind long enough to reflect on whether one internally generated word rhymes with another. This raises the question—still open—of whether inner speech itself can occur in the absence of the ability to hold and reflect upon it in working memory, or whether, instead, the lack of an ability to intentionally maintain internally generated verbal utterances in mind should be seen as the lack of inner speech itself.

Another feature of these methods for eliciting and assessing inner speech worth flagging is that they assume inner speech to carry phonetic information about the sounds of words. Some argue that "condensed" or "impoverished" versions of inner speech are common and that these either fail to carry such information, or that do so only in a diminished sense (Fernyhough, 2004; Oppenheim & Dell, 2010; Vicente & Jorba, 2019). If this is correct, then behavioural measures that assume inner speech to necessarily have a strong auditory-phonological component are only suitable for the study of a sub-set of all inner speech episodes. Moreover, researchers seeking to build on Levelt's (1989) influential model of speech production typically identify several distinct stages of inner speech generation—distinguishing representation of the semantic "message" to be communicated from elements such as the syntactic frame, phonemic structure, and articulatory plan (Grandchamp et al., 2019; Vicente & Jorba, 2019). They may hold that silent rhyming taxes a relatively late stage of inner speech—one richer than the day to day voice in the head that isn't fixated on rhyme judgments. Though see Langland-Hassan (2018) for an argument that all inner speech episodes do indeed have an auditory-phonological component.

A last hybrid method for assessing inner speech that has proven influential is to ask participants to repeat words—especially "tongue-twisters"—in the head and to ask them to indicate when they make inner-pronunciation errors (Oppenheim & Dell, 2008). On the basis of such results, Oppenheim and Dell conclude that inner speech is "impoverished at the featural level" compared to overt speech (2008, p. 536). This is classified as a hybrid method here because it involves both a (behavioural) inner speech eliciting task (repeating "tongue-twisters," silently) and introspective judgments, on the part of participants, about features of their own inner speech.

3.3 Neuroimaging

Neuroimagining has been used with increasing success to locate the neural networks responsible for inner speech (Geva, 2018b; Hurlburt et al., 2016b; Loevenbruck, Grandchamp, Rapin, Nalborczyk, & Dohen, 2018; P. K. McGuire et al., 1996b; Perrone-Bertolotti et al., 2014). Such studies typically incorporate two means for detecting inner speech simultaneously. In some versions, participants are simply asked to repeat words or phrases silently, "in the head" during scanning (P. K. McGuire et al., 1996b). In any such paradigm, participants are trusted to generate (and therefore detect) their own inner speech at the same time that the imaging device (fMRI or PET, typically) is used to detect related neural activity. Alternatively, participants may be given a task assumed to require inner speech, such as judging rhymes silently (Geva, Jones, et al., 2011b), or reading silently (Yao, Belin, & Scheepers, 2011), while scanning takes place.

The first method—asking participants to simply generate inner speech on their own—has the limitation that it presumes participants know what inner speech is and that they know when they are--or are not—generating it. This may not seem problematic if we are simply interested in the neural correlates of carefully repeating a sentence in one's head. However, questions can be raised about whether such methods provide an adequate means for discovering the neural correlates of other kinds of inner speech that may occur "in the wild"—i.e., outside of experimental contexts—which some researchers (especially followers of Vygotsky (2012/1962)) hypothesize is more "condensed" or abbreviated in nature, leaving out auditory, phonological, and syntactic elements (Grandchamp et al., 2019; Jones & Fernyhough, 2007a). It also rules out the study of any forms of inner speech that we are not well-suited to introspectively detect. The second method—conducting neuroimaging during a concurrent task presumed to require inner speech—has the above-noted limitation that the task used to elicit inner speech may exploit only one aspect of inner speech, or only one type of inner speech.

4. THE COMPONENTS OF INNER SPEECH

Inner speech appears initially, to introspection, to be a unitary phenomenon: a single *voice* in the head without obvious *parts* or *components* (other than the distinct words that make up an inner utterance). But a moment's reflection shows this picture to be too simple. Typically, when we say something in the head—e.g., "*Remember bananas, baking soda, and whipped cream…*"—there is an auditory-phonological component to the state. We are aware not just of the semantic content (or *meaning*) of the episode; we are also, in some sense, aware of the sounds or phonemes of the relevant words. Such auditory elements account for the difference between saying a phrase in one's head with an American as opposed to British accent, for example. And they also account for the usefulness of inner speech to judgments concerning rhymes and homonyms which are, at bottom, auditory judgments. These auditory-phonological features contrast to the semantic features, which are what we can think of as the meaning of the phrase. Two inner speech utterances in distinct languages—'It is snowing' and 'Il neige', for example—might have the same semantics yet quite different auditory-phonological characteristics. And while the auditory-phonological component and semantic component of an inner speech episode are perhaps the most obvious elements that combine in ordinary inner speech, there are likely others. For instance, it may be that there is a separate articulatory component to inner speech—a representation or simulation of the kinds of lip, tongue, and throat movements necessary for the corresponding overt utterance (Loevenbruck et al., 2018). The precise number and nature of the components of inner speech—and their interrelation—is an active area of research.

An important question in this vicinity is whether inner speech itself is to be considered an amalgam of these associated components or, rather, should be identified with just one of them. Jackendoff (1996) can be read as proposing that inner speech *proper*—i.e., the voice in the head that "never shuts up" (p. 10)—should be identified with the auditory-phonological component alone. Thought itself, for Jackendoff, remains unconscious, occurring in an amodal symbolic format. "We become aware of thought taking place – we catch ourselves in the act of thinking," Jackendoff explains, "only when it manifests itself in linguistic form" (1996, p. 10). And yet, for Jackendoff, this linguistic form itself lacks the semantics of thought—otherwise it would simply *be* thought.

Gauker (2011, 2018) defends a contrasting view, arguing that inner speech itself ought not to be confused with "auditory verbal imagery." Just as we distinguish between our auditory perception of another's speech and that speech itself, he proposes, we should distinguish between the representation of our own inner speech—which takes place through the use of auditory verbal imagery—and inner speech itself, which lacks any sensory component. On Gauker's view, inner speech is amodal (i.e., non-sensory) in nature. However, unlike some more traditional views, such as Fodor's (1975), where amodal mental representations occur in a proprietary language of thought, Gauker holds that inner speech is both amodal and occurs in our spoken, "natural" languages. Certain neural states, on his view, are also symbols of English sentences, though the symbols do not have or represent any auditory-phonological properties. On his view, inner speech occurs (in an English speaker) when neural states of the sort that realize symbols of English are activated in a relevant sequence. An interesting theoretical question raised by this proposal is: in virtue of what does a sequence of meaningful symbols—realized in the brain, or anywhere else—count as symbols of one natural language, as opposed to another, once they have been divorced from the phonological (and graphic) features distinctive of that language?

