An Exchange on Linguistic Accounts of Inner Speech

Ray Jackendoff & David J. Lobina

October, 2020

The following is a brief exchange on what inner speech is and does from the perspective of a linguist. It was brought about by the publication of a book review by one of the authors of the exchange (https://inference-review.com/article/just-a-thought), which in turn was brought about by the publication of a book (https://global.oup.com/academic/product/inner-speech-9780198796640?cc=gb&lang=en&#). The two pieces were meant for publication elsewhere, but they have ended up here. The presentation of the exchange is fully collaborative in that one author put the files together and uploaded them to LingBuzz and the other said 'OK'.

Inner Speech: One Linguist's View

Ray Jackendoff

In his review of Peter Langland-Hassan and Agustín Vicente's *Inner Speech: New Voices*, David Lobina expresses disappointment that the volume includes few contributions from linguists, "a missed opportunity."[1] To my knowledge, though, linguists have had little or nothing to say about inner speech, aside from Noam Chomsky's oft-repeated remark that people spend more time talking to themselves than to each other.[2] For better or worse, I happen to be an exception, having proposed an analysis of inner speech in a series of publications.[3] The present essay is an attempt to summarize my view and relate it to some of the positions laid out by Lobina and various contributors to the volume under review.

The major issues discussed in *Inner Speech* are (a) What is the *structure* of inner speech? (b) What is the *experience* of inner speech? and (c) What is the *function* of inner speech – what can you do with it? These questions are best addressed in the context of parallel questions posed of overt, spoken speech. I take them up in turn, first for overt speech and then for inner speech.

The structure of overt speech involves linked mental data structures of (at least) three types. First is phonological structure, the organization of words and sentences in terms of phoneme sequences that are grouped into syllables and feet and that are correlated with stress patterns and intonation contours. Second is syntactic structure, the organization of words into phrases in terms of parts of speech (noun, verb, etc.) and phrasal categories (noun phrase, verb phrase, etc.). Third is semantic or conceptual structure, the organization of meanings in terms of conceptualized entities such as objects, actions, and causation, as well as more abstract entities such as intention and obligation. The conceptual structure of a sentence is the message that the sentence conveys: it is over conceptual structure that semantic notions such as inference, reference, and relation to world knowledge can be computed, and it is conceptual structure that is preserved (insofar as possible) in translation.

These three kinds of structure are independent. Phonology and meaning are not *derived* from syntax, as in mainstream generative linguistics.[4] Rather, they are correlated.[5] For instance, the word *cat* is stored in memory as a link between a phonological structure (how it is pronounced), the syntactic category Noun, and the piece of conceptual structure that encodes one's knowledge of cats. A sentence is likewise a triple of linked structures: a phonological structure that arranges

---

[1] Peter Langland-Hassan and Agustín Vicente, eds., *Inner Speech: New Voices* (Oxford: Oxford University Press, 2018); David Lobina, "Just a Thought" (Review of *Inner Speech*), in *Inference*

[2] For instance in Noam Chomsky, *On Nature and Language* (Cambridge: Cambridge University Press, 2002), pp. 75-77.

[3] Ray Jackendoff, *Consciousness and the Computational Mind* (Cambridge, MA: MIT Press, 1987); "How Language Helps Us Think," in *Pragmatics and Cognition* 4 (1996), 1-24; *Language, Consciousness, Culture* (Cambridge, MA: MIT Press, 2007); *A User's Guide to Thought and Meaning* (Oxford: Oxford University Press, 2012).

[4] Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge: MIT Press, 1965); *The Minimalist Program* (Cambridge: MIT Press, 1995); Robert Berwick and Noam Chomsky, *Why Only Us?* (Cambridge: MIT Press, 2016).

[5] Ray Jackendoff, *Foundations of Language: Brain, Meaning, Grammar, Evolution* (Oxford: Oxford University Press, 2002); Peter Culicover and Ray Jackendoff, *Simpler Syntax* (Oxford: Oxford University Press, 2005); Ray Jackendoff and Jenny Audring, *The Texture of the Lexicon* (Oxford: Oxford University Press, 2020).

the phonology of the sentence's words in linear order, a syntactic structure that arranges the syntax of its words into hierarchical syntactic phrases, and a conceptual structure that combines the meanings of the words into the meaning of the sentence. Independent "interface" principles determine how the linear order in phonology corresponds to hierarchical structure in syntax and compositional meaning in conceptual structure.

In language comprehension, phonological structure is constructed on the basis of auditory input; the linkages, encoded in the words and in the interface principles, are used to construct a conceptual structure (i.e. a meaning) for the heard sentence. In language production (following Levelt's model[6]), the speaker has a conceptual structure in mind, and uses the words and the interface principles to construct syntax and phonology linked to it. In turn, phonology drives motor planning and eventual motor output, culminating in overt speech.

Crucially, the relation between the three structures is not always one-to-one, and meaning is not always precisely expressed in words. In a famous example,[7] one waitperson says to another, *The ham sandwich over in the corner wants more coffee*. The meaning conveyed is 'the person who ordered/who is eating the ham sandwich,' but the phonology and syntax omit the underlined parts of the intended meaning. Moreover, as seen in the gloss, there is more than one way to make the omitted part explicit, and there is no need to decide which version is correct. That is, the unexpressed part of the meaning is not (and need not be) completely precise. Language use is full of such situations, and competent hearers have no difficulty filling in the blanks.

