

Inner Speech in Action*

Víctor Fernández Castro

Universidad de Granada

vfernandezcastro@ugr.es

Abstract

This paper assesses two different approaches to inner speech that can be found in the literature. One of them regards inner speech as a vehicle of conscious thought. The other holds that inner speech is better characterised as an activity derived from social uses of its outer counterpart. In this paper I focus on the explanatory power of each approach to account for the control of attention and behaviour in the context of executive tasks. I will argue that the vehicle view cannot capture some central cases of inner speech in executive tasks because they cannot be described as cases of bringing thought into consciousness. Then I will offer a revised version of the activity view and I will apply it to some examples so as to show that it is better posited to account for them. I end by considering two objections to the activity view and a possible way to address them.

Keywords: inner speech, attention, activity view, vehicle view, executive functions

1. Introduction

A central challenge of philosophy of mind and cognitive sciences is to account for the cognitive roles that language plays in human mind. Language seems to be closely bound to several human cognitive capacities. On the one hand, language is vital for the acquisition of some unique human mental capabilities such as inferential capacities, sophisticated social cognition, or abstract reasoning (Astington and Baird, 2005; Gentner and Goldin-Meadow, 2003). On the other hand, it amplifies and modifies the functioning of some capacities that we share with other animals, for instance, spatial cognition, flexibility, and decision-making (Carruthers and Boucher, 1998; Evans 2008; Gomila, 2012). In any case, what seems to be far from obvious is that language possesses

* ¹ The author wishes to thank the comments by Fernando Martínez Manrique, and the members of the Philosophy department 1 at Universidad de Granada. Research for this paper was funded by the Spanish Government through Research Projects FFI2011-30074-C01-C02 and the fellowship FPI BES-2012-052157

important functions in human cognition apart from its communicative use.

Such research has brought with it the interest in the phenomenon of inner speech. Inner speech is usually defined as the phenomenon of experiencing that one is talking silently to oneself. This phenomenon seems to be related to some cognitive aspects as well, for instance, consciousness, decision-making, or attention. The interest in inner speech started with the work of the Soviet psychologist Lev Vygotsky. He noticed that children talk out loud with non-communicative purposes, a phenomenon that is usually called ‘private speech’ (Vygotsky, 1981). They use this private speech as a cognitive tool to regulate their behaviour and attention. Vygotsky postulated that this non-communicative use of language was a previous stage in the acquisition of inner speech. Since the decade of the sixties when Vygotsky's work was translated into English, socio-cultural approaches to the mind have been providing evidence about private and inner speech, and their relation to the cognitive development of children. This research program attempted to demonstrate that private speech, as a precursor of inner speech, played a mediatory function in the acquisition of psychological executive capacities such as self-regulation of behaviour and attention, or the control over some cognitive processes (Díaz and Berk, 1992; Winsler et al. 2009) A well-studied function of inner and private speech is precisely this capacity of regulating attention and behaviour.

The increase of interest in the work of Vygotsky provoked the appearance of an enormous quantity of investigation about private speech and its role in development. For example, the social origins of private speech have been tested in empirical research that compares social speech used by caregivers and by their preschool children working with them in some tasks (Bronson, 2000; Nelson, 1996). In those studies, children exhibit private speech episodes reproducing the instructions given by their caregivers. Another source of evidence supporting the relationship between private speech and attention comes from children diagnosed with Attention Deficit Hyperactivity Disorder. They use more overt private speech when compared to same-age children in tasks such as Lego construction, or selective attention (Winsler, 1998). Besides its functions in children, inner speech seems to sculpt attention by structuring perceptual experience in experts (Kirsh and Maglio, 1994). Considering this, the influence of inner speech in regulating attention and behaviour seems to be well established.

In spite of the agreement about the evidence, we can distinguish two different approaches for understanding the phenomenon of inner speech: *the vehicle view* and *the activity view*. According to the former, the cognitive functions of inner speech are grounded on its representational vehicle, which allow us to have conscious access to our thoughts (Jackendoff, 1997; Clark, 1997; Bermúdez, 2003). In this view, inner speech has a proper specific function: being the vehicle for conscious thoughts. According to the latter, inner speech is an activity derived from social uses of language in

outer speech. From this perspective, inner speech has a multiplicity of functions, which are inherited from outer speech. In other words, functions of inner speech, such as evaluating ourselves, attending thoughts, or motivating ourselves, derive from outer speech uses of evaluating, entertaining, and motivating. The main concern of this paper is to evaluate the explanatory capacity of both approaches to account for the regulation of behaviour and attention. After analysing both views, I will argue that the activity view is better posited to explain the capacity of inner speech to gain control over attention and behaviour in executive tasks. Then, I consider two possible objections to the activity view. Although I do not give a definite answer to those objections, I state a possible way to solve them.

2. The Vehicle View of Inner speech

In cognitive science and philosophy, we can find several manners of characterising the phenomenon of inner speech. A widespread characterisation claims that talking silently to oneself enhances the capacity of having access to our own thoughts. Language exhibits features that make it an appropriate *vehicle* to codify thought so that it can be object of farther thoughts. In other words, language is what Clark (1998) calls a facilitator for *second-order dynamics*. Among other things, rehearsing sentences privately allows us to self-ascribe beliefs or desires by inferring them from the meaning of the sentences. Following Jorba and Vicente (2014), I will call this approach *the vehicle view*². This view is characterised by two claims. Firstly, the basic feature of inner speech is its representational character, so that it is recruited for its cognitive purpose due to its appropriate vehicle. Secondly, the basic function of inner speech is to codify thoughts in order to make them available for consciousness. Although those two claims are shared for most of the vehicle view defenders (Clark, 1998; Jackendoff, 1998; Bermúdez, 2003), there are differences regarding what properties are relevant to make language an appropriate vehicle for conscious access.