A reviewer proposes, in response to this question, that it could be in virtue of the symbols being "amodal yet language-specific." However, the challenge remains to say what qualifies an amodal symbol as being specific to one language as opposed to another. To this end, it is worth considering a distinction put forward by many linguists between a *phoneme* and a *phone*. Phones are sometimes described as "concrete" sounds events—having a duration, spatial location, and so on—whereas phonemes are described as amodal "abstractions" over phones, such that a phoneme is a set of phones that tend to be heard as the same linguistic building-block. Whereas changing a phoneme in a word changes the word's meaning, altering a phone may not. (Two phones heard as the same phoneme are called *allophones.*) Up to this point we have simply of spoken "auditory-phonological" representations without distinguishing phones from phonemes, as the relationship between mental representations, on the one hand, and the phone/phoneme distinction, on the other, is vexed and underexplored. However, it may nevertheless seem that, while the mental representation of phones must involve representing auditory features, and thus involve sensory representations of some kind, the representation of more abstract phonemes may be amodal in nature. If that were right, representations of phonemes could then be said to be both amodal and language-specific, as phonemes are indeed the most basic building-blocks of a particular language relevant to distinguishing one word from another (there are 44 in English). However, on this interpretation of phonemes as amodal, the point remains that representations of phonemes would lack the semantics of the ordinary words of a natural language; for representations of phonemes still concern (abstractions over) sounds; thus they would not have the kind of informational content we typically attribute to ordinary words. We can thus re-pose the question that ended last paragraph as: what would qualify a set of symbols *with natural-language-like meanings* as being symbols of one language, as opposed to another, when those symbols occur distinct from any representations of the phones, phonemes, graphemes, and articulatory plans distinctive of the language?

Returning to the question of how to understand the relation among the components of inner speech, other theorists seek a middle ground where inner speech has both semantic and auditory-phonological features essentially. Carruthers (2011, 2018) argues that the auditory-phonological and semantic features of inner speech, while distinct, are "bound into" each other in roughly the way in which, in visual perception, the shape and colour properties of an object are "bound into" a single visual perception of an object, despite being processed in separate areas of visual cortex. Langland-Hassan (2014) has argued, on theoretical grounds, that the semantic and auditory-phonological components of inner speech cannot be "attached" to each other in any literal sense, and that this has implications for the kinds of cognitive roles that inner speech might serve (see below, on metacognition). The gist of this argument is that mental representations are individuated by the information they carry, or by what they represent; if an inner speech utterance of the word 'cow' both represents cows *and* the sound of the English word 'cow', then there must be two distinct representations at work, and not just one. Bermúdez (2018) responds to Langland-Hassan's arguments in defence of the view that an inner speech episode can be a single mental state that incorporates semantic and auditory-phonological components simultaneously, both as essential features. Another option in this vicinity is to hold that 'inner speech' can mark an auditory or articulatory component in some cases, and, in others, a non-sensory and non-articulatory component; Oppenheim and Dell's (2010) "flexible abstraction hypothesis" has this flavour.

Researchers in psychology and neuroscience have in general placed less emphasis on the question of which components of inner speech qualify as inner speech *proper,* or which occur consciously, seeking instead to map out in cognitive wiring diagrams the different components of inner speech and their interrelations. Such diagrams typically include, at a minimum, discrete stages for *conceptualization* (where a pre-verbal "message" is generated), *formulation* (where relevant phonemic features and a syntactic "frame" are selected), and *articulatory planning* (where motor instructions are selected and triggered) (Grandchamp et al., 2019; Levelt, 1989; Loevenbruck et al., 2018). In addition, researchers have attempted to place the components of inner speech within a larger cognitive architecture for motor control—one featuring components such as *efference copies, forward models,* and *comparators* (Pickering & Garrod, 2013). These constructs are familiar to theories on which ordinary perception and action involves a continuous prediction of sensory input, based on motor commands, and a comparison of actual sensory feedback to the prediction (Jeannerod, 1995; Miall, Weir, Wolpert, & Stein, 1993; Wolpert, Miall, & Kawato, 1998). Where mismatches (or "errors") are detected, the system is alerted to make adjustments. Pickering and Garrod (2013) have proposed that inner speech can be viewed as the sensory output of a forward model that serves as a prediction of the sensory input corresponding to planned overt utterances. On this sort of view, inner speech occurs as part of the monitoring of overt speech; what we experience as inner speech—in the absence of overt speech—is simply this capacity for overt speech prediction run "offline."

A challenge for the view of inner speech as a forward model, noted by Oppenheim (2013), is to explain how errors are detected in inner speech itself. Oppenheim and Dell (2008) have found that participants report errors in their own inner speech when asked to repeat "tongue-twisters" silently, in the head. On the basis of the kinds of errors they report, and the differences between those errors and the errors observed in overt speech, Oppenheim and Dell conclude that inner speech is phonetically "impoverished" compared to overt speech, occurring at a more "abstract" level of language processing (i.e. before full phonemic features are selected corresponding to the desired message). A number of theorists have attempted to explain how errors could be detected in inner speech, while still viewing inner speech part of a larger prediction-comparison architecture for motor control, by positing multiple stages at which the components of inner speech are predicted and compared to internally generated results (Grandchamp et al., 2019; Vicente & Jorba, 2019).

5. INNER SPEECH'S PLACE IN THE BRAIN

Inner speech's neural substrates are increasingly well understood, with the *proviso* that proposals are inevitably attached to assumptions concerning which tasks or behaviours elicit or exploit inner speech *proper* (see Geva (2018a) for a review). As we will see, language production areas, such as the left inferior frontal gyrus (lIFG), are the most significantly activated during inner speech, though, in some conditions, there is significant activation of language-perception areas (such as the superior temporal gyrus (STG)) as well (Tian & Poeppel, 2010). Some of the earliest PET and fMRI imaging work on inner speech explored differences in neural activation between an inner speech condition—where subjects were asked to silently produce simple phrases in the head, incorporating specific words—and an imagining hearing speech condition, where subjects were asked to imagine hearing someone else say particular phrases (P. K. McGuire et al., 1996a; Shergill et al., 2001). McGuire et al.'s (1996) results showed activation in the lIFG in both conditions, and added activations during the “imagining another’s speech” condition in the left premotor cortex, supplementary motor area, and left temporal cortex. Shergill et al.’s (2001) results were broadly consistent with these. Much of this work has been conducted in the process of comparing the neural activations of healthy controls to those of people suffering from auditory verbal hallucinations, with each group being asked to generate inner speech and auditory verbal imagery of various kinds (Allen, Aleman, & McGuire, 2007; Jones & Fernyhough, 2007b; P K McGuire et al., 1996; Shergill, 2003).

