

INNER SPEECH AND THE LANGUAGE OF THOUGHT

Since the influential work by Willem Levelt and colleagues (Levelt 1989; Levelt et al. 1990), one central assumption in some accounts of inner speech is that it requires or presupposes a language of thought (LoT). On this conception, LoT is the linguistic medium needed to generate the input to the entire speech production system. This view also aligns with a general conception of inner speech as involving the expression or vehicle of subpersonal amodal thought (e.g. Bermúdez 2003; Carruthers 2009, 2018; Prinz 2011). Moreover, recent developments see the LoT hypothesis as vindicated across a number of research fields and experimental paradigms in cognitive science (Mandelbaum et al. 2022; Quilty-Dunn et al. 2023). The natural presumption is that inner speech and associated capacities are clear candidates to be added to the list of LoT-supporting evidence.

However, the question of whether inner speech really requires or presupposes a particular representational format deserves careful consideration. First, recent approaches have suggested that LoT is not a commitment of phenomenological or mechanistic approaches to inner speech or to the idea that we – sometimes or often – think in language (Vicente 2022; Kompa 2023). Secondly, fresh attempts to establish LoT on the basis of empirical evidence are in further need of clarification, and have been met with (sometimes staunch) resistance, with Deep Neural Networks (DNNs) standing as well-supported alternatives (LeCun, Bengio, & Hinton 2015; Millière 2024).

The goal of this paper is twofold: (i) first, to provide conceptual clearing of the ground on the connection between inner speech and the LoT; (ii) second, to outline a functional, empirically informed answer to the question of whether inner speech presupposes or in any way underscores a LoT.

Regarding (i), articulation of the exact connection needs to take into account the characterizations of LoT and inner speech on the market. The LoT hypothesis, introduced by Jerry Fodor (1975, 2008), posits that thinking occurs in a mental language (or *Mentalese*) which is structurally similar to natural language. We will introduce the view that LoT and competing hypotheses are better seen as affording integrated multi-level explanations of cognitive phenomena in a given domain. A multi-level approach strengthens the plausibility of the LoT by integrating explanations across (roughly Marrian) levels in ways that meet the *demarcation constraint* (LoT and non-LoT formats can coexist and their explanatory import identified), the *specificity constraint* (different LoTs may be postulated depending on the specific domain of application), and the *empirical constraint* (empirical evidence must be possibly gathered in favor of, or against a particular LoT when set against alternatives). Unlike the property-based approach favoured by Quilty-Dunn et al. (2023) this account requires the proper delineation of explanatory levels, and in particular the functional, the algorithmic, and physical realization levels.

Inner speech (IS), on the other hand, broadly refers to the phenomenon of talking to oneself silently, one prominent kind of inner speech being verbal thinking or thinking in words. Yet the connection between IS and LoT needs

to heed a diversity of characterizations of IS as well. The commitment to LoT disintegrates, for instance, if IS is considered to be only, or essentially, an introspectively accessible kind of experience keyed to a particular natural language. If restricted to a personal and conscious level, IS would be compatible with *any* underlying mechanism that results in that experience.

However, IS has been also defined by appealing to the speech production mechanism, where inner speech is the result of a multi-stage process of linguistic articulation, going from the formulator level to the semantic/syntactic and auditory-phonological levels. In Levelt's model, LoT structures must be retrieved at the conceptualization stage as part of a multilayered process which gradually incorporates features of outer speech. From this perspective, LoT has a distinctive functional role to play at the "preverbal" stage. In this case, however, the proposal has to address the question of "overdetermination", given that it would imply the postulation of two different mechanisms for explaining how verbal thinking acquires its contents— a LoT and a speech production route.

On a more liberal interpretation of Levelt's model, IS only requires episodes that substantially engage the speech production system, so that episodes of thinking could result without requiring a LoT (Kompa 2023). The landscape of available characterizations of IS is however rich (Gregory and Langland-Hassan 2024). The present suggestion is that only if IS has associated a more specific functional role, can the connection with LoT structures be meaningfully stated.

In relation to (ii), therefore, we will examine a functional account of IS along these lines, where a particular cognitive function is associated with IS experience or IS production mechanisms. From this perspective, a commitment to LoT is possible depending on the representational format that is required in cognitive processing to fulfill a target cognitive function.

The idea that IS is associated with cognitive functions is widely accepted at least since Vygotsky's work (1934). The assessment of whether these functions are LoT-supporting requires minute demarcation of the functions at hand. An initial hypothesis to be considered is that LoT is more plausible in some cognitive tasks (e.g. problem solving, memory recall) than others (e.g. motivation, self-regulation).

Even then, the ultimate criterion for a significant commitment is empirical and degree-like. We may find cases in which inner speech functions are likely subpersonally realized in a LoT, as well as in alternative non-symbolic representations, including associations or computational DNN systems, which have been successfully applied to system-1 processing, language translation, or image recognition.

The resulting approach would steer a middle course between those who tend to presuppose LoT to be a condition on IS – as suggested by Levelt's influential account – and those who tend to presuppose that IS and the corresponding notion of "thinking in language" is a phenomenon to be reductively explained in terms of format-neutral natural language capacities.

References

Bermúdez, José Luis, 2003, *Thinking without Words*, (Philosophy of Mind Series), Oxford/New York: Oxford University Press.

- Carruthers, P. 2009. How We Know Our Own Minds: The Relationship between Mindreading and Metacognition”, *Behavioral and Brain Sciences*, 32(2): 121–138.
- Fodor, J. A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. A. 2008. *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.
- Gregory, C., and Langland-Hassan, P. 2024. Inner Speech, In. E.N. Zalta & U. Nodelman (eds.), *The Stanford Encyclopedia of Philosophy*., URL = <<https://plato.stanford.edu/archives/sum2024/entries/inner-speech/>>
- Kompa, N. 2023. Inner speech and ‘pure’ thought – Do we think in language?
- Kompa, N., and J. L. Mueller. 2020. How abstract (non-embodied) linguistic representations augment cognitive control. *Frontiers in Psychology*.
- LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep learning. *Nature* 521(7553), 436–444.
- Levelt, W. J. M. 1989. *Speaking. From intention to Articulation*. Cambridge MA: MIT Press.
- Levelt, W. J. M., A. Roelofs, and A. S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–38.
- Mandelbaum, E., Dunham, Y., Feiman, R., Firestone, C., Green, E.J., Harris, D.W., ... & Porot, N., Quilty-Dunn, J. (2022). Problems and mysteries of the many languages of thought. *Cognitive Science* 46, e13225.
- Millière, R. (2024). Philosophy of cognitive Science in the age of Deep Learning. *Wires Cognitive Science*, e1684.
- Prinz, J.J., 2011. The Sensory Basis of Cognitive Phenomenology. In T. Bayne and M. Montague (eds.), *Cognitive Phenomenology*. Oxford/New York: Oxford University Press, 174–196.
- Quilty-Dunn, J., N. Porot, and E. Mandelbaum. 2023. The best game in town: The re-emergence of the language of thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences* 46, e261.
- Vicente, A. 2022. Speaking for thinking: “thinking for speaking” reconsidered. In P. Fossa (ed.), *Inner Speech, Culture & Education*. Springer.
- Vygotsky, L. S. 1934. *Thought and Language*. (Revised and translated by A. Kozulin, 1986). Cambridge, MA: MIT Press.