Inner speech partakes of the same three linked structures. As Lobina, Carruthers, Langland-Hassan/Vicente,[8] and others in *Inner Speech* observe, inner speech clearly has phonological structure. As for syntax, Lobina and Hurlburt/Heavey[9] worry that the syntax of inner speech is often fragmentary. But this is true of overt speech as well, and what fragments there are in inner speech have well-formed syntactic structure. Finally, since inner speech is (typically) meaningful, it has a conceptual structure to which the phonology and syntax are linked. Inner speech differs from overt speech chiefly in that it forgoes (or inhibits) the step from phonology to motor planning and/or motor output. Anecdotally, people differ in how much inner speech they experience; Geva[10] suggests that there is also considerable individual variation in whether inner speech is accompanied by unexecuted motor planning.

Bermúdez is concerned with the fact that the message one has in mind is often more precise than the inner speech that purports to express it.[11] However, this is not an issue specific to inner speech: as noted above, even in overt speech, a speaker's intended meaning is often not perfectly or completely mirrored in the phonology and syntax. And since inner speech is not constrained by the need for successful communication, it can get away with being sloppier than overt speech.

---

[6] W. J. M. Levelt, *Speaking* (Cambridge: MIT Press, 1989).
[7] Geoffrey Nunberg, "The Non-Uniqueness of Semantic Solutions: Polysemy," in *Linguistics and Philosophy 3* (1979), 143-184.
[8] Peter Langland-Hassan and Agustín Vicente, "Introduction," in *Inner Speech*
[9] Russell Hurlburt and Christopher Heavey, "Inner Speaking as Pristine Inner Experience," in *Inner Speech*, 168-97.
[10] Sharon Geva, "Inner Speech and Mental Imagery," in *Inner Speech*.
[11] José Luis Bermúdez, "Inner Speech, Determinacy, and Thinking Consciously about Thoughts," in *Inner Speech*.

Turning to the *experience* of overt speech: Everyone acknowledges that one "hears" speech, that is, the "qualia" of speech include phonology.[12] Syntactic structure, on the other hand, is unconscious, as everyone seems to agree; that's why linguists can't simply intuit it. A more controversial question is whether the experience includes conceptual structure as well. I would contend that it does not. There is no space here to go through all the arguments, but here are two.

First, we have immediate apprehension of phonology – we can easily count syllables and appreciate rhymes, for instance. But the structure of meanings is opaque – which is why semanticists cannot agree on how to analyze even the most elementary features of conceptual structure.

Second, we can experience nonsense items that have phonology but no meaning, such as *otkin adarab utal*. So conceptual structure is not necessary in order to experience something as language(-like); phonology alone is sufficient. At the same time, consider the tip-of-the-tongue experience: we have some word's meaning in mind, yet we can't recover its phonology. We know we have the meaning in mind, because we can accept or reject proffered suggestions for the word in question ("Salacious? Um, no, that's not it.") And we may be able to recover some fragment of its phonology ("Hmm, it starts with a *k*, and maybe its stress pattern is *da-DA-da*"). But its overall conscious form is a blank, plus perhaps a feeling of a meaning being present.

In other words, phonology is necessary and sufficient for linguistic experience to have a form, and semantics alone is neither necessary nor sufficient. In short, the unconscious conceptual structure component of speech *constitutes* the thought being conveyed; the conscious phonological component *expresses* or *points to* it.

Now consider the experience of inner speech. It too depends on phonology. First of all, as observed by Langland-Hassan,[13] inner speech is "in a language." This points toward a locus in phonological form, where languages differ, rather than toward conceptual structure, which is language-independent. Second, one can "hear" the meaningless *otkin adarab utal* in imagination (as I am doing as I write this). And one can have "tip-of-the-mind" experiences, in which one struggles to recover a word or name, but without any intention of saying it out loud (think of doing a crossword puzzle, for instance). Thus, in opposition to Bermúdez and Wilkinson/Fernyhough,[14] I conclude that the experience of inner speech, like that of overt speech, reflects phonological structure, plus what might be called a "feeling of meaningfulness," to which I return in a moment.

This conclusion, for both overt and inner speech, is highly counterintuitive. We are accustomed to thinking of consciousness as the highest form of mental activity, deeply intertwined with intelligence and reasoning. Of course it is natural to think this, given that we have no conscious access to conceptual structure, where the real substance of reasoning takes place. What we do

[12] A puzzle is why the experience of speech is auditory rather than motor or proprioceptive – why we don't (normally) *feel* ourselves speaking; I set this aside.
[13] Peter Langland-Hassan, "From Introspection to Essence," in *Inner Speech*.
[14] Bermúdez, op. cit.; Sam Wilkinson and Charles Fernyhough, "When Inner Speech Misleads," in *Inner Speech*.

have conscious access to is phonological form, which is just a structured string of phonemes. But this is made of the wrong units for supporting reasoning.[15]

Still, there is some truth in the intuition that conscious language, overt or inner, plays a role in shaping thought. The phonological form linked to a conceptual structure serves as a conscious "handle" that can be attended to. This stabilizes the thought, gives it an identity, and thereby makes it something that can be deliberately recalled, compared to other thoughts, questioned, or treated as hypothetical.[16] Moreover, conscious linguistic support is necessary in order to attain many important aspects of the human conceptual system, for instance rules of games, the past, organized religion, the number system, and concepts that intrinsically involve numbers, such as finance and science – and let us not forget philosophy. Much of the power of human thought, then, comes from linking conscious phonological structure, which makes it possible to individuate and identify concepts, to *un*conscious conceptual structure, where the actual processes of reasoning are carried out. In turn, an inference in conceptual structure – a new thought – can be made consciously manifest by linking it to a phonological form, which may either be spoken or experienced as inner speech. (As pointed out by Frankish, this process is roughly what Daniel Kahneman calls "System 2."[17])

Which brings us to the *function* of inner speech. Lobina, Frankish, and Machery observe that one can do with inner speech just about all the things one can do with overt speech: comment, question, evaluate, express desires, do arithmetic, rehearse what one is going to say, and so on.[18] As observed above, we can even recite nonsense syllables in inner speech. What we *cannot* do with inner speech is communicate these messages to other people. What we *can* do with inner speech but not overt speech is keep our thoughts to ourselves. There are no great mysteries here.