Other imaging studies have used silent rhyme judgments, of the sort described above, as a means for eliciting inner speech. In a typical study of this kind, participants are asked to press a button to indicate whether a pair of presented words rhyme, while neuroimagining (typically fMRI) occurs. In a similar paradigm, Kell et al. (2017) used fMRI to investigate the neural correlates of silent reading, finding strong bilateral similarities to those of reading aloud and evidence that inner speech during reading codes detail as fine-grained as consonant voicing. Working with patients preparing for surgery for intractable epilepsy, experimenters have taken recordings directly from the surface of the brain (electrocorticograpy) to assess the neural correlates of silently voicing words (Martin et al., 2014; Martin, Iturrate, Millán, Knight, & Pasley, 2018; Pei et al., 2011). Across a number of such studies, the areas most commonly implicated remain the lIFG and areas thought to be involved in phonological assembly and processing, such as the supramarginal gyrus (Kell et al., 2017; Owen, Borowsky, & Sarty, 2004; Pei et al., 2011; Poldrack et al., 2001; Pugh et al., 1996; Tian & Poeppel, 2010). Geva et al. (2011b) found compatible evidence on a population with post-stroke aphasia; lesions to the left inferior frontal gyrus pars opercularis (Brodmann Area 44) and its adjacent white matter (the arcuate fasciculus) were predictive of performance on silent rhyming tasks.

As noted, theories concerning the neural substrates of inner speech are tied to assumptions about the kinds of tasks that elicit inner speech. While many of the above studies have focused on “monologic” inner speech—where one simply speaks or repeats single words or phrases—some have emphasized the need to explore the neural correlates of “dialogic” inner speech, where participants are asked to imagine a conversation among multiple parties (Alderson-Day et al., 2016). It has been argued that such studies have greater ecological validity, insofar as a central role for inner speech in everyday life is negotiating interpersonal dynamics (Fernyhough, 2013). (Though see Gregory (2017) for an argument that inner speech is not by nature dialogic.) Alderson-Day *et al.* (2016) found that, in contrast to monologic inner speech, dialogic inner speech conditions activated the superior temporal gyrus (STG) bilaterally, the left inferior and medial frontal gyri, and various posterior midline structures. Unlike monologic inner speech, the extended networks activated during dialogic inner speech showed significant overlap with networks activated during Theory of Mind tasks (including the right posterior STG), suggesting that dialogic inner speech has a closer connection to social cognition than the monologic inner speech explored in most other imaging studies.

In another attempt to increase the ecological validity of inner speech neuroimaging, Hurlburt et al. (2016a) compared activation during traditional elicited (monologic) inner speech tasks to that during “spontaneously” generated inner speech (i.e. inner speech generated without an explicit prompt). They were able to measure the latter by first training participants in the Descriptive Experience Sampling (DES) paradigm described above. Participants then heard beeps at random intervals while being scanned and indicated the instances when they had been engaged in inner speech just prior to the beep. This allowed the neuroimaging recording at those moments to be compared to that recorded during elicited tasks. The experimenters found distinct neural signatures associated with elicited versus spontaneous inner speech, with elicited inner speech decreasing activation in Heschl’s gyrus and increasing activation in the lIFG, while spontaneous inner speech showed increased activation in Heschl’s gyrus and no significant effect in the lIFG. Such results raise the possibility that what is studied in traditional elicited inner speech neuroimagining studies is a somewhat distinct phenomenon from the voice in the head familiar to everyday experience. The matter remains unsettled, however, as Grandchamp et al. (2019) failed to find the same differential relation between lIFG and auditory cortex in spontaneous versus elicited inner speech.

6. THE FUNCTIONS OF INNER SPEECH

However common inner speech may be to everyday experience, the answer to *why* it is that we engage in inner speech is less clear. One way to focus attention on the question of inner speech's function is to ask how our lives would be different if we were no longer able to generate inner speech. Aside from the obvious—such as that we could no longer plan just what to say—which cognitive tasks would become more difficult, or impossible? Further, how would our awareness of our own thoughts change? Would we be in a situation where we could still think all the thoughts we used to think, while not being able to find the words with which to express them (even "in the head")? Or would it be better described as a situation where we simply could not think some of the thoughts we used to think? These questions all concern the various cognitive—and metacognitive—functions inner speech might play (see also Alderson-Day and Fernyhough (2015)).

One of the most well-known and widely researched roles for inner speech is in maintaining information in working (or "short term") memory. Early research on working memory found that the phonemic features of words, together with word length, influence how easily they are remembered, with lists of phonetically similar words being more difficult to remember than phonetically-dissimilar words (Baddeley, 1966a, 1966b). This result meshes with the intuitive datum that, when presented with a list of words to remember, we repeat them in inner speech, and that such inner speech has auditory-phonological features. Baddeley and Hitch (1974) posited a "phonological loop" as a sub-component of a larger working memory system, through which verbal material could be kept cognitively available (Baddeley, 1992, 2007). Predictably, people with aphasia—unable to easily produce language, usually as a result of stroke—have corresponding working memory deficits (Caspari, Parkinson, LaPointe, & Katz, 1998; Friedmann & Gvion, 2003). However, it has been noted that, if working memory is required for language comprehension and production (e.g. in resolving contextual ambiguity, or planning speech), then language could be impaired due to a general working memory deficit, and not the other way around (Wright & Shisler, 2005). Thus, the relation of inner speech to working memory—and the dependency on one of the other—remains an open research question.

Inner speech has also been identified as a resource supporting executive function, most prominently in tasks that require switching between different responses and rules (Baddeley, Chincotta, & Adlam, 2001; Cragg & Nation, 2010; Miyake, Emerson, Padilla, & Ahn, 2004). Such studies often exploit a dual task structure involving *articulatory suppression*—a method common to experiments seeking to identify a linguistic contribution to task performance. Participants are asked to engage in a simple linguistic task assumed to tax the language system, such as repeating a string of words, or repeating back recently heard sentences or words, while concurrently completing another task, such as sorting according to a rule. In a control condition, participants complete the same sorting task while concurrently completing non-linguistic distractor task thought to impose a comparable cognitive load to the linguistic distractor task in the other condition (Alloway, Kerr, & Langheinrich, 2010; Saeki & Saito, 2004). Decreased performance in the linguistic interference task is thought to result from the inability to draw on inner speech as a resource. Closely related to executive function are processes supporting planning, problem solving, delay of gratification, and other forms of self-regulation. Here, too, inner speech has been argued to play a role. Barkley (1997) proposes that delays in the development of self-directed speech may be partly to blame for problems with self-regulation seen in attention deficit hyperactivity disorder (ADHD). More generally, (and as discussed by Petrolini, Vicente, and Jorba (in draft)), cognitive-behavioural therapies employing forms of structured "self-talk" have been relatively successful in supporting controlled behaviour (Meichenbaum & Goodman, 1971; Miranda, Presentación, Siegenthaler, & Jara, 2011). It is natural to infer from the efficacy of self-directed outer speech that inner speech often plays the same supporting role. The role of inner speech executive function and self-regulation remain open and contentious areas of research, however, in part due to a lack of uniformity in how each of these notions—viz., 'executive function', 'self-regulation', and 'inner speech'—are understood by researchers (Petrolini, Vicente, & Jorba, in draft).