For instance, Jackendoff and Bermúdez defend that language is the only accurate vehicle to bring propositional thoughts to consciousness. As Jackendoff put it:

“Language is the only modality that can present to consciousness abstract parts of thought like kinship relations, reasons, hypothetical situations and the notion of inference. Only through language can such concepts form parts of experience rather than just being the sources of intuitive urges” (Jackendoff, 1997: 200)

² Jorba and Vicente (2014), and Martínez-Manrique and Vicente (2015) employ the label ‘format view’ to characterise this approach. Although the terms have different connotations, the distinction is irrelevant for the purposes of this paper

Although Jackendoff and Bermúdez share a similar view about the general function of inner speech, they focus on different linguistic features that underlie this function. According to Jackendoff, linguistic items are accurate for conscious access because of its phonetic nature, since all consciousness is perceptual. On the other hand, Bermúdez asserts that language is the only vehicle that allows us to access propositional thought among its competitors. Contrary to the language of thought, natural language is a personal-level vehicle, that is, something that can be consciously accessed and employed by an agent.. And contrary to images or cognitive maps, it is appropriately structured.

In another version, inner speech is the appropriate vehicle for second-order dynamics because of its context-independency and neutral-modality:

“[Language is] a type of code which minimizes contextuality (most words retain more-or-less the same meaning in different sentences in which they occur), is effectively modality-neutral (an idea may be prompted by visual, auditory or tactile input and yet be preserved using the same verbal formula), and allows easy rote memorization of simple strings” (Clark 1997: 210)

According to Clark, regular cognition is amorphous and context-dependent. Anchoring sentences in our working memory makes thoughts context-independent and introspectable. That is what makes language the social scaffolding for cognition³.

Notwithstanding differences, those views share the claim that language allows us to access and manipulate our own thoughts. The virtue of the vehicle view resides in its capacity for explaining straightforwardly some well-known cognitive functions of inner speech. When linguistic information is anchored in our working memory, we can gain control over our attention and behaviour by manipulating the thoughts that this information conveys. Gaining control over attention and behaviour provides the agent with some cognitive advantages: Firstly, inner speech retrieves decontextualised information, thus, we can reason with, or consider, mental states which are independent of the proximal environment. As such, rehearsing inner utterances contributes to cognitive flexibility and anticipation. Furthermore, the vehicle view accounts for the improvement

³ Clark's view differs from those of Jackendoff and Bermúdez. In fact, his view could be considered an activity view as far as he considers that inner speech helps to reproduce the dynamic of external communication in order to enhance cognitive processes. (Clark 2004). In addition, he states a variety of uses of inner speech apart from anchoring thoughts in working memory. However, his concerns with the extended mind hypothesis motivated his emphasis in the vehicle approach which I attack here.

of behavioural control. Accessing self-generated thoughts enhances our focus of attention on belief and desires that can be altered in order to cause specific patterns of behaviour. By producing phonological representations we bring thoughts to our consciousness because our comprehension mechanisms access the semantic properties of those representations⁴. Then we can feed those contents into our cognitive mechanisms in order to specific behaviours.

Secondly, the information linguistically articulated can be compared with visual information. In coherence with the standard theory of working memory (Baddeley, 2007), the *central executive system* can compare the information retrieved by the *phonological loop* with visual information of the visual-sketchpad. This offers the capacity of testing our own belief with information from the environment, or with visual information retrieved from long-term memory. The vehicle of inner speech is appropriate to be hold by working memory and, therefore, it can be contrasted with other types of information.

Finally, the above quotation of Jackendoff emphasises how language can present subparts of a thought to consciousness. Focusing attention on a specific subpart of a sentence permits us to pay attention to its mental counterpart. This characteristic of the vehicle view is especially important for defenders of parallel distributed processing. According to them, our mind processes information in a distributed manner; mind is basically a pattern-completing device which cannot identify, or work with, discrete parts (Rumelhart & McClelland, 1986). However, as long as we master language and we are able to rehearse utterances internally, we can freeze those patterns in our working memory and pay attention to its different parts through its linguistic structure. In sum, the vehicle view considers that the only function of inner speech, conscious access to our thoughts, is sufficient to explain the roles of inner speech in cognition. In particular, regulation of behaviour and attention seems to be *prima facie* accounted when we consider that inner speech has the appropriate representational vehicle.

3. The Activity View of Inner Speech

In contrast to the vehicle view, the *activity view* holds that inner speech is a form of *activity* (Ryle, 1979; Hurlburt et al. 2013; Jorba and Vicente, 2014; Martinez-Manrique and Vicente, forthcoming). According to this approach, inner speech is something that we do, in contrast to something that happens to us, such as mistakes or slips (Jorba and Vicente, 2014). Considering inner speech as an activity implies that when we speak to ourselves we are doing something with a specific purpose. Inner speech is the inner counterpart of outer speech; therefore, we can say that inner speech has as

⁴ Langland-Hassan (2014) argues that this is problematic for the vehicle view because the phonological representations of inner speech would produce informational content related with the meaning of such phonological string but not information about the world (beliefs). Although I agree with this criticism, I will consider different arguments against the vehicle view.

many functions as outer speech has. Moreover, given that many functions of outer speech cannot be characterised as informing or exchanging information, inner speech cannot be characterised only in terms of recruiting a representational vehicle.

The main precedent of the activity view is Vygotsky's seminar work on thought and language. He defends that inner speech is an internalised version of outer speech. Vygotsky thinks that all high-level mental functions are acquired from social interactions during development. Similarly, inner speech is the internalised version of social linguistic interactions:

"The very mechanism underlying higher mental functions is a copy from social interaction; all higher mental functions are internalized social relationships... Their composition, genetic structure, and means of action – in a word, their whole natures – is social. Even when we turn to mental processes, their nature remains quasi-social. In their own private sphere, human beings retain the functions of social interaction" (Vygotsky 1981: 164).

