Trends in Cognitive Sciences

CellPress OPEN ACCESS

Review

Inner speech as language process and cognitive tool

Charles Fernyhough [1,*,@] and Anna M. Borghi[2]

Many people report a form of internal language known as inner speech (IS). This review examines recent growth of research interest in the phenomenon, which has broadly supported a theoretical model in which IS is a functional language process that can confer benefits for cognition in a range of domains. A key insight to have emerged in recent years is that IS is an embodied experience characterized by varied subjective qualities, which can be usefully modeled in artificial systems and whose neural signals have the potential to be decoded through advancing brain–computer interface technologies. Challenges for future research include understanding individual differences in IS and mapping form to function across IS subtypes.

Highlights

IS, namely silent speech directed to the self, has become a topic of increased attention in recent years, partly because of evidence of its involvement in key cognitive and emotional processes.

Because of its private, subjective nature, IS has traditionally been resistant to scientific study, but new cognitive and neuroscientific techniques are making it more tractable to empirical research.

IS can be seen as a cognitive tool that allows one to manipulate inner representations. It enhances performance in various cognitive tasks, especially if challenging, ranging from attention to memory to categorization.

Analysis of the phenomenology of IS shows that there are different subtypes of IS that perform various functions.

Thinking about how IS develops within an individual lifespan can provide a conceptual framework allowing researchers to account for various empirical findings.

Voices in our heads

IS is the form of internal language that many people report as a significant feature of their subjective experience. IS has been proposed to serve a range of cognitive functions, including action planning [1], emotion regulation [2], mediation of past- and future-oriented cognitions [3], and creative thinking [4]. It can be considered to be a form of **auditory imagery** (see Glossary) in which the self actively communicates with itself in a way that preserves some of the dynamics of social exchange [5]. IS can also be understood as a covert form of self-talk [6] – the self-directed speech (both silent and out-loud) that has been particularly linked to performance on skilled motor tasks such as sports [7].

This review concerns experiences variously described as inner (or internal) monologue, inner (or internal) dialogue, inner voice, covert (or silent) self-talk, internal narrative, verbal thinking, endophasia, autocommunication, and so on. The term 'inner speech' is preferred because it emphasizes the active nature of the phenomenon and does not prejudge featural and functional issues such as the dialogicality or plurality of IS voices. Researchers in this area have recognized the importance of distinguishing this covert (silent) form of self-directed speech from its overt (out-loud) form [7]. In the developmental psychology literature (Box 1), the latter is termed private speech [8] (Figure 1). Within the category of IS is a further important distinction between inner speaking (where the subjective experience is of producing an utterance in IS, analogous to speaking into an audio recorder [9,10]) and inner hearing (where the experience is of perceiving an utterance that has been innerly spoken, such as in listening to one's own voice on a recording). IS has been considered to be a form of auditory imagery in which language is produced for functional ends [5], thus distinguishing it from the less functionally significant process of imagining one's own speech [11].

The fact that a person's IS is objectively unobservable has made it notoriously resistant to scientific study. The most direct method for finding out about a person's IS is through self-report measures such as questionnaires [12], although responses on such instruments (Box 2) have the weakness of potentially being distorted by the preconceptions of the participants about what type of mind

[1]Department of Psychology and Centre for Research into Inner Experience, Durham University, Durham DH1 3LE, UK
[2]Department of Dynamic and Clinical Psychology, and Health Studies, Sapienza University of Rome and Institute of Cognitive Sciences and Technologies, Italian National Research Council, 00185 Rome, Italy

*Correspondence: c.p.fernyhough@durham.ac.uk (C. Fernyhough).
@Twitter: @cfernyhough

Box 1. Inner speech across the life-course

In Vygotsky's theory [6,93], IS develops in childhood through the gradual internalization of linguistically mediated exchanges with others, allowing the individual to begin to employ language previously used to regulate the behavior of others in the autoregulation of cognition and behavior Figure 1. One implication is that IS can transition flexibly between condensed and expanded forms and, under conditions of stress and cognitive challenge, be re-externalized as overt self-talk or private speech [77]. Vygotsky's theory thus provides a conceptual framework for integrating findings on covert and overt forms of self-talk, particularly across the timecourse of its emergence in childhood [5].

Vygotsky's theory has implications for accounts of IS as the prediction of the sensory consequences of an articulatory action (discussed in the section on Inner speech and the brain). One implication relates to the processes of syntactic and semantic abbreviation that are proposed to accompany the internalization processes [6], which mean that IS cannot merely be conceptualized as external speech minus the articulation. A related question concerns what form of speech is attenuated in predictive processing models. If IS develops according to the scheme proposed by Vygotsky, the type of speech whose articulatory components are attenuated in IS should be private speech, not social speech. Owing to the transformations accompanying internalization [94], private speech will already be functionally and formally distinct from its precursor, social speech, thus necessitating the refinement of predictive processing models of IS that conceive of it as attenuated social speech.

This view of IS has gained support from studies of neural development in the preschool years, particularly findings on the anatomical development of subcomponents of the dorsal language pathway, including the arcuate fasciculus [95]. Vygotsky's account of IS development also aligns with findings linking distinct subtypes of IS (including dialogic and other voices IS; Box 2) to engagement with non-real social partners such as imaginary companions [96]. The possibility that IS might provide a medium for interacting with non-present others [97] is also consistent with findings that preschool children with imaginary companions are more likely to engage in private speech, suggesting that such speech may have a role in simulating and facilitating social exchanges [98].

they have [13]. Overt (out-loud) self-directed (or private) speech is generally easier to measure than its covert counterpart, although collapsing it together with IS (as has frequently happened in the sports self-talk literature; Box 3) risks obscuring important **phenomenological** and functional differences between IS and private speech (Box 1). A further methodological issue arises when studies instruct participants to use IS in particular ways but do not actually measure the speech that is produced [14]. Another interesting method to assess IS, specifically by impeding its production, is the use of verbal interference tasks such as **articulatory** suppression, which typically involves repeating words out loud so as to block language processing [15]. Also useful, particularly in the neuropsychological domain, are tasks such as rhyme judgments, homophone judgments, and judgments of lexical stress of written words (recently combined into a specific IS test battery [16]). Neuroscientific techniques such as brain imaging and neurostimulation have added detail to the picture of how IS arises functionally in the brain [17,18]. Various **experience sampling** methods have also been employed to study IS [19]. Particularly productive has been work with Descriptive Experience Sampling (DES), a method that helps participants to gain skill in providing detailed reports on moments of experience [20,21], resulting in phenomenological descriptions which can subsequently be used as inputs for more quantitative analyses [22] (Box 2).

Such methodological innovations have contributed to a recent resurgence of interest in IS, and many outstanding questions about the phenomenon are now more tractable to empirical inquiry. A further reason for the growth of interest in IS has been the recent wave of studies emphasizing the importance of language in shaping cognition [23–26]. As we explore in more detail below, this renewed interest is exemplified by current debates on the capabilities of **large language models (LLMs)** and **linguistic distributional knowledge** for reproducing human cognitive abilities [27].

Many of these advances can be understood within a framework of understanding IS as a **cognitive tool** that shares most (if not all) of the functions of overt language but does not require a specific social setting. In this review we assess the present utility of this framework and its potential to stimulate further research in the area. We first ask about the status of IS as a form of language,

Glossary

Articulatory: relating to how humans produce speech sounds.
Auditory imagery: the subjective experience of an auditory percept in the absence of any direct sensory cause of that experience.
Cognitive tool: the idea, derived from Vygotsky's theory, that mental entities such as words in natural language can augment pre-existing cognitive capacities.
Electrocorticography: an electrophysiological technique that involves placing electrodes directly onto the surface of the cortex to record its electrical activity.
Electromyography (EMG): a technique used to measure the electrical activity generated by the muscles of the body.
Embodied: constrained by the type of body an agent possesses.
Experience sampling: a family of methods that ask participants to report introspectively on their own thoughts, feelings, and behaviors.
Flanker task: a class of cognitive tasks that assess the ability to suppress responses that are inappropriate in a particular context.
Large language models (LMMs): a class of machine learning models that can perform a variety of natural language tasks.
Linguistic bootstrapping: the hypothesis that language facilitates complex cognitive tasks by allowing labels to serve as substitutes for bundles of sensorimotor information.
Linguistic distributional knowledge: knowledge concerning how words are related to one another as a result of their co-occurrence.
Phenomenological: relating to the subjective, qualitative aspects of experience.
Predictive processing: a framework for understanding the brain as predicting what exists in the environment instead of passively receiving sensory data from it.
Rumination: a style of thinking that occurs as a maladaptive response to distress and negative mood.

