What the neurocognitive study of inner language reveals about our inner space

Hélène Lœvenbruck

Université Grenoble Alpes, CNRS, Laboratoire de Psychologie et NeuroCognition (LPNC),

UMR 5105, 38000, Grenoble France

Abstract

Our inner space is furnished, and sometimes even stuffed, with verbal material. The nature of inner language has long been under the careful scrutiny of scholars, philosophers and writers, through the practice of introspection. The use of recent experimental methods in the field of cognitive neuroscience provides a new window of insight into the format, properties, qualities and mechanisms of inner language. Gathering findings from introspection and empirical works, this article first assesses the proportion of language in our inner space. Several variants of inner language are then described, including wilful vs spontaneous instances, condensed vs expanded forms, silent vocalisation during reading or writing, contained vs ruminative occurrences, and self-controlled vs hallucinatory cases. The nature of these variants and their embodied multisensory qualities are examined. Finally, a neurocognitive model of the production of inner language is drawn, in the framework of predictive control, speculating on the neural mechanisms that underlie one of the most significant components of our inner space.

Keywords: Inner language, Verbal mind wandering, Rumination, Hallucination, Sensorimotor representation, Embodiment, Predictive Control

Introduction

Our inner space is furnished with verbal material which contributes to enriching our inner life. Our internal voice plays a central role in self-awareness, helping us to remember our past, to imagine and plan our future, to interpret our present environment, to get a better control of situations, to solve problems, to encourage, comfort or regulate ourselves. Engaging in mental verbalisation shapes our inner existence and is instrumental in the maintenance of a coherent self-narrative.

The claim that our mental space contains inner verbalisation can be traced back to Ancient Egyptian sages and Ancient Greek philosophers. Early Medieval scholars were inspired from these ancient views on inner experience. Augustine's *Confessions*, which appeared in 397-398, are considered as the first book using a subjective tone and focusing on inner experience. In the *Confessions* and in many of Augustine's later works, it seems that language invades the author's inner space.

Since then, inner language has been under the scrutiny of philosophers, writers, poets, filmmakers, artists, literary scholars, psychoanalysts, and linguists, through the practice of careful introspection and reflection. Their investigations suggest that silent self-talk and inner dialogues or conversations take an important part of our inner space. Inner language is often reported as pervasive or even ubiquitous. The French philosopher, psychologist and epistemologist Victor Egger for instance claimed:

A tout instant, l'âme parle intérieurement sa pensée. Ce fait, méconnu par la plupart des psychologues, est un des éléments les plus importants de notre existence. Il accompagne la presque totalité de nos actes ; la série des mots intérieurs forme une succession presque continue, parallèle à la succession des autres faits psychiques ; à elle seule elle retient donc une partie considérable de la conscience de chacun de nous. [1] (Egger, 1881, p. 1).

John Locke (1970) similarly asserted: "it is in constant use, accompanying many language-related activities such as writing, silent reading, learning, thinking, listening and,

---

[1] At every moment, the soul is speaking its thought internally. This fact, ill-recognized by many psychologists, is one of the most important elements of our existence. It accompanies nearly all of our acts; the series of interior words forms a nearly continuous succession, in parallel with the succession of other psychic facts; it thus retains, in itself, a considerable part of our consciousness.

possibly, dreaming." (Locke, 1970, p. 7). Such a stance is also taken by the linguist Gabriel Bergounioux (2001, p. 107):

Pas d'activité vigile qui ne soit accompagnée d'une sonorisation intériorisée, fût-elle réduite aux inepties de l'avant-sommeil, aux remembrances du vieillard idiot, à un ressassement ou une ritournelle, et pas non plus d'activité onirique.[2]

A similarly extreme position is taken by Baars (2003) who claimed that "Overt speech takes up perhaps a tenth of the waking day; but inner speech goes on all the time". Even Chomsky affirmed: "Now let us take language. What is its characteristic use? Well, probably 99.9% of its use is internal to the mind. You can't go a minute without talking to yourself. It takes an incredible act of will not to talk to yourself". (Chomsky, 2012, p. 11).

Some of these introspective accounts can be examined, tested and complemented using recent experimental methods and technology developed in psychology and cognitive neuroscience. Findings from these latter fields may provide a new window of insight into the format, properties, qualities and mechanisms of inner language and may allow us to better describe what our inner space consists of.

In this article, introspective views of inner language are juxtaposed with empirical data, many of which are reviewed elsewhere, e.g. in Perrone-Bertolotti et al. (2014), Alderson-Day & Fernyhough (2015) or Lœvenbruck et al. (2018).

This article first assesses the significance of language in our inner space and the proportion it takes. Several variants of inner language are then described, including wilful vs spontaneous instances, condensed vs expanded forms, silent vocalisation during reading or writing, contained vs ruminative occurrences, and self-controlled vs hallucinatory cases. It is then shown that some variants of inner language have multisensory qualities, with the presence of auditory, somatosensory and visual elements. It is argued that wilful versions of inner language may recruit also motor processes. Finally, a neurocognitive model of the production of inner language is drawn, in the framework of predictive control, speculating on the neural mechanisms that underlie one of the most significant components of our inner

---

[2] No wakeful activity is unaccompanied by some interiorised sound, be it pre-sleep nonsense, a foolish old man's reminiscing, brooding or a music earworm, and no oneiric activity either.

space.

1. The importance of inner language in our inner space

1.1 Evidence for the occurrence of inner language

Inner language is claimed to be an essential part of our inner space. But is our inner space filled with language, with all of its clothes, or rather by meaning? Can we find evidence that this mental activity that we refer to as inner language is indeed spelled out in a linguistic form, that it does use ordinary words and syntax? Augustine himself made a distinction between the *verbum in corde,* the interior verbalisation, and the *locutio*, the exteriorised oral form (Cary, 2011; Panaccio, 2014). If the latter obviously uses words and is expressed in a given language, the former, according to Augustine (in *De Trinitate*) is universal, it precedes overt language production, and it is not associated with any particular language. As we will see below (in 1.2) several introspective works and language impairment case reports suggest indeed that cognitive activity can occur without using natural language. Yet, introspectively, we can sometimes hear a little voice in our head. Is these instances, is our little voice expressed in a given language, is it shaped by the language(s) we use?

Examining inner experience in bilingual and multilingual speakers can contribute to better qualifying our inner space. In her book entitled "The Bilingual Mind", Pavlenko (2014) dedicated a chapter to bilingualism and inner speech, in which she reviews several studies of inner language use in bilingual and multilingual speakers. Another survey by Dewaele (2015) provides further interesting data. These reviews show that the age of acquisition of the second language (L2) (or of the third, fourth, fifth, or more languages, LX) is a strong factor in determining which language is dominant in the inner speech of participants. Another factor is the linguistic context of the cognitive event entertained in the inner speech. Autobiographical memory retrievals tend to be mentally uttered in the language used when the event took place. Context of acquisition and socialisation are factors that facilitate the shift from L1 to LX in inner speech. Acquiring the LX in a naturalistic environment (rather than an instructed setting) will increase its use in inner speech. Self-perceived proficiency also influences the language choice in inner speech and mental activities (mental calculation, reasoning). Higher frequency of social speech in the LX also increases the likelihood of its use

for inner speech. Inner LX use is also proportional to the size of the speaker's social network in LX. The dominant language in inner speech is also predicted by length of residence in the new country and the language predominantly used in overt speech. In a very recent study, Resnik (2018) has found that in addition to these factors (high frequency of L2 use, naturalistic exposure to L2, high self-reported proficiency in L2), a high bilingualism index and the overall number of languages known all contributed to boost L2 use in inner speech.

As shown in the growing number of studies on inner speech in multilingual speakers, our inner space does seem to contain verbal material that is not abstract, or purely semantic, but expressed in given natural languages, with full syntactic and lexical clothing. Our inner space incorporates our cultural, social and linguistic environment.

1.2 Other components of inner space: images, emotions, sensations, unsymbolised elements

It seems safe to conclude that our inner space is furnished with an important amount of verbal material. But how much? Although the contention that inner language is pervasive is widely held, quantitative descriptions of its occurrence in the general population are more nuanced. In a thought sampling study, Klinger & Cox (1987) had students carry a beeper for up to seven days and describe properties of their mental content at each random beeper signal. They found that 51% of thought samples contained some interior monologue. Even lower occurrences are found when using more careful inner experience sampling. Hurlburt has created a method, called Descriptive Experience Sampling (DES), designed to obtain more accurate accounts of inner experience (Hurlburt, 1993; 2011). Traditional questionnaires on inner experience are biased, since participants use language to describe their experience, which places them in a verbal thinking mode and leads them to overestimate the amount of inner speech. In addition, questionnaires contain pre-defined questions, which can orient the participants' descriptions. Instead, DES does not specify in advance what characteristics to explore. After having carried the beeper for a day and having jotted down notes about their inner experience before each beep, subjects participate in an "expositional interview" with the investigators, during which they are guided to describe their inner experience with the highest possible fidelity. This sequence of beeped reports and detailed interviewing is repeated, leading to increasing skill in the reporting. Using DES with

hundreds of people for more than thirty years, Hurlburt and colleagues have routinely met participants who reported no moments of inner speaking at all. They conclude that the frequency of inner speaking displays a large inter-individual variability, ranging from about zero to nearly 100%, with a mean of about 26% of sampled moments (Heavey & Hurlburt, 2008). They suggest that the rest of our inner experience consists of four other components: inner seeing, feeling, sensory awareness and unsymbolised thinking. Inner seeing is visually imagining something that is not present. Feeling corresponds to "affective experiences, such as sadness, happiness, humor, anxiety, fear, joy, nervousness, anger, embarrassment, etc". Sensory awareness is paying attention to a sensory aspect of the environment (cold, wind, hunger). Unsymbolised thinking is "thinking a particular definite thought without the awareness of that thought being conveyed in words, images or any other symbols". It is argued to be a distinct phenomenon, not a precursor to any other phenomenon, and to be clearly articulated and specific. Even if this classification can be debated and if some authors reject the existence of unsymbolised thinking (Carruthers, 2009), it has been acknowledged that non-verbal thinking may exist. Hurlburt's notion of unsymbolised thinking is for instance reminiscent of Paivio's (1990) Dual Coding Theory, according to which stimuli can be coded and mentally manipulated in a verbal or visual mode.

