



How not to decide whether inner speech is speech: Two common mistakes

Daniel Gregory¹

Accepted: 24 March 2022 / Published online: 25 April 2022
© The Author(s) 2022

Abstract

Philosophical interest in inner speech has grown in recent years. In seeking to understand the phenomenon, many philosophers have drawn heavily on two theories from neighbouring disciplines: Lev Vygotsky's theory on the development of inner speech in children and a cognitive-scientific theory about speech production. I argue that they have been too uncritical in their acceptance of these theories, which has prevented a proper analysis of inner speech.

Keywords Inner speech · Vygotsky · Forward model theory

1 Introduction

Until recently, inner speech – the silent voice in the mind¹ – received little attention from analytic philosophers. This started to change around ten years ago and there is now a recognizable literature on the topic.² One of the questions about inner speech which has received the most attention in this period is the question of its ontology. What is the nature of the phenomenon?

The consensus which has emerged among philosophers working on the topic is that inner speech is a kind of actual speech. When you produce external speech, you are speaking audibly; when you produce inner speech, you are speaking silently.

¹ As Langland-Hassan (2021) says, there is no 'commonly agreed operational definition' for the term, 'inner speech' (p. 2). The difficulty is that no characterization is likely to be 'theory neutral' (p. 2), i.e., to avoid relying on or entailing commitments that others might reject. Other terms sometimes used to identify the phenomenon, rather than to define it, include 'internal verbalization' and 'talking in your head' (p. 1.). Note that it cannot even be said that inner speech is inherently linguistic. It is certainly open to say that someone is speaking when they imitate the sounds of a foreign language, even if they are not producing actual words of that language; it is certainly open to say that they are producing inner speech if they do something similar internally.)

² The topic now has its own PhilPapers category.

✉ Daniel Gregory
daniel.gregory@uni-tuebingen.de

¹ Department of Philosophy, University of Tübingen, Tübingen, Germany

In both cases, though, you really are speaking; the only difference is the volume. Proponents of this view include Martínez-Manrique and Vicente (2010), Jorba and Vicente (2014), Gerrans (2015), and Wilkinson (2020); I have also argued for the view (Gregory, 2016, 2018), though no longer hold it.³ An alternative theory would be that inner speech is a kind of imagined speech. On this hypothesis, producing inner speech is like imagining swimming or imagining eating or imagining sewing. The words that we seem to be speaking in our minds are instead auditory images of the sounds which are produced when we speak aloud. Inner speech is a representation of speech rather than a kind of speech itself. No one has unequivocally endorsed this theory, though Roessler (2016) shows some sympathy for it. It is also open to hold that inner speech is neither actual speech nor imagined speech but something else altogether.⁴

My objective is to show that work on this question has been hampered by the largely uncritical acceptance of two theories about inner speech from neighbouring disciplines. The first of these is the theory of the famous psychologist, Lev Vygotsky, about the development of inner speech in children. On Vygotsky's theory, inner speech is a developmental successor to the external self-directed speech which children produce early in life. The second is a theory from cognitive science about the production of inner speech. It holds that inner speech is produced when the process involved in producing external speech is terminated before completion. Insofar as both of these theories connect inner speech so closely to external speech, they are certainly suggestive of the view that inner speech is a kind of actual speech. My contention is that taking these theories largely for granted, despite grounds for caution, has led some philosophers to accept that view too hastily. This is a significant problem, for there is really no more fundamental question about inner speech than the question of its ontology.

In Section 2, I will review Vygotsky's theory. My primary objective there will be to highlight a problem in Vygotsky's argument but I will also take up a secondary task. This is to show that Vygotsky may not actually have had the empirical support he claimed for his theory. In Section 3, I will review the dominant theory of inner speech production and show that there are challenges for it as well. In Section 4, I will show how widely these two theories have been accepted by philosophers working on inner speech and explain how this has interfered with a proper analysis of the ontology of the phenomenon.

2 Inner speech development

Thought and Language, the major source in which Vygotsky sets out his theory of inner speech, was first translated into English in 1962 (Vygotski, 1934/1962). This translation includes a curtailed version of the critical second chapter where

³ I will not argue for an alternative in this paper.

⁴ I thank Daniel Stoljar and Inga Vermeulen for first drawing my attention to this issue, before it had received significant attention in the literature.

Vygotsky's discussion of the development of inner speech is located (Hanfmann & Vakar, 1962). Two more translations appeared in the 1980s, Vygotsky (1934/1986) and Vygotsky (1934/1987). It appears that both are deficient in various ways (van der Veer & Yasnitsky, 2011) but both include fuller versions of the second chapter. A revised and expanded edition of the 1986 translation appeared in 2012 (Vygotsky 1934/2012).

In the following, I will refer primarily to the 1986 translation. At one point, I will refer to the 1987 translation. There is nothing in either the 1962 translation or the 2012 edition that would introduce complications.

2.1 Vygotsky's theory

According to Vygotsky, '[t]he primary function of [external] speech, in both children and adults, is communication, social contact' (1934/1986, p. 34). For this reason, Vygotsky believes, the first speech which children engage in – presumably communication with caregivers – is 'essentially social' (p. 35). How does *inner* speech develop? Like Piaget before him, Vygotsky observed that young children have a tendency to speak audibly about their present activities (p. 25 ff.). Children engage in this practice, which Vygotsky, following Piaget, called 'egocentric' speech, until around the age of seven or eight (p. 29, citing Piaget).⁵ (Piaget referred to this type of speech as 'egocentric speech' primarily because the child engaging in it 'does not attempt to place himself at the point of view of his hearer' (Vygotsky 1934/1986, p. 26, quoting Piaget, 1923/1959, p. 9).) Elaborating Piaget, Vygotsky writes that '[t]he child does not try to communicate, expects no answers, and often does not even care whether anyone listens to him' (1934/1986, p. 26). Piaget thought that egocentric speech simply disappears at around age seven or eight (Vygotsky 1934/1986, pp. 29, 32). Vygotsky, on the other hand, thought that it continues but that it is internalised – that children cease producing utterances in egocentric speech externally and begin to do so internally (pp. 30–33). As he puts it: egocentric speech 'does not simply atrophy but "goes underground," i.e., turns into inner speech' (p. 33; see also pp. 86–88).

Vygotsky offers two considerations in support of his conclusion. First, he points to what he takes to be two features which egocentric speech and inner speech share. For one thing, he believes that both serve the 'function' of 'speech-for-oneself' (p. 32). If one 'asks [an adult] subject to solve some problem thinking aloud', Vygotsky believes, 'one would find a striking similarity to the egocentric speech of children' (p. 32). For another, he claims that children's egocentric speech and adults' inner speech have the 'same structural characteristics' (p. 32). Utterances in either egocentric speech or inner speech taken 'out of context ... would be incomprehensible to others because they omit to mention what is obvious to the speaker' (p. 32).