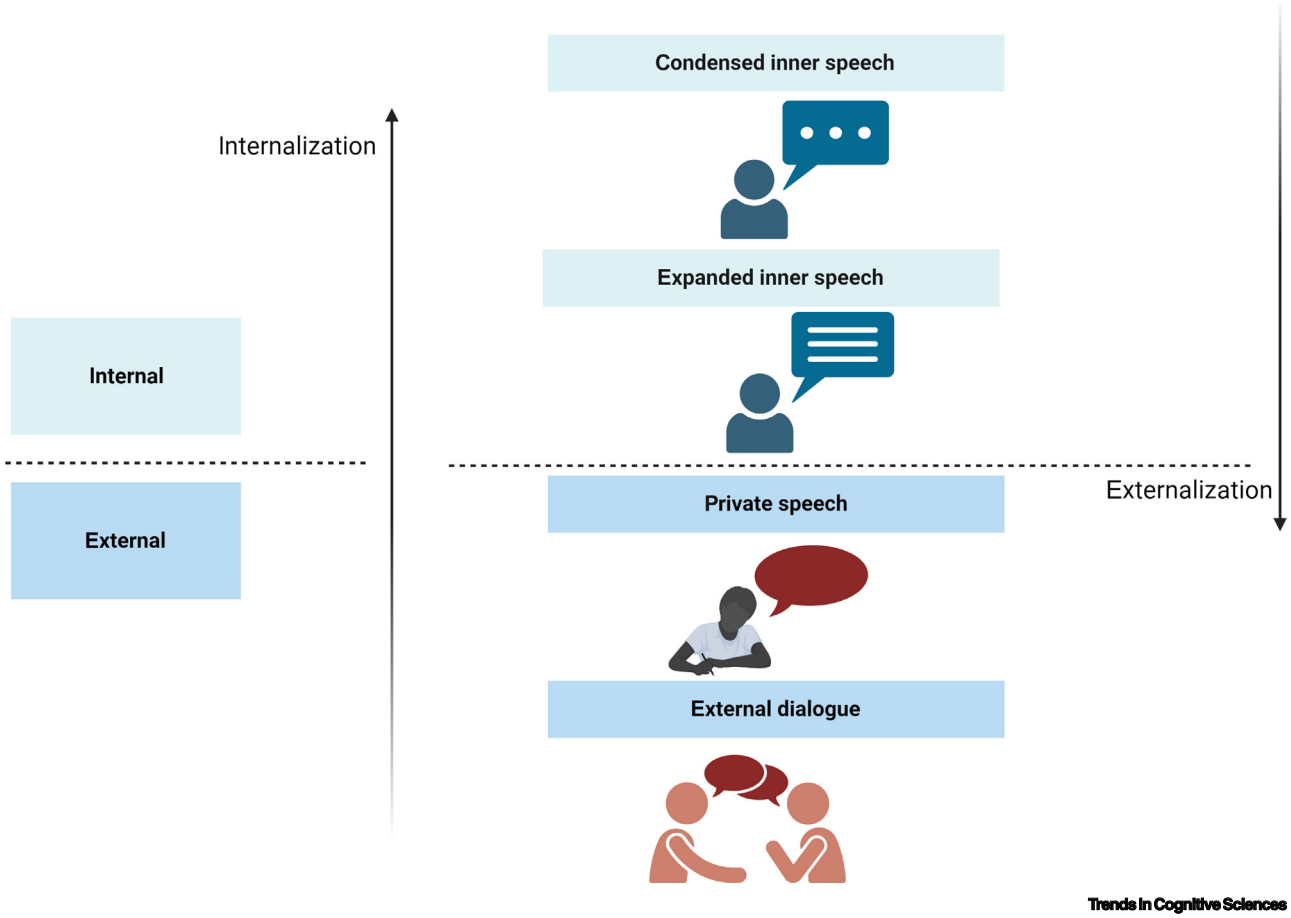

Figure 1. From external dialogue to inner speech (IS). This figure illustrates the processes of internalization described by Vygotsky [6]. From the earliest days of life, infants are engaged in social exchanges which become mediated by speech as language is acquired (external dialogue). These social exchanges become internalized to form private speech or overt self-talk. Further transformations accompany the internalization of private speech into IS, which can take varied positions on a spectrum of expanded and condensed forms. Re-externalization of internalized speech means that people can move flexibly between condensed and expanded IS, particularly with changing cognitive demands, and can also re-externalize their IS as overt private speech [77].

particularly its functional properties as a medium for communicating with the self. We then consider how IS is instantiated in the brain, including how new research is beginning to understand the neural signatures of the varied phenomenological features of IS. We review recent findings on the capacity of IS to enhance cognitive performance in a range of domains before considering what its **embodied** nature means for its potential to be modeled in artificial systems, drawing on advances in brain–computer interface technology, robotics, and artificial intelligence.

Inner speech as a linguistic form

Researchers have long been interested in how to characterize IS as a language process, particularly the extent to which it is similar to external speech. One way to address this question is by asking whether IS retains the acoustic and **articulatory** features of spoken language. Theoretical considerations [6] and phenomenological evidence [9] (Box 2) indicate that IS can vary on a spectrum of featural richness: from fully phonologically featured (in terms of tone, timbre, accent, pitch, etc.) to highly abstract (where phonological features are reduced or absent). Two theoretical models occupy opposite ends of this spectrum. At one end, motor simulation views [28,29] hold that IS is a mental simulation of overt speech that falls short of producing the articulatory actions

Box 2. What is inner speech like?

A focus of recent research on IS has been on its phenomenology or subjective qualities. Vygotsky's [6] theory entails that IS should take different forms (e.g., relating to the extent to which utterances are articulated in full sentential structure). However, likely owing to difficulties in studying IS, scientific research on the phenomenon has tended to treat it as a monolithic entity – a single 'voice in the head'. The systematic study of variations in IS quality began with the development of the Varieties of Inner Speech Questionnaire (VISQ) [99], which in its revised form [34] captures five dimensions of IS quality. These include dialogicality (postulated by Vygotsky to result from the origin of IS in social, conversational speech), condensation (following from Vygotsky's theory of how IS is syntactically and semantically transformed as it is internalized), the presence of other voices in IS (involving the hearing of the voices of others, also betraying the social origin of IS), evaluative/critical (associated with a relatively negative self-concept), and positive/regulatory (roughly mapping onto the motivational self-talk examined in the sports literature). The subtypes of IS identified by the VISQ have been replicated in Spanish [100] and Hebrew [36]. As summarized elsewhere in this review, recent research has begun to link specific subtypes of IS with particular cognitive and emotional processes, including psychopathological variables such as hallucination-proneness [34].

The phenomenological diversity of IS is also evident in research using Descriptive Experience Sampling (DES), although such studies have not always provided strong support for features such as condensation [9]. Given that different methods can produce such discrepancies in data, research using multiple methods for assessing IS quality within a single study seems to be particularly advisable [101]. Beyond the traditional boundaries of cognitive science, studies that incorporate humanities (including literary) methods and concepts [102] may represent a fruitful line of future interdisciplinary inquiry into the subjective nature of IS.

necessary for producing audible speech. At the other end of the spectrum, abstraction accounts [30] view IS as occurring independently of bodily actions, including articulatory processes.