Hurlburt is not alone in defending that our inner space is not entirely filled with language. Other authors have estimated that approximately one quarter of our conscious waking life consists of inner language (e.g. Uttl, Morin, Faulds, Hall & Wilson, 2012). Furthermore, inner language, when it occurs, can be intertwined with non-linguistic fragments. According to Wiley (2014), words can be combined with "sounds, numbers, visuals, colors, tastes and odors, tactile feelings, kinesthetics and emotions." These elements can be placed into syntactical slots, producing inner utterances that are only partially verbal (e.g. the words "I'd like" followed by the image of a hamburger).

Our inner space is therefore not fully occupied by language: on average, approximately a quarter of our inner space consists of inner verbalisation, the rest is made up of images, emotions, sensations and unsymbolised elements.

2. Various instances of inner language

2.1 Wilful inner language vs verbal mind wandering

Inner language manifests in various ways. We often deliberately engage in short instances of inner speech, for instance when we count, make a list, or schedule our weekly objectives. We can engage in longer sophisticated inner talk, carried out in full sentences, when we prepare a lecture, think hard about an argument, or imagine possible future conversations. These short and long instances of inner language can be referred to as "wilful" or "deliberate" inner language. But sometimes our internal monologue is less wilful, less active and "more passive" (Bonald, cited in Egger, 1881). The more passive form of inner language was seized by Plato who drew a distinction between opinion and fancy: "[thought] is a silent inner conversation of the soul with itself. [...] when this arises in the soul silently by way of thought, can you give it any other name than opinion? [...] And when such a condition is brought about in anyone, not independently, but through sensation, can it properly be called anything but seeming, or fancy?" (Plato, *Sophist*, 263e-264a). This fanciful inner language has been referred to as "verbal mind wandering" (Perrone-Bertolotti et al., 2014), and often occurs during "resting states" (mind wandering can also take a non-verbal form, such as visual imagery, hence the adjective "verbal"). Verbal mind wandering consists of flowing, spontaneous, unconstrained, external-stimulus-independent verbal thoughts. Mind wandering (MW) has been the focus of several recent neurocognitive studies. Smallwood and Schooler (2015) have recognized several beneficial outcomes of mind wandering (see also Mooneyham and Schooler, 2013, Schooler et al., 2014, Smallwood and Andrews-Hanna, 2013). According to them, the first benefit of MW is prospection. The thoughts that occur during MW are often future-oriented, helping to improve daily life. A second beneficial outcome is creativity. MW seems to be associated with the capacity to generate novel, creative thoughts. A third value is to add meaning to personal experience. By engaging in mental time travel, MW enables people to integrate past and future events into a meaningful life narrative. Other beneficial outcomes include mental breaks to relieve boredom from monotonous activities, and day-dreaming to prepare for potential obstacles or threats. It can be speculated that among the many forms of VM, the verbal one (unwilful inner language) is the most likely to play the first three of these roles: prospection, creativity

and meaning.

2.2 Condensed vs expanded inner language

In addition to the various intentional degrees of inner language, various degrees of unfolding have been identified. Some variants are condensed relative to other fully formed, or expanded, versions (e.g. Fernyhough, 2004; Alderson-Day and Fernyhough, 2015 or Geva, Jones, Crinion, Price, Baron & Warburton, 2011b). Condensation seems to operate at different levels: articulation, phonology, lexicon and syntax. Introspective accounts of condensation are abundant. It has been argued for instance that because inner language is directed to oneself, it is shortened compared to overt speech addressed to an external listener. Egger (1881, p. 69-71) was the first to clearly list physiological and social constraints for why inner language may be shorter. First, he argued that we cannot overtly articulate as quickly; the speed of our tongue movements being physiologically limited ("*à parler trop vite la langue s'embarrasse*"[3]). Moreover, when we speak aloud, we need to take breath between fragments of speech, as speech only occurs during expiratory phases. Because it is not subjected to these physiological constraints, inner speech is accelerated compared with overt speech. Secondly, Egger noted social constraints. When we speak to someone, we need to articulate clearly and slowly, in order to be understood. When we use inner speech, this social constraint can be abandoned and our articulation can be more "sketchy". Furthermore, according to Egger (1881, p.71), some expressions that we use mentally bear meanings that are explicit only to ourselves. In order for them to be understood, we would need to provide contextual information and to replace these expressions by a detailed and explicit discourse.

Therefore, inner speech is not only physically shortened with respect to overt speech, it can also be syntactically condensed, or left elliptical. Vygotsky further developed this notion (Vygotsky, 1934/1986). His theory is based on introspection, and on examination of children's private speech (or egocentric speech), in which children talk to themselves aloud, and which he claimed to be a precursor of adult inner speech. He asserted that important words or grammatical affixes may be dropped. According to him, the syntax of inner speech is "predicated," in the sense that only necessary information is supplied. The subject, and

---

[3] "when speaking too quickly the tongue gets tangled up"

even the verb, might be omitted. Bergounioux follows this notion of abbreviation, condensation and predication, providing detailed linguistic descriptions of the phenomenon. According to Bergounioux (2001, p. 120), "l'endophasie ne semble différer de la parole explicite ni par sa grammaire ni par son lexique, à la réserve d'un emploi généralisé de l'asyndète et de l'anaphore, et d'une surreprésentation de la prédicativité."[4]

Examples of such linguistic operations can be found in literary or artistic works, typically those associated with the "monologue intérieur" (Dujardin, 1887, 1931) or "stream of consciousness" movement, initiated by Dujardin (Smadja, in press). Although Dujardin depicted internal monologue as swarming with syntactically expanded sentences, later renditions – such as Molly Bloom's monologue in Joyce's *Ulysses*, the disjointed monologue in Samuel Beckett's *The Unnamable* or the human monologues overheard by angels in Wim Wender's *Wings of Desire* – are closer to the introspective descriptions of Egger, Vygotsky or Bergounioux, in that they are more fragmentary, abbreviated, predicated and condensed, at both the syntactic and lexical levels.

Empirical grounding for the condensed quality of inner speech at the syntactic and lexical levels can be found in an astute study of the rate of spontaneous covert speech production (Korba, 1990). Participants were asked to mentally solve a series of short verbal problems. They reported the elliptical inner speech used to solve each problem, which gave an estimation of the number of words used in this type of mentation. They were then instructed to expand the same volume of words into a full ostensive statement of their internal problem-solving strategies, which provided an extended word count. The extended word count represented an equivalent rate of speech in excess of 4000 words per minute, which cannot possibly be reached in overt mode. These findings are in favour of the introspective claim that inner verbalisation is condensed with respect to overt public speech, at least at the syntactic and lexical levels.

In line with Egger's claim that inner speech is less articulated, it has further been suggested that the phonological form of inner words may itself be abbreviated. Several soviet psychologists suggested that inner speech is phonologically reduced, with many phonemes

---

[4] "endophasia does not seem to differ from explicit speech, neither by its grammar, nor its lexicon, except perhaps in a generalised use of asyndeton, anaphora and an over-representation of predication"

being dropped, typically vowels, and only the word-initial sounds being clearly produced (e.g. Vygotsky, 1934/1986, p. 237, 244; Anan'ev cited in Sokolov, 1972, p. 50). This claim receives support from studies that show that word production is faster in inner than overt mode, even when lexical content is controlled (Marshall and Cartwright, 1978; Anderson, 1982). Marshall and Cartwright (1980) examined both word and sentence productions in a recitation task. They found that silent recitation was faster than overt recitation for lists of one- and three-syllable words as well as lists of grammatical and ungrammatical sentences. MacKay (1981) also examined sentence production. Participants were asked to produce identical sentences as rapidly as possible, either overtly or covertly. Results indicated that both internal and overt speech improved with practice and that overt sentence production took longer. These controlled studies imply that inner speech is abbreviated with respect to overt speech, even at the phonological level, in line with introspective views. This could suggest that some of the phonological or articulatory processes involved in overt speech are absent in covert mode. An alternative interpretation, described in Section 3.4, is that inner speech involves the same operations as overt speech but that the execution of articulator movements takes longer than their simulation.

Another source of information on the lower levels (phonological and articulatory) of covert production are error patterns. As explained by Oppenheim and Dell (2008), speech errors display a lexical bias, a bias towards creating words (e.g. REEF LEECH leading to LEAF REACH) rather than nonwords (e.g. WREATH LEAGUE leading to LEATH REEG), in both overt and covert speech modes. This bias has been interpreted by these authors as evidence for the spreading-activation model of Dell (1986), which posits an interactive flow of activation between phonological and lexical levels. The lexical bias suggests that inner speech activates not only conceptual but also lexical as well as phonological representations.

A second bias has been reported, called the "phonemic similarity effect," a tendency to exchange phonemes with common subphonemic articulatory features (e.g. REEF slips more often to LEAF, with /r/ and /l/ sharing many features such as voicing and approximant, than REEF to BEEF, with /r/ and /b/ only sharing voicing). This effect is explained by Oppenheim and Dell with reference to reciprocal activations between articulatory and phonological levels. Oppenheim and Dell (2008, 2010) have only found this effect in overt mode or with

inner speech accompanied with mouthing. This has led them to claim that inner speech is fully specified at the lexical and phonological levels, but that it is impoverished at the lower subphonemic (or articulatory) level. As we will see in Section 3.4, there are several empirical arguments against this claim, however, so the conclusion on the articulatory poverty of inner speech should be considered with caution.