Without diminishing the importance of the above results, one might ask whether inner speech also supports cognition in some less peripheral way—serving, for instance, as a medium for conceptual or "abstract" thought itself, and not merely as a cognitive aid to memory or executive function. While the hypothesis that inner speech serves as the primary medium for conceptual thought or reasoning in general has had advocates over the years (Frankish, 2018; Gauker, 2011; Paivio, 1971) experimental support for its serving as a *necessary* resource for such remains slight. The evident capacities for complex cognition shown by non-human animals, pre-linguistic humans, and language-impaired humans all speak against the strong hypothesis that conceptual thought—however it is to be understood—requires the active use of inner (or outer) speech. Nevertheless, there is increasing support for the idea that *abstract* thought—defined as thought about objects or properties not easily perceived—is supported by language, and by inner speech in particular (Borghi et al., 2017; Dove, 2014; Gentner & Boroditsky, 2001; Lupyan & Bergen, 2016). Researchers working within the embodied cognition tradition, in particular, have seen language perception and production—and its internalization of each, via inner speech—as a resource for explaining the human ability to form and make use of abstract concepts in broadly sensorimotor terms (Barsalou, Santos, Simmons, & Wilson, 2008; Dove, 2018, 2019).

Of course, human thought—even "abstract thought"—is not a monolith but a constellation of distinct and mutually-supporting cognitive capacities; there is reason to see inner speech as just one resource in that larger group (Lupyan & Bergen, 2016). It has been proposed that inner speech serves as a means for integrating information across distinct cognitive modules (Carruthers, 2002); and there is evidence from experiments employing verbal shadowing that suppressing inner speech interferes with the ability to integrate shape and colour information (Hermer-Vazquez, Spelke, & Katsnelson, 1999). An alternative proposal, with support from verbal interference studies and studies with people with aphasia, is that inner speech serves as an aid to fixing attention on a particular feature shared by multiple items—such as colour—when they have little else in common (Lupyan, 2009; Lupyan & Mirman, 2013). More generally, psychologists have emphasized the capacity for inner speech to serve as a kind of "control system" or tool for augmenting and manipulating non-verbal thought, without language or inner speech itself constituting a representational medium for thought (Roebuck & Lupyan, 2020).

A possibility that has interested empirically-oriented philosophers, in particular, is that inner speech might constitute the central means by which people become conscious of (or, alternatively, gain metacognitive awareness or) their own thoughts (Clark, 1998; Frankish, 2004; Langland-Hassan, 2014; Machery, 2018; Martinez-Manrique & Vicente, 2010; Morin, 2005). This thesis is typically defended with respect to a certain sub-class of thought—such as conceptual or "propositional" thoughts of the sort neatly expressible in language (Bermudez, 2003, 2018; Carruthers, 2011; Machery, 2018). The idea here is that, while inner speech may not be necessary simply for having thoughts of a particular kind, it is the primary means by which we are able to internally reflect upon—and become conscious of—those thoughts. This awareness is hypothesized to allow for reflection upon the logical or inferential properties that hold among thoughts (Bermudez, 2003; Clark, 1998), or to serve as data for feats of self-interpretation about one's other mental states (Carruthers, 2011; Gregory, 2018). Empirical support for these theoretical proposals remains relatively indirect (see Carruthers (2011) and Morin (2018) for reviews). However, in one recent study, people with aphasia—who were determined, through a silent rhyming task, to have inner speech deficits—were found to be unreliable (compared to controls) in their metacognitive judgments concerning whether they had correctly categorized objects by an abstract commonality, such as being energy sources, or being alive (Langland-Hassan, Gauker, Richardson, Dietz, & Faries, 2017). The subjects with aphasia did not show comparable deficits in the categorizations themselves, insofar as they were able to correctly identify which objects go together as reliably as controls. For this reason, the experimenters concluded that inner speech plays a distinctively *meta*cognitive role.

Scientifically assessing whether inner speech plays an important role in metacognition—or in making one conscious of one's own thoughts—remains challenging due to the general difficulty in assessing when people are or are not aware of their own thoughts. More powerful and more creative experimental paradigms are still needed.

7. INNER SPEECH AND AUDITORY VERBAL HALLUCINATIONS

A common symptom of psychosis is the experience of "hearing voices," or auditory verbal hallucinations (AVHs). A hotly debated question in the cognitive science of inner speech is how to understand the role of inner speech in the generation of AVHs. One way to think about AVHs is that the person experiencing them is generating episodes of inner speech without realizing that they are doing so. A number of competing theories have been developed to explain how it is that patients create complex inner utterances without feeling as though they are doing so.

Before looking more closely at some of those proposals, it is worth considering a more straightforward hypothesis. Instead of conceiving of AVHs as episodes of *inner speech* that, somehow, are misidentified as being created by someone else, one might alternatively view them as fundamentally *perceptual* in nature—as cases of having a perceptual experience just as if someone else is speaking, caused by spontaneous (pathological) activation of speech perception areas of the brain, when, in fact, no one is speaking (Wu, 2012; Wu & Cho, 2014)? After all, if AVHs are more like *receiving* speech than *producing* speech, we have a more straightforward explanation for why patients do not take themselves to be the agents responsible for the episodes. While this may indeed be a plausible account of some episodes that are reported as AVHs, researchers have largely pursued other possibilities, for several reasons. One traces to 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). Many AVHs are described by patients as being more like their own "verbal thought" (Hoffman et al., 2008) or even as "soundless" (Junginger & Frame, 1985). This gives some reason to pursue a single style of explanation that will account for the fact that patients do not feel in control of their AVHs in both these cases and those that present with stronger perceptual qualities (Langland-Hassan, 2008, 2016).

A second reason is that neuroimaging shows language production areas are indeed triggered during AVHs, in addition to language perception areas (Allen et al., 2007; Allen et al., 2012; Bohlken, Hugdahl, & Sommer, 2016). This suggests that AVHs are not a purely perceptual phenomenon and that language episodes are indeed being formulated by the patient's language-production mechanisms without the patient having the sense that he or she is doing so. Simply conceiving AVHs as akin to hearing speech without an actual speaker overlooks this constitutive role of the language production system in their genesis. Though it bears noting that ordinary language perception involves *some* activation of language-production areas as well. In any case, a powerful confirmation of whether AVHs are an inherently productive phenomenon, as opposed to primarily perceptual, would be to assess whether a patient with intractable AVHs experiences a significant decrease in their AVHs after suffering a stroke and acquiring a non-fluent (i.e. productive) form of aphasia, such as Broca's or anomic aphasia. We know of no published reports of this combination of conditions.

The most common style of cognitive explanation for AVHs, pioneered by Christopher Frith (1992), sees AVHs as just one in a constellation of passivity symptoms experienced by people with schizophrenia (Blakemore, Smith, Steel, Johnstone, & Frith, 2000; Spence et al., 1997). "Passivity" symptoms are disruptions in one's sense of being the agent responsible for an action or, in the present case, mental episode. Many of these theories draw on the prediction-comparison framework described above, originally developed by theories of motor control to explain how bodily movements are self-monitored and corrected (Miall et al., 1993; Wolpert et al., 1998). The general idea behind these "comparator" approaches is that if an episode of inner speech is not appropriately predicted by internal monitoring mechanisms, it will be experienced as resulting from some agency other than one's own—and that this is the case regardless of the subjective "loudness" of the experience, and regardless of whether the voice seems to come from inside or outside the head. Neuroimaging evidence has been provided in validation of some of the theoretical posits of these models, such as the idea that inner speech involves the generation of a sensory prediction that is also generated during overt speech production (Tian & Poeppel, 2010, 2012). Such models continue to be developed and expanded (Grandchamp et al., 2019).