Instead of providing exclusive support for either of these positions, recent evidence suggests that IS varies on a continuum between extremes of featural richness and abstraction [5]. Evidence for motor activation in IS includes activity recorded by **electromyography (EMG)** of articulatory muscles during IS production [31,32] and disruptions of IS resulting from articulatory suppression [33]. Evidence for abstraction comes from findings that IS can occur in more condensed and less featurally rich forms [34–36]. A recent example supporting abstraction showed a dissociation between speeds in reading aloud and silent reading in adult participants. Whereas speed of reading aloud correlated with articulation speed, silent reading (assumed to involve IS) correlated more strongly with vocabulary and conceptual knowledge [37].

Foremost among the properties of language is its function for communication. When considering IS as a linguistic form, a paradox arises: how can language be used to communicate information to the same self that is generating that language? This paradox is particularly problematic in the context of a conventional 'conduit' [38] view of language as the transmission of information, because this seems to exclude the possibility that an agent can communicate to itself information that it does not already know. Accepting the various proposed roles for IS in self-communication seems to entail a view of language as having a broader range of functionality than is allowed by conventional conduit views [39]. In particular, understanding IS as a productive, active process rooted developmentally in social processes (Box 1) allows self-directed language to have functions beyond the transmission of information: for example, in regulating the self's activities in an analogous way to how speech regulates the behavior of others [6,11] (see also the section on Inner speech as a cognitive tool below).

Inner speech in the brain

IS can thus be understood as a form of language, directed towards the self, which retains the range of functional possibilities of external speech. Linking to the question of whether IS includes an articulatory component, researchers have asked how its neural signatures associate selectively with activity in brain networks linked to speech production and auditory imagery [40]. Relevant evidence comes from findings of distinct neural signatures for listening to speech versus imagining speech [41]. Although both conditions showed overlapping activation in bilateral superior temporal gyrus and supplementary motor areas, imagined speech showed distinct activation

Box 3. Self-talk in motor performance

The term self-talk collapses the distinction between IS and overt speech. A considerable body of research has examined how self-talk facilitates sports performance. Dual-process theories distinguish between strategic and organic self-talk, the latter of which can be spontaneous or goal-directed [103], and between instructional and motivational self-talk [104]. In elite basketball players, instructional but not motivational self-talk improves shooting accuracy and reduces coordination movement variability [105]. Male basketball players use self-talk to reduce anxiety and lower psychophysiological arousal; during competitions, self-talk helps to maintain activation, regulate behavior, and prevent disengagement [106]. In novice golf players, instructional self-talk is associated with better technique and action control, and motivational self-talk is associated with increased heart-rate and reduced heart-rate variability, testifying to increased effort [107]. Supporting dual-process theories, in competitive tennis matches emotional intensity is lower with exclusively goal-directed self-talk (which regulates emotions) than with spontaneous self-talk [108].

Machine learning methods have been used to analyze the quality of participant self-talk to predict which sport (badminton vs running) is being engaged in [109]. Runners, engaged in a solitary sport, use IS to prevent disengagement, whereas badminton players use it to express worries and control anxiety. During marathon and half-marathon running, self-talk is more valenced in high-pressure situations, and becomes shorter and more repetitive in low-pressure contexts [109].

Questionnaires have limitations as a guide to IS experiences. For example, in tennis athletes the correspondence between self-talk accessed through questionnaires and observation is often weak [110]. Methods such as Descriptive Experience Sampling (DES) may be more powerful in providing information on specific IS episodes. DES shows that covert self-talk (IS) in golfers is tenfold more frequent than overt self-talk and occurs more during tournaments [20]. Experimental interventions offer a further method for assessing the effects of self-talk on motor performance. Verbal interference affects cycling more than visuospatial interference or control tasks, suggesting that self-talk fosters physical endurance [111]. Self-talk interventions influence the performance of cyclists in hot environments [112] and induce physiological changes (e.g., breath frequency alterations) in runners [113].

Self-talk studies also provide evidence on individual differences in IS [114]. The social isolation hypothesis, which associates self-talk with greater isolation, is challenged by findings that self-talk does not differ substantially between solo travelers and non-travelers [115]. More evidence supports the cognitive disruption hypothesis, which predicts that more frequent self-talk is associated with increased experiences of negative or stressful self-related events [114,116].

in left inferior frontal gyrus (Broca's area), which was taken as evidence for the involvement of an articulatory component in the latter, combined with auditory imagery (*cf* [42]). One problem in interpreting these findings is that spontaneous IS is best understood as a functional linguistic act rather than an act that is simply imagined [11]. In addition, although Broca's area typically activates during performance of IS tasks [43], there are reasons for caution in assuming that neural signatures elicited by task instructions will align with IS that arises more spontaneously [17]. Simply instructing participants to generate or imagine IS in experimental contexts may result in experiences that are some way removed from natural, functional IS, with important implications for the scientific lessons that can be learned from their study.

Related to the question of the functionality of spontaneous IS is the idea that it is a consequence of simulated speech actions which are not converted into articulatory activations. Viewing IS as a sensory consequence of a simulated action allows it to retain an active (and thus functional) dimension while also resonating more generally with popular **predictive processing** accounts of perception [44]. Predictive processing accounts of IS are not yet highly developed, perhaps owing to the greater challenges that such models face in understanding 'offline' cognition (such as internal imagery and mind-wandering) in comparison to processes where perceptual inputs and action outputs are tightly linked [45]. However, predictive processing accounts of IS have begun to emerge [46], and developmental questions about how IS is internalized from social and private speech promise to shape new understandings of what should be predicted in such models (Box 1). Thinking about IS as the sensory product of a simulated action has also illuminated how predictions of speech actions are neurally instantiated. The corollary discharge that accompanies an IS utterance (and suppresses neural and perceptual responses to self-generated speech) has recently been shown to carry precise information about both the timing

and content of the IS utterance [47]. Participants generated inner phonemes either 300 ms before, concurrently with, or 300 ms after they heard an audible phoneme that either matched or did not match the content of the inner phoneme. Production of inner phonemes attentuated the N1 component of the auditory cortex event-related potential, but only when the two phonemes occurred at the same moment in time and had the same content, suggesting close commonalities between the neural processes involved in generating internal and external speech.

Predictive processing models have also proved their value in accounting for the varied phenomenological features of IS (Box 2). One recent fMRI study [35] presented a comprehensive attempt to relate featural variation in IS to neural processes, focusing on three dimensions of dialogality (relating to the extent to which utterances preserve the conversational structure of social dialogue), intentionality (referring to the spontaneous vs intentional nature of the utterances), and condensation (relating to the extent to which utterances are abbreviated relative to external speech). Following previous modeling of dialogic IS in terms of neural substrates of social cognition [18], the dialogic nature of IS was explored through a perspective-switching (first- vs third-person) task, and was found to relate to neural activation in areas previously associated with such tasks, including precuneus and parietal lobules. Intentionality was examined by comparing intentional IS with verbal mind-wandering, with the latter associated with greater activation in right-hemisphere areas including parieto-fronto-temporal regions. Although not targeted empirically in this study, variation in condensation was modeled by specifying inhibitory control at the levels of utterance conceptualization, formulation, and articulatory planning. Although preliminary, these findings represent an important first step in specifying the neural signatures of different subtypes of IS. The idea that varieties of IS are subserved by different neural mechanisms gains further backing from an activation likelihood estimation (ALE) meta-analysis [48] supporting a separation between speech production and speech reception areas underlying different types of IS.

Inner speech as a cognitive tool

Interest in IS has been heightened by evidence that language augments and transforms cognition [23,49,50]. The next section summarizes recent studies showing that IS can improve attention and cognitive control, enhance cognitive flexibility, facilitate working and long-term memory, influence categorization, and enhance the ability of humans to reflect on themselves and monitor their own cognitive processes.

Attention and cognitive control

IS can help in conflict resolution and cognitive control, as evidenced by studies requiring focused attention. Participants who use more evaluative and motivational IS, as assessed through questionnaires including the Varieties of Inner Speech Questionnaire (VISQ; Box 2), perform better on non-verbal tasks requiring focused attention, specifically in the Simon task (in which participants are less influenced by the position of stimuli on a screen when not relevant to the task) and in the **flanker task** with arrows as stimuli. The facilitatory effects are more consistent with these non-verbal tasks than with verbal attention tasks such as the Stroop task and the flanker task with letters [51,52]. Importantly, IS explains patterns of findings over and above constructs such as intelligence and working memory capacity. Other evidence suggests that IS sustains focused attention over time. In one study, participants performed a boring task in which response times to infrequent stimuli were recorded, and reported on their experiences when the stimuli appeared [53]. Task-relevant IS improved focused attention and processing efficiency.