To summarise, some variants of inner language have been claimed to be sketchy and impoverished at many levels, including syntactic, lexical, phonological and articulatory. This abbreviation bestows inner language an abstract quality that has led some researchers to consider it as an amodal phenomenon. MacKay (1992) stated that inner speech is nonarticulatory and nonauditory. According to him, articulatory movements "are irrelevant to inner speech. Even the lowest level units for inner speech are highly abstract." For MacKay, "[t]he seemingly auditory quality of our internal speech cannot be automatically attributed to events within an auditory or acoustic system, or even, as we will see, to *any* strictly sensory system." This strong stance is in line with Bergounioux' (2001) claim that "endophasia, phenomenologically speaking, is speech without a signal", i.e. without a "dépense d'énergie quantifiable et capturable" ("quantifiable and capturable energy expenditure"). This would implie that inner language is divorced from bodily experience and includes, at most, faded auditory representations.

As we will see in the following sections, inner language is not always condensed and some variants of inner language are in fact fully expanded. Neurocognitive data do not corroborate the strong claim that inner language is impoverished and lacks articulatory or auditory specification (see 3). Some instances of inner language, including mental sentence repetition, instructed mental sentence generation, silent reading, verbal working memory are dependent on perceptuo-motor processes and their operational details. These instances may be considered as embodied, involving physical processes that unfold over time and leading to the creation of articulatory and auditory percepts. They presumably integrate a variety of representations, from semantic concepts to articulatory features, via lexical items and phonological representations.

These seemingly opposite views on the condensed and expanded quality of inner language are not mutually exclusive, however. Fernyhough (2004) has suggested that inner

speech varies with cognitive and emotional conditions between these two forms. The expanded form can even be considered as an outcome of the condensed form, which itself can be construed as the conceptual message cast in a pre-verbal form, that involves lemmas[5], linearly ordered, but that does not yet have the full phonological (articulatory, acoustic) specification that expanded inner language has (see e.g. Levelt, 1989). A similar position is defended in Vicente & Martínez-Manrique (2016). Inner language can be defined as truncated overt verbalisation, but the level at which the production process is interrupted (abstract linguistic representation vs. articulatory specification) depends on which variant of inner language is at play. The less wilful variants of inner language, such as verbal mind wandering, might take on a more condensed format, whereas wilful forms might be more expanded.

Interestingly, the variant of inner language that is most studied in experimental psychology and cognitive neuroscience is the wilful form, because it can be examined in controlled settings, in a replicable fashion. But the variant that has been the focus of most introspective works by literary scholars, philosophers, artists, is unbidden interior monologue, a subset of verbal mind wandering (which also includes interior conversations). This variant is the most difficult to examine experimentally, as it is spontaneous. This could explain why introspective and empirical accounts sometimes differ. A few neuroimaging studies have endeavoured to compare wilful and spontaneous inner language, but they are far too rare for any conclusion to be drawn (Hurlburt, Alderson-Day, Kühn & Fernyhough, 2016; Grandchamp and colleagues, in preparation).

2.3   Inner language during reading

Another variant of inner language is the silent vocalisation that accompanies reading. According to Egger (1881), inner speech is in fact easiest to notice when one reads. Egger (1881) and Ballet (1886) suggested that during reading, one does not only mentally articulate the words that are read, but one can experience hearing them too. Several experimental psychology studies have examined this claim and have shown that reading does involve an inner voice. It has been shown for instance that silent reading is influenced by

---

[5] The term lemma in Levelt and colleagues' terminology refers to the word's syntax, see Levelt et al. (1999). It is different from the lexeme which denotes the word's phonological features and from the lexical concept which refers to the word's semantics.

pronunciation characteristics. Alexander & Nygaard (2008) have shown that silent reading of a text is influenced by the knowledge we have of its author's speaking rate. Participants take longer to silently read a text when they are told that it was written by a slow talker rather than a fast talker. Silent reading is also modulated by the reader's regional accent. Filik & Barber (2011) compared the eye movements of English participants with different regional accents who were reading limericks. Limericks are short poems in which the final word rhymes with the end words of the first two lines. The authors created limericks in which the final word would rhyme or not, depending on the regional accent. When the final word did not rhyme in the reader's accent, a disruption in the eye movement record was observed, compared to when it rhymed. This finding suggest that silent reading includes properties of the reader's own pronunciation habits.

The occurrence of inner speech during silent reading has also been confirmed by recent intracranial electroencephalography recordings. Perrone-Bertolotti and her colleagues (2012) measured activity in the temporal voice area (TVA) of epileptic patients. This region in auditory cortex is selectively activated during human voice perception (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000). The patients were instructed to silently read words. The results show that silent reading activates the TVA. Moreover, TVA activity was found to strongly increase when participants were reading attentively. This suggest that the inner voice heard during silent reading is not an automatic process, which would be triggered in response to any written word, but that it is modulated by attention.

It should be noted, however, that silent reading is not systematically associated with inner speech production, even when attention is high. Levine, Calvanio & Popovics (1982) have reported the case of a patient, who as a result of a stroke, became mute and was unable to speak covertly. He could not tell whether two words rhymed, which suggests he could not evoke the auditory representations associated with words. Yet his reading abilities remained intact. His visual imagery was strongly developed (he could make highly accurate drawings from memory). This case report therefore suggests that when visual imagery is proficient, then reading processes may sometimes bypass phonological mediation (covert pronouncing and inner voice hearing), directly linking the word's written form to its semantic content.

2.4 Inner language during writing

Writing is generally considered to involve an inner voice. In the film *Paterson*, Jim Jarmusch lets us follow Paterson, a bus driver who writes poetry. The film suggests that Paterson elaborates his poems mentally before writing them down. We can hear his inner voice as he composes each poem, and then as he writes the lines in his notebook. Studies in experimental psychology have shown that writing involves many processes, including idea generation, concept and word retrieval from semantic and lexical memory, syntactic processing and access to graphemic forms (letters in the words). A debated question is whether the transformation from lexical to graphemic forms recruits inner speech. According to the "phonological mediation hypothesis," spoken forms of words are retrieved before graphemic forms can be accessed. This hypothesis is supported by studies of brain-lesioned patients, which show that deficits in spoken language are associated with impairments in written language production (e.g. Luria, 1966).

An alternative view is that orthographic forms can be accessed from abstract lexical knowledge without phonological mediation. A few brain-lesioned patient studies have reported dissociations between writing and speaking impairments. The patient in the study by Levine et al. (1982) discussed above was deprived of inner speech but could still read and write. Rapp *et al.* (1997) presented the case of a stroke patient who suffered from speech deficits. He was often unable to provide the correct spoken name of an object, although he could write it. These cases therefore seem to argue against the phonological mediation hypothesis, in that writing can be achieved without spoken language mediation, when word production is impaired, due to a stroke. It cannot be ruled out, however, that the seemingly direct link between grapheme and meaning, was initially (before the stroke) mediated by covert speech and that the direct connection was gradually learned. The recent study of a child with congenital oral apraxia is more compelling (Cossu, 2003). Despite his inability to produce any articulation, this child had normal reading and writing skills. He presumably did not rely on a covert version of his own articulation to read and write. Therefore, articulatory mediation is not always necessary during writing. Nevertheless, since this child had preserved auditory capacities, we could argue that writing is still associated with auditory representations, and that the lexicon-grapheme transformation may well rely on auditory-phonological representations. Therefore, this study does not fully contradict the

phonological mediation hypothesis, even though it makes it weaker. The representations at play in this child are auditory at best, and not articulatory.

2.5   Contained vs ruminative inner language

Inner speech plays a central role in human consciousness at the interplay of language and thought (Morin, 2005) and is beneficial to many cognitive operations. It interacts with working memory to encode new material (Baddeley & Hitch, 1974). It is involved in remembering personal past episodes, including conversations, situations and emotions, i.e. it plays a role in autobiographical memories (Morin, 2012). It is used in future planning, in reasoning, in problem solving (Sokolov, 1972; Baldo et al., 2015), in cognitive control, executive function, cognitive flexibility (Emerson & Miyake, 2003), in consciousness, self-awareness, self-regulation, self-motivation (Morin, 2009), in self-encouraging and self-comforting (Pavlenko, 2014).

Inner language can sometimes play a detrimental role, however, when it becomes repetitive and negative. Dostoyevsky's *Crime and Punishment* provides illustrative descriptions of how ruminative forms of inner language may become intrusive, such as when the ex-student and future murderer Raskolnikov, sitting in a tavern, reflects upon the mysteries of chance and destiny: "A strange idea was pecking at his brain like a chicken in the egg, and very, very much absorbed him" (part 1, chapter 6).

Self-reflection, pondering about ourselves, our feelings, thoughts and behaviours, can contribute to clarifying the meaning of past and present experiences (Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). However, it can lead to unconstructive consequences when self-referent thoughts transform into verbal rumination, i.e. repetitive and self-critical inner speech (Watkins, 2008; Nalborczyk et al., 2017). It has been shown that rumination alters cognitive performance in depressed or dysphoric patients and that it can predict and exacerbate the maintenance of dysphoric or depressive states (Davis & Nolen-Hoeksema, 2000).

It still remains to be understood why excessive, negative inner speech impairs performance whereas more contained and positive inner speech improves cognitive performance.

2.6   Self-controlled vs hallucinatory forms

Another dysfunctional case of inner language are auditory verbal hallucinations. As argued by Fernyhough (1996, 2004) and Alderson-Day and colleagues (2016), inner language is often dialogic, mirroring the external experience of communication. We can have imaginary conversations and we can then hear the others' voices, their timbre, their pitch. When we do so, we usually know that these voices are self-generated and we do not mistake these imaginary voices for external voices. This is because we are endowed with a self-monitoring mechanism. It has been suggested that when this mechanism is defective, auditory verbal hallucination may occur.