Some, including Frith himself (2012), have argued that this comparator-style of explanation, as applied to AHVs, has considerable limitations, or is altogether misguided (Synofzik, Vosgerau, & Newen, 2008; Vicente, 2014; Wilkinson, 2014a; Wu, 2012). Others have responded that, suitably amended, something like the comparator hypothesis can indeed be extended to explain AVHs and possibly related phenomena as well, including the experience of "thought insertion" (Ford & Mathalon, 2005; Jones & Fernyhough, 2007a; Langland-Hassan, 2008; Loevenbruck et al., 2018; Stephan, Friston, & Frith, 2009; Swiney, 2018; Swiney & Sousa, 2014). These debates remain active and unsettled. Neuroimaging of patients experiencing AVH shows activation both in speech production and speech perception areas, and, taken as a whole, is equivocal with respect to most theories of the finer-grained cognitive bases of AVH (see Allen et al. (2007) and Bohlken et al. (2016) for reviews; see also Anthony (2004)). Recently, a number of theorists have moved forward from traditional comparator models to try to understand AVH in similar ways via the use of predictive processing models of cognition (Fletcher & Frith, 2009; Swiney, 2018; Wilkinson, 2014b; Wilkinson & Fernyhough, 2017).

A lingering challenge for all of these approaches is to explain the possibility of comprehensible *dialog* between patients and their AVHs, which has been reported in a number of cases (Langdon, Jones, Connaughton, & Fernyhough, 2009; Wu & Cho, 2013). Whether we view AVHs as resulting from breakdowns in self-monitoring, or as spontaneous episodes of activation in auditory cortex, it is not immediately evident why such breakdowns would be timed in such a way as to allow for a comprehensible give-and-take discussion between the patient and his or her "voices."

8. NEW DIRECTIONS

As earlier noted, an important feature of inner speech—one that distinguishes it, as a psychological phenomenon, from neurocognitive language-processing *in general*—is its familiarity to first person experience and its resulting susceptibility to conscious control. We know how to generate detailed linguistic utterances "in our heads," and this unremarkable fact becomes a powerful tool when paired with technologies for detecting and transcribing inner utterances on the basis of neural or other bodily indices. For instance, Kapur et al. (2018) created a wearable interface that detects neuromuscular signals in the jaw and throat that are generated when people silently voice words and phrases in inner speech. Through machine learning, they trained an artificial neural network to translate this neuromuscular information back into matching linguistic items (see also Meltzner et al.

(2008) for similar results). This allows wearers of the device to send their inner speech utterances electronically to a computer interface or other user, enabling a new form of silent communication (with 92% mean accuracy on digit recognition tasks—despite no discernible mouth movements on the part of the user). Applications include sending messages when speaking aloud or typing is not possible (e.g. during military or law enforcement operations), interacting seamlessly with smart devices, and communicating with other language users in noisy environments. One can also imagine therapeutic applications, including facilitating speech for people with mouth, tongue, or vocal cord abnormalities, and experimental interventions on people suffering from AVHs. Perhaps if one's AVHs were projected aloud by such a device, this would facilitate great control of AVHs by the patient, and greater insight into their aetiology.

In a related line of research, Anumanchipalli et al. (2019) used machine learning to develop a "neural speech prosthetic" capable of directly translating inner speech-related neural activity to written text., or synthesized speech through a speaker (see also Martin et al. (2018) for related research). The researchers used electrocorticography (ECoG) to take direct neural readings from five participants undergoing intracranial monitoring for epilepsy treatment. Readings were taken from the ventral sensorimotor cortex (vSMC), superior temporal gyrus, and the inferior frontal gyrus, while participants spoke hundreds of sentences aloud. Recurrent artificial neural networks were then trained on these readings to associate the neural activity with specific word sounds, such that the network could later be used to generate the word sounds itself, given the same set of neural readings as input. The experimenters further demonstrated that the network could, with some reliability, generate words and phrases from neural activity that occurred when sentences were simply "mimed" (where "miming" involves making the relevant mouth movements for saying a phrase, but without producing sound). Of note is the high reliance that the network placed on detecting patterns in the neural representation of "speech kinematics" and articulatory plans (pp. 494-6). While obviously more invasive than the device developed by Kapur et al., this technology looks to be an important stepping-stone on the way to restoring speech for patients with paralysis.

Conclusion

Researchers are drawn to inner speech for many reasons. Some see in it the key to grasping the relationship of abstract thought to language. Others emphasize its importance as a support to more general cognitive capacities, such as metacognition, memory, and executive function. Still others hope that better grasping the cognitive architecture and neural substrates of inner speech will shed light on and facilitate intervention upon auditory verbal hallucinations in schizophrenia. Researchers in neuroscience are now learning to exploit inner speech, through its neural and neuromuscular signatures, as a means to creating new forms of communication that may be of aid to populations with disabilities.

Theories of the components and stages of inner speech have grown more detailed in recent years and are now supported by concurrent work in neuroimaging. Substantive theoretical questions remain open, including which components of inner speech can occur in isolation, whether there are importantly different forms of inner speech, which experimental tasks are best suited for eliciting inner speech, and how our scientific understanding of inner speech can be further leveraged to improve lives. While there remains open texture in the definition of what constitutes inner speech, this has not impeded scientific progress and is to be expected as we achieve a more detailed image of little voice in the head.

Acknowledgments

Many thanks to Shivam Patel, Agustín Vicente, Wayne Wu, and two anonymous reviewers for this journal for comments and suggestions that improved this overview.

Further Reading

Alderson-Day, B., & Fernyhough, C. (2015). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology. *Psychol Bull, 141*(5), 931-965. doi:10.1037/bul0000021

Langland-Hassan, P. & Vicente, A. (2018) *Inner Speech: New Voices* (Oxford: Oxford University Press).

Fernyhough, C. (2016) *The Voices Within: The History and Science of How we Talk to Ourselves,* (New York: Basic Books).

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

Alderson-Day, B., & Fernyhough, C. (2015). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology. *Psychol Bull, 141*(5), 931-965. doi:10.1037/bul0000021

Alderson-Day, B., McCarthy-Jones, S., Bedford, S., Collins, H., Dunne, H., Rooke, C., & Fernyhough, C. (2014). Shot through with voices: Dissociation mediates the relationship between varieties of inner speech and auditory hallucination proneness. *Consciousness and Cognition, 27*, 288-296. doi:10.1016/j.concog.2014.05.010

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C. (2018). The varieties of inner speech questionnaire–revised (VISQ-R): replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition, 65*, 48-58.

Alderson-Day, B., Weis, S., McCarthy-Jones, S., Moseley, P., Smailes, D., & Fernyhough, C. (2016). The brain's conversation with itself: neural substrates of dialogic inner speech. *Social Cognitive and Affective Neuroscience, 11*(1), 110-120.