Whilst Vygotsky applies this approach to every higher mental function, my interest here concerns only those high mental functions which seem to derive from linguistic interaction. They seem to follow a developmental path that can be roughly summarised thus: At the age of 2, children start to communicate with others and to tag their environment. At some time after starting to communicate, they start to talk out loud to themselves in public. This phenomenon is pervasive in pre-school and primary school children and it is what researchers call *private speech*. Private speech highlights the social origin of inner speech, since it seems to be a previous stage in our complete internalisation of language (inner speech). From this developmental trajectory, we can draw the hypothesis that the cognitive functions of inner speech derive from our regular communicative practices. Thus, we can approach the cognitive functions of inner speech by elucidating how some of those speech acts are performed in social contexts.

Appearances notwithstanding, communication is not a unitary action. Although communication has usually been understood as transmission of information, we can perform a highly diverse set of linguistic actions with different purposes. This claim was famously supported by Austin (1962), who baptises as 'the descriptive fallacy' the inclination of philosophers to understand any statement as being a description of a state of affairs. Ryle (1979) captures a similar point in this way:

"Our purpose in asking questions differs from our purposes in praying, scolding, play-acting, multiplying,

and will-drafting. Sometimes philosophers speak as if behind all or most of our sayings there is the monolithic intention to induce beliefs in our interlocutor. But there is no such uniform intention. Sometimes we mean amuse, reassure, or scandalize him, or to evoke sympathy from him; or to reprimand him; and sometimes it is to ourselves, instead of an interlocutor to who we say things, and so on and so on.” (Ryle 1979: 89)

Our saying is not a question of transmitting information from one head to another. Linguistic practices enhance our social coordination, which involves a variety of social actions.

In the above quotation, Ryle associates the purpose of our speech acts with the intentions or means to provoke something in our interlocutor. This emphasises the variety of actions we can perform with language. The conventional outcomes of a linguistic interaction can be very different sorts of things: an emotional response, a specific course of action, a belief, etc. In order to produce a given conventional outcome with a speech act it is necessary to take into account a variety of norms and conventions, including intentions and grammar surface, but also other norms stemming from authority, contextual information, social relationships, or rituals. Leaving aside all the aspects involved in the speech act, it is remarkable that the functions of outer speech cannot be characterised in terms of purposes of a single individual, but in terms of the interactions involved in a given communicative exchange. As Millikan (2005) put it:

“The functions of public language forms are not on the same level as either speaker purposes or hearer purposes taken alone. The conventional functions of language forms are not, for example, merely standard speaker purposes. Conventional language forms are selected for performing services satisfactory at once to both partners in communication. Their functions must balance speaker with hearer interests. Because the conventional function of a linguistic form will remain stable only if it continues to serve the interests of both speakers and hearers often enough, I call it a ‘stabilizing function’”(Millikan 2005: pp. 58).

That means that the function of outer speech is to implement different dispositions in the speaker and hearer. That is, different tendencies of acting, feeling, or believing in certain ways. The importance of this connection between a linguistic action and a subsequent conventional outcome is vital to emphasise that outer speech functions cannot be characterised in terms of the vehicle or semantic content alone. Different types of outer speech pay different cash value. Indeed, it seems that acts of outer speech can be typified in terms of the different dispositions to behave that they are meant to implement both in speaker and hearer. As Austin (1962) demonstrates outer speech functions cannot be reduced to or typified by surface grammatical conventions. We can perform different speech act types with the same grammatical forms, for instance, I can use declarative forms to perform an order ‘Anthony will study biology’ [and that’s it!], a promise ‘I’m visiting you

in a few weeks' or an assertion 'Anthony wears a red T-shirt' [and looks gorgeous]. Thus, we are interested in the function of the outer speech and how it produces some dispositions in speaker and hearers. For instance, how an order provokes certain dispositions to act in the hearer, how an insult provokes irritability or agitation, or how an assertion produces tendencies to infer further beliefs.

Keeping in mind those characteristics, we can address inner speech from a different perspective. According to the activity view, inner speech has a multiplicity of cognitive functions that can be tracked to the functions that outer speech exhibits in social contexts. In the same way that talking to others implements behavioural and cognitive responses in hearers, talking to oneself implements behavioural and cognitive responses in the agent. When I give myself an order, I trigger some patterns of behaviour, or when I am motivating myself I trigger a specific emotional response. In talking to ourselves, we reproduce the same activity with conventional outcomes we usually produce in social actions. We find the same functions of inner speech that we find in outer speech functions. As Martínez-Manrique and Vicente (2015) put it:

“OS [Outer Speech] can play the same cognitive roles as IS [Inner Speech], including the alleged roles related to consciousness [...]. In principle, anything that we tell in IS could be told in OS, and for exactly the same purposes. So if IS had the function of making thoughts conscious, it would certainly not be its proper function but a function of language in general” (Martinez-Manrique and Vicente, 2015:6)

The importance of understanding inner speech in those terms is that it denies that inner speech has a unique proper function. Instead, inner speech has a multiplicity of functions that are related to outer speech functions. Moreover, characterising those functions in terms of a representational vehicle does not capture all this multiplicity because some inner speech functions are not tied to beliefs, desires or mental content in general.

Prima facie, the activity view opens the possibility of covering a bigger range of cases than the vehicle view. Presenting inner speech in those terms implies that we can alter our cognitive capacities with different purposes: we can give orders to ourselves so as to trigger specific patterns of behaviour, we can call our attention to different things, or we can evaluate our own behaviour and mental states. The vehicle view seems to be restricted to the case of accessing our mental states and the cognitive functions that we can derive from it. However, as soon as we take into account the different types of outer speech functions and investigate the different conventional outcomes they perform, we are better positioned to explain more cases of cognitive functions of inner speech. Notice that considering that inner speech has many functions that are derived from outer speech does not imply that all the functions of the outer speech have an inner counterpart. Some functions

like promising or inviting may require being public and having more than one interlocutor involved. I will say something else about this in the last section.