Aside from spontaneous IS, strategic self-talk can enhance attentional processes and processing time, as shown in a study on different attentional functions including alertness and vigilance [54]. Distanced self-talk (referring to oneself by using one's own name instead of the first-person singular

pronoun) improves control and is effective in the choice of healthy food in dieters [55]. Compared to immersed (i.e., first-person) self-talk, distanced self-talk also facilitates emotion regulation when reflecting on both future and past negative personal experiences [56]. Results from a dictator game (in which participants in a 'decider' role determine whether and how much money to assign to 'recipients') in both laboratory and online settings revealed that participants using distanced self-talk made the most rational choice in contexts where the decision was one-shot and there was no opportunity for reciprocity [14]. Using distanced self-talk is also associated with more abstract self-representations. In two studies, participants responded to the question 'Who am I?' (starting from their own or their best friend's perspective: self-distance vs social-distance) or 'Who is [participant's name]?' Self-distance provided a more detached perspective than social-distance: self-distancing participants produced more abstract person characterizations and used more abstract statements [57].

Cognitive flexibility and creativity

In addition to enhancing cognitive control, IS also enhances cognitive flexibility. A recent computational model [58] replicated human data on the Wisconsin Card Sorting Test (WCST), a neuropsychological test of task-switching ability. Improvements in efficiency of manipulating internal representations were achieved by decoupling of internal representations from specific visual features, making the agent more capable of engaging or disengaging attention after positive or negative outcomes to its action. The model also reproduced experimental findings that autistic children (Box 4) and older adults (but not younger adults) show poorer performance than controls in rule-switching [59]. The cognitive flexibility granted by IS [60] may account for evidence that forms of condensed and evaluative/critical IS correlate with creative potential [61] (Box 2).

Box 4. Individual differences in inner speech

There is considerable inter-individual variation in IS frequency and phenomenology [34,117]. In atypically developing samples [118], adolescents with Social (Pragmatic) Communication Disorder self-reported IS in both monologic and dialogic forms, and the two forms correlated with each other [119]. A large online community sample study found that higher scores on the evaluative/motivational IS subtype were associated with a stronger negative relation between autism traits and emotion regulation strategies [120]. Patterns of IS use have been linked to atypical cognitive flexibility in the autism spectrum [121], although it remains difficult to establish the direction of any causal relations. With regard to psychopathological variables, an association has been described between other voices IS and proneness to auditory verbal hallucinations in a community sample [34]. In schizophrenia patients, cognitive performance was found to moderate this relation, but this was not observed in patients with high cognitive function, with implications for interventions to improve cognitive performance in such patients [122].

A developmental perspective can also help in understanding the implications of damage to the language system such as occurs in acquired aphasia. Contradictory findings have emerged on how different forms of language loss affect IS, such as evidence that IS can remain intact alongside failures of overt speech production [123]. In a sample of 53 aphasic patients, levels of self-reported IS related to phonological retrieval but not to speech production, further evidence that at least some forms of IS are independent of articulatory processes [124]. A functional systems approach [94,125] makes it possible to understand how typical development before the acquisition of aphasia can allow the emergence of structures necessary to support particular cognitive functions, such that alternative cognitive processes can be recruited post-insult to perform the function concerned [5]. More generally, conceptualizing IS as a functional system provides a neurocognitive model for understanding how the core language system can interact with other neural systems to flexibly augment a range of processes such as social cognition, executive functioning, and autobiographical memory [94,126,127].

A lifespan approach to IS can shed light on emerging findings of cultural differences. The vehicle for IS is natural language, including signed languages known to operate as a medium for private and inner self-directed signing [128]. To date, transcultural differences in IS have not been systematically investigated, although there is relevant evidence concerning its developmental precursor, private speech [129]. Further research into multilingualism, where more than one language can potentially be a medium for IS, will continue to illuminate how IS confers benefits as a cognitive tool [130].

Memorization

IS can enhance working and long-term memory, likely owing to **linguistic bootstrapping**. Compared to a spatial suppression condition – in which participants had to tap five red marks sequentially – articulatory suppression impaired performance in the Luria hand test (LHT), a non-verbal neuropsychological test of serial recall in which participants have to reproduce sequential movements [62]. In another study, participants watched videos and then were tasked to build simple models (e.g., a birdhouse) from memory [63]. Disrupting IS through articulatory suppression during encoding affected the accuracy and number of recalled events more than other dual tasks that did not require access to language.

Categorization

Results pertaining to IS and categorization are mixed. In Mongolian (but not in Chinese), two different words refer to light and dark blue. The presence of the two labels influenced free sorting and visual search by speakers of Mongolian, which was faster when the items belonged to two different categories [64]. A verbal (but not spatial) interference task disrupted the effect, suggesting that participants covertly named the color.

Other studies suggest that IS has no specific role in conceptualization. When a verbal interference task was used during the naming of simple objects and while responding to the visual depiction of actions, occupying the articulatory buffer did not impact on conceptualization [65]. The role of language has been examined in the recognition of wine odors, wine-related odors, and common odors in wine experts and novices [66]. In one experiment, participants smelled odors and half were asked to name them; in the second, participants performed a verbal task that should inhibit IS. Experts were better at recognizing odors of wines, but not wine-related and common odors, and their expertise-specific advantage was not mediated by language. IS may be particularly important when conceptualization and naming are challenging, thus explaining the absence of a verbal mediation effect when dealing with simple objects or elements on which participants are experts.

Consistent with this idea are studies showing that IS might be more important in operating with abstract than with concrete concepts. Several studies suggest that IS can facilitate the acquisition and processing of abstract concepts such as 'freedom' or 'justice' compared to concrete concepts, which have more obvious perceptual dimensions and typically take a single object as referent. Various evidence reveals that, during abstract concept processing, the mouth motor system is activated: for example, people are faster when responding to abstract concepts with the mouth than with the hands, unlike with concrete concepts (reviewed in [67,68]). Developmentally, longer use of a pacifier slows down processing of abstract concepts at a later age [69]. Articulatory suppression selectively delays response times to abstract (compared to concrete) words in a task that requires sorting words into concrete versus abstract categories [33]. Mouth activation in such studies may betray the involvement of IS, proposed to be crucial both during the acquisition and processing of abstract concepts [25,67].

Different subtypes of IS may play different roles during abstract concept acquisition and processing. IS has been proposed to support inner social metacognition, an internal monitoring mechanism through which people interact with themselves dialogically to access an abstract concept and find a possible word meaning [70]. Similarly, IS can underpin social metacognition, a mechanism for preparing to interact with others to ask for information, understand what they intend (i.e., accessing the content of their concepts), or negotiate concepts and word meanings [23,67,71] (Figure 2).