⁵ Vygotsky provides no citation here but he is generally discussing Piaget (1923/1959).

Second, Vygotsky refers to an experimental finding that, at the age when, it is hypothesised, egocentric speech is being internalised, ‘children facing difficult situations resort now to egocentric speech, now to silent reflection’, surmising that ‘the two can be functionally equivalent’ (p. 33; see also p. 30). What exactly does Vygotsky mean by ‘functionally equivalent’? He now means something more than ‘speech-for-oneself’. He writes that, ‘[b]esides being a means of expression and of release of tension, [egocentric speech] soon becomes an instrument of thought in the proper sense – in seeking and planning the solution of a problem’ (p. 31). The reason for this is a finding that the amount of egocentric speech young children produce increases significantly if some difficulty is introduced to a task they are carrying out and that it also seems to assist in working through the difficulty.⁶

In one case Vygotsky reports, for example, a child was preparing to draw, only to find that no pencil of the right colour was available (pp. 29–30). The child then produced the following egocentric speech: ‘Where’s the pencil? I need a blue pencil. Never mind, I’ll draw with the red one and wet it with water; it will become dark and look like blue’ – and acted accordingly. Older children given tasks and confronted with difficulties alternate between producing task-relevant egocentric speech and thinking silently, presumably producing inner speech. Moreover, ‘[w]hen asked what he was thinking about, such a child answered more in line with the “thinking aloud” of a preschooler’ (p. 31). This is taken to support the conclusion that inner speech plays the same kind of problem-solving and action-planning functions as egocentric speech.

In brief, then, Vygotsky adduces evidence that inner speech is a developmental successor to egocentric speech, pointing to their shared ‘structural’ characteristics and functions, and concludes that inner speech just *is* egocentric speech, gone ‘underground’.

2.2 Vygotsky’s evidence

How strong is Vygotsky’s evidence? The first thing to say is that Vygotsky’s referencing in *Thought and Language* is wanting. Many references are omitted; others are incomplete. The reason, an editor explains, is that Vygotsky was succumbing to tuberculosis while writing *Thought and Language* and ‘had no time for the luxury of including well-prepared references’ (Kozulin, 1986, p. lvi). Consequently, identifying some of Vygotsky’s sources requires persistent investigation and, even then, total certainty that one has located the material that he had in mind is impossible. Once the sources which Vygotsky (was likely) referring to in support of the theory just outlined are located and checked, it emerges that his evidence is not as strong as one might suppose.⁷

⁶ The preceding three paragraphs are largely reproduced from Gregory (2017).

⁷ The difficulties with Vygotsky’s referencing are well known to Vygotsky scholars. Regarding the final chapter of *Thought and Language*, for example, van der Veer and Zaverschneva (2018) write: ‘It turn[s] out that Vygotsky’s chapter was a veritable patchwork of unidentified quotations and that large parts of his text have to be attributed to other authors’ (p. 108).

Let us begin with Vygotsky's evidence that the speech of an adult thinking aloud while solving a problem would bear a 'striking similarity to the egocentric speech of children'. Vygotsky refers to an experiment performed by John B. Watson but does not cite a particular source. There is a subsection in Watson's (1919) *Psychology from the Standpoint of a Behaviorist* dedicated to the development of inner speech which Vygotsky cites elsewhere in *Thought and Language* (Vygotsky 1934/1986, pp. 84–85), so it seems likely that he has in mind something from this subsection. Two passages seem like candidates. First:

The shift [from explicit to implicit language] is not complete even in the adult. This is clear from the observation of individuals while they are reading and thinking. Many persons never get to the point where they can read without articulating the words sufficiently for the process to become overt – the lips are moved in unison with the eyes While thinking many use articulate speech or even lip speech much as do the readers just described. Again, certain people who talk to themselves incessantly when alone or when in the presence of one greatly inferior never complete the transition stage. (Watson, 1919, p. 323).

Although Watson mentions observations about adults reading and thinking, these observations do not really support Vygotsky's claim that there is a 'striking similarity' between the speech of adults asked to think aloud and the egocentric speech of children. Watson only says that some adults 'use articulate speech' when they are thinking. This articulate speech might bear little similarity to egocentric speech – even if Watson also believed that inner speech is a developmental successor to egocentric speech.

Second:

In the acquisition of general bodily acts of skill we have found by experimentation that every short-cut possible which would abbreviate action and increase speed and skill is finally hit upon by the individual in a trial and error way. ... The same thing undoubtedly takes place in silent talking and thinking. Even if we could roll out the implicit processes and record them on a sensitive plate or phonograph cylinder it is possible that they would be so abbreviated, short-circuited and economized that they would be unrecognizable (Watson, 1919, p. 323).

Watson is not reporting an experiment on inner speech; he is speculating that inner speech would follow a developmental trajectory identified in experiments on other phenomena. This is not what one would expect to find on the basis of Vygotsky's text. So, it is not clear that the evidence which Vygotsky refers to in support of his claim that the speech of an adult thinking aloud while solving a problem would bear a 'striking similarity to the egocentric speech of children' is anywhere to be found.⁸

What about the evidence for the claim that inner speech and egocentric speech can be 'functionally equivalent', where this means that both can be 'an instrument

⁸ I thank René van der Veer for extremely helpful correspondence about Vygotsky's reference to Watson (1919).

of thought in the proper sense – in seeking and planning the solution of a problem’ (Vygotsky 1934/1986, p. 31)? This evidence was the experimental finding that young children given tasks and confronted with difficulties would produce egocentric speech which was relevant to overcoming those difficulties; older children would alternate between producing such egocentric speech and, it seemed, inner speech. Asked what they were thinking about, the older children would also produce something resembling the younger children’s egocentric speech. It was inferred that both kinds of speech played the same roles in problem-solving and planning.

I described above an example Vygotsky provides of experimental evidence of a child’s egocentric speech apparently influencing their behaviour – the case of the child who wanted to draw with a blue pencil but, finding none available, decided to draw with a red one. Vygotsky provides one other example of this phenomenon. It is a striking event which happened accidentally during an experiment:

A child of five-and-a-half was drawing a streetcar when the point of his pencil broke. He tried, nevertheless, to finish the circle of [the] wheel, pressing down very hard, but nothing showed on the paper except a deep colorless line. The child muttered to himself, ‘It’s broken,’ put aside the pencil, took watercolors instead, and began drawing a broken streetcar after an accident, continuing to talk to himself from time to time about the change in his picture. The child’s accidentally provoked egocentric utterance so manifestly affected his activity that it is impossible to mistake it for a mere byproduct [of his activity]. (p. 31).