Auditory verbal hallucination (AVH) or "hearing voices" can be considered as speech perceptions in the absence of any relevant external acoustic input. It affects 50-80% of the patients who suffer from schizophrenia (Nayani & David, 1996). Patients report hearing voices, which are often distressing and engender suffering and functional disability as well as social marginalisation (Franck, 2006). Auditory verbal hallucination is, however, a complex phenomenon with multiple forms and causes (Larøi & Woodward, 2007). It also occurs in non-psychiatric populations, and it is estimated that 4-10 % of the healthy population experience it (Linden et al., 2011).

Many theoretical models have been proposed to explain AVHs in schizophrenia (see, David, 2004, for a review). An influential model formulates AVHs as dysfunctions of the monitoring of inner speech (Feinberg, 1978; Frith, 1992). The model claims that due to a failure of the self-monitoring mechanism, the inner speech of the patient is not identified as self-generated and is experienced as coming from an external source.

However, voice hearers (patients with schizophrenia as well as healthy individuals) can also use inner language deliberately, without experiencing voices. Inner language in voice hearers is not always dysfunctional. Some researchers have focused on the distinction between AVH and inner speech in non-clinical hallucinators. Linden and colleagues (2011) have argued that the distinction is related to subjective control: AVH occurs spontaneously, while wilful inner speech occurs under volitional control. This claim is in line with studies by Rapin *et al*. (2012) and Lavigne *et al*. (2015) who suggest that the supervisory processes that are at play during willful inner speech can serve to normalise the activity in sensory cortex.

The absence of such processes could explain why hyperactivity in sensory cortex is observed in hallucinatory experience. It can therefore be speculated that wilful inner speech engages supervisory control that modulates sensory activity, whereas more spontaneous forms of inner language, deprived of supervisory and self-monitoring processes, may end up being attributed to external sources.

3. Various formats of inner language

3.1 Auditory sensations

Early introspective works (Egger, 1881, Ballet, 1886) have claimed that inner speech is endowed with auditory qualities. Egger (1881) wrote that "[l]a parole intérieure a l'apparence d'un son"[6] and that "[l]es caractères de la parole [rythme, hauteur, intensité, timbre] […] se retrouvent tous dans la parole intérieure"[7].

The concept of a mind's ear finds support in psycholinguistic data. The "Verbal Transformation Effect" (VTE) refers to the perceptual phenomenon in which listeners report hearing a new percept when an ambiguous stimulus is repeated rapidly (Warren, 1961). Rapid repetitions of the word "life", for example, produce a soundstream fully compatible with segmentations into "life" or "fly". Smith, Wilson, and Reisberg (1995) further examined the VTE, and found that it also occurs in a covert mode. In addition, they observed a reduction of the effect during auditory interference. These findings suggest that subjects rely on the mind's ear to detect transformations. The neural correlates of the VTE have been examined by Sato and colleagues (2004). Participants were asked to silently repeat pseudo-words. Active search for verbal transformation increased activity in several brain regions, including auditory cortex.

Findings of error detection during covert tongue-twister repetition also indicate that inner speech has auditory qualities that can be attended to. Several studies (see Dell & Oppenheim, 2015 for a review) have investigated error slips reports. They show that participants are able to attend to and report the "errors that they hear," like they do with slips produced in audible speech. This can be interpreted as a role for the mind's ear in inner

---

[6] "Inner speech has the appearance of a sound."
[7] "The characteristics of speech [rhythm, pitch, intensity, timbre] (…) are all found in inner speech. "

speech monitoring.

Further empirical arguments for the auditory nature of inner speech come from neuroimaging studies. Several fMRI (functional Magnetic Resonance Imaging) studies of covert speech production reveal auditory cortex, specifically superior temporal gyrus, activation (Perrone-Bertolotti et al., 2014 for a review). Although this activation is lesser than the one observed in overt speech, it implies that an auditory experience accompanies inner speech. In that line, an fMRI study by Lœvenbruck, Baciu, Segebarth and Abry (2005) suggested that covertly produced speech can include prosodic characteristics, with distinctive auditory features that correspond to objectively measurable cerebral correlates.

To sum up, behavioural and neuroimaging data suggest that auditory sensations are present during several variants of inner language.

3.2 Somatosensory sensations

The phenomenological intuition that inner language involves a voice that can be heard in the mind's ear is not controversial and meets with empirical findings. But other sensory qualities may be attributed to inner language, typically imaginary proprioceptive and tactile sensations. Taine (1870) himself was a precursor of that idea when he wrote: "A l'état normal nous pensons tout bas par des mots mentalement entendus, ou lus, ou prononcés, et ce qui est en nous c'est l'image de tels sons, de telles lettres, ou de telles *sensations musculaires et tactiles du gosier, de la langue et des lèvres*. (je souligne)"[8]. Paulhan (1886) wrote that lengthy verbal thinking can cause fatigue in articulatory muscles, which implies that inner speech involves somatosensory sensations. According to Lackner and Tuller (1979), overt speech errors can be detected by means of proprioceptive information on articulatory configurations as well as tactile information about labial or lingual contacts. It has been suggested that proprioceptive and tactile feedback play a role in speech motor control (Levelt, 1989; Postma, 2000). It can therefore be speculated that imagined proprioceptive and tactile feedback could be part of inner speech, just as imagined voice is. In addition to the mind's ear, the "mind's touch" should also be considered. Neuroimaging studies corroborate

---

[8] In the normal state, we silently think with words that are mentally heard, read or uttered, and what is inside of us is the image of certain sounds, of certain letters, or of *certain muscular and tactile sensations in the throat, tongue and lips*. (emphasis is mine)

this assumption. Several studies reviewed by Perrone-Bertolotti and colleagues (2014) show somatosensory cortex activation during tasks that involve inner speech.

3.3   Visual images

Introspection suggest that the "mind's eye" also plays a role in inner language. Paulhan (1886) claimed that inner speech may sometimes include visual images. By visual images he meant the form, shape and colour of the letters that compose written words. But other visual elements may also be included.

Recent works on inner verbalisation in deaf individuals suggest that it may contain visual elements related to articulation or sign. Bellugi, Klima, and Siple (1975) compared the properties of short term memory in normal hearing participants and deaf participants whose native language was American Sign Language (ASL). Lists of words were presented to the deaf participants in the visual modality as signs on a videotape. The same words were presented in the auditory modality on an audiotape to the hearing controls. The task was to recall the signed or spoken words and to write them in English orthography. The errors made by hearing subjects were mainly sound-based (e.g. "vote" misrecalled as "boat"). This suggests that hearing subjects had been encoding and remembering the words in terms of their phonological properties. In signing subjects, many substitution errors coincided with words that were visually (not auditorily) close to the target, such as "noon" replaced by "tree," which corresponds to a similar arm position in ASL. Other behavioural studies of verbal working memory in deaf signers similarly reflect a transfer from the auditory to the visual modality. Wilson and Emmorey (1998) observed a sign length effect in deaf users of ASL, analogous to the auditory word length effect in spoken language. Poorer memory performance was found for long signs compared to short signs, independently of the auditory word length. Manual suppression (repetitive movements of the hands) produced a drop in performance, just like articulatory suppression (repetitive syllable production) disrupts verbal working memory in hearing subjects. These studies suggest that sign language is stored in memory in terms of its gestural properties. Therefore, inner language in deaf signers presumably involves visual representations.

Gestures are not only used in the deaf population. They accompany speech in normal

hearers and play a fundamental role in thought and speech (De Ruiter, 2007). Gesture and speech are coordinated to form coherent multimodal messages. Moreover, speech is audiovisual: lip reading enhances speech comprehension when the acoustic signal is degraded by noise (Sumby & Pollack, 1954). Lip reading occurs even with nondegraded acoustic signals, such as in the McGurk effect (McGurk & MacDonald, 1976). This illusionary effect occurs when an auditory syllable (such as /ba/) is synchronously presented with the video of a face uttering a discrepant visual syllable (such as /ga/). Most participants report hearing a syllable corresponding to the fusion of the auditory and visual channels (/da/ or /ða/). Based on this audiovisual integration effect and other studies, it has been argued that auditory and visual speech information include common stages of processing (Nahorna, Berthommier, & Schwartz, 2015). It can therefore be assumed that visual information (facial and manual) may be involved in inner speech, even in hearing subjects. A preliminary work by Arnaud, Schwartz, Lœvenbruck, and Savariaux (2008) provides tentative suggestions that speakers can have visual representations of their own lip movements. More research is needed to confirm that inner language involves visual (labial, facial, manual, written) representations, even in the hearing population.

To sum up, inner verbalising appears to involve the reception of imaginary sensory signals, including auditory, somatosensory and visual elements, handled by the mind's ear, touch and eye. The format of inner language can therefore be described as multisensory.

3.4 Motor representation

In parallel with the sensory accounts, it has been suggested that inner speech requires motor processes. The earliest claims concerning the motor quality of inner speech probably date back to Erdmann (1851) and Geiger (1868), who, as cited by Stricker (1885), introspectively observed that inner speech is accompanied by feelings of tension in the speech musculature. Bain himself wrote in 1855: "When we recall the impression of a word or a sentence, if we do not speak it out, we feel the twitter of the organs just about to come to that point. The articulating parts, — the larynx, the tongue, the lips, — are all sensibly excited; a suppressed articulation is in fact the material of our recollection."