4.Inner Speech and Attention

Once I stated both views, I will analyse the explanatory power of both theories in specific contexts. As I said before, this paper is concerned with the role of inner speech in contexts of executive tasks. Executive tasks are those where we need to deploy executive psychological functions such as planning, control of attention, or monitoring. In those tasks we need to fix goals, change attention from task to task, or inhibit external stimuli. Paradigmatic examples are everyday practices such as cooking, repairing a car, or doing a jigsaw puzzle. Let me take one of those examples, for instance, cooking. Inner speech helps us to guide our steps over the recipe and to maintain our attention on the relevant objects of our working space; we tell things to ourselves such as: ‘cut the fruit’ or ‘now, the eggs’. At some point you see the sugar on the table and say ‘oh, the sugar’ because you forgot to put it into the dough. Or you just forget a step in the recipe and wonder ‘what is next?’ Those examples of executive tasks are clear contexts where inner speech makes a difference in regulating our attentional capacities. In this section, I will address the issue of how the two competing views of inner speech deal with ordinary episodes of regulating attention in those contexts.

4.1 Two problems of the Vehicle view

When applying the vehicle view to specific cases, we have to take into account that this approach considers conscious access as the proper function of inner speech. From this view, the regulation of attention or any other function must rely on this proper function in some way or another. At first glance, this is not problematic for the cases of attention since attention and conscious access have been traditionally related⁵ (Chun and Wolfe, 2000; O’Regan and Noe, 2001; Prinz, 2012). For instance, anchoring a string of words in our working memory with a specific semantic content would allow us to retain the goal of a specific task. Coming again to the case of cooking, I can retain my attention to the next step of a recipe by muttering the names of the ingredients I should use next. Or to take another example, I can repeat an instruction to myself (cut the onion smaller) in order to perform the action slowly.

Despite this, when we analyse carefully cases like those we can find two problems for the vehicle

⁵ Recently, many scholars have been questioning this assumption and considering attention and consciousness as different processes which can be manipulated using different tools (see Koch and Tsuchiya, 2007; **Montemayor and Haladjian, 2015**).

view: inner speech is too fragmentary to retrieve complete thoughts that we can attend; and the general approach of the vehicle view does not allow to explain different attentional capacities that seem to be involved in executive tasks. Let me explain this criticism.

In the first place, the cases we have been treating have something in common: the inner speech involved is typically, highly fragmentary. As Martínez-Manrique and Vicente (2010) argue, any theory of inner speech should explain two truisms about our experience:

- (i) “We have phenomenological acquaintance with our inner voice even when we are not rehearsing linguistic actions, and in many cases we do not experience phonological representations of sentences but dispersed linguistic items.
- (ii) We can experience richer, more sentence-like inner speech, typically but not exclusively related to linguistic activities.” (Martínez-Manrique and Vicente, 2010: 143).

In this particular case, our interest concerns the first truism. In our outer speech, regular utterances are affected by a high grade of underdeterminacy (Recanati, 2004; Carston, 2002), that is, the truth-conditional meaning of a sentence varies from context to context. If this phenomenon affects regular complete utterances, we can expect the underdeterminacy to be much higher in fragmentary items that we use in inner speech ('yes', 'help', or 'goes there'). However, we experience such fragmentary items in our inner episodes as being meaningful (Vicente and Martínez Manrique, 2008). Therefore, a theory of inner speech must explain why we understand our inner speech episodes in spite of being so fragmentary. The reason of why this problem is more pressing for the vehicle view is because it considers inner speech as a vehicle for thoughts. In other words, those fragmentary episodes are phonological representations of complete thoughts we access consciously. From this perspective, the only way that our interpretative mechanisms can make sense of our inner speech is through phonological representations. This implies that the mental mechanisms responsible for understanding utterances would need a great amount of contextual information in order to determinate the meaning of such expressions. For instance, if I say to myself ‘the screws’ in a context where I am repairing a socket, this could mean different things: ‘I forgot to buy the screws’, ‘I must use the screws, now’, or ‘Where are the screws?’ Even in the context of an executive task where the contextual information is restricted, the meaning of the expression can vary, and therefore, it is impossible to specify the thought that is brought into consciousness. Notice that this problem can be extrapolated to the cases where we use words to focus our attentional capacities. If the regulation of attention were related to the second-order dynamics, then, the semantic underdeterminacy of inner speech would produce attention mistakes that actually do not occur. For instances, failing to determinate if I am focusing on an action (driving the screws to

close the carburator) or a salient object (the right screw).

The second problem of the vehicle view concerns some intuitions of attention. At this point, one may notice that some of the cases presented above are difficult to describe under the slogan ‘bringing thoughts into consciousness’. When we say to ourselves ‘the red piece’ while doing a jigsaw puzzle, it seems that we are focusing on the piece itself, instead of on our thoughts about the piece. Intuitively, using inner speech lets us gain control over attention in two ways: putting our thoughts under attention, and calling our attention to events or objects in the environment. However, the latter conflicts with the proper function of inner speech that the vehicle view presents. Restricting the inner speech function to mental state access leaves aside cases where inner speech does not seem to function with this purpose, or even where beliefs and desires do not seem to be involved, for instance, those cases where inner speech triggers specific patterns of behaviour or reinforces emotional responses. In general, inner speech alters psychological mechanisms besides manipulating mental states. Concerning attentional tasks, calling our attention to specific aspects of the environment does not necessitate accessing mental states, but rehearsing expressions with the *ostensive power* of pointing out certain objects and events. If inner speech plays a role by increasing our capacity for attention, then it must be by maintaining, changing or calling out our attention to our environment. Let us see an example of private speech where a child is trying to solve a jigsaw puzzle:

“A ver. El triángulo... éste. Este triángulo. Éste. ¿Éste? creo que sí. Aquí. Éste aquí. Mm...Sí. No, el cuadrado. Aquí. Y ahora va esto aquí. ¡Bien! ya tengo un lado. Y ahora pongo este cuadrado aquí. Así. Y... ya está. Ésta es fácil” (De Dios and Montero, 2014: 31)

“Let see. The triangle...this one. This triangle. This. ¿This? I think so. Here. This one here. Mm...Yes. No, the square. Here. And now this goes here. ¡Good! I already did a side. And now I put the square here. Like that. And... that’s all. This one is easy” (De Dios and Montero, 2014: 31)

In this piece of private speech, the references to objects ‘triangle’, ‘square’ do not seem to be cases of attending to thoughts. Because of the situation, we can infer that the agent is maintaining her attention over the triangle or the square. What seems to be at issue is the attentional capacity itself. The function of those linguistic items is to call the attention of the agent to something.

This second problem of the vehicle view concerning attention is an example of a more general problem. According to the vehicle view, when we experience an episode of inner speech we are accessing a semantic content from which we infer a thought we have access to. However, this does

not account of what we do with this thought. Having conscious access to a thought is not thinking this thought. If inner speech has cognitive functions, those functions must be related to cognitive capacities in itself, that is, it must affect thoughts, inferential capacities, attention and so on. However, the vehicle view presents a perspective where inner speech is only a way to attend those thoughts, without any capacity to alter them or alter the cognitive capacities of the subjects (see Martínez-Manrique and Vicente (2015) for a similar point).

4.2 The activity view and attention

In the last section, I presented two problems of the vehicle view to account for the capacity of inner speech to regulate attention. The first problem emphasises the problem of accessing mental states through the fragmentary representational content of inner speech. The second problem is the incapacity of the vehicle view to account for the distinction between paying attention to a thought and paying attention to the environment. In this section, my aim is to explicate how the activity view avoids those difficulties.

As we saw before, maintaining attention to our workspace requires accounting for the ostensive power of the linguistic expressions such as ‘the triangle’ or ‘screws’. According to the activity view, talking to ourselves is something that we do in order to perform a conventional outcome. That is, we tell things to ourselves in order to produce certain responses in ourselves, in the same way that we tell things to others in order to produce certain responses in them. The first problem is how we distinguish those cases of paying attention to a thought from cases where we pay attention to the environment through linguistic items. If inner speech inherits its functions from outer speech, it is expectable to find a parallel distinction in our social linguistic exchanges. In other words, understanding the relation between inner speech and attention requires characterising those expressions with which are able to call attention in social contexts. When we want someone to pay attention to something we usually do things such as describing or asserting: ‘The triangle is red’ or ‘there is a screw there’. However, when we use such expressions in our inner speech it seems that we only call our attention to the semantic content, as the vehicle view defends.

However, there is a different kind of activity that seems to exhibit the features that we need in order to explain how inner speech focuses our attention on the environment. This activity is what Kukla and Lance (2010) call ‘observatives’. Observatives are speech acts which have the function to *express a speaker recognition episode*. Although observatives may be performed with a declarative grammar surface, we have some conventional English expressions to mark them, such as ‘Yo!!’ and ‘Lo!!’ The paradigmatic example used by Kukla and Lance is:

(1) As I see a rabbit dart into a bush I call out, 'Lo, a rabbit!'

Observatives differ from assertions in several respects and it is interesting to see those contrasts in order to emphasise their specific characteristic. Firstly, performing an assertion does not require having perceptual access to an event. I can say 'the cat is on the mat' even if I have not seen the cat. However, someone who has not experienced the perceptual episode cannot perform felicitous observatives:

"Entitlement to the observative (the 'Lo!' speech act) is inherently one's own. In uttering this speech act, I express my receptive recognition of a rabbit. This speech act does not merely make the declarative claim that a rabbit is present. Nor does it merely make the declarative claim that I see a rabbit. Rather, it serves a special recognitive function: it marks or expresses my detection of a rabbit. It is the recognizing, and not just what is recognised or who is recognizing, that is given expression in such a claim, and since what is expressed is the indexed recognition itself, this entitlement is not generalizable, even in the ideal" (Kukla and Lance 2010: 47)

Although sometimes we may perform observations in declarative form, as 'There is a crocodile in the bathroom!', this speech act differs from an assertion in the expression of recognition. However, what it is really important for our purpose is that some observatives do not only express recognition but they also *ostend*: They call the attention of the other agent to notice and recognise what is recognised by the speaker. It is irrelevant whether we are trying to induce a belief in the hearer or not. The primary effect we are trying to obtain is to call her attention to something. In the same vein, when considering this function in the case of inner speech, what we are doing in cases of executive tasks where we need to manipulate to our attention in the environment is *observing things to ourselves*. Notice that when we describe the situations where an observative of this type is at stake, it is weird to say that we are trying to induce a belief in the hearer. We are trying to call her attention to something. In the same vein, it is weird to describe the situation where I say to myself 'the sugar!' as a case where I am bringing a thought to consciousness. What I am doing is something different; it is to call my attention to something, observing something to myself. If an agent is trying to gain control over what she is doing, what she needs is to call her attention to the objects in the workspace, and this requires triggering our attentional capacities instead of accessing specific mental states.