Not addressed by the contributors to *Inner Speech* (nor, to my knowledge, by almost *any* writers on consciousness) are questions of this sort: If inner speech and overt speech are made of the same elements and principles of combination, what differentiates the two? How does your brain make this distinction a part of your experience? And within inner speech, how is the experience of deliberately talking to yourself different from the experience of unbidden, apparently spontaneous inner speech? These aspects of experience don't have an articulate form, the way phonology does. Rather, they impart a "feel" to one's conscious percepts. Crucially, they cannot come out of nowhere: like every other aspect of experience, the brain has to compute them somehow or another.[19]

---

[15] The independence of thought from inner speech was observed already by the nineteenth century philosopher and psychologist Heymann Steinthal, in his 1881 *Abriss der Sprachwissenschaft* (see quotation in my *User's Guide to Thought and Meaning*, 90). For more extensive discussion of Steinthal's view, see W. J. M. Levelt, "Sleeping Beauties," in I. Toivonen, P. Csúri, and E. van der Zee (eds.), *Structures in the Mind*, MIT Press, 2015, pp. 235-255.

[16] Keith Frankish (almost) makes this point: "Inner Speech and Outer Thought," in *Inner Speech*, 221-43.

[17] Keith Frankish, op. cit.; Daniel Kahneman, "Thinking, Fast and Slow" (New York: Farrar, Straus and Giroux, 2011).

[18] David Lobina, op. cit.; Keith Frankish, op. cit.; Edouard Machery, "Know Thyself," in *Inner Speech*.

[19] More precisely: The brain has to compute a mental structure that is capable of giving rise to experience. I leave aside the huge mystery of how mental structures give rise to conscious experience, i.e. the computational version of the traditional mind-body problem. See my *Consciousness and the Computational Mind* for discussion of my position.

There doesn't seem to be a convenient word for these features of experience; I have called them variously "affects," "valuation features," or "character tags." The immediate question is: What are these features, and how are they computed?

A first example has already been alluded to: the feeling attached to a piece of phonology when it is perceived as a meaningful piece of language. The previous sentence carries such a feeling; *otkin adarab* does not. *Colorless green ideas sleep furiously* identifies each word as meaningful but the whole expression as meaningless; a tip-of-the-tongue experience is identified as meaningful, but it lacks conscious content. These feelings can be computed an internal monitor that detects whether the phonology in question is linked to a conceptual structure. Let's call the output of such a monitor a feature that we'll call [±Meaningful].

Next compare the experience of an actual utterance to that of inner speech. They are both built from elements of phonological form. However, an actual utterance is driven by auditory input, whether from one's own uttered speech or someone else's, while an inner utterance is purely internally generated. This difference can be checked by monitoring the connection between auditory input and phonological form. Let's call this feature [±External]. "Real" utterances are [+External] and inner speech is [–External].

For a third case, consider the experience of producing speech voluntarily versus hearing someone else speak. In either case, the conscious linguistic structure is still phonological form. The distinction can be made by a feature [±Self-controlled]. One's own speech is [+Self-controlled] and someone else's is [–Self-controlled]. This feature has to be computed by whatever brain mechanism is responsible for the volitional initiation of action.

These features can be combined in eight different ways, to characterize different kinds of experience.

[±Meaningful]: "I'm hearing language" vs. "I'm hearing nonsense."
[+External, +Self-controlled]: "I'm speaking"
[+External, –Self-controlled]: "Someone else is speaking"
[–External, +Self-controlled]: "I'm talking to myself" (inner speech)
[–External, –Self-controlled]: "I'm hearing unbidden inner speech"

Auditory hallucinations occur when one is producing unbidden inner speech but assigning it the character tags [+External, –Self-controlled], that is, experiencing the inner speech as originating from someone else. A case like this would be a schizophrenic's "hearing voices." Similarly, speech in a dream is obviously produced internally, without conscious volition. But it is experienced as [+External], i.e. real, and as [+Self-controlled] if one is speaking in the dream and [–Self-controlled] if someone else is.

These features are not just made up to account for experiential properties of language. They appear in every domain of perception and action. For instance, the visual appearance of a natural scene is experienced as [+Meaningful], while that of a Jackson Pollock painting is (for me at least) [–Meaningful]. The difference between the experience of seeing reality and experiencing visual imagery is [±External]. A visual image that one has conjured up deliberately is [+Self-controlled];

one that pops into one's head is [–Self-controlled]. Visual hallucinations and the visual component of dreaming work just like their linguistic counterparts: misassignment of character tags. (Vision of course lacks phenomena that are [+External, +Self-controlled], since one cannot create visual environments by looking at them.)