Stricker (1885, chapter II) designed a clever introspective exercise to experience this

orofacial activity. He hinted that, when one's mouth is positioned into the rounded shape required to pronounce the sound of an "o," if one tries to imagine uttering that of an "m," a slight contraction is felt in the lip muscles, as if one was pressing lips to pronounce the labial sound. Stricker (1885) claimed from several introspective exercises that inner speech is accompanied by sensations in the oral musculature similar to those driving the actual pronunciation of articulated sounds. He introduced the notion of motor representations associated with inner speech and speculated that word representations consist in the awareness of impulsions driven from cerebral speech centres to speech muscles.

In the same vein, Watson (1919) described inner speech (which he referred to as "implicit language") as a weakened form of overt speech. He explicitly considered inner speech (which he equated with thought) as a "highly integrated bodily activity" (Watson, 1919, 325). According to him, inner speech involves "abbreviated, short-circuited and economised processes" (323), but it is not clear whether he actually postulated that inner speech systematically involves overt movement, or rather motor programs, i.e. simulated actions. The extreme view that inner speech requires actual movement has been refuted by Smith and colleagues (1947). Curare was administered to a healthy volunteer, inducing a temporary skeletal muscular paralysis. Although the volunteer became incapable of mouth movement and of overt speech, he was still aware of the questions asked and was able to correctly report them after recovery. This experiment suggests that some form of inner speech must have been present during muscular paralysis. Therefore, verbal thinking, memory storage and presumably inner speech can take place even when articulation is completely prevented. Thus, the extreme version of Watson's view cannot be upheld. A more nuanced view, referred to as the *Motor Simulation* hypothesis, is that inner speech is a mental simulation of articulation, without actual movement. As such, it may feature physiological correlates with recordable physical signals. In this physicalist or embodied view, inner speech production is described as similar to overt speech production, except that the motor execution process is blocked and no sound is produced (Grèzes & Decety, 2001; Postma & Noordanus, 1996). Under the Motor Simulation hypothesis, a continuum exists between overt and covert speech, in line with the continuum between imagined and actual actions proposed by Decety and Jeannerod (1996). This hypothesis has led some authors to claim that inner speech by essence should share features with speech motor actions (Feinberg,

1978; Jones & Fernyhough, 2007). The Motor Simulation hypothesis is supported by empirical findings, including physiological measurements, neural evidence and psycholinguistic data.

Physiological and neural evidence

Objective measurements of respiratory rate, speaking rate, muscular activity and cerebral patterns all suggest that inner speech involves motor processes.

As concerns respiratory rate, Conrad and Schönle (1979) have shown that the respiratory cycle varies along a continuum, from rest to overt speech, via inner speech. During rest, the breathing cycle is symmetrical, with inspiration and expiration phases displaying similar durations. In overt speech, the cycle is strongly asymmetrical with a short inspiration and a long expiration during which speech is emitted. Inner speech is also characterised by a prolonged expiratory phase. They concluded that this modification of the respiratory cycle from rest to inner speech suggests that motor processes are at play during inner speech (see Chapell, 1994, for similar findings).

Speaking rate findings are more debated. As mentioned in Section 2.2, silent recitation has been found to be faster than overt recitation by many researchers. (Anderson, 1982; Korba, 1990; MacKay, 1981; Marshall and Cartwright, 1978, 1980). Some studies of inner speech rate have found similar results for recitation in covert and overt modes, however (Landauer, 1962; Weber & Bach, 1969; Weber & Castleman, 1970). This would suggest that inner and aloud speech may involve common central processes, at least during recitation of stored words, sentences or discourses (alphabet, numbers, pledges). Netsell and colleagues have examined more spontaneous sentence production in both covert and overt modes (Netsell, Kleinsasser, & Daniel, 2016). Participants generated full sentences by saying the first thing that came to their mind. Spontaneous sentence generation involves conceptual preparation and formulation (including morphological, phonological, and phonetic encoding) before articulation can take place (e.g. Levelt, 1989). In inner speech, articulation is inhibited, but conceptual preparation and formulation involve processes that unfold over time. Using spontaneous sentence generation, Netsell and colleagues found that the rate of inner speech (5.8 syllables per second) was significantly faster (5.2 syllables per second) than that of overt speech. But the fact that the difference is relatively small implies that speaking

aloud may only differ from inner speech by the additional time needed to overtly articulate, once the speech motor plan is fully designed. As advocated by Netsell and colleagues, more research is needed to provide precise measures of speaking rate during covert and overt speech, and to allow for informative conclusions on the time course of the two processes.

Concerning muscular activity, Stricker's introspective observation that inner speech is accompanied with muscular contraction finds support from a few electromyographic (EMG) studies during controlled tasks involving inner speech. Using electrodes inserted in the tongue tip or lips of participants, Jacobson (1931) was able to detect EMG activity during several inner speech tasks. Sokolov (1972) carried out surface EMG measurements of lip and tongue muscles. He recorded more intense muscle activation when participants had to perform complex tasks, such as problem solving, which, according to him, necessitated substantial inner speech production. Surface EMG recordings carried out by McGuigan and Dollins (1989) indicated that the lips were significantly active when silently reading the letter "P" (bilabial articulation), but not when reading the letter "T" (alveolar articulation) or a nonlinguistic control stimulus. The reverse pattern was observed for the tongue. The authors concluded that the speech musculature used for overt production of specific phonemes is selectively active in covert production of these phonemes. Livesay, Liebke, Samaras, and Stanley (1996) measured EMG activity in the lips of participants during rest and several mental tasks. They found a significant increase in EMG activity during silent recitation compared to rest, but no increase during non-verbal visualisation. A study on speech muscle activity during dreamed speech using inserted electrodes suggests that the silent (non-phonated) speech that occurs in dream is associated with EMG activity in orbicularis oris and mentalis muscles (Shimizu & Inoue, 1986). Surface EMG activity has also been detected in orbicularis oris during auditory verbal hallucination (which has been described as inner speech attributed to an external source, see Section 2.6) in patients with schizophrenia (Rapin, Dohen, Polosan, Perrier, & Lœvenbruck, 2013). A study by Nalborczyk and colleagues (2017) on induced mental rumination, which can be viewed as a form of excessive negative inner speech (see Section 2.5), has also found an increase in labial EMG activity during rumination compared with a relaxed state. In addition, after rumination induction, an orofacial relaxation session reduced labial EMG activity and had a beneficial (decreasing) effect on mental ruminations. Although more work needs to be carried out to disentangle the

factors that modulate lip activity during rumination (negative affects may influence labial activity), this study suggests that the motor system is involved during mental rumination.

A further argument for the motor nature of inner language comes from cerebral patterns. As reviewed in Perrone-Bertolotti and colleagues (2014, see also Perrone-Bertolotti et al., 2016 and Lœvenbruck et al., 2018), covert and overt speech production both recruit essential language areas in the left hemisphere. These include motor and premotor cortex in the frontal lobe including Broca's area (left inferior frontal gyrus), sensory areas (bilateral auditory areas and Wernicke's area in superior temporal gyrus), and an associative region, the left inferior parietal lobule, including the left supramarginal gyrus. However, there are differences. Consistent with the Motor Simulation hypothesis and the notion of a continuum between covert and overt speech, overt speech is associated with more activity in motor and premotor cortices than inner speech (e.g. Palmer et al., 2001). Moreover, overt speech more strongly activates sensory areas, and typically auditory areas (Shuster & Lemieux, 2005). This suggests that overt speech includes sensory activation associated with the processing of one's own uttered speech. Reciprocally, inner speech involves cerebral areas that are not activated during overt speech (Basho, Palmer, Rubio, Wulfeck, & Müller, 2007). Some of these activations (cingulate gyrus, left middle frontal gyrus) can be attributed to the inhibition of overt response. Overall, these findings support the claim that inner speech is a motor simulation of speech, and that, as such, it shares most of the processes dedicated to overt speech production, including motor planning but excluding motor execution. The processes involved in overt speech therefore include those required for inner speech (except for inhibition). Some brain lesion studies support this view: when overt speech is impaired, inner speech is either intact (Baddeley & Wilson, 1985; Vallar & Cappa, 1987) or altered (e.g. Levine, Calvanio & Popovics, 1982; Martin & Caramazza, 1982), depending on the processes impacted.

A few studies have reported a dissociation that goes against this view, however (e.g. Geva, Bennett, Warburton, and Patterson, 2011a; Langland-Hassan, Faries, Richardson, and Dietz, 2015). They found that the patients' performance in covert speech tasks was poorer than in overt speech tasks. As explained in Lœvenbruck et al. (2018), this dissociation can be explained by limitations in the tasks used. Covert speech was only tested using rhyme

judgment, which does not reflect genuine speech production and which may well be easier overtly (even in healthy patients).

Psycholinguistic data

Psycholinguistic data further indicate that motor processes and articulatory representations are part of inner speech production.

As explained in Section 2.2, some researchers have suggested that inner speech is impoverished at the articulatory level. This claim is still debated however, since a phonemic similarity effect has in fact been found by Corley, Brocklehurst and Moat (2011) during tongue-twister production, even in a covert mode. Furthermore, a study by Smith, Hillenbrand, Wasowicz, & Preston (1986) shows that articulatory content influences speaking rate in both overt and covert modes. Certain repeated stimuli required more time to produce because they included articulatorily complex sequences, typically alternations of similar phonemes in the same syllable position (e.g. "wristwatch" longer than "wristband", because involving two gestures with the same articulator /r/-/w/ instead of two gestures with two different articulators /r/-/b/, which are easier to anticipate and coordinate). The finding that articulatorily complex stimuli also took longer to produce covertly suggests that subphonemic coordination and anticipation principles are at play during inner speech.