What do we know about the experiments Vygotsky is referring to? He indicates in a note (p. 263, Endnote 10) that he carried out the relevant experiments with Alexander Luria, Alexei Leontiev, Rosa Levine and others, directing the reader to Vygotsky and Luria (1930). The paper referred to is scarcely more than a page and it describes only in general terms the different ways in which young children and older children responded when presented with some challenge in a task they were carrying out. Even the ages of the children are not specified. It does not provide details of any particular experiment and there is no mention of asking older children what they were thinking about while they appeared to be reflecting silently. In (Vygotsky 1934/1987), Vygotsky indicates that the experiments have been described in detail elsewhere, but there is no indication where. So, it cannot really be said that Vygotsky provides solid empirical support for his claim about the functional equivalence of egocentric speech and inner speech in *Thought and Language* – at least as that work has been presented in English.⁹

This is really just exegesis, though it might concern those who have taken Vygotsky’s claims on faith. One might say that it does not matter if Vygotsky did not have the evidence he claimed, so long as such evidence is now available. And, to

⁹ For completeness, Vygotsky mentions one other source which Kozulin (editor of Vygotsky (1934/1986)) identifies as Lemaître (1905) (see Vygotsky (1934/1986), p. 263, Note 11). Vygotsky writes: ‘It was shown that the inner speech of schoolchildren is labile and unsettled. This is an indication that what we observe is a developmentally immature process that is still unstable and indefinite’ (p. 33). It is a passing reference, which I will not explore here.

a considerable extent, it is available – more recent researchers testing his claims have generally found support for them – but the picture is, unsurprisingly, complex and nuanced. Summarising the situation, Winsler and Naglieri (2003) wrote the following:

A primary goal of much [research on Vygotskian theory] has been to explore the early developmental trajectory of children's private speech use and internalization [references omitted]. Investigators have largely confirmed Vygotsky's (1934/1986, Vygotsky, 1930/1978) original observations of an overall curvilinear developmental trend for overt private speech with such speech increasing in frequency of use during the preschool period, peaking around the ages of 4 to 6, and then becoming less common later as it is gradually replaced with more covert forms of self-talk, including whispers, inaudible muttering, and eventually silent, inner speech. *Of course, this developmental pattern is a global one, averaging across many tasks, individual differences among children, and a variety of task conditions and social contexts that are also known to influence children's spontaneous use of private speech during problem solving* [references omitted]. (p. 660, emphasis added.)

Reporting their own study of speech activity among participants aged from 5 to 17 presented with a problem-solving task, Winsler and Naglieri (2003) also found evidence of this pattern but with a striking complication: 'partially covert but still observable private speech use (whispers and muttering) showed a clear inverted U-shape, nonlinear pattern starting off at 13.4% for the 5-year-olds, peaking at 28.2% for the 9-year-olds, and becoming less common again for the older children' (p. 665). There is evidence as well that some people may not experience inner speech or at least experience very little of it (Heavey & Hurlburt, 2008). It is not clear if Vygotsky's developmental theory could account for this.

So, while there is evidence which supports Vygotsky's empirical claims, those claims have not been vindicated unequivocally. There is more complexity to the matter than one would think from reading Vygotsky alone. That there are individual differences among children, in particular, cannot be very surprising to anyone. At this point, though, I want to move on from the question of empirical support and turn to another problem for Vygotsky: his argument.

2.3 Vygotsky's argument

For the purposes of analysing Vygotsky's argument, I will set aside concerns about his evidence and assume that his empirical claims are true.

Vygotsky's argument is abductive. He takes it that there is evidence that the incidence of egocentric speech diminishes at a certain age; that the incidence of inner speech increases at the same age; and that egocentric speech and inner speech have the same structural characteristics and function. His claim is that the best explanation of this data is provided by the theory that egocentric speech turns into inner speech. No fundamental change takes place during the transition,

so inner speech is the same phenomenon as egocentric speech, except that it is silent. It goes ‘underground’.

This theory has two tenets, which need to be distinguished: 1) that inner speech replaces egocentric speech during child development; and 2) that no fundamental change takes place during the transition. One can consistently accept the first and deny the second. The problem for Vygotsky is that a particular theory which does precisely this is no less successful in accounting for his data than his own theory.

Recall the theory, foreshadowed in the Introduction and to be discussed more extensively later, that inner speech is a kind of imagined speech. On this theory, producing inner speech is like imagining swimming or imagining eating or imagining sewing. What seem to be words that we are speaking silently are in fact auditory images representing word sounds. Inner speech is not actually a kind of speech – as egocentric speech obviously is – but a representation of speech. With respect to the development of inner speech, a proponent of this rival theory need not contest Vygotsky’s analysis. They can allow that inner speech replaces egocentric speech during a child’s development, but they will say that inner speech is *not* the same phenomenon as egocentric speech. Egocentric speech is replaced by something which is quite different from it.

To be sure, if inner speech is a type of imagining, it is a special type of imagining. For example, it feels more similar to the action it represents than any other case of imagining. Still, there is no *prima facie* reason that producing inner speech could not consist of imagining speaking. And, critically, the theory that inner speech involves imagining speaking, as just described to include the commitment that inner speech replaces egocentric speech, accommodates Vygotsky’s data just as well as his internalisation theory does. It could still be that the incidence of egocentric speech decreases and the incidence of inner speech increases at the same stage in child development and it could still be that inner speech has the same structural and functional profile as egocentric speech. These findings are no less likely on the hypothesis that mature inner speech is imagined speech (and thus a fundamentally different phenomenon from the egocentric speech which it replaces) than on the hypothesis that mature inner speech is silent egocentric speech (and thus fundamentally the same).

What is really going on, of course, is that Vygotsky does not sufficiently distinguish the two tenets of his theory. Once we do separate them, we can see that, while his data (again, assuming their legitimacy) might provide support for the claim that inner speech replaces egocentric speech, they do not provide support for the claim that no fundamental change takes place during the transition. This is why a rival theory, which accepts the claim about the development of inner speech but offers a different (again, *prima facie* plausible) account of the ontology of mature inner speech explains Vygotsky’s own data equally well.

This does not show that Vygotsky’s theory, understood as involving both the developmental claim and the ontological claim, is wrong but it does show that his argument for that theory does not succeed. An abductive argument only succeeds if it shows that a particular theory explains the relevant data more successfully than any other theory. Vygotsky’s argument does not do this.

This issue about the ontology of inner speech will later take centre stage. At this point, though, I want to introduce the second theory which has had a major influence on philosophers thinking about inner speech.

3 Inner speech production

3.1 The forward model theory of inner speech

On an extremely influential theory, when we issue the motor commands necessary to perform a physical action, a kind of copy of those motor commands is also produced, often called an ‘efference copy’. This efference copy allows for a prediction to be produced of the sensory feedback that one will receive when the motor command is executed, via a process called ‘forward modelling’. A ‘comparator’ system compares the anticipated sensory feedback and the actual sensory feedback. If a mismatch between the anticipated feedback and the actual feedback is detected, the movement which one is trying to make can be corrected. It is also thought that, if a motor command is generated but not fully executed, the anticipated feedback becomes conscious, in the form of mental imagery representing the sensory feedback which would have been received if the command had been executed. On one application of this theory, if you produce the motor commands necessary to say something externally, but then abort the commands before they are executed, you will experience an imagistic representation of the speech sounds you would have produced and, therefore, heard if you had executed the commands. This imagistic representation is inner speech.