In short, the very same set of character tags appear in language and vision, mental faculties that are otherwise quite distinct. I take this to be a significant result of the analysis. In other publications, I have proposed several additional character tags and have shown that they pertain not only to language and vision, but also proprioception and action.[20] Individual character tags have been occasionally discussed in the literature, under rubrics such as the "feeling of knowing" and the "feeling of pastness."[21] However, I know of no other treatments of *ensembles* of character tags as suggested here. It is a matter for future research to determine the full repertoire of character tags and the range of domains to which they apply, as well as to discover means of studying them experimentally.

If this analysis is correct, the *content* of conscious experience is determined by particular levels of representation (phonological structure in the case of language), while what might be called the "mode of presentation" is determined by the character tags. This is another aspect of the analysis that is highly counterintuitive. It is hard to wrap one's head around the idea that the huge phenomenological differences between overt and inner speech are modulated by a small collection of binary parameters. However, I am unaware of any alternative proposals of the same scope, and I comfort myself with the thought that quantum mechanics is just as difficult to reconcile with our ordinary grasp of the world. On the other hand, the analysis offered here may feel more threatening, because it reflects deeply on our natural conception of Us – how we humans are put together. For my part, that's why I find it fascinating.

[20] See especially *Language, Consciousness, Culture* and *A User's Guide to Thought and Meaning*, op. cit.

[21] Asher Koriat, "How do we know that we know? The accessibility model of the feeling of knowing," *Psychological Review* 100 (1993), 609-639; Denis Perrin, Kourken Michaelian, and André Sant'Anna, "The phenomenology of remembering is an epistemic feeling," *Frontiers in Psychology*, 3 July 2020, https://doi.org/10.3389/fpsyg.2020.01531

My Own View: Response to Ray Jackendoff

David J. Lobina

I'm very grateful to Ray Jackendoff for sending in *a linguist's* contribution to the study of inner speech. I like to think his letter functions as a complement to my review as well as a response to my lamentation that the *contribution* of a linguist wouldn't have gone amiss in *Inner Speech: New Voices*. My lament was much more general than Jackendoff takes it to be, though; not so much a complaint that the editors hadn't included any account or theory by a linguist on what inner speech is (and is for), but a criticism that the editors hadn't invited any linguist to the debate, which is a different matter. I aimed to rectify this shortcoming in my review by pointing to some features of linguistic production which linguists and only linguists (actually, Noam Chomsky; *the* linguist?) have identified and which are rather relevant for the question at hand; namely, the apparent fact that language use is (infinitely) productive, appropriate to circumstances (in a pragmatic sense), and effectively stimulus independent. As explained in my review, Chomsky subsumes these three properties under what he calls "the creative use of language" and has argued that the study of this phenomenon may lie beyond the remit (and tools) of cognitive science.

I'm not as pessimistic, but I do think that the stimulus independence of language use, in particular, along with what is involved in language production in general,[1] casts significant doubt on some of the more substantial claims regarding how inner speech is used in thinking processes, some of which (the claims, that is) appear in the book I reviewed. Jackendoff himself makes reference to such claims when he points to the role of inner speech in Type 2 reasoning processes (via Keith Frankish, one of the contributors to the book), and it is this very contention which I regard rather misguided. I shall concentrate on this particular issue in this response to Jackendoff's letter, both expanding and reclaiming what I originally wrote in my review. Having said that, it was of course rather flippant of me to suggest that *any* linguist could have contributed to *Inner Speech* in the manner I outlined in the review, but to be fair by "a linguist's contribution" I naturally meant my own.[2]

There's plenty to agree with Jackendoff in his response, and I had stressed some of the points he makes in the original review as well, such as the evident commonality between outer or

[1] Willem Levelt's model of linguistic production, which I referenced in my review and which Jackendoff cites too in his letter, is the usual starting point. In simple terms, this model has it that production starts with the formulation of a message (a thought) to be communicated, which is then followed by the selection of the appropriate words, a way to put these words together into a sentence, and the issuing of motor commands to the organs in charge of producing the message, be this in speech or through hand gestures. This model has often come under some criticism for being perhaps too straightforward and simple. Speech is full of false starts and changes of perspective, and as a matter of fact we don't always express whatever message we initially entertain. There is also some evidence from the field of psycholinguistics that hearing your own sentences as you utter them has an effect on what you say next. That is, language comprehension, often described as a phenomenon in which we receive linguistic input and recover its meaning and thus the message-thought being communicated, would affect what at first sight would be the opposite phenomenon, language production. See, Martin Pickering and Simon Garrod "An integrated theory of language production and comprehension" *Behavioral and Brain Sciences* 36 (2013), especially the comments on the target article and the overall discussion, which showcase the actual state of affairs (i.e., it's not so straightforward). See, also, footnote 7 *infra*.

[2] The phrase *a linguist* is used in both the specific and generic sense in the first two paragraphs, and purposely so. I should add that one of the editors of *Inner Speech* also pointed out to me that very few linguists had written about inner speech when I sent him a copy of a longer and rather different version of my review (and I gather that Jackendoff's contribution was a missed opportunity in itself). I have written about inner speech, in the terms I employed in the review, in David J. Lobina & José E. García-Albea, "On Language and Thought: A question of Format", in *On Language and Thought: A Question of Format*, eds. Roberto De Almeida & Lila Gleitman (Oxford, England: Oxford University Press, 2017), 249-273.

overt speech and inner speech, not only in terms of use (speech in general is often sketchy) but more importantly in terms of the general structure of speech (which includes, as Jackendoff outlines, phonology, syntax, and semantics/concepts).[3] I also agree with Jackendoff, with some provisos, that inner speech clearly has phonological structure (though only some philosophers could have doubted this) and that what differentiates inner speech from outer speech is, chiefly, that motor instructions are not executed in the case of inner kind of speech. My qualifications regarding the last two points are a matter of nuance rather than substance.