Moreover, Scott, Yeung, Gick and Werker (2013) have examined the influence of concurrent inner speech production on speech perception. Scott and colleagues showed that the content of inner speech orients the perception of ambiguous syllables. In a first experiment, they found that ambiguous /ɑ'bɑ/ - /ɑ'vɑ/ sequences were perceived differently depending on the concurrent inner production (more perception of /ɑ'bɑ/ when inner producing /ɑ'bɑ/ and the opposite pull when producing /ɑ'vɑ/). In a second experiment on the same ambiguous syllables, they tested subphonemic effects. They found that inner production of /ɑ'fɑ/ biased perception towards /ɑ'vɑ/, and imagining /ɑ'pɑ/ biased perception towards /ɑ'bɑ/. This suggests that subphonemic content is still present in inner speech. Overall, these findings suggest that, contrary to Oppenheim and Dell's (2010) findings and in line with Corley and colleagues's (2011), inner speech can be specified at the articulatory level. A recent fMRI study suggests that inner speech during reading codes detail as fine as voicing (Kell, Darquea, Behrens, Cordani, Keller & Fuchs, 2017). In this study, the

number of voiceless and voiced consonants in the silently read sentences was systematically varied. Increased voicing modulated voice-selective regions in auditory cortex. Overall, these data suggest that inner speech may be specified at the articulatory level.

To wrap up the arguments presented in Section 3, the format of some variants of inner language (at least the expanded deliberate form, see Section 2) is both motor and sensory. It can be construed that imaginary acts give rise to multisensory percepts. But these acts themselves could stem from prior sensory goals, as Paulhan hinted in 1886, which could themselves be derived from more abstract representations (condensed inner speech).

4. Neural mechanisms of inner speech: simulation, prediction and the feeling of agency

These many facets of inner language, one of the most significant components of our inner space, can be accounted for when a predictive control perspective is taken. In predictive control accounts, any action is accompanied with a prediction of its sensory consequences. Motor and sensory aspects are thus tightly linked. The "Action" view (Jones & Fernyhough, 2007) and the "Activity" view (Martinez-Manrique & Vicente, 2015) hold that inner language is itself an action. In line with these views, and in the framework of Frith and colleagues' predictive account of action control (Frith, 1992; Frith, Blakemore & Wolpert, 2000), we have designed a neurocognitive predictive model of the last stage of inner speech production (i.e. articulatory programming: from phonetic goals to the motor program), which accounts for the sensory as well as motor qualities of inner speech (Lœvenbruck et al., 2018).

It can be speculated that predictive control also operates at the earlier stages of inner language production. Hierarchical predictive control has been applied to overt speech control by Pickering and Garrod (2013, 2014). Pickering and Garrod's model includes pairs of controllers and predictors that use efference copy mechanisms to implement monitoring at each level of speech production (semantics, syntax, phonology). Vicente & Martínez-Manrique (2016) suggest that this type of modelling can be applied to inner language production. In Lœvenbruck (in preparation), I elaborate on these suggestions and I propose a hierarchical predictive control model of language production, from communicative intention to articulatory program, that includes a detailed account of inner speech production (see also Grandchamp et al., in preparation). This model, illustrated in Figure 1,

includes semantic, syntactic and articulatory levels.

At the lowest hierarchical level, i.e. articulatory programming, wilful inner speech is considered as deriving from desired phonetic goals, in a heteromodal format that integrates multiple sensory representations. As explained in Lœvenbruck et al. (2018), these desired goals are transformed into motor commands by a controller (or "internal inverse model"). The motor commands are inhibited and their efference copy is assigned as input to a predictor (or "simulator", "forward internal model of the vocal apparatus") that generates simulated acts, which themselves provide predicted multisensory percepts (voices, somatosensory sensations, facial visemes). These predicted percepts unfold over time. The inner voice in wilful, expanded inner speech, precisely consists of these predicted signals. This simulated experience occurs earlier than the actual experience would, which explains why inner speech may take shorter to be delivered than overt speech (see Section 2.2). An integrator transforms these multisensory percepts into a heteromodal representation. A compararison between predicted heteromodal states and desired phonetic goals provides an error signal which can be used to monitor inner speech. It has been claimed that the comparison between desired goals and predicted states also contributes to the sense of agency, of feeling in control of one's inner speech (Rapin et al., 2013, 2016, revised from Frith, 1992). If desired and predicted states match, then the perceived stimuli are self-generated. A defect in this mechanism can explain the phenomenon of auditory verbal hallucination. If the prediction is faulty, there is no match between predicted and desired states, agency is defective and the inner voice (predicted experience) can feel alien.

At the higher levels, the predictors are not simulators of the vocal apparatus (contrary to Pickering & Garrod's account), because there is no physical apparatus to simulate. Instead, I speculate that predictors are computational procedures that transform one type of mental representation into another. Comparisons between desired and predicted states play a role in monitoring at each level in the hierarchy. They presumably also play a role in agency.

At the formulating level (syntax-phonology encoding), the desired pre-verbal message is transformed into a phonetic plan by a controller. A predictive transformation converts this plan into a predicted pre-verbal message, which can be compared with the desired pre-verbal message. If the prediction does not match the goal, then the controller receives an error signal

and is adjusted, and the lower (articulatory planning) level is affected (i.e. before articulatory programming even takes place).

Similarly, at the conceptualising level, a predicted communicative intention is generated by the highest controller-predictor pair in the hierarchy. This prediction is compared with the original desired communicative intention. If they do not match, then the controller can be adjusted and lower levels in the hierarchy are affected.
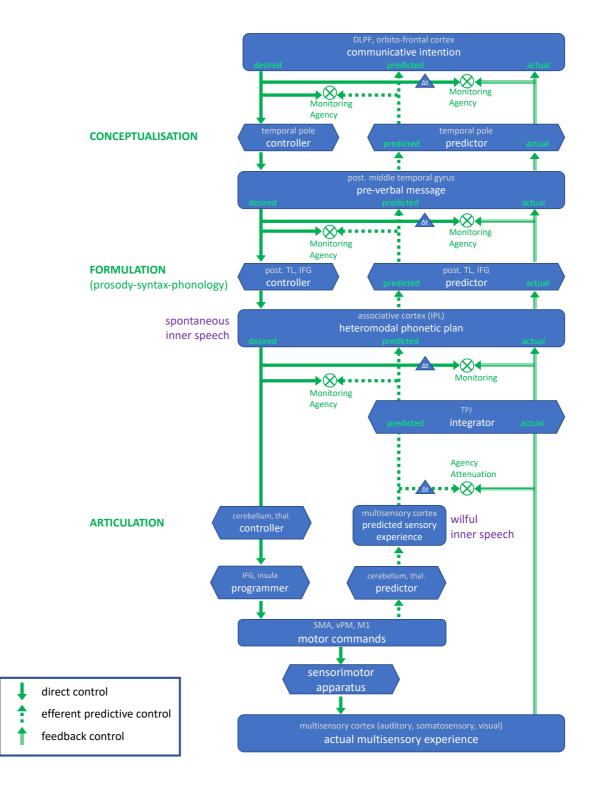
Figure 1. *A hierarchical predictive model of speech production, inspired from suggestions by Haruno et al. (2003), Pacherie (2008), Pickering & Garrod (2013) and Duffau et al. (2014).*

As shown in Figure 1, following suggestions and models by Indefrey (2011), Guenther and Vladusich (2012), Hickok (2012) Tian and Poeppel (2013), Duffau and colleagues (2014), I speculate that the articulation level engages the auditory cortex (posterior superior temporal gyrus, superior temporal sulcus), as well as the somatosensory cortex (anterior supramarginal gyrus and primary sensory cortex, in parietal lobe), together with the temporo-parietal junction, cerebellum, left inferior frontal gyrus, insula, supplementary motor area, ventral premotor cortex and lower primary motor cortex (see Lœvenbruck et al., 2018). Similarly, I propose that the formulating level involves the arcuate fasciculus, left inferior frontal gyrus, posterior part of the temporal lobe and of the superior longitudinal fasciculus and posterior middle temporal gyrus. Finally, the conceptualising level presumably engages the dorsolateral prefrontal cortex, orbito-frontal cortex and temporal pole.

This hierarchical model accounts for the difference between wilful and spontaneous inner speech. Wilful inner speech consists of predicted multisensory percepts that unfold over time and that result from computations of pairs of controller and predictor models, all through the hierarchy, down to the lowest articulatory level. Spontaneous inner speech (unbidden thoughts) is subjectively more evanescent and tenuous. I speculate that it corresponds to mere desired heteromodal states, phonetic goals, deriving from higher levels (semantic and syntactic). In that case, inner speech production is cut short before articulatory programming. Therefore, the phonetic goals are not transformed into simulated acts and their predicted sensory consequences, resulting in a more fleeting experience. I further assume that during wilful inner speech, top-down executive signals may be issued in prefrontal cortex to launch the last prediction mechanism as well as well as to inhibit motor execution. These signals are hypothesized to be absent in spontaneous inner speech, hence the absence of simulated acts and their predictions. The absence of a prediction itself makes for the weaker feeling of agency which characterizes spontaneous inner speech (Gallagher, 2004). As this model shows, the predictive control mechanism, when functional, therefore contributes to creating the rich sensory qualities of inner speech, as well as the feeling of agency, of awareness of our wilful verbal thoughts. Flaws in the prediction or in the comparison processes could explain the disruption in agency observed in auditory verbal hallucination. Further research needs to be carried out to better describe how top-down

signals and comparator mechanisms at different hierarchical levels all contribute to agency.

Conclusion

Inner language takes a significant part of our inner space, with many beneficial outcomes, which span from improving cognitive performance to contributing to autonoetic consciousness. It can become excessive (in verbal rumination), and even run amok (in auditory verbal hallucination) thus becoming detrimental, and engendering suffering and functional disability. The integrated approach presented here, in which inner language is conceived of as a multimodal act with multisensory percepts, offers interesting insights into the various forms of beneficial and detrimental inner language. But many issues still need to be resolved. A deeper understanding of how the oscillations between wilful and spontaneous forms of inner language may enhance cognitive performance could help people with high concentration needs. It could also be beneficial to the understanding of verbal rumination as well as auditory verbal hallucinations. In addition, although many of the subcomponents of inner language production can be associated with specific neural networks (see Section 4), several operations remain ill-described. It is still unclear which networks process the outcomes of the comparisons supposed to occur after predictions are made at each level and how an efficient cognitive control mechanism might integrate these outcomes. More research is needed also on the processes by which we can generate inner speech with someone else's voice. Do we have a predictor for each of the voices we know?