This theory of action monitoring has been developed by countless researchers and endless references could be provided. von Holst and Mittelstaedt (1950) and Sperry (1950) initiated the modern development of the theory but it has a very long history (see Grüsser (1995)). The work of Feinberg (1978) and Frith (e.g., Frith, 1992), applying the theory to explain symptoms of certain psychiatric disorders, increased its prominence. Jeannerod (1994, 1995) was highly significant in connecting aborted motor commands to mental imagery. For some applications of the theory to inner speech in particular, see Tian and Poeppel (2012) and Grandchamp et al. (2019).

In the following, I will refer to the theory just reviewed, as applied to inner speech, as the ‘forward model theory of inner speech’¹⁰ – or ‘FMT-IS’.

3.2 Problems for FMT-IS

Various challenges have been raised for FMT-IS. I am going to review three from Gauker (2018) and one from Oppenheim (2013) (which Gauker cites approvingly). Two of Gauker’s objections do no damage to FMT-IS but Oppenheim’s challenge

¹⁰ Gauker (2018) uses the same term.

and Gauker's remaining one do present difficulties. I will also present a new problem for FMT-IS.

First, Gauker says that he does not 'see any reason to assume that every act of inner speech ... has to be the product of a plan to speak of the sort that puts a forward model into action', i.e., a plan or intention to speak aloud (p. 65). As he points out, you can 'engage in inner speech when there is no one [you] intend to speak to and when [your] vocal apparatus is fully engaged in whistling' (p. 65). Regarding the latter, Gauker doubts that one could intend, even subconsciously, to speak externally, if one also has an intention to perform some other motor action with the same parts of the body at the same time.

On the observation that we can produce inner speech when the vocal apparatus is otherwise engaged: I think Gauker is mistakenly assuming that what is true of conscious intentions is true of subconscious intentions. But what is true of conscious mental states is not necessarily true of corresponding subconscious mental states. Compare: we cannot hold inconsistent conscious beliefs, but everyone holds inconsistent beliefs subconsciously. Lewis (1982) famously confessed that he had once thought that Nassau Street in Princeton ran approximately east-west; that a close-by railroad ran approximately north-south; and that they were more or less parallel. But he obviously held these beliefs subconsciously; it would not be possible to be conscious of all three at a particular time because the inconsistency would become apparent. Things could be similar with intentions. You cannot hold two inconsistent intentions consciously, but it does not follow that you cannot do so subconsciously – even if you cannot actually execute both. A proponent of FMT-IS will presumably say that, while you are whistling, you can nonetheless have a subconscious intention to speak externally 'of the sort that puts a forward model into action'. If motor commands to speak externally are produced but not fully executed, then you will experience inner speech.

On the observation that we can engage in inner speech in the absence of anyone else: a similar issue arises. It is rare to have conscious intentions to speak externally in the absence of anyone else, but it is difficult to know what subconscious intentions we might have at *any* particular point in time, precisely because they are subconscious. A proponent of FMT-IS can say that we have subconscious intentions to speak externally far more often than we realise, including when no one else is present.

Gauker's second argument appeals to the fact that there are other forms of mental imagery which presumably do not result from aborted motor commands. One can, he points out, imagine a flying turtle. Moreover,

both kinds of imagery [i.e. visual imagery of, e.g., a flying turtle, and the imagery involved in inner speech] can arise spontaneously ... and they can even interact. (I might address my inner speech to the imaginary flying turtle.) This suggests that at some level they are the same kind of thing. (2018, p. 66).

So, if visual imagery of a flying turtle does not result from aborted motor actions, then likely nor does inner speech.

It is true that visual imagery of a flying turtle and inner speech are *at some level* the same kind of thing. But they are different things at a slightly lower level and they

may be produced in quite different ways. Compare: a pencil and a pen are *at some level* the same kind of thing, i.e., writing implements. At another, slightly lower level, they are different kinds of thing: one is a graphite-based writing implement and the other is an ink-based writing implement. They are also manufactured differently. This pattern is common. You can create the same kind of thing at some level by either painting or taking a photograph, viz., a picture, but you will be creating different kinds of thing at a slightly lower level, viz., a painting or a photograph. You can create the same kind of thing at some level by steeping tea leaves or dissolving coffee powder, viz., a hot beverage, but you will be creating different kinds of thing at a slightly lower level, viz., tea or coffee. So, it simply does not matter if visual imagery of a flying turtle and inner speech are at some level the same kind of thing. Presumably, visual imagery of a flying turtle does not result from aborted motor commands but this does not tell against FMT-IS.

The issue that Oppenheim raises (and which Gauker cites approvingly) has at its centre the fact that ‘people easily detect their inner speech errors’ (Oppenheim, 2013, p. 369; Oppenheim cites nine sources in support of this claim). This is hard to account for if inner speech consists in the imagistic realisation of predictions produced by forward models. The predictions produced by forward models, recall, are just representations of the feedback one will receive if motor commands which are generated in service of an intention to perform a physical action are executed. If motor commands issued in service of an intention to speak aloud are executed but with some small error, such that the individual produces, for example, an incorrect phoneme, then the sensory feedback received can be compared to the prediction and the error can be detected. But if the motor commands are not actually executed at all, there will be no feedback that can be compared to the prediction, even if the prediction has become conscious in the form of mental imagery.¹¹

This is a clever argument but there is room to challenge it. One might suspect that individuals attend to their inner speech in an experimental context much more than is ordinarily the case. If so, it may be the case that slips in our inner speech ordinarily go unnoticed, precisely because there is no comparison which can take place, but that they are noticed in an experimental context because of the increased attention which participants are directing to their inner speech. This would mean that inner speech could, after all, be the conscious manifestation of predictions produced by forward models.¹²

This explanation is plausible but a question can be asked of it too. Assume that FMT-IS is right: inner speech does involve predictions of sensory feedback which become conscious in the form of auditory imagery if motor commands to speak externally are not executed. Now suppose that someone in an experimental setting notices an incorrect phoneme in their inner speech, though only because they are

¹¹ Oppenheim himself believes that forward models are involved in inner speech but he thinks that this is not the whole story. Consistent with this, the particular issue which he raises here does not imply that forward models play no role in the production of inner speech, but it does cast doubt on one particular theory about the role they play, viz., FMT-IS.

¹² I thank a referee for this point.

attending carefully to their inner speech. If they had not been attending carefully to their inner speech, they would not have noticed. This seems to imply that, if the relevant motor commands were actually executed and an external utterance including the incorrect phoneme was produced, then no mismatch between the predicted sensory feedback and the actual sensory feedback could be detected by the comparator. The reason is that the prediction itself already included the incorrect phoneme. (We know that the prediction included the incorrect phoneme because the individual noticed the incorrect phoneme in their inner speech in the actual case in which they did not fully execute the motor commands – and inner speech (on FMT-IS) is nothing other than the imagistic realisation of this kind of prediction.) The predicted feedback and the actual feedback would match, so there would be no discrepancy to detect – and the incorrect phoneme would evade the comparator. This is a surprising consequence of combining FMT-IS with the suggestion that slips in inner speech are only detected when we attend to our inner speech.