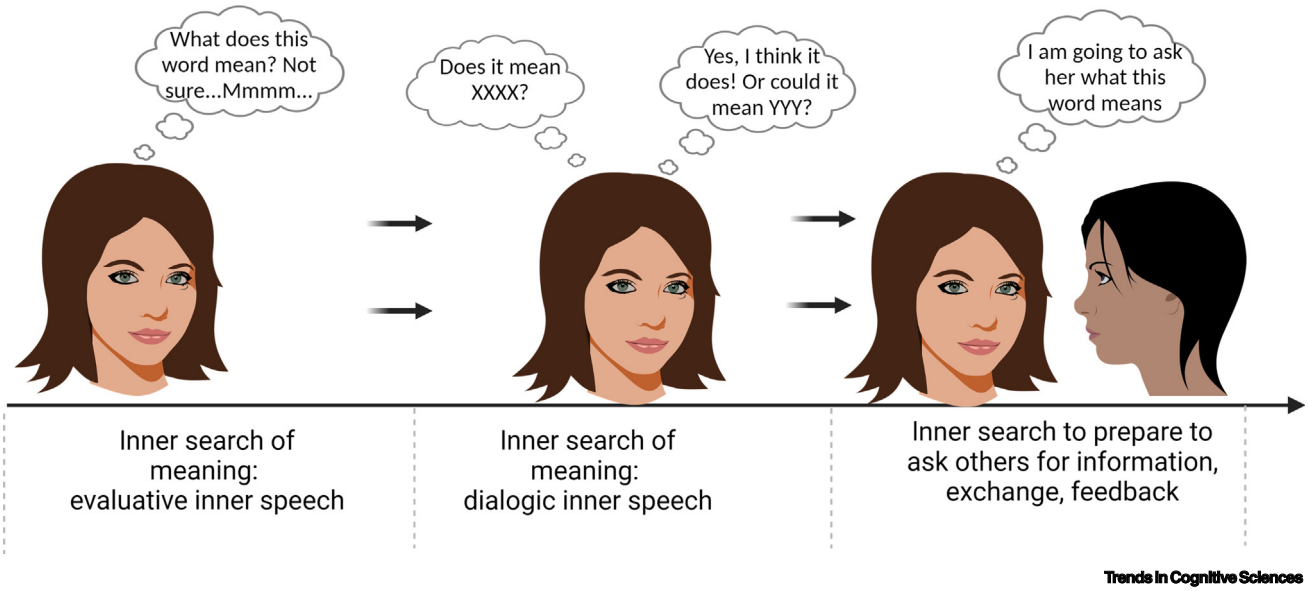

Figure 2. Inner social metacognition. This figure illustrates how the process of inner social metacognition might work. When processing abstract concepts, people need to monitor their own knowledge in search of the meaning of a word and to discover its eventual limitations. They might use evaluative inner speech (IS, first panel) to do this. Subsequently, dialogic IS might be used to home in internally on relevant information (second panel). Finally, people might use IS to prepare themselves to refer to others to find information and support (third panel).

Inner speech and the self

Studying IS sheds new light on relations between the self and the outside world, including wider social processes. IS can provide an opportunity to reflect on ourselves [30] and form a coherent self-concept [72], thereby enhancing human capacity for self-observation, introspection, and metacognition. In this sense, IS can work as a cognitive tool that is useful in exploring and monitoring the self's activities. Using an open thought listing method (that has an advantage over questionnaires because it captures self-generated rather than predefined forms of IS), students were asked when they used IS, why, and about what [73]. IS was associated with everyday activities (taking public transport, meeting people, performing hygiene tasks, etc.) and when talking about preferences, friends, school, and general life, with the most frequent uses concerning self-regulation, self-reflection, and cognition. Such reports suggest that we can experience ourselves as subject ('I') when we are minimally aware of ourselves and focus instead on the surrounding world, and as object ('Me') when the self is the focus of our thoughts [74]. A self-report study of IS phenomenology in university students showed that, although self-as-subject experiences were more frequent than self-as-object, both were associated with everyday activities such as problem-solving, planning, and self-motivation [75]. Further investigation of such aspects of IS phenomenology may illuminate questions of form and function in IS subtypes: for example, whether non-dialogic (monologic) IS is particularly associated with immersed, self-as-subject experiences, or whether condensed and expanded IS play specific roles in reducing cognitive load [76] and outsourcing cognitive demands [77], respectively. Finally, not all literature converges on a foundational role of IS for self-awareness and regulation. Analysis of IS phenomenology through questionnaires has shown that self-awareness is of more value than IS in explaining self-regulation, thus questioning the importance of IS in explaining this aspect of behavior [78].

This section has reviewed evidence that IS can augment cognition by focusing attention, providing a distance from the self that enhances the rationality of decision-making, enabling

greater cognitive flexibility, improving memory in non-verbal tasks, and giving its users a firmer grasp of abstract concepts. More mixed findings have been obtained for categorization and self-awareness. A task for future research will be to understand whether IS performs these functions in a similar way to overt self-directed (private) speech. Another important issue concerns whether IS is crucial only during development or whether it continues to influence cognition online in adulthood (Boxes 1 and 4).

Inner speech in artificial systems

There is a long history of understanding IS as an embodied process rooted in real-world social interactions [6,74]. Thinking about how IS emerges through human development (Box 1) raises interesting questions of how embodied IS can be modeled in non-human (and thus differently embodied) systems, one reason for the interest in the topic among artificial intelligence (AI) and robotics researchers. Further reasons for this interest stem from the evidence, reviewed in the previous section, that IS confers cognitive benefits on its users, raising the possibility that similar benefits might accrue to IS in artificial systems.

Research on IS in such systems has been an area of particular growth in recent years. In the example of the computational model described above [58], systematic lesioning showed that IS enhanced the cognitive flexibility of an agent when tested on the WCST. Modeling IS in an artificial agent has also been shown to enhance learning from external information. Intra-agent speech was modeled in an agent exploring a 3D virtual world, showing that the agent was able to manipulate and answer questions about a novel object without any direct experience of it [79]. LLMs operating in both virtual and real-world environments (such as in a table-top rearrangement task) are able to use a form of inner monologue to improve the following of instructions to arrange objects, suggesting that inner monologue presents a natural framework for LLMs to incorporate information about their actions, thus enabling closed-loop feedback for planning on a range of long-horizon tasks [80].

Data on the use of IS in LLMs are susceptible to well-known problems of interpretation, including whether intra-agent speech as modeled in an artificial system has a genuinely causal role in any observed enhancements of cognition. The question of whether intra-agent speech in such systems has anything like the phenomenology or agentive force of human IS will likely have to await the appearance of a conscious AI, with all the reasons for skepticism that this idea entails. In the meantime, however, studies of artificial agents may shed light on some of the factors that constrain both biological and non-biological systems from attaining various forms of self-awareness. Presenting a basic cognitive architecture that allows a robot to use self-talk to explore simple arrangements of everyday objects [81], one team proposed that the robot's use of speech to generate descriptions of its own actions might in turn enable a rudimentary form of self-awareness, by allowing the robot to observe and describe its own actions. Consistent with this speculation, an artificial agent implementing this basic architecture showed some success on the classic 'mirror test' of self-recognition [82]. Advantages of the model include the fact that it did not require training on the task, and that the agent's self-talk was transparent (or overt), thus allowing researchers to examine which specific aspects of the model's reasoning related to success.

Another area of growth has been the study of brain–computer interfaces as a means of decoding IS. Recent work has investigated whether IS can be decoded through surface EMG, with some findings providing reason for skepticism. An automatic classification approach to EMG signals was unable to discriminate the phonetic content of IS in a way that was possible with overt speech [31]. A study of experimentally induced **rumination** (compared to a distraction condition)

combined facial EMG with self-report questionnaires [32]. EMG signals did not differ between the two conditions, although self-report measures showed that rumination involved more IS. This supports an interpretation that IS may in some cases be accompanied by articulatory muscle activation (discussed in the section on Inner speech as a linguistic form), but that an absence of any such activation should not be taken as evidence for an absence of IS.

The conclusion that the presence of IS may not be reflected by gross differences in muscular electrical activity points to the need for more sophisticated methods to analyze electrophysiological signatures of IS. In one such approach, computational processing of EEG signals was investigated as part of a publicly available IS dataset [83]. A convolutional neural network classified small target groups of vowels and words, and obtained accuracy scores of 35.20% and 29.21% respectively. High (>90%) levels of accuracy have been reported for decoding higher-level phoneme categories from an EEG database using a deep capsule neural network [84]. Phase-locking of IS to theta-band oscillations has shown potential for decoding IS elicited by reading [85], mirroring similar findings for overt speech [86]. An encoding model trained on fMRI signals from participants listening to continuous language has had considerable success in decoding cortical signals of imagined speech generated in an entirely separate task [87]. Training brain–computer interface systems on word classification holds particular promise for speech prosthesis and neuropsychological rehabilitation [88], particularly when based on methods such as **electrocorticography** [89] and the use of bimodal datasets (e.g., fMRI combined with EEG) [90]. Less invasive methods for IS decoding, such as those based on EEG, may eventually have everyday applications in smart homes and healthcare devices [83], including wearable technologies [91].