Allen, P., Aleman, A., & McGuire, P. K. (2007). Inner speech models of auditory verbal hallucinations: Evidence from behavioural and neuroimaging studies. *International Review of Psychiatry, 19*(4), 407-415. doi:10.1080/09540260701486498

Allen, P., Modinos, G., Hubl, D., Shields, G., Cachia, A., Jardri, R., . . . Hoffman, R. (2012). Neuroimaging Auditory Hallucinations in Schizophrenia: From Neuroanatomy to Neurochemistry and Beyond. *Schizophrenia Bulletin, 38*(4), 695-703. doi:10.1093/schbul/sbs066

Alloway, T. P., Kerr, I., & Langheinrich, T. (2010). The effect of articulatory suppression and manual tapping on serial recall. *European Journal of Cognitive Psychology, 22*(2), 297-305. doi:10.1080/09541440902793731

Anthony, D. (2004). The cognitive neuropsychiatry of auditory verbal hallucinations: An overview. *Cogn Neuropsychiatry, 9*(1-2), 107-123. doi:10.1080/13546800344000183

Anumanchipalli, G. K., Chartier, J., & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature, 568*(7753), 493-498.

Baddeley, A. (1966a). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quaerterly Journal of Experimental Psychology, 18*, 302-309.

Baddeley, A. (1966b). 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. (1992). Working memory. *Science, 255*(5044), 556-559. doi:10.1126/science.1736359

Baddeley, A. (1994). The magical number seven: Still magic after all these years? *Psychological review, 101*(2), 353-356. doi:10.1037/0033-295X.101.2.353

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

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130*(4), 641-657. doi:10.1037/0096-3445.130.4.641

Baddeley, A., & Hitch, G. (1974). Working memory. The psychology of learning and motivation. *New York, NY: Academicp*.

Bain, A. (1855). The senses and the intellect. London, England: John W Parker & Son, West Strand.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull, 121*(1), 65-94. doi:10.1037/0033-2909.121.1.65

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. *Symbols, embodiment, and meaning*, 245-283.
Bermudez, J. L. (2003). *Thinking without Words*. Oxford: Oxford University Press.
Bermudez, J. L. (2018). Inner Speech, Determinacy, and Thinking Consciously About Thoughts. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices*. Oxford: Oxford University Press.
Bermúdez, J. L. (2018). Inner Speech, Determinacy, and Thinking Consciously about Thoughts. *Inner Speech: New Voices*, 199.
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.
Block, N. (1981). *Imagery*. Cambridge, Mass.: MIT Press.
Bohlken, M. M., Hugdahl, K., & Sommer, I. E. C. (2016). Auditory verbal hallucinations: neuroimaging and treatment. *Psychological Medicine, 47*(2), 199-208. doi:10.1017/S003329171600115X
Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychol Bull, 143*(3), 263-292. doi:10.1037/bul0000089
Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin et Memoires de la Societe anatomique de Paris, 6*, 330-357.
Burnett, P. C. (1996). Children's self-talk and significant others' positive and negative statements. *Educational Psychology, 16*(1), 57-67.
Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences, 25*, 657-726.
Carruthers, P. (2011). *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.
Carruthers, P. (2018). The Causes and Contents of Inner Speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: Nature and Functions*. Oxford: Oxford University Press.
Caspari, I., Parkinson, S. R., LaPointe, L. L., & Katz, R. C. (1998). Working Memory and Aphasia. *Brain and Cognition, 37*(2), 205-223. doi:http://dx.doi.org/10.1006/brcg.1997.0970
Chihara, C. S., & Fodor, J. A. (1965). Operationalism and Ordinary Language: A Critique of Wittgenstein. *American Philosophical Quarterly, 2*(4), 281-295.
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.
Cooley, C. H. (1992). *Human nature and the social order*: Transaction Publishers.
Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics in cognitive science, 2*(4), 631-642. doi:10.1111/j.1756-8765.2009.01080.x
de Sousa, P., Sellwood, W., Spray, A., Fernyhough, C., & Bentall, R. P. (2016). Inner Speech and Clarity of Self-Concept in Thought Disorder and Auditory-Verbal Hallucinations. *The Journal of Nervous and Mental Disease, 204*(12), 885-893. doi:10.1097/nmd.0000000000000584
Dennett, D. C. (1991). Real patterns. *Journal of Philosophy, 88*(1), 27-51.
Di Biase, M. A., Zhang, F., Lyall, A., Kubicki, M., Mandl, R. C. W., Sommer, I. E., & Pasternak, O. (2019). Neuroimaging auditory verbal hallucinations in schizophrenia patient and healthy populations. *Psychological Medicine*, 1-10. doi:10.1017/S0033291719000205
Dove, G. (2014). Thinking in words: language as an embodied medium of thought. *Topics in cognitive science, 6*(3), 371-389.
Dove, G. (2018). Language as a disruptive technology: abstract concepts, embodiment and the flexible mind. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1752), 20170135.

Dove, G. (2019). More than a scaffold: Language is a neuroenhancement. *Cognitive Neuropsychology*, 1-24.
Fama, M. E., Hayward, W., Snider, S. F., Friedman, R. B., & Turkeltaub, P. E. (2017). Subjective experience of inner speech in aphasia: Preliminary behavioral relationships and neural correlates. *Brain Lang, 164*(Supplement C), 32-42. doi:https://doi.org/10.1016/j.bandl.2016.09.009
Fama, M. E., Henderson, M. P., Snider, S. F., Hayward, W., Friedman, R. B., & Turkeltaub, P. E. (2019). Self-reported inner speech relates to phonological retrieval ability in people with aphasia. *Consciousness and Cognition, 71*, 18-29. doi:https://doi.org/10.1016/j.concog.2019.03.005
Fernyhough, C. (2004). Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. *New ideas in Psychology, 22*(1), 49-68. doi:https://doi.org/10.1016/j.newideapsych.2004.09.001
Fernyhough, C. (2013). Inner speech. In H. Pashler (Ed.), *The encyclopedia of the mind* (Vol. 9, pp. 418-420). Thousand Oaks, CA: Sage Publications.
Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience, 10*(1), 48-58.
Fodor, J. A. (1975). *The Language of Thought*. New York: Crowell.
Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*: The MIT Press.
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.
Frankish, K. (2018). Inner Speech and Outer Thought. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 221-243). Oxford: Oxford University Press.
Friedmann, N., & Gvion, A. (2003). Sentence comprehension and working memory limitation in aphasia: A dissociation between semantic-syntactic and phonological reactivation. *Brain Lang, 86*(1), 23-39. doi:https://doi.org/10.1016/S0093-934X(02)00530-8
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.
Gauker, C. (2018). Inner Speech as the Internalization of Outer Speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: Nature and Functions*. Oxford: Oxford University Press.
Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language acquisition and conceptual development, 3*, 215-256.
Geva, S. (2018a). Inner Speech and Mental Imagery: A Neuroscientific Perspective. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 105-130). Oxford: Oxford University Press.
Geva, S. (2018b). Inner Speech and Mental Imagery: A Neuroscientific Perspective. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 105-130). 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
Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J.-C., & Warburton, E. A. (2011a). *The neural correlates of inner speech defined by voxel-based lesion–symptom mapping* (Vol. 134).
Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J. C., & Warburton, E. A. (2011b). The neural correlates of inner speech defined by voxel-based lesion-symptom mapping. *Brain, 134*(Pt 10), 3071-3082. doi:10.1093/brain/awr232