The second problem of the vehicle view has to do with the semantic underdeterminacy of those terms. However, this is not problematic if we abandon the idea that such fragmentary items have complete truth-conditional contents. According to the activity view, we do many things in speaking to ourselves and it is not necessary to have determinate thoughts in order to perform determinate actions. In outer speech, our conventional outcome can be an automatic behaviour, an emotional response or, as in this case, changing the focus of attention. In executive tasks, inner speech is especially fragmentary because we do not need to make explicit a semantic content, i.e., we do not need to attend to a belief or a desire. In outer speech, when we call the attention of someone we can do it without putting explicit linguistic representations into play; in fact, sometimes we only need to point out an object. The same happens with our inner speech. That is the reason why we experience our inner speech as being meaningful in spite of its fragmentary nature in such contexts. Moreover, the activity view is not restricted to the information that we can draw from the phonetic representation of inner speech. On the contrary, the activity view points out to what we do with this speaking, and therefore, pragmatic considerations about intentions, situations and actions are considered to make sense of the speaking episode.

5. Regulation of behaviour

Regulating behaviour is another capacity intimately related to inner speech. Both theoretical positions presented in this paper must contribute to the explanation of how inner speech helps us to control behaviour. The concern with this capacity has been central for inner speech research since the seminal works of Vygotsky. As in the case of attention, the vehicle view approaches the issue through mental state access. Given that the proper function of inner speech is bringing mental states into consciousness, the vehicle view proposes that altering the causes of behaviour such as beliefs, desires and other mental states, is the only manner to control behaviour. The emphasis of the vehicle view resides in how we access mental states that we were not considering before, and thus, how this provokes changes in our mental states that trigger specific patterns of behaviour. For instance, let us imagine I decide to practice some cycling tomorrow because I believe I should start to do some sport. However, when I wake up and I am preparing myself to go for cycling, I tell to myself ‘I have a lot of work this week’. This sentence makes me access the thought: I believe that I have a lot of work this week. Having conscious access to this thought alters my behaviour (I go to get some work done instead of going for cycling). This is an example of how the vehicle view explains the control of behaviour. In a similar vein, rehearsing utterances can modify our behaviour by altering our desires or fixing new objectives (e.g., I want a coffee).

However, as in the case of attention, the vehicle view is flawed as soon as we start to analyse executive contexts. The expression ‘regulating behaviour’ is an umbrella term to cover many different things that we do. For instance, evaluating an action or procedure, guiding, changing our targets, and so on. However, the vehicle view cannot capture all those manners for gaining control over our behaviour. Executive contexts usually require triggering complex courses of behaviour and modulate attentional capacities from task to task. Inner speech plays the role of structuring those complex behaviours, evaluating our behaviour, or triggering specific pieces of behaviour. In order to state my argument against the vehicle view, let us see another example of private speech in an executive context (De Dios and Montero, 2014):

“A lo mejor es así. Pues a lo mejor así. ¡Uf! (Murmura). ¡No sé, es que éste no lo sé! Ay. A ver. Así. Cuadrado. (Murmura). Así me parece que va bien. Los zapatos parecen. (Murmura). Ahí. (Murmura). No sé, es que claro, uno mediano....” (De Dios and Montero, 2014: 31).

“it may be like that. Or it may be like that. Uf! (Mutters). I don't know, I don't know this one! Ay. Let's see. Square. (Mutters). I think it is ok like that, the shoes seem (Mutters). There (Mutters). I don't know, sure, a middle-sized one...”

In this quotation, the infant gains control over her actions by using *guides* (‘It may be like that’) and *evaluations* (‘I think it is ok like that’). As De Dios and Montero argue, high-level tasks increase the use of evaluative expressions in private speech. In outer speech, evaluations and guides are expressions we use to modify the others’ behaviours. As many scholars have been defending in metaethics (Hare, 1952; Blackburn, 1998), what is most important for the use of those expressions is that they are used to guide choice and action. That is, the use of those terms is to provoke actions and behavioural response in the hearer. For instance, if someone says that eating meat is wrong, you do not expect to see her eating a T-bone steak. In fact, the connection with the behaviour entitles you to protest if you see her eating meat.

However, what is interesting for our purposes is not the moral character of those expressions but its connection with the behaviour. In situations where those expressions are used, it is not always expected that the other will have a specific mental content, but that she will respond in certain ways. This is especially obvious in a pedagogic environment where you are learning a specific skill, for instance, handcrafting or sewing. If a teacher corrects you saying ‘this is wrong’ he does not expect

you to change any mental states but changing your way to proceed or acting. This is characteristic of what we do by instructing, prescribing, evaluating, or guiding. Imagine I attend a martial art course where my teacher gives me some instructions: 'don't put the hand on the floor when you are in someone's guard', 'maintain a short distance with your opponent', and so on. In those cases, it seems weird to say that the intention of the teacher is to induce a mental state in the hearer. He is not trying to induce mental states in you, or inform you about something. The teacher is trying to influence the behaviour of the pupil, trying to make you acquire some dispositions and tendencies to act. In section 2, I defined the activity view as maintaining that the function of outer speech is to produce a conventional outcome. This idea is parallel to the Rylean idea that expressions of natural language are enabling-instruments. In other words, when I learn to use linguistic expressions, I am learning to perform some possible transactions, some possible actions that my use of the word enables in certain contexts: informing, ordering, or commanding. The cases of prescriptives, commands, or instructions are paradigmatic of this view: giving an instruction or a command marks what behaviours are at stake. When my cooking-teacher gives me the instruction of cutting the onion smaller, he is pointing out which pattern of behaviour I must follow. The purpose of some instances of outer speech is not to inform, but to show the hearer which behaviour is in accord with the situation.