Concluding remarks

Traditionally seen as resistant to scientific study, IS has recently emerged as a growth area in research [92]. This likely stems from its apparent involvement in a range of cognitive and affective processes in which self-directed language appears to have an enhancing effect. Vygotsky's [6,93] original conception of such language as a 'psychological tool' has been supported by a wealth of empirical evidence that talking silently to oneself can enhance cognition. At the same time, our review of recent evidence suggests that only some cognitive domains benefit from accompanying self-directed speech, and much empirical work will be necessary to test, put pressure on, and potentially refute the dominant theoretical ideas in this area. Future research on IS (see Outstanding questions) will likely benefit from continued methodological advances, particularly in understanding individual differences in IS and in improving techniques for linking subtypes of IS to specific functional benefits. The varied voices of IS still have much to reveal about their instantiation in the nervous system and their potential for transforming human thought and non-human intelligence.

Outstanding questions

How can different methods for assessing IS (questionnaires, experience sampling, verbal interference paradigms, neurophysiological methods, etc.) best be integrated into a comprehensive toolkit for advancing the science of IS?

How can individual differences in IS be better understood, including through the development of new methods for assessing IS? What developmental, personal, and sociocultural factors might help to explain these individual differences? What methods are suitable for investigating the question of whether some people do not use IS at all (sometimes termed anauralia or anendophasia)?

How do different subtypes of IS (condensed, dialogical, evaluative, etc.) confer specific functional benefits for cognition, and how do they differ across varied cognitive profiles?

How can IS be modeled in artificial systems? Can IS be decoded for practical purposes by using electrophysiological and other methods?

Can IS be a tool for metacognition, including as a medium for reflecting on our own cognitive processes? What methods would allow the use of IS for such purposes to be addressed empirically?

How are different subtypes of IS involved in the processes of categorization? Which subtypes of IS are employed in different conceptual processes such as concept acquisition, negotiation, and use?

Has human IS changed, and will it continue to change, with the use of interactive social media? How will it change with the use of new communicative systems made possible by the employment of LLMs?

Acknowledgments

This research was funded in whole, or in part, by the Wellcome Trust (grant WT108720), by National Recovery and Resilience Plan (PNRR) PE 08, age.it, spoke 4, Healthy Aging, and by Concepts in Interaction with Others and with Ourselves, Sapienza Excellence Projects grant RG12117A5D1EB0B3. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) public copyright license to any Author Accepted Manuscript version arising from this submission.

References

1. Lidstone, J.S.M. *et al.* (2010) The roles of private speech and inner speech in planning during middle childhood: evidence from a dual task paradigm. *J. Exp. Child Psychol.* 107, 38–51
2. Morin, A. *et al.* (2011) Self-reported frequency, content, and functions of inner speech. *Procedia Soc. Behav.Sci.* 30, 1714–1718
3. Larsen, S.F. *et al.* (2002) Inner speech and bilingual autobiographical memory: a Polish–Danish cross-cultural study. *Memory* 10, 45–54
4. Fernyhough, C. (1996) The dialogic mind: a dialogic approach to the higher mental functions. *New Ideas Psychol.* 14, 47–62
5. Alderson-Day, B. and Fernyhough, C. (2015) Inner speech: development, cognitive functions, phenomenology, and neurobiology. *Psychol. Bull.* 141, 931–965

6. Vygotsky, L.S. (1987) Thinking and Speech. In *The Collected Works of L.S. Vygotsk* (Vol. 1) (Rieber, R.W. and Carton, A.S., eds)
7. Winsler, A. (2009) Still talking to ourselves after all these years: a review of current research on private speech. In *Private Speech, Executive Functioning, and the Development of Verbal Self-Regulation* (Winsler, A. *et al*., eds), pp. 3–41, Cambridge University Press
8. Flavell, J.H. (1966) Le langage privé. *Bull. Psychol.* 19, 698–701
9. Hurlburt, R.T. *et al.* (2013) Toward a phenomenology of inner speaking. *Conscious. Cogn.* 22, 1477–1494
10. Kühn, S. *et al.* (2014) Inner experience in the scanner: can high fidelity apprehensions of inner experience be integrated with fMRI? *Front. Psychol.* 5, 1393
11. Wilkinson, S. and Fernyhough, C. (2018) When inner speech misleads. In *Inner Speech: New Voices* (Langland-Hassan, P. and Vicente, A., eds), pp. 244–260, Oxford University Press
12. Racy, F. *et al.* (2022) Concurrent validity and reliability of representative inner speech questionnaires. *Psychol. Rep.* Published online June 13, 2022. https://doi.org/10.1177/003329412211091
13. Hurlburt, R.T. *et al.* (2022) Measuring the frequency of inner-experience characteristics. *Perspect. Psychol. Sci.* 17, 559–571
14. Gainsburg, I. *et al.* (2022) Distanced self-talk increases rational self-interest. *Sci. Rep.* 12, 511
15. Nedergaard, J.S.K. *et al.* (2023) Verbal interference paradigms: a systematic review investigating the role of language in cognition. *Psychon. Bull. Rev.* 30, 464–488
16. Geva, S. and Warburton, E.A. (2019) A test battery for inner speech functions. *Arch. Clin. Neuropsychol.* 34, 97–113
17. Hurlburt, R.T. *et al.* (2016) Exploring the ecological validity of thinking on demand: neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS ONE* 11, e0147932
18. Alderson-Day, B. *et al.* (2016) The brain's conversation with itself: neural substrates of dialogic inner speech. *Soc. Cogn. Affect. Neurosci.* 11, 110–120
19. Alderson-Day, B. and Fernyhough, C. (2015) Relations among trait and state measures of inner speech quality using smartphone app experience sampling. *Front. Psychol.* 6, 517
20. Dickens, Y.L. *et al.* (2018) On investigating self-talk: a descriptive experience sampling study of inner experience during golf performance. *Sport Psychol.* 32, 66–73
21. Van Raalte, J.L. *et al.* (2019) Dialogical consciousness and descriptive experience sampling: implications for the study of intrapersonal communication in sport. *Front. Psychol.* 10, 653
22. Fernyhough, C. *et al.* (2018) Investigating multiple streams of consciousness: using descriptive experience sampling to explore internally and externally directed streams of thought. *Front. Hum. Neurosci.* 12, 494
23. Borghi, A.M. (2023) *The Freedom of Words: Abstractness and the Power of Language*, Cambridge University Press
24. Dove, G. (2022) *Abstract Concepts and the Embodied Mind: Rethinking Grounded Cognition*, Oxford University Press, Oxford, New York
25. Dove, G. *et al.* (2020) Words have a weight: language as a source of inner grounding and flexibility in abstract concepts. *Psychol. Res.* 8, 2451–2467
26. Lupyan, G. *et al.* (2020) Effects of language on visual perception. *Trends Cogn. Sci.* 24, 930–944
27. Wingfield, C. and Connell, L. (2022) Understanding the role of linguistic distributional knowledge in cognition. *Lang. Cogn. Neurosci.* 37, 1220–1270
28. Corley, M. *et al.* (2011) Error biases in inner and overt speech: evidence from tongue twisters. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 162–175
29. Watson, J.B. (1913) Psychology as the behaviorist views it. *Psychol. Rev.* 20, 158–177
30. Bermúdez, J.L. (2018) Inner speech, determinacy, and thinking consciously about thoughts. In *Inner Speech: New Voices* (Langland-Hassan, P. and Vicente, A., eds), pp. 199–220, Oxford University Press
31. Nalborczyk, L. *et al.* (2020) Can we decode phonetic features in inner speech using surface electromyography? *PLoS ONE* 15, e0233282
32. Moffatt, J. *et al.* (2020) Inner experience differs in rumination and distraction without a change in electromyographical correlates of inner speech. *PLoS ONE* 15, e0238920
33. Fini, C. *et al.* (2022) Articulatory suppression delays processing of abstract words: the role of inner speech. *Q. J. Exp. Psychol.* 75, 1343–1354
34. Alderson-Day, B. *et al.* (2018) The varieties of inner speech questionnaire – revised (VISQ-R): replicating and refining links between inner speech and psychopathology. *Conscious. Cogn.* 65, 48–58
35. Grandchamp, R. *et al.* (2019) The ConDialInt model: condensation, dialogality, and intentionality dimensions of inner speech within a hierarchical predictive control framework. *Front. Psychol.* 10, 2019
36. Sabag, T. *et al.* (2023) The psychometric properties of the varieties of inner speech questionnaire-revised in Hebrew. *Front. Psychol.* 13, 1092223
37. Brysbaert, M. and Vantieghem, A. (2023) No correlation between articulation speed and silent reading rate when adults read short texts. *Psychol. Belg.* 63, 82–91
38. Reddy, M. (1979) The conduit metaphor – a case of frame conflict in our language about languate. In *Metaphor and Thought* (Ortony, A., ed.), pp. 284–324, Cambridge University Press
39. Deamer, F. (2021) Why do we talk to ourselves? *Rev. Phil. Psychol.* 12, 425–433
40. Tian, X. *et al.* (2016) Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex* 77, 1–12
41. Lu, L. *et al.* (2023) Common and distinct neural representations of imagined and perceived speech. *Cereb. Cortex* 33, 6486–6493
42. Nalborczyk, L. *et al.* (2021) Distinct neural mechanisms support inner speaking and inner hearing. *PsyArXiv* Published online November 18, 2021. https://psyarxiv.com/x7emk/
43. Aleman, A. *et al.* (2005) The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words. *Cereb. Cortex* 15, 221–228
44. Clark, A. (2016) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford University Press
45. Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204
46. Wilkinson, S. and Fernyhough, C. (2017) Auditory verbal hallucinations and inner speech: a predictive processing perspective. In *Before Consciousness: In Search of the Fundamentals of Mind* (Radman, Z., ed.), pp. 285–304, Imprint Academic
47. Jack, B.N. *et al.* (2019) Inner speech is accompanied by a temporally-precise and content-specific corollary discharge. *NeuroImage* 198, 170–180
48. Pratts, J. *et al.* (2023) Bridging phenomenology and neural mechanisms of inner speech: ALE meta-analysis on egocentricity and intentionality in a dual-mechanistic framework. *OSFpreprints* Published online March 31, 2023. https://osf.io/gaum4/
49. Lupyan, G. (2016) The centrality of language in human cognition. *Lang. Learn.* 66, 516–553
50. Roebuck, H. and Lupyan, G. (2020) The internal representations auestionnaire: measuring modes of thinking. *Behav. Res.* 52, 2053–2070
51. Gade, M. and Paelecke, M. (2019) Talking matters – evaluative and motivational inner speech use predicts performance in conflict tasks. *Sci. Rep.* 9, 9531
52. Gade, M. *et al.* (2020) Simon says – on the influence of stimulus arrangement, stimulus material and inner speech habits on the Simon effect. *J. Exp. Psychol. Learn. Mem. Cogn.* 46, 1349
53. Nedergaard, J. *et al.* (2023) 'Stay focused!': the role of inner speech in maintaining attention during a boring task. *J. Exp. Psychol. Hum. Percept. Perform.* 49, 451
54. Galanis, E. *et al.* (2022) Effects of a strategic self-talk intervention on attention functions. *Int. J. Sport Exerc. Psychol.* 20, 1368–1382
55. Furman, C.R. *et al.* (2020) Distanced self-talk enhances goal pursuit to eat healthier. *Clin. Psychol. Sci.* 8, 366–373
56. Orvell, A. *et al.* (2021) Does distanced self-talk facilitate emotion regulation across a range of emotionally intense experiences? *Clin. Psychol. Sci.* 9, 68–78