This response is not conclusive. The proponent of FMT-IS might ultimately say that a slip in external speech of the kind described should *not* be detected by the comparator, though, again, it might be detected if, say, the individual attends carefully to their external speech. But there is clearly a difficult issue here which proponents of FMT-IS need to address.¹³

Before coming to Gauker's remaining argument, I want to suggest another problem for FMT-IS. FMT-IS seems to imply that we frequently continue generating the motor commands necessary to produce an utterance aloud after the motor commands necessary to produce earlier parts of the utterance have been abandoned. Suppose you produce the inner speech utterance, 'The dog ran out the door'. On FMT-IS, this starts with the intention to say aloud, 'The dog ran out the door'. The motor commands necessary to enunciate the first sound, 'the', issue. The commands are abandoned at some point prior to production of the sound, which results in an auditory image of the sound occurring in your mind. Why, at this point, do the motor commands necessary to produce aloud the second sound, 'dog', issue? The intended action of saying aloud, 'The dog ran out the door', has already failed because enunciating the first sound was essential to the action. It is strange that one should simply continue, as if the first phoneme had successfully been produced aloud. (To make the point more vivid: why do the motor commands necessary to produce externally the final phoneme issue once *everything* preceding it has been represented in your mind imagistically but not said aloud?) This is not how we ordinarily respond when intentional actions fail. We might start again but we do not ordinarily just persist with an action if the first part of it (or almost all of it) has already failed.

Of course, it could just be a fact that motor commands to speak externally often continue to issue – and that we continue to abandon them – well after an intentional action of speaking externally has manifestly failed. But it would be a surprising fact. It does not sit comfortably alongside any notion that intentional actions are performed to achieve goals. This is a cost for FMT-IS.

¹³ See also Martínez-Manrique & Vicente (2015, p. 11, Footnote 21) on this issue.

Finally, Gauker queries why it should be that representations of anticipated feedback become conscious if motor commands are not executed and he calls on proponents of the theory to explain this. He also claims that:

it is not true anyway that aborting an action normally brings to consciousness the product of a forward model. Even if I start to open a door and then change my mind, I do not necessarily experience an imagistic representation of myself opening the door. And if I do picture to myself what I was aiming to achieve, it is not obvious that that picturing is a case of a product of a forward model that becomes conscious. Rather, picturing to myself what I aimed to achieve might belong to a more deliberate reflection on my intended but aborted act. (2018, pp. 65-66).

The questions Gauker is asking here do highlight difficulties for FMT-IS. It is not obvious why the prediction of sensory feedback generated by a forward model should become conscious if motor commands are aborted; it is even less obvious why this should happen in some contexts (such as when motor commands to speak externally are aborted) but not in every context. However, we should note how demanding Gauker's questions are. We know very little about why some mental states are conscious and others are not. No theory in cognitive science provides anything approaching an adequate explanation. We should not be too critical of FMT-IS because it does not do so either.

3.3 Being realistic

This brings us to a key issue. What is the status of FMT-IS? It is a *theory*. It does have significant appeal, elegantly and parsimoniously connecting the production of inner speech to a broader theory of action monitoring. But FMT-IS is not settled science. As the foregoing discussion shows, there are reasons to take caution before embracing it. Some of the doubts which have been raised can be dispelled and those which remain vary in their forcefulness. But we can certainly say that the theory is in need of revision and refinement – and this is an ongoing project in cognitive science. This informs how we should make use of FMT-IS. FMT-IS is extremely suggestive and we might expect of anyone who straightforwardly denies it that they can provide a solid motivation for doing so. But it is also a mistake to treat FMT-IS as given.¹⁴

¹⁴ A second philosopher (in addition to Gauker) who challenges FMT-IS is Shivam Patel, who argues that the theory cannot account for all of the various kinds of inner speech (2021). Patel is operating with a broad notion of inner speech which includes, e.g., a phenomenon detected by empirical researchers (he cites Hurlburt et al. (2013) and Hurlburt and Heavey (2018)) in which subjects feel like they are producing inner speech but without actually producing any words, referred to as 'unworded inner speaking'. It is not obvious to me that we should think of this as a kind of inner speech or that we should expect that it is produced in a way related to the production of *worded* inner speaking. Settling this might require significant argumentation and the fact that Patel presents another argument against FMT-IS does no damage to my claim that the theory should be treated with caution, so I will not go into this further here.

4 Bad influences

A theory about the development of inner speech in children and a theory about the production of inner speech are less secure than they might seem. The reason this matters in a philosophical context is that the two theories have had an immense influence on philosophers thinking about inner speech. In the next sub-section, I will give some examples to show how pervasive this influence has been. I will then demonstrate how this has interfered with a proper investigation of the ontology of inner speech.

4.1 Widespread acceptance

Let us begin with some examples of Vygotsky's influence. Clowes (2007) writes the following:

The key to understanding the properties, both functional and phenomenal, of inner speech [is] given in understanding the aetiology of how they take up their inner role. This idea, first articulated by Vygotsky (1934/1986) I call the *internalization hypothesis*. (p. 63; emphasis original.)

Langland-Hassan (2008) writes:

Following Jones and Fernyhough (2007) – themselves drawing on the empirical work of developmental psychologist Vygotsky (1934/1987) – we can see the very ability to engage in inner speech as grounded in the ability to simulate acts of auditory perception and speech ‘offline’. On this view, we *first* learn to grasp linguistic meaning through hearing and watching others interact, then acquire the ability to speak. Only later are we able to engage in inner speech through simulating these learned abilities ... (p. 347; emphasis original.)

From Jorba and Vicente (2014): ‘The ... explanation of inner speech that we want to advance develops Vygotsky's ... general idea that inner speech is overt speech internalized.’ (p. 93.) Martínez-Manrique and Vicente (2010) explicitly ‘[endorse] the Vygotskian view that inner speech is internalized outer speech that is used as an aid to cognition’ (though they add that it may have other roles) (p. 164). Wilkinson (2020) calls it a ‘very attractive theory’ (p. 12; see also Wilkinson & Fernyhough, 2018). Fernández Castro (2016) also discusses it in favourable terms.

Now FMT-IS. As we have seen, there has been some resistance to FMT-IS, specifically from Gauker,¹⁵ but evidence of its extensive influence is easy to find. Gersans (2015) writes the following:

In inner speech, the mechanisms of speech production are activated, but output or translation into overt action is inhibited. The motor instruction to produce a phoneme should produce [an efference copy] [B]ecause the

¹⁵ And Patel (see Footnote 14).

overt component of speech is inhibited, there is no sensory reafference [i.e., feedback] (p. 296).