In summary, although an integrative neurocognitive model, gathering findings from introspection and empirical works, can shed light on the format of inner language, many issues are far from resolved. I believe that endeavouring to further combine introspective efforts with objective behavioural and neurophysiological measurements, should help to better portray our inner linguistic space.

References

Alderson-Day, B. & Fernyhough, C. (2015). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology. *Psychological Bulletin,* 141, 931-965.

Alderson-Day, B., Weis, S.; McCarthy-Jones, S., Moseley, P., Smailes, D. & Fernyhough, C. (2016). The Brain's Conversation with Itself: Neural Substrates of Dialogic Inner Speech. *Social Cognitive and Affective Neuroscience.*

Alexander, J. D., & Nygaard, L. C. (2008). Reading voices and hearing text: talker-specific auditory imagery in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 446-459.andersoin

Anderson, R. E. (1982). Speech imagery is not always faster than visual imagery. *Memory & Cognition,* 10, 371-380.

Arnaud, L., Schwartz, J.-L., Lœvenbruck, H., & Savariaux, C. (2008). Perception as a (Shaped) Mirror of Action: It Seems Easier to Lipread One's Own Speech Gestures than those of Somebody Else. *Workshop on Speech and Face to Face Communication*, Grenoble, 27-29 Oct. 2008.

Baars, B. (2003). How brain reveals mind neural studies support the fundamental role of conscious experience. *Journal of Consciousness Studies,* 10, 100-114.

Baddeley, A.D., Hitch, G.J. (1974). *Working memory*. London: Academic Press.

Baddeley, A., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language; 24*(4), 490–502.

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

Baldo, J. V., Paulraj, S. R., Curran, B. C. & Dronkers, N. F. (2015). Impaired reasoning and problem-solving in individuals with language impairment due to aphasia or language delay. *Frontiers in Psychology, 6*, 1523.

Ballet, G. (1886). Le langage intérieur et les diverses formes de l'aphasie. *Félix Alcan, éditeur, Paris.*

Basho, S., Palmer, E. D., Rubio, M. A., Wulfeck, B., & Müller, R. A. (2007). Effects of generation mode in fMRI adaptations of semantic fluency: paced production and overt speech. *Neuropsychologia, 45*(8), 1697-1706.

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312.

Bellugi, U., Klima, E., & Siple, P. (1975). Remembering in signs. *Cognition*, 3 (2), 93-125.

Bergounioux, G. (2001). Endophasie et linguistique [Décomptes, quotes et squelette]. *Langue française*, N°132, *La parole intérieure*, sous la direction de Gabriel Bergounioux, 106-124.

Campbell, R. & Dodd, B. (198°). Hearing by eye. *Quarterly Journal of Experimental Psychology,* 32, 85-99.

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and brain sciences,* 32, 121-138.

Cary, P. (2011). The Inner Word Prior to Language. *Philosophy Today*, 192–198.

Chapell, M. S. (1994). Inner Speech and Respiration: Toward a Possible Mechanism of Stress Reduction. *Perceptual and Motor Skills,* 79, 803-811.

Chomsky, N. (2012*). The science of language: Interviews with James McGilvray*. Cambridge University Press.

Conrad, B., & Schönle, P. (1979). Speech and respiration. *Archiv für Psychiatrie und Nervenkrankheiten, 226*, 251–68.

Corley, M., Brocklehurst, P. H., & Moat, H. S. (2011). Error biases in inner and overt speech: Evidence from tongue twisters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 162–175.

Cossu, G. (2003). The role of output speech in literacy acquisition: Evidence from congenital anarthria. *Reading and Writing*, *16*(1–2), 99–122.

David, A. S. (2004). The cognitive neuropsychiatry of auditory verbal hallucinations: An overview. *Cognitive Neuropsychiatry*, 9, 107–123.

Davis, R. N., & Nolen-Hoeksema, S. (2000). Cognitive inflexibility among ruminators and nonruminators. *Cognitive Therapy and Research, 24,* 699-711.

Decety, J., & Jeannerod, M. (1996). Mentally simulated movements in virtual reality: does Fitt's law hold in motor imagery? *Behavioral Brain Research, 72*, 127-134.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283-321.

Dell, G., & Oppenheim, G. M. (2015). Insights for Speech Production Planning from Errors in Inner Speech. *The Handbook of Speech Production,* Redford, M. (Ed.), John Wiley & Sons, West Sussex, UK, 404-418.

De Ruiter, J. P. (2007). Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture,* 7(1), 21-38.

Dewaele, J.-M. (2015). From obscure echo to language of the heart: Multilinguals' language choices for (emotional) inner speech. *Journal of Pragmatics*, *87*, 1–17.

Dujardin, E. (1887). Les lauriers sont coupés. *Revue indépendante.*

Dujardin, E. (1931). *Le Monologue intérieur*, Paris.

Egger, V. (1881). *La Parole intérieure. Essai de psychologie descriptive*. (G. Baillière, Ed.). Paris.

Emerson, M.J., Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, 48, 148-68.

Erdmann, J. E. (1851). *Psychologische Briefe,* Aufl. Leipzig: Reichardt.

Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin, 4*, 636–640.

Fernyhough, C. (1996). The dialogic mind: a dialogic approach to the higher mental functions. *New Ideas in Psychology*, 14(1), 47–62.

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

Fernyhough, C. (2016). *The Voices Within: The history and science of how we talk to ourselves*. Basic Books.

Filik, R., & Barber, E. (2011). Inner speech during silent reading reflects the reader's regional accent. *PloS One*, *6*(10), e25782.

Franck, N. (2006). *La schizophrénie: La reconnaître et la soigner* [Schizophrenia: Detection and care]. Paris: Odile Jacob Publishing.

Frith, C. D. (1992). The cognitive neuropsychology of schizophrenia. *Psychology Press.*

Frith, C., D., Blakemore, S., & Wolpert, D. (2000). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Research Reviews,* 31, 357-363.

Gallagher, S. (2004). Neurocognitive models of schizophrenia: a neurophenomenological critique. *Psychopathology, 37*(1), 8-19.

Geiger, L. (1868). *Ursprung und Entwickelung der menschlichen Sprache und Vernunft,* Stuttgart.

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

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J.-C., & Warburton, E. A. (2011b). The neural correlates of inner speech defined by voxel-based lesion-symptom mapping. *Brain*, 134(10), 3071-3082.

Grandchamp R., Rapin L., Perrone-Bertolotti M., Pichat C., Lachaux J.P., Baciu M., Lœvenbruck H. et al. (in preparation). Cerebral correlates of deliberate inner speech and of instances of reported verbal mind wandering.

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping, 12*, 1-19.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25, 408 - 422.

Heavey, C. L. & Hurlburt, R. T. (2008). The phenomena of inner experience *Consciousness and cognition,* 17, 798-810.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience,* 13(2), 135-145.

Hurlburt, R. T. (1993). *Sampling inner experience in disturbed affect*. Springer Science & Business Media.

Hurlburt, R. T. (2011). *Investigating pristine inner experience: Moments of truth*. Cambridge University Press.

Hurlburt, R. T.; Alderson-Day, B.; Kühn, S. & Fernyhough, C. (2016). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLoS ONE,* 11, e0147932.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in psychology*, 2, 255.

Jacobson, E. (1931). Electrical measurements of neuromuscular states during mental activities. V. Variation of specific muscles contracting during imagination. *American Journal of Physiology*.

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition,* 16, 391-399.

Kell, C. A., Darquea, M., Behrens, M., Cordani, L., Keller, C. & Fuchs, S. (2017). Phonetic detail and lateralization of reading-related inner speech and of auditory and somatosensory feedback processing during overt reading. *Human Brain Mapping, 38*, 493-508.

Klinger, E. & Cox, W. M. (1987). Dimensions of thought flow in everyday life *Imagination, Cognition and Personality,* 7, 105-128.

Korba, R. J. (1990). The Rate of Inner Speech. *Perceptual and Motor Skills,* 71, 1043-1052.

Lackner, J. R., & Tuller, B. H. (1979). Role of efference monitoring in the detection of self-produced speech errors. In W. E. Cooper, & E. C. T. Walker (Eds.), *Sentence processing: psycholinguistic studies presented to Merril Garret*. Hillsdale, NJ: Lawrence Erlbaum.

Landauer, T. K. (1962). Rate of implicit speech. *Perceptual and motor skills*, 15, 646-646.

Langland-Hassan, P., Faries, F. R., Richardson, M. J., & Dietz, A. (2015). Inner speech deficits in people with aphasia. *Frontiers in Psychology*, 6, 1-10.

Larøi, F., & Woodward, T. S. (2007). Hallucinations from a cognitive perspective. *Harvard Reviews of Psychiatry*, 15, 109–117.

Lavigne K. M., Rapin L. A. ; Metzak P. D., Whitman J. C., Jung K., Dohen M., Lœvenbruck H., Woodward T. S. (2015). Left-dominant temporal-frontal hypercoupling in schizophrenia patients with hallucinations during speech perception. *Schizophrenia Bulletin*, 41 (1), 259-267.

Levelt, W. J. M. (1989). *Speaking: from intention to articulatio*n. Cambridge, MA: MIT Press.

Levine, D. N., Calvanio, R., & Popovics, A. (1982). Language in the absence of inner speech. *Neuropsychologia*, 20(4), 391-409.