Extrapolating all this to inner speech provides us with the theoretical apparatus to approach the control of inner speech over behaviour. Again, in the same way that it would be unsatisfactory to describe those episodes as cases of informing or inducing a thought in the hearer, in the case of inner speech it would be weird to say that we are bringing thoughts into consciousness. What we are performing is an activity. For instance, imagine I am repairing my car. At some point, I realise that I am doing something wrong and I say 'that is wrong'. This speech act makes me stop and reconsider my way to proceed. However, it does not alter my desire to repair the car or even the belief that by doing x I will repair the car. What is modified is the way I am doing x. What I am doing is evaluating my own behaviour. In the same vein, I can give me guides, prescriptions, or commands. When I evaluate my behaviour I do not expect to have access to any cognitive states, I am influencing my behaviour. When I perform a guide or instruction, I expect my speech to have the same conventional outcome that it has in social contexts –in this case, influencing the way to proceed. The vehicle view cannot capture those different functions of inner speech because we need a notion of performative force: we need to distinguish different types of actions we perform with our linguistic inner utterances.

The moral of this section is quite similar to section 4. The vehicle view cannot capture some central

cases of inner speech in executive tasks because they cannot be described as cases of bringing thought into consciousness. Speaking to ourselves is something that we do with specific purposes; we command things to ourselves or evaluate our behaviour in order to achieve goals. Understanding inner speech in those terms explains the multiplicity of functions that inner speech has and why they all cannot be explained by appealing to the representational vehicle of language.

6. Two problems of the activity view

The main goal of this paper is to defend how the activity view pays more dividend than the vehicle view when it is tested in executive tasks contexts. However, the activity view must respond to some obvious problems that arise when we consider that the function of inner speech is derived from outer speech functions. I do not mean to give a complete answer to those difficulties here. Nevertheless, I would like to elucidate some possible ways to think about them.

The cornerstone of my view relies on approaching cognitive functions of inner speech from outer speech. Putting inner talk to work requires recruiting the same cognitive dynamics of outer speech. However, there are some cases of inner speech that do not have an obvious social counterpart. Ryle puts the finger on this weakness when considering the case of asking. Asking something to ourselves is an often practice in our everyday life, we usually wonder where we left the keys or when we should go for vacation. Nevertheless, in our outer speech, making questions is usually a way to demand information. This seems to be a counter-example to our view; after all, it seems weird to demand information to oneself. Ryle noticed that our inner questions have a great similarity with the questions we perform with heuristic intentions. Ryle proposes to consider cases where a teacher asks questions to his students with a heuristic purpose. When a teacher asks questions to a pupil, he is not demanding information but trying to test her abilities. If she fails to respond, he can spell out a different question to give cues or lead her to find the error of her answers. The reason is that, in some sense, the teacher knows that his pupils can find the correct answer. Something similar happens when we ask a question to ourselves, we know that we are able to get the answer or even that we know it but we have forgotten the answer for a moment. The questions are used as a heuristic or cue for retrieving our memories.

Another possible objection to the view is that some functions we deploy in outer speech are not deployed in inner speech because there are some material conditions that are required for outer speech that cannot be reproduced in our inner speech. For instances, while performing an order in social context requires the speaker to have certain authority, no authority is at issue when we order

something to ourselves. This could cast some doubts on the activity view. Either we accept that when order something to ourselves we are doing something different, and then, we have counterexamples against the idea that inner speech is derived from outer speech; or we need an explanation of what happens in those cases where the material conditions required in outer speech are not the same in inner speech.

In order to discern a possible solution to those problems I would like to emphasise some ideas concerning the nature of inner speech. Firstly, I argued that inner speech is derivative from outer speech. This does not mean that we can use inner speech immediately after starting to talk. If inner speech is an activity, it is required to pass some period of training and drilling. This is probably the reason why the process of internalisation takes time. Performing an activity in different contexts requires adjusting some norms and requirements of the action to the new contexts. Imagine a group of children playing football in a park. Some institutional norms of football require being modified or adjusted to the conditions where the children play. The role of the referee may rely on the group; offside could be abolished because of the dimension of the field; or the goal height needs to be an imagined line because there is no goal itself. Those kinds of adjustments do not make the children play a different game. In the same sense, when we are talking to ourselves we are not doing something radically different from talking to others. We are just reconfiguring and adjusting the practice of speaking in order to enhance our cognitive functions.

Understanding internalisation as a process of adjusting explains why we can expect to have the cases described above: Firstly, functions of inner speech which are derived from marginal cases of outer speech. That is because the cognitive necessities that we have can be very different from those that are necessary to trigger in our hearers. Secondly, we can expect cases of inner speech that undergo drastic changes in the process of internalisation. Again, the inter-psychological plane is quite different from the intra-psychological plane. Therefore, adapting a function from one plane to another requires adapting some characteristic of the activity.

7. Conclusion

The aim of this paper was to analyse the explanatory power of two theories of inner speech. In particular, how they explain the capacity of inner speech for regulating behaviour and attention in the context of executive tasks. Although the vehicle view explains a wide range of cases by pointing a proper function of inner speech, it fails to account for some important distinctions of what we do when we talk to ourselves. Firstly, it cannot capture the intuitive distinction between using inner speech for paying attention to a thought and using it for paying attention to our environment.

Secondly, it cannot capture the distinction between modifying our mental states in order to produce a given behaviour, and modifying this behaviour or the way to proceed. The diagnosis in both cases seems to be same; inner speech has a variety of functions that cannot be described with the slogan ‘bringing thought into consciousness’. On the contrary, understanding inner speech as an activity that is parallel to outer speech allows us to understand this variety of functions. Talking to oneself is something that we do with the purpose of observing, prescribing, commanding, attending, etc.

The activity view is not exempt from difficulties, it has to face two types of counterexamples. Firstly, there are cases where the function of the inner speech does not have an obvious counterpart in outer speech. Secondly, some outer speech functions rely on some norms and situations that cannot be reproduced at the intra-psychological level; however, we can find inner speech functions similar to these ones. This seems to show that not all functions of inner speech derive from outer speech. Although I do not have a definite answer to those problems, I argued that understanding internalisation as a process of adjusting functions from the inter-psychological level to the intra-psychological level can shed light to those issues.