He also cites Jeannerod (1995) for the proposition that abandonment of motor commands could result in the experience of conscious imagery representing the feedback which one would have received if the motor commands had been fully executed. This is FMT-IS.

Vicente and Martínez-Manrique (2016) write that they ‘endorse the idea that inner speech involves generating ... predictions ... by the monitoring system’ (p. 180). The predictions which they are referencing are the predictions of sensory feedback which will result if motor commands to produce external speech, once initiated, are executed; correspondingly, they are the predictions which are converted into imagistic representations of that feedback if those motor commands are not executed. This, again, is FMT-IS. Vicente & Martínez-Manrique emphasize that they think there is more than this to inner speech. They think that inner speech is best conceived of as an activity that involves forming an intention to express a certain thought content; formulating a conceptual or semantic representation of that content; formulating a linguistic structure (words arranged syntactically) which expresses the conceptual or semantic representation; and then issuing (and finally aborting) the motor commands necessary to articulate that linguistic structure (see also Martínez-Manrique & Vicente, 2015). Regarding the voice-like phenomenon in the mind which is ordinarily denoted by the term, ‘inner speech’, however, it is clear that they accept FMT-IS. See also Vicente and Jorba (2019).

Carruthers (2018) writes that ‘[i]t seems that most inner speech results from rehearsed motor plans’ (p. 37). It is clear from context that, by ‘rehearsed motor plans’, Carruthers means the initiated then aborted motor commands which are described by FMT-IS.¹⁶ (For Carruthers, the instances of inner speech which do not result from rehearsed motor plans are episodic memories of speech, whether the speech of oneself or of others. As he speculates, surely rightly, such instances are ‘comparatively rare’ (p. 32).) Frankish (2018), citing several of Carruthers’s publications, also presents FMT-IS in a very favourable light.

In a word, Vygotsky’s developmental theory and FMT-IS permeate the philosophical literature on inner speech.

4.2 The wrong direction

I now come, finally, to show how the widespread acceptance of Vygotsky’s developmental theory and of FMT-IS has had a deleterious effect on philosophical thinking about the fundamental question of the ontology of inner speech: whether inner speech is a kind of actual speech; or a kind of imagined speech; or something else altogether.

¹⁶ Carruthers’s support for FMT-IS is evident in a long series of publications, such as Carruthers (2008, 2009).

The view that inner speech is a kind of actual speech is basically a consensus position. Martínez-Manrique and Vicente (2010) write this:

To put it in a slogan-like manner, we will say that, in our view, inner talk is just inner talk. By this apparently tautological statement, we mean that inner speech shares fundamental properties and computational demands with outer speech. Inner speech is like outer speech, but used in a very particular context. (p. 160).

Jorba and Vicente (2014) write that ‘[i]n a first approach, inner speech can be characterized as the phenomenon of silently talking to ourselves’ (p. 87). Wilkinson (2020) writes that ‘[i]nner speech is an action, and, more specifically, speech’ (p. 16). In Gregory (2016), I wrote that ‘the best way to think of inner speech is that it is a type of actual speech as much as external speech is a type of actual speech’ (p. 666; see also Gregory (2018)).¹⁷

By contrast, although the view that inner speech is a kind of imagined speech is often noted as a possible position (Gregory, 2016, 2018; Wilkinson & Fernyhough, 2018; Wilkinson, 2020), no one has wholeheartedly endorsed it. Roessler (2016) comes closest (see also Roessler, 2013). Roessler is open to the idea that inner speech *can* involve actual speech. He writes: ‘I’ll ... assume that there can be acts of speech that are barely audible and perhaps completely inaudible’ (2016, p. 547). However, he also holds that a lot of inner speech is imagined speech.

Roessler is clear that not all instances of imagined speech are instances of inner speech. If you imagine Margaret Thatcher saying, ‘There’s no such thing as society’, this is not inner speech, but it is imagined speech (Roessler’s example, p. 545). Almost all philosophers working on inner speech would endorse the analysis of this example.¹⁸ (It does not matter that the speech production system might also be heavily involved in producing this kind of imagined speech, taking for granted that it is also somehow involved in the production of inner speech. Almost all philosophers working on the topic take them to be clearly distinct phenomena, notwithstanding this.) But Roessler also thinks there can be cases where we imagine speaking and treat our imagined speech as subject to the same norms as actual speech. He offers this example:

Suppose you try to recall Austin’s middle name, and after a while the name comes to mind: you find yourself ‘saying in inner speech’—in the mode of *imagining* saying, let’s assume—‘Langshaw’. Now while an imagined assertion is not the same as a real assertion, it can in some ways be tantamount to one, provided the imaginative exercise is informed by suitable intentions. Suppose the intention controlling your act of inner speech is to express your knowledge of Austin’s middle name. Then the act of imagining saying ‘it’s Langshaw’ incurs the same liabilities as would a real assertion that his name

¹⁷ Though, again, I no longer hold this view.

¹⁸ Possible exceptions being Patel, with his wider notion of inner speech (see Footnote 12), and Caruthers, who includes episodic memories of speech within the category of inner speech (see Subsection 4.1).

was Langshaw. For example, if the imagined assertion turned out to be false, one would react to one's act of imagination in much the same way as when a real assertion is exposed as erroneous. (pp. 547-548)

For Roessler, cases like this are 'an important part of the extension of what's ordinarily called "inner speech"' (p. 547). (Nothing turns on the fact that Roessler's example involves an imagined assertion. An imagined exhortation – 'Come on!' – can 'incur the same liabilities' as an audible exhortation if 'informed by suitable intentions', e.g., to motivate oneself.)

Those who hold that inner speech is actual speech would presumably say that cases like the Langshaw case are not really instances of inner speech but are more like the case where one imagines Margaret Thatcher saying 'There's no such thing as society'. For these philosophers, such cases are certainly episodes of imagination involving speech, but they are not within the extension of what is ordinarily called 'inner speech'. Accordingly, it may be that the disagreement between Roessler and those who hold that all inner speech is actual speech is only about the extension of the term, 'inner speech'. The key point for present purposes is that Roessler does not claim that paradigm examples of inner speech – those which everyone would take to be within the extension of the term, 'inner speech' – involve imagined speech. He appears open to the position that these are instances of silent, actual speech.

Gerrans (2015) describes inner speech as involving imagination but it is clear that, in doing so, he only means to say that inner speech is covert. At one point, for example, he writes that 'inner speech is a form of imaginary action' (p. 296) but he then explains in a footnote that '[e]quating inner speech with *imaginary action* limits the type of imagination involved to that described by Jeannerod (2006): the covert rehearsal of an action without an overt component' (Gerrans, 2015, p. 296, Footnote 4; emphasis original). He also writes that, '[i]n essence, [inner speech] is covert speech' (p. 296). So, despite his terminology, Gerrans is clearly a proponent of the view that inner speech is silent (or covert), actual speech.