To begin with, there are, I think, two types of inner speech. The first kind is one I engage in quite often when I go for a walk by myself, though I am only aware of doing it after the fact – or after I notice passers-by staring at me and smiling. What I do is imagine conversations I could be involved with, and as I do so, I can't help but articulate my own interventions, which I guess is what makes other people smile (thankfully this *parlar da soli* is not treated as a pathology any more). Once I realise what I was doing, I have two seemingly incompatible feelings about it: that I was producing what felt like normal speaking, and at the same time that I was the only person capable of 'hearing' this speech. I'm not describing the experience of talking to ourselves in a very low voice, which clearly isn't a case of inner speech. In fact, if I set out to imagine these conversations consciously, it is unquestionable that I can articulate speech without producing any sound, and yet I can 'hear' my own speech as clearly as in the more unconscious case (this might be similar to what signers do when interpreting for the deaf and hard of hearing). The second kind of inner speech is the paradigmatic example of what is usually meant by inner speech: the experience of talking to ourselves without carrying out any articulation of any kind, even though here too we can 'hear' our own speech.

The two types of inner speech might differ in more ways than one, but what they seem to have in common, apart from the fact that the utterer is exclusively privy to either inner voice, is that in both cases there clearly is phonological structure. One way in which they do seem to differ is in the motor instructions that are executed, or not (and if so, which ones). In my review, I referred to some recent, and relevant, work showing that what the brain encodes during the production of speech is not sound itself, as customarily thought, but the motor commands to produce speech – the movements of the vocal tract (lips, tongue, jaw, etc.). Thus, in the case where I articulate but no external sound is produced even though I can hear my inner voice, the one motor command that is not executed would be phonation, the phenomenon of air being expelled from the lungs through the glottis, producing sound, but all other organs involved in speaking are seemingly set in motion (mouth, tongue, etc.), whereas in the paradigmatic case of inner speech the latter would also be inhibited.

This minor point out of the way, what I do disagree with Jackendoff about regards the role of inner speech in thinking; as mentioned, Jackendoff alludes to inner speech as the possible medium in which Type 2 reasoning is conducted, and I think this is far from what is at all warranted.[4] The thought that to inner speech is to be thinking is certainly ubiquitous, from

[3] Jackendoff says that I worry about the syntax of inner speech being fragmentary and adds that this is also true of outer speech. I don't actually worry about this at all and, as a frequent reader of *Language Log*, I am very aware of how sketchy speech is. I allude to this again in footnote 7, *infra*.

[4] I shall put aside some other disagreements in this response, such as the contrast Jackendoff draws between his parallel architecture framework and minimalist approaches to linguistic derivations, which I think is exaggerated; his conflation between semantic and conceptual representations, which I think is a mistake; and his explanation of certain functions of inner speech in terms of "features", which I think lacks explanatory force. I should add that I am surprised by some of the things Jackendoff has left out from his own work on inner speech – for instance, how inner speech may help bring attention to certain percepts within what he has in the past called the Intermediate Theory of Consciousness – as I find this much more compelling and interesting (and something that ought to have

literature (where it is usually known as the interior monologue) to ancient and medieval philosophy, and I would say it is worth our time detailing why this is problematic.

In my review, I started by quoting an exchange from James Joyce's *Ulysses* in which one of the main characters (Stephen Dedalus) appears to conduct a rather extensive bit of inner speech *while* in conversation with someone else, and I wondered whether this was a psychologically realistic account of his thoughts – not whether Stephen had these thoughts, but whether he had them in such a way. This question regarding Stephen's train of thoughts is of course a rather subjective one, for interior monologues are private events. I am myself certain that during my waking life I am constantly having thoughts, and that most of these thoughts are entirely unconscious in that they do not turn up in my inner speech. The apparently simple act of crossing the street, for instance, may well involve various inferences – e.g., if I cross the street now I will be run over by the incoming car and there is enough space between this car and the next, therefore… – and none of this is usually put into words, aloud or to myself. Or at least they don't need to be for me to be able to carry out such reasoning and not get run over.

To say that such thoughts are unconscious doesn't mean that I have no access to them, nor that my behaviour is in any way unwilled or irrational. I really do intend not to get run over. And it is of course perfectly possible to put one's thoughts into verbal form – perhaps it is by putting words to thoughts that we become aware of the thoughts we are having. But if this is the case, then language may often simply reflect our thoughts rather than being the vehicle in which we have them. We can use language to describe and identify thought, but the thoughts themselves might be of a different nature.