57. Gainsburg, I. and Kross, E. (2020) Distanced self-talk changes how people conceptualize the self. *J. Exp. Soc. Psychol.* 88, 103969
58. Granato, G. *et al.* (2020) A computational model of language functions in flexible goal-directed behaviour. *Sci. Rep.* 10, 21623
59. Granato, G. *et al.* (2022) A computational model of inner speech supporting flexible goal-directed behaviour in autism. *Sci. Rep.* 12, 14198
60. Fernyhough, C. (2016) *The voices within: the history and science of how we talk to ourselves*, Basic Books, New York
61. de Rooij, A. (2022) Varieties of inner speech and creative potential. *Imagi. Cogn. Pers* 460–489
62. Mitsuhashi, S. *et al.* (2018) Role of inner speech on the Luria hand test. *Cogent Psychol.* 5, 1449485
63. Banks, B. and Connell, L. (2023) Access to inner language enhances memory for event. *PsyArXiv* Published online Feburary 1, 2023. https://doi.org/10.31234/osf.io/64udy
64. He, H. *et al.* (2019) Language and color perception: evidence from Mongolian and Chinese speakers. *Front. Psychol.* 10, 551
65. Gerwien, J. *et al.* (2022) What is the interference in 'verbal interference'? *Acta Psychol.* 230, 103774
66. Croijmans, I. *et al.* (2021) Wine experts' recognition of wine odors is not verbally mediated. *J. Exp. Psychol. Gen.* 150, 545
67. Borghi, A.M. *et al.* (2018) Abstract concepts, language and sociality: from acquisition to inner speech. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373, 20170134
68. Borghi, A.M. *et al.* (2019) Words as social tools: language, sociality and inner grounding in abstract concepts. *Phys Life Rev* 29, 120–153
69. Barca, L. *et al.* (2020) Overusing the pacifier during infancy sets a footprint on abstract words processing. *J. Child Lang.* 47, 1084–1099
70. Borghi, A.M. and Fernyhough, C. (1870) Concepts, abstractness and inner speech. *Philos. Trans. R. Soc. B* 2023, 20210371
71. Borghi, A.M. (2022) Concepts for which we need others more: the case of abstract concepts. *Curr. Dir. Psychol. Sci.* 31, 238–246
72. Morin, A. (2018) The self-reflective functions of inner speech: Thirteen years later. In *Inner Speech: New Voices* (Langland-Hassan, P. and Vicente, A., eds), pp. 276–298, Oxford University Press
73. Morin, A. and Racy, F. (2022) Frequency, content, and functions of self-reported inner speech in young adults: a synthesis. In *Inner Speech, Culture & Education* (Fossa, P., ed.), pp. 147–170, Springer
74. Mead, G.H. (1934) *Mind, Self, and Society from the Standpoint of a Social Behaviorist*, University of Chicago Press
75. Morin, A. (2022) Self-reported inner speech illuminates the frequency and content of self-as-subject and self-as-object experiences. *Psychol. Conscious. Theory Res. Pract.* 9, 93
76. Kompa, N.A. and Mueller, J.L. (2022) Inner speech as a cognitive tool – or what is the point of talking to oneself? *Philos. Psychol.* 1–24
77. Fernyhough, C. (2004) Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. *New Ideas Psychol.* 22, 49–68
78. Verhaeghen, P. and Mirabito, G. (2021) When you are talking to yourself, is anybody listening? The relationship between inner speech, self-awareness, wellbeing, and multiple aspects of self-regulation. *Int. J. Pers. Psychol.* 7, 8–24
79. Yan, C. *et al.* (2022) Intra-agent speech permits zero-shot task acquisition. *Adv. Neural Inf. Process Syst.* 32, 2423–2438
80. Huang, W. (2022) Inner monologue: embodied reasoning through planning with language models. In *Proceedings of the 6th Conference on Robot Learning* (Li, K. *et al.*, eds), pp. 1769–1782, ML Research Press
81. Chella, A. *et al.* (2020) Developing self-awareness in robots via inner speech. *Front. Robot. AI* 7, 16
82. Pipitone, A. and Chella, A. (2021) Robot passes the mirror test by inner speech. *Robot. Auton. Syst.* 144, 103838
83. Simistira Liwicki, F. *et al.* (2022) Rethinking the methods and algorithms for inner speech decoding and making them reproducible. *Neuro Sci.* 3, 226–244
84. Macías-Macías, J.M. *et al.* (2023) Interpretation of a deep analysis of speech imagery features extracted by a capsule neural network. *Comput. Biol. Med.* 159, 106909
85. Yao, B. *et al.* (2021) Reading direct speech quotes increases theta phase-locking: evidence for cortical tracking of inner speech? *NeuroImage* 239, 118313
86. Assaneo, M.F. and Poeppel, D. (2018) The coupling between auditory and motor cortices is rate-restricted: evidence for an intrinsic speech-motor rhythm. *Sci. Adv.* 4, eaao3842
87. Tang, J. *et al.* (2023) Semantic reconstruction of continuous language from non-invasive brain recordings. *Nat. Neurosci.* 26, 858–866
88. Martin, S. *et al.* (2018) Decoding inner speech using electrocorticography: progress and challenges toward a speech prosthesis. *Front. Neurosci.* 12, 422
89. Wandelt, S.K. *et al.* (2022) Online internal speech decoding from single neurons in a human participant. *MedRxiv* Published online November 5, 2022. https://doi.org/10.1101/2022.11.02.22281775
90. Simistira Liwicki, F. *et al.* (2023) Bimodal electroencephalography-functional magnetic resonance imaging dataset for inner-speech recognition. *Sci. Data* 10, 378
91. Kapur, A. *et al.* (2018) AlterEgo: a personalized wearable silent speech interface. In *23rd International Conference on Intelligent User Interfaces*, pp. 43–53, Association for Computing Machinery
92. Langland-Hassan, P., Vicente, A., eds (2018) *Inner Speech: New Voices*, Oxford University Press
93. Cole, M. *et al.*, eds (1978) *L.S. Vygotsky Mind in Society: The Development of Higher Mental Processes*, Harvard University Press
94. Fernyhough, C. (2010) Vygotsky, Luria, and the social brain. In *Self and Social Regulation: Social Interaction and the Development of Social Understanding and Executive Functions* (Sokol, B. *et al.*, eds), pp. 56–79, Oxford University Press
95. Geva, S. and Fernyhough, C. (2019) A penny for your thoughts: children's inner speech and its neuro-development. *Front. Psychol.* 10, 1708
96. Fernyhough, C. *et al.* (2019) Imaginary companions, inner speech, and auditory verbal hallucinations: what are the relations? *Front. Psychol.* 10, 1665
97. Honeycutt, J.M. (2020) Imagined interactions and inner speech. *Imagin. Cogn. Pers.* 39, 386–396
98. Davis, P.E. *et al.* (2013) Individual differences in children's private speech: the role of imaginary companions. *J. Exp. Child Psychol.* 116, 561–571
99. McCarthy-Jones, S. and Fernyhough, C. (2011) The varieties of inner speech: links between quality of inner speech and psychopathological variables in a sample of young adults. *Conscious. Cogn.* 20, 1586–1593
100. Perona-Garcelán, S. *et al.* (2012) Dissociative experiences as mediators between childhood trauma and auditory hallucinations. *J. Trauma. Stress.* 25, 323–329
101. Alderson-Day, B. and Fernyhough, C. (2014) More than one voice: investigating the phenomenological properties of inner speech requires a variety of methods. *Conscious. Cogn.* 24, 113–114
102. Bernini, M. and Fernyhough, C. (2022) Resampling (narrative) stream of consciousness: mind wandering, inner speech, and reading as reversed introspection. *Mod. Fict. Stud.* 68, 639–667
103. Latinjak, A.T. *et al.* (2019) Speaking clearly... 10 years on: the case for an integrative perspective of self-talk in sport. *Sport Exerc. Perform. Psychol.* 8, 353
104. Hardy, J. *et al.* (2018) Reflections on the maturing research literature of self-talk in sport: contextualizing the special issue. *Sport Psychol.* 32, 1–8
105. Abdoli, B. *et al.* (2018) A closer look at how self-talk influences skilled basketball performance. *Sport Psychol.* 32, 9–15
106. Latinjak, A.T. *et al.* (2019) Goal-directed self-talk used to self-regulate in male basketball competitions. *J. Sports Sci.* 37 (12), 1429–1433
107. Bellomo, E. *et al.* (2020) Mind and body: psychophysiological profiles of instructional and motivational self-talk. *Psychophysiology* 57, e13586