The third possible position, that inner speech is neither actual speech nor imagined speech but something else altogether, does not appear to have been developed in any way in the recent philosophical work on the topic. So, the view that inner speech is a kind of actual speech really is the dominant position.

What I want to suggest is that convergence on the view that inner speech is a kind of actual speech has largely resulted from general acceptance of Vygotsky's theory on the development of inner speech and of FMT-IS. The influence of Vygotsky's theory in this connection is easy to detect. Wilkinson, in a collaborative paper written with a psychologist, suggests that, in order to understand the nature of inner speech as a kind of actual speech, 'we need to ask not just, "What is inner speech?" or "What does it look like once developed?" but also: "How and why did it develop?"' (Wilkinson & Fernyhough, 2018, p. 247). They offer Vygotsky's theory as an answer to that last question. Martínez-Manrique and Vicente (2010) and Jorba and Vicente (2014) are also explicit in deriving the view that inner speech is actual speech from Vygotsky's developmental theory. Recall the quotations above: Martínez-Manrique & Vicente '[endorse] the Vygotskian view that inner speech is internalized outer speech that is used as an aid to cognition' (2010, p. 164); Jorba

& Vicente write that '[t]he ... explanation of inner speech that we want to advance develops Vygotsky's ... general idea that inner speech is overt speech internalized' (2014, p. 93.)

There is not such clear textual evidence that acceptance of FMT-IS has influenced philosophers in arriving at the view that inner speech is a kind of actual speech, but it seems highly likely that it has played a role, for two reasons. First, it is a short step from FMT-IS to the view that inner speech is a kind of actual speech. If one thinks that producing inner speech involves the same process as producing external speech minus the final step of articulation, then it would be very natural to suppose that inner speech is the same kind of thing as external speech except that it is not made audible. Second, there is significant overlap between those who show sympathy for the Forward Model Theory and those who hold that inner speech is a kind of actual speech, e.g., Vicente and Martínez-Manrique (see Martínez-Manrique and Vicente (2010, 2015) and Vicente and Martínez-Manrique (2016); Gerrans (2015); and Gregory (2016, 2018). This is unsurprising, given the first point.¹⁹

Of course, it does not matter in itself if philosophers have converged on a particular view because of similar influences. The reason that there is a problem in this case is that, as we have seen, the theories which have apparently led many philosophers to adopt a particular view about the nature of inner speech are theories which should not be taken for granted. It may be that Vygotsky's developmental theory and FMT-IS will eventually be vindicated in full. But if it turns out that they have defects which cannot be overcome, then it may be that philosophers thinking about inner speech have been led in a wrong direction entirely. Specifically, they may have been led to a totally incorrect understanding of the fundamental nature of the phenomenon.

It must be said that some philosophers who subscribe to Vygotsky's developmental theory and / or FMT-IS and who endorse the view that inner speech is actual speech have given additional arguments for it. As well as making use of Vygotsky's developmental theory, for example, Wilkinson and Fernyhough (2018) argue that inner speech is a kind of actual speech on the basis that it seems like we can perform speech acts in inner speech and that this could only be possible if inner speech is a kind of actual speech. There is room for argument here. For one thing, not all speech acts are performed with speech. You can, for example, perform the speech act of greeting someone with a gesture, e.g., waving. So, it is at least open to accept that we perform speech acts in inner speech but deny that inner speech is actually speech. In any case, what Wilkinson & Fernyhough offer is clearly an independent

¹⁹ One analysis of inner speech which I have not considered here is that of Langland-Hassan (2014). Langland-Hassan is concerned with the question of whether it is possible that inner speech both represents speech sounds *and* expresses propositional content. He concludes that this is not possible and suggests that what we have been calling 'inner speech' must in fact comprise two separate mental states (a representation of the sound of an utterance and a 'propositional, amodal thought[] with related content[]' (p. 530). He then explores how we should use the term, 'inner speech': to refer to just one or both of the two mental states. He does not directly comment on the matter of whether inner speech (on any understanding of 'inner speech') can be considered either imagined speech or actual speech. Determining whether his view implies this is a complicated matter, which I will not take up here.

argument for the actual speech view. (I tried to provide some further such arguments in Gregory (2016, 2018).) So, the influence of Vygotsky's developmental theory and FMT-IS on thinking about the ontology of inner speech has not been total, but it has certainly been powerful.

5 Conclusion

Where does all of this leave us? It is always salutary to be apprised of relevant theories from neighbouring disciplines. What I have tried to show, however, is that in this instance, the most relevant theories from neighbouring disciplines are less secure than one might think – and thus not of great assistance. Vygotsky's abductive argument for his developmental theory fails because it does not account for his data (if we assume the legitimacy of that data) uniquely well. His theory is also not fully supported by the data which is now available. FMT-IS faces a series of theoretical challenges. None of these challenges is decisive and some of them are weak. It may also be that the stronger challenges will, in time, be met. At present, though, the case against FMT-IS should at least give us pause before we embrace the theory.

I have also shown that Vygotsky's developmental theory and FMT-IS have had an immense influence on philosophical thinking about inner speech, notwithstanding the reasons to treat them with caution. This includes philosophical thinking about the ontology of inner speech, which is perhaps the most central issue for understanding the phenomenon in general.

How, then, should we investigate the ontology of inner speech? The fact that two theories from neighbouring fields do not provide such assistance as we might hope does not mean that there is no material from neighbouring fields which can assist. There may be. But a question about the ontology of a mental state will inevitably be a primarily philosophical question, to be answered through philosophical argumentation. What matters in the first instance is the properties of inner speech as we find it, regardless of how it develops and regardless of how it is produced. So philosophers might now do best to ask themselves: how would we think of inner speech if we had never heard of Vygotsky's developmental theory or of FMT-IS? This approach is not guaranteed to deliver the best characterisation of inner speech, but it might allow a wider perspective on the question.

Acknowledgements I thank Martine Nida-Rümelin for discussing an earlier version of this paper with me. I thank Daniel Stoljar and Inga Vermeulen for raising with me the critical issue of the relationship between inner speech and imagined speech. I thank the anonymous referees who reviewed the article.