Something along these lines must be what's going on in some of the most famous bouts of inner speech in literature, such as Molly Bloom's interior monologue at the end of *Ulysses* or the self-soliloquy of Tolstoy's protagonist in part 7 of *Anna Karenina*. More a method for a writer to describe the richness of a character's thoughts and feelings than the accurate reporting of a literal event in inner speaking, these two famous monologues are far too extensive, sophisticated, and detailed to be typical examples of inner speech. Streams of consciousness but not streams of thought, these particular examples feel (and read) more like "essays", certainly a world apart from my own experience of using inner speech to imagine future conversations, reimagine past ones or rehearse sentences for an article I am working on. Molly's interior monologue is almost 25,000 words long, after all; Leopold would have been fast asleep by the time she came to reminisce saying Yes.[5]

Many philosophers have also seen much thought in inner speech. Plato may have been alluding to inner speech when he described thinking as 'a talk which the soul has with itself' in the *Thaetetus*, whilst in the *Philebus* he talks of how we can use an interior dialogue to form opinions by formulating questions and answers to ourselves. And in medieval times we find St Augustine supposing that inner speech is the voice of imagination and St Aquinas regarding it

made it into *Inner Speech*; along with my own contribution, of course). See, for instance, Ray Jackendoff, "How language helps us think" *Pragmatics & Cognition,* 4 (1996), 1-34.

[5] The introspective feeling that talking to ourselves is to actually be thinking really is quite widespread, and there's plenty of anecdotal evidence around. I can certainly confirm that it is common for bilingual speakers to be asked what language we think in, by which it is really meant what language we use to talk to ourselves. It has to be one apparently, and if that doesn't quite settle it, the next question is often what language we dream in. I'm told that the sentiment is present in certain strands of Buddhism too, where practitioners actively seek to block their inner speech in order to stop the process of thinking, supposedly allowing them to perceive the world as it is, unfiltered by sentence-thoughts. In this case to not inner speak is, I suppose, to not think.

as a way to practice outer speech.[6] Some contemporary philosophers have also seen thinking as some sort of interior dialogue, but what I suspect is happening here, as I mentioned in passing in my review, is that these thinkers are over-intellectualising what goes on in inner speech – it may not be a surprise that individuals who devote most of their time to careful reasoning and reflection hold such views. There is no doubt that inner speech may take the form of an interior dialogue and sometimes even be rather expansive (though unlike the written word), but this is surely more common an experience for academics than it is for most people – who else spends so much time articulating their own thoughts for debate, presentation, and publication? Both Molly and Anna come to various conclusions during their long soliloquies, and there's plenty of serendipity and suddenness in each of their realisations (which is, I think, one of the points the writers of these monologues wanted to convey).

Be that as it may, the main point of contention here is exemplified by a suggestion from Peter Carruthers, one of the contributors to *Inner Speech*, on what inner speech does – an idea which Frankish, in his own contribution to the book, fraternises with.[7] This is the proposal that the meaning of inner speech sentences may be 'broadcast' to other mental abilities such as problem-solving, effectively using language as a vehicle to set in motion a thinking process. Other contributors to the book I reviewed have defended similar ideas. Christopher Gauker outlines what he calls a Lockean view of communication, which goes against much of mainstream philosophy of language, and in so doing identifies problem solving with conversing with ourselves (the interior dialogue once again). Conversation, Gauker argues, is the medium in which thought is revealed, to the point that an act of speech is itself the production of an act of thought, a reference to Wilfrid Sellars's 'thinking out loud' phrase, a former teacher of Gauker's and the fountain for much of his against-the-grain theory of communication.[8]

Frankish, for his part, associates speech to psychological theories of reasoning, specifically to thinking that is intentional and conscious, known as Type 2 reasoning in the literature, which

---

[6] See, John M. Cooper, ed., *Plato: Complete Works* (Cambridge: Hackett Publishing Company, 1997), on Plato, and Claude Panaccio, *Mental Language* (New York: Fordham University Press, 2017) on the medieval philosophers I mention. The more modern Jerry Fodor was obviously half-joking when he reported what his inner speech experience felt like – 'I can't solve this; it's too hard. I'm not smart enough to solve this. If Kant couldn't solve this, how can they possibly expect me to' – but he was also alluding to the fact that different people may well report different types of inner speech and, with these, different opinions on what inner speech is and does. See, Jerry Fodor, *In Critical Condition* (Cambridge, MA: The MIT Press, 1998), 68.

[7] Peter Carruthers, "The Causes and Contents of Inner Speech," in *Inner Speech: New Voices*, eds. Peter Langland-Hassan and Agustín Vicente (Oxford, UK: Oxford University Press, 2018), 31-52. In his contribution to the book, Carruthers argues further that inner speech may have specifically arisen in evolution to enable the rehearsal and evaluation of overt speech actions. A position not too dissimilar to Aquinas's own take on the matter, the ability to inner speak would be nature's way to make us better out-loud speakers. Carruthers's case is partly based on the psycholinguistic evidence I alluded to in footnote 1, *supra*, though it is nonetheless rather speculative. I suppose (some) speculation is unavoidable in the construction of evolutionary scenarios for mental abilities such as language and speech, but in this case the argument doesn't stand on very solid ground. In particular, Carruthers takes the evidence from psycholinguistics too far and far too seriously – there is no consensus in the literature as to how widespread the effects of comprehension on production are or what the evidence actually says about the connection between these two processes. Further, in any case, the evidence is perfectly compatible with Levelt's model of language production (briefly outline in footnote 1, and in the text). False starts and changes in what one is saying may simply point to the adjustments speakers commonly carry out in order to put the message across in the clearest possible way – or it may in fact point to the very plausible possibility that one doesn't stop thinking when producing a sentence. After all, it is doubtful that one has a thought and then speaks it machine-like, as if one stops thinking while in the middle of speaking – surely to be speaking is not to be in the mental vacuum of a parrot. If you change your mind as you speak, this should be reflected in what you don't end up saying.

[8] Christopher Gauker, "Inner Speech as the Internalization of Outer Speech," in *Inner Speech*, 53-77.