108. Fritsch, J. *et al.* (2022) Self-talk and emotions in tennis players during competitive matches. *J. Appl. Sport Psychol.* 34, 518–538
109. Nedergaard, J. *et al.* (2021) Valence, form, and content of self-talk predict sport type and level of performance. *Conscious. Cogn.* 89, 103102
110. Thibodeaux, J. and Winsler, A. (2018) What do youth tennis athletes say to themselves? Observed and self-reported self-talk on the court. *Psychol. Sport Exerc.* 38, 126–136
111. Nedergaard, J. *et al.* (2023) Mind over body: interfering with the inner voice is detrimental to endurance performance. *Psychol. Sport Exerc.* 68, 102472
112. Hatzigeorgiadis, A. *et al.* (2018) Beat the heat: effects of a motivational self-talk intervention on endurance performance. *J. Appl. Sport Psychol.* 30, 388–401
113. Basset, F.A. *et al.* (2022) Type of self-talk matters: its effects on perceived exertion, cardiorespiratory, and cortisol responses during an iso-metabolic endurance exercise. *Psychophysiology* 59, e13980
114. Brinthaupt, T.M. (2019) Individual differences in self-talk frequency: social isolation and cognitive disruption. *Front. Psychol.* 10, 1088
115. Pasternak, K. and Oleś, P. (2023) Self-talk functions in travelers. *J. Constr. Psychol.* 1–9
116. Oleś, P.K. *et al.* (2020) Types of inner dialogues and functions of self-talk: comparisons and implications. *Front. Psychol.* 11, 227
117. Nedergaard, J. and Lupyan, G. (2023) Not everyone has an inner voice: behavioral consequences of anendophasia. In *Proceedings of the 45th Annual Meeting of the Cognitive Science Society* (Goldwater, M. *et al.*, eds), pp. 617–624, Cognitive Science Society
118. Vissers, CThWM *et al.* (2020) The emergence of inner speech and its measurement in atypically developing children. *Front. Psychol.* 11, 279
119. Siedler, A. *et al.* (2019) Self-reported individual differences in inner speech (internal monologue and dialogue) in adolescents with social (pragmatic) communication disorder. *Acta Neuropsychol.* 17, 39–53
120. Albein-Urios, N. *et al.* (2021) Inner speech moderates the relationship between autism spectrum traits and emotion regulation. *J. Autism Dev. Disord.* 51, 3322–3330
121. Petrolini, V. *et al.* (2020) The role of inner speech in executive functioning tasks: schizophrenia with auditory verbal hallucinations and autistic spectrum conditions as case studies. *Front. Psychol.* 11, 572035
122. Mahfoud, D. *et al.* (2023) The moderating effect of cognitive impairment on the relationship between inner speech and auditory verbal hallucinations among chronic patients with schizophrenia. *BMC Psychiatry* 23, 431
123. Sierpowska, J. *et al.* (2020) The black box of global aphasia: neuroanatomical underpinnings of remission from acute global aphasia with preserved inner language function. *Cortex* 130, 340–350
124. Fama, M.E. *et al.* (2019) Self-reported inner speech relates to phonological retrieval ability in people with aphasia. *Conscious. Cogn.* 71, 18–29
125. Luria, A.R. (1965) L.S. Vygotsky and the problem of localization of functions. *Neuropsychologia* 3, 387–392
126. Fernyhough, C. (2019) Modality-general and modality-specific processes in hallucinations. *Psychol. Med.* 49, 2639–2645
127. Simons, J.S. *et al.* (2022) Brain mechanisms underlying the subjective experience of remembering. *Annu. Rev. Psychol.* 73, 159–186
128. Zimmermann, K. and Brugger, P. (2013) Signed soliloquy: visible private speech. *J. Deaf. Stud. Deaf. Educ.* 18, 261–270
129. Al-Namlah, A.S. *et al.* (2006) Sociocultural influences on the development of verbal mediation: private speech and phonological recoding in Saudi Arabian and British samples. *Dev. Psychol.* 42, 117–131
130. Resnik, P. (2021) Multilinguals' use of L1 and L2 inner speech. *Int. J. Biling. Educ. Biling.* 24, 72–90