Graham, G., & Stephens, G. L. (2000). *When Self-Consciousness Breaks*. Cambridge, MA: MIT Press.
Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., . . . Lœvenbruck, H. (2019). The ConDialInt Model: Condensation, Dialogality, and Intentionality Dimensions of Inner Speech Within a Hierarchical Predictive Control Framework. *Frontiers in Psychology, 10*(2019). doi:10.3389/fpsyg.2019.02019
Gregory, D. (2017). Is inner speech dialogic? *Journal of Consciousness Studies, 24*(1-2), 111-137.
Gregory, D. (2018). The Feeling of Sincerity: Inner Speech and the Phenomenology of Assertion. *Thought: A Journal of Philosophy, 7*(4), 225-236. doi:10.1002/tht3.391
Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and Cognition, 17*, 798-810.
Heavey, C. L., Moynihan, S. A., Brouwers, V. P., Lapping-Carr, L., Krumm, A. E., Kelsey, J. M., . . . Hurlburt, R. T. (2019). Measuring the Frequency of Inner-Experience Characteristics by Self-Report: The Nevada Inner Experience Questionnaire. *Frontiers in Psychology, 9*(2615). doi:10.3389/fpsyg.2018.02615
Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of Flexibility in Human Cognition: Dual-Task Studies of Space and Language. *Cognitive Psychology, 39*(1), 3-36. doi:https://doi.org/10.1006/cogp.1998.0713
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.
Hurlburt, R. T., & Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences, 5*(3-4), 271-301.
Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016a). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLoS ONE, 11*(2), e0147932-e0147932. doi:10.1371/journal.pone.0147932
Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016b). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLoS ONE, 11*(2), e0147932. doi:10.1371/journal.pone.0147932
Jackendoff, R. (1996). How language helps us think. *Pragmatics and Cognition, 4*(1), 1-34.
Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia, 33*(11), 1419-1432.
Jones, S., & Fernyhough, C. (2007a). Neural correlates of inner speech and auditory verbal hallucinations: a critical review and theoretical integration. *Clinical Psychology Review, 27*(2), 140-154.
Jones, S., & Fernyhough, C. (2007b). Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration. *Clinical Psychology Review, 27*, 140-154.
Jones, S., & Fernyhough, C. (2007c). 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.
Kapur, A., Kapur, S., & Maes, P. (2018). *Alterego: A personalized wearable silent speech interface.* Paper presented at the 23rd International Conference on Intelligent User Interfaces.
Kell, C. A., Darquea, M., Behrens, M., Cordani, L., Keller, C., & Fuchs, S. (2017). Phonetic detail and lateralization of reading-related inner speech and of auditory and somatosensory feedback processing during overt reading. *Human Brain Mapping, 38*(1), 493-508. doi:10.1002/hbm.23398
Kosslyn, S., Thompson, W. L., & Ganis, G. (2006). *The Case of Mental Imagery*. New York: Oxford University Press.

Kühn, S., Fernyhough, C., Alderson-Day, B., & Hurlburt, R. T. (2014). Inner experience in the scanner: can high fidelity apprehensions of inner experience be integrated with fMRI? *Frontiers in Psychology, 5*(1393). doi:10.3389/fpsyg.2014.01393

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. *Psychol Med, 39*(4), 655-663. doi:10.1017/s0033291708003978

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. (2016). Hearing a Voice as one's own: Two Views of Inner Speech Self-Monitoring Deficits in Schizophrenia. *Review of Philosophy and Psychology, 7*(3), 675-699. doi:10.1007/s13164-015-0250-7

Langland-Hassan, P. (2018). From Introspection to Essence: The Auditory Nature of Inner Speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 71). Oxford: Oxford University Press.

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

Langland-Hassan, P., Gauker, C., Richardson, M. J., Dietz, A., & Faries, F. R. (2017). Metacognitive deficits in categorization tasks in a population with impaired inner speech. *Acta Psychologica, 181*, 62-74. doi:https://doi.org/10.1016/j.actpsy.2017.10.004

Langland-Hassan, P., & Vicente, A. (2018). *Inner Speech: New Voices*. Oxford: Oxford University Press.

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.

Levine, D. N., Calvanio, R., & Popovics, A. (1982). Language in the absence of inner speech. *Neuropsychologia, 20*(4), 391-409. doi:https://doi.org/10.1016/0028-3932(82)90039-2

Loevenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., & Dohen, M. (2018). A Cognitive Neuroscience View of Inner Language. *Inner Speech: New Voices*, 131.

Lordat, J. (1843). Analyse de la parole pour servir à la théorie de divers cas d'Alalie et de Paralalie (de mutisme et d'imperfection du parler) que les Nosologistes ont mal connus. *Journal de la Société de médecine pratique de Montpellier, 7*, 417-433.

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review, 16*(4), 711-718.

Lupyan, G., & Bergen, B. (2016). How Language Programs the Mind. *Topics in cognitive science, 8*(2), 408-424. doi:10.1111/tops.12155

Lupyan, G., & Mirman, D. (2013). Linking language and categorization: Evidence from aphasia. *Cortex, 49*, 1187-1194.

Machery, E. (2018). Know Thyself: Beliefs vs. Desires in Inner Speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 261-275). Oxford: Oxford University Press.

Martin, S., Brunner, P., Holdgraf, C., Heinze, H.-J., Crone, N. E., Rieger, J., . . . Pasley, B. N. (2014). Decoding spectrotemporal features of overt and covert speech from the human cortex. *Frontiers in Neuroengineering, 7*(14). doi:10.3389/fneng.2014.00014

Martin, S., Iturrate, I., Millán, J. d. R., Knight, R. T., & Pasley, B. N. (2018). Decoding inner speech using electrocorticography: Progress and challenges toward a speech prosthesis. *Frontiers in neuroscience, 12*, 422.