Frankish sees as being largely language involving, as Jackendoff mentions.[9] According to Frankish, problem solving is often a matter of breaking down a problem into sub-problems, and this is typically conducted in a questioning and prompting manner in language, much as one does when questioning a friend in a social context. Frankish uses the specific example of how we may come to decide whether to go to a party we have been invited to, an event that might well start in inner speech by literally self-questioning ourselves 'do I want to go to the party?'. Such a question would set in motion a process of posing and answering questions, eventually reaching a conclusion – a case of reasoning.

Putting aside the worry that this may well be, as mentioned, a case of over-intellectualising what goes on in speech, the general idea, whilst at first sight plausible, is not only unlikely but unnecessary. First of all, the claim that Type 2 reasoning is conducted in language is Frankish's own take on the matter; it is certainly not the standard view of defenders of dual-processing reasoning. There are in fact many accounts of the representations and processes involved in reasoning, such as mental logic, mental models, or bayesianism, and none of these explicitly manipulate linguistic vehicles (or inner speech).[10] This is most evident in the sort of explanations that have been put forward to account for specific cases of problem-solving tasks that would engage Type 2 reasoning mechanisms, even when such tasks are presented in linguistic form and some kind of logical reasoning is involved. In tasks such as the famous Wason selection task or the representation of a negated thought or syllogism, what seems to be crucial is how the problems are represented, and this typically involves translating linguistic information into the relevant format, be this a mental model or else.[11]

In addition, what seems to be the standard view in the field is that Type 2 reasoning correlates with both general intelligence and working memory, and it is doubtful that language constitutes the causal factor in either case, let alone that inner speech is literally being employed to lay out the arguments and inferences customarily carried out during a problem-solving task.[12] The role of working memory is an interesting one, given that even though this kind of memory does employ a language-mediated "phonological loop", working memory makes use of other components, including a central executive unit as well as an episodic buffer and a visuo-spatial scratchpad, and all these components are central and operative when working memory is being

[9] Keith Frankish, "Inner Speech and Outer Thought," in *Inner* Speech, 221-43.

[10] See, Martin Braine & David O'Brien (eds.), *Mental Logic* (New Jersey: Lawrence Erlbaum Associates, 1998), on mental logic; Philip Johnson-Laird, "Against Logical Form", *Psychologica Belgica* 50 (2010)193-221, on mental models; and Mike Oaksford & Nick Chater, *Bayesian rationality: The probabilistic approach to human reasoning* (Oxford, England: Oxford University Press, 2007), on bayesianism. These accounts are often meant to apply to both Type 1 and Type 2 forms of reasoning (Type 1 involves fast, automatic processes), with the proviso that the features operative at each level differ (e.g., biases and heuristics in Type 1, issues to do with attention and processing effort in Type 2).

[11] The Wason selection task is specifically discussed in the context of dual theories of reasoning in Jonathan Evans, "In two minds: dual-process accounts of reasoning" *Trends in cognitive sciences* 7 (2003), 454-459. The task itself is very dependent on the context it is presented, as Evans shows. As for the negation of syllogism and the line, see Isabel Orenes, David Beltrán & Carlos Santamaría, "How negation is understood: Evidence from the visual world paradigm" *Journal of Memory and Language* 74 (2014), 36-45, for a study on the representation and processing of negation in terms of mental models.

[12] Jonathan Evans, "In two minds: dual-process accounts of reasoning" *Trends in cognitive sciences* 7 (2003), 454, 456. It seems, in fact, that reasoning in general is heavily constrained by working memory capacity, and here too the role of language (or inner speech) would be limited.

employed (and, in any case, it is not at all obvious how implicated inner speech actually is in the phonological loop).[13]

Finally, even if speech were to be involved in Type 2 reasoning, it could not have the role scholars such as Carruthers and Frankish ascribe to it, precisely because of *the* linguist's insistence that language use is creative. The crucial feature in this respect, as noted, is the claim that language production does not depend on any stimulus – no matter what circumstances we face, there really is no telling what exactly we'll be saying about them. As I described in my review, Chomsky first discussed the stimulus independency of language production in a well-known review of BF Skinner's *Verbal Behaviour*, and one of his main arguments in that paper was that there is no causal connection between stimuli and what one may choose to say about them. As Chomsky puts it, a given context can only incite us to say *something*, it can't compel us to say anything *in particular*.[14]

This is not entirely true of thinking (or of having thoughts). If you see a car you can't help but entertain the idea of a car, though you need not say 'car' at all, to yourself or to others. If you see a person attempting to cross the street between two moving cars and you see that there is not enough space to do so, you can't help but recognise that this is a dangerous situation. You may say so to someone or to yourself, but by the time you say it you have already thought the thought. More likely, you would shout 'watch out', and in this case the thought that you are witnessing a dangerous situation would go unsaid. What's more, in this and many other cases, there is a lot of thinking going on prior to what you end up saying, most of which goes entirely unsaid. You did 'see' that there wasn't enough space between the two cars to cross the street safely, and this required an act of thinking, but you need not say anything to think it.