Linden, D.; Thornton, K.; Kuswanto, C.; Johnston, S.; van de Ven, V. & Jackson, M. (2011). The Brain's Voices: Comparing Nonclinical Auditory Hallucinations and Imagery. *Cerebral Cortex,* 21, 330-337.

Livesay, J., Liebke, A., Samaras, M., & Stanley, A. (1996). Covert speech behavior during a silent language recitation task. *Perception and Motor Skills, 83*, 1355–1362.

Locke, J. (1970). Subvocal speech and speech. *Asha,* 12, 7-14.

Lœvenbruck, H., Baciu, M., Segebarth, C., & Abry, C. (2005). The left inferior frontal gyrus under focus: an fMRI study of the production of deixis via syntactic extraction and prosodic focus. *Journal of Neurolinguistics*, *18*(3), 237–258.

Lœvenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P., Baciu, M. & Perrone-Bertolotti M. (in press). A cognitive neuroscience view of inner language: to predict and to hear, see, feel. In *Inner Speech: new voices*, Peter Langland-Hassan & Agustín Vicente (eds.), Oxford University Press.

Luria, A. (1966). *Higher Cortical Functions in Man*. New York: Basic Books.

MacKay, D. G. (1981). The problem of rehearsal or mental practice. *Journal of motor behavior,* 13, 274-285.

MacKay, D. G. (1992). *Constraints of theories of inner speech.* In D. Reisberg (Ed.), Auditory imagery (pp. 121–142). Hillsdale, NJ: Erlbaum.

Marshall, P. H., & Cartwright, S. A. (1978). Failure to replicate a reported implicit-explicit speech equivalence. *Perceptual and Motor Skills,* 46, 1197-1198.

Marshall, P. H., & Cartwright, S. A. (1980). A final (?) note on implicit/explicit speech equivalence. *Bulletin of the Psychonomic Society,* 15, 409-409.

Martin, R. C., & Caramazza, A. (1982). Short-term memory performance in the absence of phonological coding. *Brain and Cognition,* 1, 50 - 70.

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

McGuigan, F. J., & Dollins, A. B. (1989). Patterns of covert speech behavior and phonetic coding. *Pavlovian Journal of Biological Science, 24*, 19–26.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature,* 264, 746-748.

Morin, A. (2005). Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence. *Journal of Consciousness Studies,* 12, 4-5.

Morin, A. (2009. Self-awareness deficits following loss of inner speech: Dr. Jill Bolte Taylor's case study. *Consciousness and cognition*,18, 524-9.

Morin, A. (2012). *Inner Speech Encyclopedia of Human Behavior*. Elsevier, 436-43.

Nahorna, O., Berthommier, F., & Schwartz, J.-L. (2015). Audio-visual speech scene analysis: characterization of the dynamics of unbinding and rebinding the McGurk effect. *The Journal of the Acoustical Society of America,* 137 (1), 362-377

Nalborczyk, L., Perrone-Bertolotti, M., Baeyens, C., Grandchamp, R., Polosan, M., Spinelli, E., Koster, E. & Lœvenbruck, H. (2017). Orofacial electromyographic correlates of induced verbal rumination. *Biological Psychology, 127*, 53-63.

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

Netsell, R., Kleinsasser, S., & Daniel, T. (2016). The Rate of Expanded Inner Speech During Spontaneous Sentence Productions. *Perceptual and Motor Skills,* 123, 383–393.

Nolen-Hoeksema, S., Parker, L. E., & Larson, J. (1994). Ruminative coping with depressed mood following loss. *Journal of Personality and Social Psychology, 67*, 92–104.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science, 3*(5), 400–424.

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

Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: the flexible abstractness of inner speech. *Mem Cognit., 38*(8), 1147–1160.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*(1), 179-217.

Paivio, A. (1990). *Mental representations: A dual coding approach.* Oxford University Press.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *Neuroimage, 14*(1), 182-193.

Panaccio, C. (2014). *Le Discours Intérieur. De Platon à Guillaume d'Ockham*. Le Seuil.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Vol. 18). John Benjamins Publishing.

Paulhan, F. (1886). Le langage intérieur et la pensée. *Revue Philosophique de la France et de l'Étranger, PUF,* 21, 26-58.

Pavlenko, A. (2014). *The bilingual mind: And what it tells us about language and thought*. Cambridge University Press.

Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., ... & Lachaux, J.-P. (2012). How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *J Neurosci*, *32*(49), 17554–17562.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Lœvenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, *261*, 220–239.

Perrone-Bertolotti, M., Grandchamp, R., Rapin, L., Baciu, M., Lachaux, J. P., & Lœvenbruck, H. (2016). Langage intérieur. In Pinto, S. & Sato, M. (eds.). *Traité de Neurolinguistique*. De Boeck-Solal, 109-124.

Pickering, M. J., & Garrod, S. (2013). Forward models and their implications for production, comprehension, and dialogue. *Behavioral and Brain Sciences*, *36*(4), 1–19.

Pickering, M. J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in Human Neuroscience*, *8*(132).

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*, 97-132.

Postma, A., & Noordanus, C. (1996). The production and detection of speech errors in silent, mouthed, noise- masked, and normal auditory feedback speech. *Language and Speech,* 39, 375-392.

Rapin, L., Dohen, M., Lœvenbruck, H., Whitman, J.C., Metzak, P., Woodward, T. (2012). Hyperintensity of functional networks involving voice-selective cortical regions during silent thought in schizophrenia. *Psychiatry Research: Neuroimaging,* 102 (2),110-117.

Rapin, L., Dohen, M., Polosan, M., Perrier, P., & Lœvenbruck, H. (2013). An EMG study of the lip muscles during covert auditory verbal hallucinations in schizophrenia. *Journal of Speech, Language, and Hearing Research : JSLHR, 56*(6), S1882–93.

Rapin, L., Dohen, M., & Lœvenbruck, H. (2016). Les hallucinations auditives verbales. In Pinto, S. & Sato, M. (eds.). *Traité de Neurolinguistique.* De Boeck-Solal, 347-370.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology*, *14*(1), 71–104.

Resnik, P. (2018). Multilinguals' use of L1 and L2 inner speech. *International Journal of Bilingual Education and Bilingualism*, *0*(0), 1–19.

Sato, M., Baciu, M., Lœvenbruck, H., Schwartz, J.-L., Cathiard, M.-A., Segebarth, C., & Abry, C. (2004). Multistable representation of speech forms: a functional MRI study of verbal transformations. *Neuroimage,* 23 (3),1143–51.

Scott, M., Yeung, H. H., Gick, B., & Werker, J. F. (2013). Inner speech captures the perception of external speech. *The Journal of the Acoustical Society of America,* 133, EL286-EL292.

Shimizu, A., & Inoue, T. (1986). Dreamed speech and speech muscle activity. *Psychophysiology*, 23(2), 210-214.

Shuster, L. I., & Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language,* 93, 20-31.

Smadja, S. (à paraître). *La Recherche en littérature : approches stylistiques du monologue intérieur*. Paris: Hermann.

Smith, B., Hillenbrand, J., Wasowicz, J. & Preston, J. (1986). Durational characteristics of vocal and subvocal speech-implications concerning phonological organization and articulatory difficulty. *Journal of Phonetics,* 14, 265-281.

Smith, J., Wilson, M., & Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia, 33*, 1433-1454.

Smith, Scott; Brown, Hugh; Toman, James; Googman, L. (1947). The Lack of Cerebral Effects of d-Tubocurarine. *Anesthesiology*, 8(1), 1-14.

Sokolov, A. N., Onischenko, G. T., & Lindsley, D. B. (1972). *Inner speech and thought*. (Plenum Press, Ed.). New York: Plenum Press.

Stricker, S. (1885). *Du langage et de la musique*. (Traduit de l'allemand par Frédéric Schwiedland, Ed.). Paris: Alcan.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America, 26*, 212-215.

Taine, H. (1870). *De l'intelligence*, 2 vols. Paris: Hachette.

Tian, X., & Poeppel, D. (2013). The Effect of Imagination on Stimulation: The Functional Specificity of Efference Copies in Speech Processing. *Journal of Cognitive Neuroscience, 25* (7), 1020-1036.

Uttl, B., Morin, A., Faulds, T. J., Hall, T. & Wilson, J. M. (2012). Sampling Inner Speech Using Text Messaging. *Canadian Society for Brain, Behavior, and Cognitive Science, Kingston, On, June 7-9 2012, 66*, 287-287.

Vallar, G., & Cappa, S. F. (1987). Articulation and verbal short-term memory: Evidence from anarthria. *Cognitive Neuropsychology, 4*(1), 55-77.

Vercueil, L., & Perrone-Bertolotti, M. (2013). Ictal inner speech jargon. *Epilepsy & Behavior*, *27*(2), 307-309.

Vicente, A., & Martínez-Manrique, F. (2016). The Nature of Unsymbolized Thinking. *Philosophical Explorations*, 19, 173-187.

Vygotsky, L. S. (1934/1986). *Thought and Language*. English Translation by Alex Kozulin. The MIT Press, Cambridge, Massachussetts.

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206.

Watson, J. B. (1919). *Psychology from the standpoint of a behaviorist*. (J. B. Lippincott, Ed.). Philadelphia.

Weber, R. J., & Bach, M. (1969). Visual and speech imagery. *British Journal of Psychology, 60*, 199-202.

Weber, R. J., & Castleman, J. (1970). The time it takes to imagine. *Perception & Psychophysics, 8*, 165-168.

Wiley, N. (2014). Chomsky's Anomaly: Inner Speech. *International Journal for Dialogical Science, 8*, 1-11.

Wilson, M., & Emmorey, K. (1998). A "word length effect" for sign language: Further evidence for the role of language in structuring working memory. *Memory & Cognition*, 26, 584-590.