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.
McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and Cognition, 20*(4), 1586-1593. doi:10.1016/j.concog.2011.08.005
McGuire, P. K., David, A. S., Murray, R. M., Frackowiak, R. S. J., Frith, C. D., Wright, I., & Silbersweig, D. A. (1995). Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. *The Lancet, 346*(8975), 596-600. doi:http://dx.doi.org/10.1016/S0140-6736(95)91435-8
McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996a). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine, 26*(01), 29-38. doi:doi:10.1017/S0033291700033699
McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996b). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine, 26*(1), 29-38. doi:10.1017/S0033291700033699
McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., Frackowiak, R. S., & Frith, C. D. (1996). The neural correlates of inner speech and auditory verbal imagery in schizophrenia: relationship to auditory verbal hallucinations. *The British Journal of Psychiatry, 169*(2), 148-159. doi:10.1192/bjp.169.2.148
Meichenbaum, D. H., & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77*(2), 115-126. doi:10.1037/h0030773
Meltzner, G. S., Sroka, J., Heaton, J. T., Gilmore, L. D., Colby, G., Roy, S., . . . Luca, C. J. D. (2008). *Speech recognition for vocalized and subvocal modes of production using surface EMG signals from the neck and face.* Paper presented at the Ninth Annual Conference of the International Speech Communication Association.
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.
Miranda, A., Presentación, M. J., Siegenthaler, R., & Jara, P. (2011). Effects of a Psychosocial Intervention on the Executive Functioning in Children With ADHD. *Journal of Learning Disabilities, 46*(4), 363-376. doi:10.1177/0022219411427349
Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychol (Amst), 115*(2-3), 123-142. doi:10.1016/j.actpsy.2003.12.004
Morin, A. (2005). Possible links between self-awareness and inner speech: Theoretical background, underlying mechanisms, and empirical evidence. *Journal of Consciousness Studies, 12*(4-5), 115-134.
Morin, A. (2018). The self-reflective functions of inner speech: Thirteen years later. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices* (pp. 276-298). Oxford: Oxford University Press Oxford.
Nayani, T. H., & David, A. (1996). The auditory hallucination: a phenomenological survey. *Psychological Medicine, 26*, 177-189.
Oppenheim, G. M. (2013). Inner speech as a forward model? *Behavioral and Brain Sciences, 36*(04), 369-370. doi:doi:10.1017/S0140525X12002798
Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*, 528-537.
Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: the flexible abstractness of inner speech. *Mem Cognit, 38*(8), 1147-1160. doi:10.3758/MC.38.8.1147

Owen, W. J., Borowsky, R., & Sarty, G. E. (2004). FMRI of two measures of phonological processing in visual word recognition: Ecological validity matters. *Brain Lang, 90*(1), 40-46. doi:https://doi.org/10.1016/S0093-934X(03)00418-8
Paivio, A. (1971). *Imagery and verbal processes*. New York: Hold, Rinehart, and Winston.
Paivio, A. (1990). *Mental representations: A dual coding approach* (Vol. 9): Oxford University Press.
Pei, X., Leuthardt, E. C., Gaona, C. M., Brunner, P., Wolpaw, J. R., & Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *NeuroImage, 54*(4), 2960-2972. doi:https://doi.org/10.1016/j.neuroimage.2010.10.029
Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Lœvenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behav Brain Res, 261*, 220-239. doi:10.1016/j.bbr.2013.12.034
Petrolini, V., Vicente, A., & Jorba, M. (in draft). The Role of Inner Speech in Executive Functioning Tasks: Autistic Spectrum Conditions and Auditory Verbal Hallucinations as Case Studies.
Piaget, J. (1923). La pensée symbolique et le pensée de l'enfant. *Archives de psychologie*.
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
Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., & Gabrieli, J. D. E. (2001). Relations between the Neural Bases of Dynamic Auditory Processing and Phonological Processing: Evidence from fMRI. *13*(5), 687-697. doi:10.1162/089892901750363235
Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. *Cognition, 77*(2), 97-132. doi:https://doi.org/10.1016/S0010-0277(00)00090-1
Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., . . . Gore, J. C. (1996). Cerebral organization of component processes in reading. *Brain, 119*(4), 1221-1238. doi:10.1093/brain/119.4.1221
Pylyshyn, Z. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences, 25*(2), 157-237.
Roebuck, H., & Lupyan, G. (2020). The Internal Representations Questionnaire: Measuring modes of thinking. *Behavior Research Methods*. doi:10.3758/s13428-020-01354-y
Saeki, E., & Saito, S. (2004). Effect of articulatory suppression on task-switching performance: Implications for models of working memory. *Memory, 12*(3), 257-271. doi:10.1080/09658210244000649
Sellars, W. (1956). Empiricism and the Philosophy of Mind. *Minnesota Studies in the Philosophy of Science, 1*(19), 253-329.
Shergill, S. S. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *The British Journal of Psychiatry, 182*(>6), 525-531. doi:10.1192/bjp.182.6.525
Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31*(2), 241-253. doi:10.1017/s003329170100335x
Simons, C. J., Tracy, D. K., Sanghera, K. K., O'Daly, O., Gilleen, J., Dominguez, M. D., . . . Shergill, S. S. (2010). Functional magnetic resonance imaging of inner speech in schizophrenia. *Biological Psychiatry, 67*(3), 232-237. doi:10.1016/j.biopsych.2009.09.007
Sokolov, A. (1972). *Inner speech and thought* (G. T. Onischenko, Trans. D. B. Lindsley Ed.). New York: Plenum 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.
Stark, B. C., Geva, S., & Warburton, E. A. (2017). Inner Speech's Relationship With Overt Speech in Poststroke Aphasia. *Journal of Speech, Language, and Hearing Research*, 1-10. doi:10.1044/2017_JSLHR-S-16-0270
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. (2018). Activity, Agency, and Inner Speech Pathology. In P. Langland-Hassan & A. Vicente (Eds.), Inner Speech: New Voices (pp. 288-331). Oxford: Oxford University Press.
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:10.1016/j.concog.2007.03.010
Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology, 1*(166). doi:10.3389/fpsyg.2010.00166
Tian, X., & Poeppel, D. (2012). Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation. *Front Hum Neurosci, 6*, 314. doi:10.3389/fnhum.2012.00314
Vicente, A. (2014). The comparator account on thought insertion, alien voices and inner speech: some open questions. *Phenomenology and the Cognitive Sciences, 13*(2), 335-353.
Vicente, A., & Jorba, M. (2019). The Linguistic Determination of Conscious Thought Contents. *Nous, 53*(3), 737-759. doi:10.1111/nous.12239
Vygotsky, L. S. (2012/1962). *Thought and Langauge, revised and expanded edition*. Cambridge, MA: MIT Press.
Wilkinson, S. (2014a). 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
Wilkinson, S. (2014b). Accounting for the phenomenology and varieties of auditory verbal hallucination within a predictive processing framework. *Consciousness and Cognition, 30*, 142-155.
Wilkinson, S., & Fernyhough, C. (2017). *Auditory verbal hallucinations and inner speech: a predictive processing perspective*: Imprint Academic, Limited.
Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal Models in the cerebellum. *TRENDS in Cognitive Science, 2*, 338-347.
Wright, H. H., & Shisler, R. J. (2005). Working Memory in Aphasia. *American Journal of Speech-Language Pathology, 14*(2), 107-118. doi:doi:10.1044/1058-0360(2005/012)
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*.
Wu, W., & Cho, R. (2014). Is inner speech the basis of auditory verbal hallucination in schizophrenia? *Frontiers in Psychiatry, 14*, 1-3.
Yao, B., Belin, P., & Scheepers, C. (2011). Silent Reading of Direct versus Indirect Speech Activates Voice-selective Areas in the Auditory Cortex. *Journal of Cognitive Neuroscience, 23*(10), 3146-3152. doi:10.1162/jocn_a_00022