The situation is mirrored in the laboratory when running experiments on reasoning and problem-solving that specifically require participants to reflect upon the solution they come up with during a task. As mentioned in my book review, I can personally testify to this. I once tested a problem-solving experiment, devised by Sangeet Khemlani and Johnson-Laird, in which you would be presented with a railway track and your task was to reorder the coaches of a train one by one from one side of a track to another.[15] You could move the coaches to a side track too, where the coaches could be stored in order to move them later on. The task had to be followed according to a set of rules and the idea was that if a letter is assigned to each coach and you had to reorder an A-B-C-D-E-F train into a C-A-B-D-F-E one, you would solve the task by dividing the main problem (reordering the whole train) into sub-problems (moving two coaches first, moving some coaches to the side track, some to the other side of the track, bringing one coach back from the side track, etc.). As they go about solving the task, most participants act as if they are following a specific plan of action, but when asked post-experiment to delineate the strategy they followed, hardly anyone is able to explicitly formulate the actual strategy the data suggest they carried out, nor do the participants think that they laid out any strategy to themselves in inner speech. The explanation Khemlani & Johnson-Laird have offered for the data is based on the mental models participants put together to plan a

[13] Alan Baddeley, "Memory", in *The MIT Encyclopedia of the Cognitive Sciences*, eds. Robert A. Wilson & Frank C. Keil (Harvard, MA: The MIT Press, 1999), 514-517, offers a concise description of the study of memory in general.
[14] Noam Chomsky "A Review of B. F. Skinner's *Verbal Behavior*" *Language* 35 (1959), 26-57. See, also, Noam Chomsky, *Language and Mind* (Cambridge, UK: Cambridge University Press, 2000).
[15] Sangeet Khemlani & Philip Johnson-Laird, "Mental Simulation and the construction of informal algorithms", *Proceedings of the Annual Meeting of the Cognitive Science Society* 35 (2013), 2698-2703.

strategy, which, as mentioned, are not linguistic at all; indeed, such plans need not be put into words for one to be able to entertain them and then execute them.

Similarly with some of the rather popular puzzles available on the internet these days, such as the 'how many triangles are there in the picture' variety, where you are shown a big triangle composed of triangles inside triangles of different sizes and you have to work out the overall number of triangles. As you count them all, you may certainly do so in language, but the realisation, if you happen to 'see' the solution, that there are triangles of different sizes, and that in some versions two triangles can combine to form another (a third), is not something you would typically work out by questioning and prompting yourself in inner speech. What this puzzle and the rail track task have in common is that though you can certainly tell yourself the solution, aloud or in inner speech, you *mustn't* do so in order to be able to entertain it. Participants typically 'see' the solution and then just carry it out.[16]

This state of affairs raises well-known problems for psychologists. A psychologist typically designs experiments in the rarefied atmosphere of a laboratory in such a way that irrelevant and unwanted factors are eliminated and participants are exposed to a very constrained set of stimuli. In such a setting, participants can be expected to behave in a specific way – that is, psychologists will make a prediction as to how the participants will respond to the stimuli. This places rather stringent conditions on experiments, but it also yields the most reliable results. Remarkably, the same can't be said of experiments on language production, which have tended to ask participants to repeat well-chosen sentences rather than asking them to produce whatever language comes to mind in the face of whatever stimulus is presented to them. And this is because it is not possible to control the volition or motivation of participants in such conditions. No matter how well-chosen a stimulus, it is well-nigh impossible to predict what a person will say in response, if indeed anything at all.

What may be more productive is the study of the effects of language use while undertaking problem-solving tasks that may depend on linguistic abilities. Particularly relevant here are recent studies employing the speech shadowing paradigm to work out what sort of thoughts can be entertained during an experimental setting in which participants have to repeat linguistic material as they complete a reasoning task. The idea has been that speech shadowing and the resolution of a non-linguistic task argued to depend on linguistic abilities could be vying for the same representations when concurrently undertaken, thereby affecting performance. The methodologies so far employed have varied and the data have been somewhat equivocal, but there is great potential here.[17]

Be that as it may, the notion that inner/outer speech plays a central role in formulating plans of action really must be abandoned – speech cannot be *the* key factor in how we reason or approach problem-solving situations. This remains such a counter-intuitive thought; the feeling

[16] This is not to deny that inner speech may help focusing attention on a specific aspect of the problem, as Jackendoff briefly mentions, but this is obviously a different matter altogether and a much weaker proposition to what Carruthers and Frankish defend.

[17] I discuss various studies on speech-shadowing and problem-solving in David J. Lobina, "Conceptual structure and the emergence of language" *International Journal of Philosophical Studies* 20 (2012), 519-539, and in David J. Lobina & José E. García-Albea, "On Language and Thought: A question of Format", in *On Language and Thought: A Question of Format*, eds. Roberto De Almeida & Lila Gleitman (Oxford, England: Oxford University Press, 2017), 249-273. In recent times, in addition, I have put together a research project that involves a number of experiments on how speech shadowing may affect the representation of logical relations (conjunction, disjunction, negation) by paying special attention to the underlying mental architecture at play in each case, especially the processing cost involved. All I need is proper employment to carry out the research!

that we think, and reflect about what we think, in our own language is both vivid and obvious. ‘Of course I came to this decision by talking to myself about it’, and it is always possible to come up with detailed examples of the dialogues, questionings, and promptings that would accurately underlie many cases of reasoning. But there’s no escaping the fact that putting words to thoughts is not a necessary condition to have such thoughts, and in any case we would have thought it all up before speaking it all, either overtly or covertly.

Just like Chomsky has intimated himself, I ended the review by suggesting that novelists may yield better information on what inner speech is used for, a comment that was meant to be tongue-in-cheek. Nevertheless, it is certainly true that literature has given us stunning examples of streams of consciousness chronicling life-defining and -affirming epiphanies. Did Saul have his in inner speech?

London, 2020