References

- Astington, J. W. and Baird, J. A. (eds.). 2005. *Why language matters for theory of mind*. New York: Oxford University Press.
- Austin, J. 1962. *How to do things with words*. Cambridge: Harvard University Press.
- Baddeley, A. D. 2007. *Working Memory, Thought and Action*. Oxford: Oxford University Press.
- Bermúdez, J. 2003. *Thinking without words*. Oxford: Oxford University Press.
- Blackburn, S. 1998. *Ruling passions*. Oxford: Clarendon Press.
- Bronson, M. 2000. *Self-regulation in early childhood*. New York: Guilford Press.
- Carruthers, P. and Boucher, J. (eds.) 1998. *Language and Thought: interdisciplinary themes*. Cambridge: Cambridge University Press.
- Carston, R. 2002. *Thoughts and utterances*. Oxford, U.K.: Blackwell Pub.
- Chun, M. M. and Wolfe, J. M. 2000. “Visual attention”. In E. B. Goldestein (eds), *Blackwell’s Handbook of Perception*. Oxford: Blackwell publications, 272–310.
- Clark, A. 1997. *Being there*. Cambridge, Mass.: MIT Press.
- Clark, A. 1998. “Magic words: how language augments human computation”. In P. Carruthers and J. Boucher (eds.), *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press, 162–83.
- Clark, A. 2004. “Is language special? Some remarks on control, coding, and co-ordination”. *Language Sciences*, 26(6): 717-726.

- De Dios, M.J. and Montero, I.G. 2014. “¿Qué se dicen los niños al afrontar tareas? Análisis lexicométrico del discurso motivacional y variables diferenciadoras”. *Avances en el estudio de la motivación y de la emoción: VIII Simposio de la Asociación de Motivación y Emoción*, Publisher: Universidad de Granada: 27-32
- Diaz, R. and Berk, L. (eds.). 1992. *Private speech: From social interaction to self-regulation*. Hillsdale, N.J.: L. Erlbaum.
- Evans, J. 2008. “Dual-processing accounts of reasoning, judgment, and social cognition”. *Annual Review of Psychology* 59: 255–78.
- Gentner, D. and Goldin-Meadow, S. (eds.). 2003. *Language in mind*. Cambridge, Mass.: MIT Press.
- Gomila, A. 2012. *Verbal minds*. Amsterdam: Elsevier.
- Hare, R. 1952. *The language of morals*. Oxford: Clarendon Press.
- Hurlburt, R. T., Heavey, C. L., and Kelsey, J. M. 2013. “Toward a phenomenology of inner speaking”. *Consciousness and Cognition*, 22: 1477-1494.
- Jackendoff, R. 1997. *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.
- Jorba, M., and Vicente, A. 2014. “Cognitive phenomenology, access to contents, and inner speech”. *Journal of Consciousness Studies*, 21(9-10): 74-99.
- Kirsh, D. and Maglio, P. 1994. “On Distinguishing Epistemic from Pragmatic Action”. *Cognitive Science*, 18(4): 513-549.
- Koch, C. and Tsuchiya, N. 2007. “Attention and consciousness: two distinct brain processes”. *Trends in Cognitive Sciences*, 11(1):16-22.
- Kukla, R. and Lance, M. 2009. “‘Yo!’ and ‘Lo!’”. Cambridge, Mass.: Harvard University Press.
- Langland-Hassan, P. 2014. “Inner Speech and Metacognition: In Search of a Connection”. *Mind & Language*, 29(5): 511-533.
- Martínez-Manrique, F. and Vicente, A. 2010. “What the...! The role of inner speech in conscious thought”. *Journal of Consciousness Studies* (17) 141–167.
- Martínez-Manrique, F. and Vicente, A. 2015. “The activity view of inner speech”. *Frontiers in psychology* 6 (252).
- Millikan, R. 2005. *Language: a biological model*. Oxford: Clarendon Press.
- Montemayor, C. and Haladjian, H. H. 2015. *Consciousness, Attention, and Conscious Attention*. Cambridge: MIT Press
- Nelson, K. 1996. *Language in cognitive development*. New York: Cambridge University Press.
- O’Regan, J. K., and Noe, A. 2001. “A sensorimotor account of vision and visual consciousness”. *Behavioral and Brain Sciences*. 24: 939–1031.
- Prinz, J. J. 2012. *The Conscious Brain: How Attention Engenders Experience*. Oxford University Press.

[SUBMITTED VERSION] Fernández Castro, V. (2016) Inner Speech in Action *Pragmatics & Cognition* 23(2): 238-258. DOI: 10.1075/pc.23.2.02cas.

Recanati, F. 2004. *Literal meaning*. Cambridge: Cambridge University Press.

Rumelhart, D. and McClelland, J. 1986. *Parallel distributed processing*. Cambridge, Mass.: MIT Press.

Ryle, G. 1979. *On thinking*. Totowa, N.J.: Rowman and Littlefield.

Vicente, A. and Martínez Manrique, F. 2011. "Inner Speech: Nature and Functions". *Philosophy Compass*, 6(3): 209-219.

Vicente, A. and Martínez-Manrique, F. 2008. "Thought, Language and the Argument for Explicitness". *Metaphilosophy*, 39(3): 381-401.

Vygotsky, L. 1981. *Thought and language*. Cambridge: MIT Press. (Original work published 1934)

Winsler, A. (1998). "Parent-Child Interaction and Private Speech in Boys With ADHD". *Applied Developmental Science*. 2(1): 17-39.

Winsler, A., Fernyhough, C. and Montero, I. (eds.). 2009. *Private speech, executive functioning, and the development of verbal self-regulation*. Cambridge: Cambridge University Press.