Funding Open Access funding enabled and organized by Projekt DEAL. This research has been carried out with the support of an Australian Postgraduate Award from the Commonwealth Government of Australia; a Research Grant for Doctoral Candidates and Young Academics and Scientists from the German Academic Exchange Service; a Swiss Government Excellence Postdoctoral Scholarship from the Swiss State Secretariat for Education, Research and Innovation; a Bridging Grant from the Excellence Initiative of the University of Tübingen; and a Humboldt Postdoctoral Fellowship.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Conflicts of interest / competing interests Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Carruthers, P. (2008). The creative action theory of creativity. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind, volume 3: Foundations and the future*. Oxford University Press.
- Carruthers, P. (2009). Action-awareness and the active mind. *Philosophical Papers*, 38, 133–156.
- Carruthers, P. (2018). The causes and contents of inner speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices*. Oxford University Press.
- Clowes, R. (2007). A self-regulation model of inner speech and its role in the organisation of human conscious experience. *Journal of Consciousness Studies*, 14(7), 59–71.
- Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, 4, 636–640.
- Fernández Castro, V. (2016). Inner speech in action. *Pragmatics & Cognition*, 23(2), 238–258.
- Frankish, K. (2018). Inner speech and outer thought. In P. Langland-Hassan & A. Vicente (Eds.), *Inner Speech: New Voices*. Oxford University Press.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Erlbaum.
- Gauker, C. (2018). Inner speech as the internalization of outer speech. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices*. Oxford University Press.
- Gerrans, P. (2015). The feeling of thinking: Sense of agency in delusions of thought insertion. *Psychology of Consciousness: Theory, Research and Practice*, 15(2), 291–300.
- Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., Lachaux, J.-P., Dohen, M., Perrier, P., Garnier, M., Baciu, M., & Lœvenbruck, H. (2019). The ConDialInt model: Condensation, dialogality, and intentionality dimensions of inner speech within a hierarchical predictive control framework. *Frontiers in Psychology*, 10, art 2019.
- Gregory, D. (2016). Inner speech, imagined speech, and auditory verbal hallucinations. *Review of Philosophy and Psychology*, 7, 653–673.
- Gregory, D. (2017). *Inner speech: A philosophical analysis*. Doctoral dissertation. Australian National University.
- Gregory, D. (2018). The feeling of sincerity: Inner speech and the phenomenology of assertion. *Thought: A Journal of Philosophy*, 7, 225–236.
- Grüsser, O. (1995). On the history of the ideas of efference copy and reafference. In C. Debru (Ed.), *Essays in the history of the physiological sciences: Proceedings of a symposium held at the university Louis Pasteur, on march 26-27th, 1993*. Rodopi.
- Hanfmann, E. & G. Vakar (1962), 'Translators' preface' in L. S. Vygotski (1934/1962), *Thought and language*, E. Hanfmann & G. Vakar (trans. and Ed.), MIT Press.
- Heavey, C. L., & Hurlburt, R. (2008). The phenomena of inner experience. *Consciousness and Cognition*, 17, 798–810.

- von Holst, E., & Mittelstaedt, H. (1950). Das Refferenzprinzip. *Die Naturwissenschaften*, 20, 464–476.
- Hurlburt, R. T., & Heavey, C. L. (2018). Inner speech as pristine inner experience. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices*. Oxford University Press.
- Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, 22, 1477–1494.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17, 187–202.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33, 1419–1432.
- Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford University Press.
- Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition*, 16, 391–399.
- Jorba, V., & Vicente, A. (2014). Cognitive phenomenology, access to contents, and inner speech. *Journal of Consciousness Studies*, 21(9–10), 74–99.
- Kozulin, A. (1986). 'Vygotsky in context', in L. S. Vygotsky (1934/1986), *Thought and language*, A. Kozulin (trans. and ed.), MIT Press.
- Langland-Hassan, P. (2008). Fractured phenomenologies: Thought insertion, inner speech, and the puzzle of extraneity. *Mind & Language*, 23, 369–401.
- Langland-Hassan, P. (2014). Inner speech and metacognition: In search of a connection. *Mind & Language*, 29(5), 511–533.
- Langland-Hassan, P. (2021). Inner speech. *WIREs Cognitive Science*, 12(2), e1544.
- Lemaître, A. (1905). Observations sur le langage intérieur des enfants. *Archives de Psychologie*, 4, 1–43.
- Lewis, D. (1982). Logic for equivocators. *Noûs*, 16, 431–441.
- Martínez-Manrique, F., & Vicente, A. (2010). "What the...!" the role of inner speech in conscious thought. *Journal of Consciousness Studies*, 17(9–10), 141–167.
- Martínez-Manrique, F., & Vicente, A. (2015). The activity view of inner speech. *Frontiers in Psychology*, 6, 232.
- Oppenheim, G. M. (2013). Inner speech as a forward model. *Behavioral and Brain Sciences*, 86(4), 369–370.
- Patel, S. (2021). From speech to voice: On the content of inner speech. *Synthese*, 199, 10929–10952.
- Piaget, J. (1923/1959). *The language and thought of the child*. Routledge & Kegan Paul.
- Roessler, J. (2013). Thought insertion, self-awareness, and rationality. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *The Oxford handbook of philosophy and psychiatry*. Oxford University Press.
- Roessler, J. (2016). Thinking, inner speech, and self-awareness. *Review of Philosophy and Psychology*, 7, 541–557.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43, 482–489.
- Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, 6, art. 314.
- van der Veer, R., & A. Yasnitsky, A. (2011). 'Vygotsky in English: What still needs to be done', *Integrative Psychological & Behavioral Science*, 45: 475–493.
- van der Veer, R., & Zavershneva, E. (2018). The final chapter of Vygotsky's *Thinking and Speech: A reader's guide*. *Journal of the History of the Behavioral Sciences*, 55, 101–116.
- Vicente, A., & Jorba, M. (2019). The linguistic determination of conscious thought contents. *Noûs*, 53, 737–759.
- Vicente, A., & Martínez-Manrique, F. (2016). The nature of unsymbolized thinking. *Philosophical Explorations*, 19(2), 173–187.
- Vygotski, L. S. (1934/1962), *Thought and language*, E. Hanfmann & G. Vakar (trans. and Eds.), MIT Press.
- Vygotski, L. S., & Luria, A. R. (1930). The function and fate of egocentric speech. In *Ninth International Congress of Psychology: Proceedings and Papers*. Psychological Review Company.
- Vygotski, L. S. (1930/1978). *Mind in society: The development of higher mental processes*, M. Cole, V. John-Steiner, S. Scribner & E. Souberman (trans. & Eds.), Harvard University Press.
- Vygotski, L. S. (1934/1986). *Thought and language*, A. Kozulin (trans. and Ed.), MIT Press.
- Vygotski, L. S. (1934/1987). 'Thinking and speech', N. Minick, (trans.). In R. W. Rieber & A. S. Carton (eds.), *The collected works of L. S. Vygotski* (vol. 1). Plenum.

- Vygotsky, L. S. (1934/2012). *Thought and language*. E. Hanfmann, G. Vakar & A. Kozulin (trans.). In A. Kozulin (Ed.), MIT Press.
- Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist*. J. B. Lippincott Company.
- Wilkinson, S. (2020). The agentive role of inner speech in self-knowledge. *Teorema*, 39(2), 7–26.
- Wilkinson, S., & Fernyhough, C. (2018). When inner speech misleads. In P. Langland-Hassan & A. Vicente (Eds.), *Inner speech: New voices*. Oxford University Press.
- Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 27. *Child Development*, 74, 659–678.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.