How Far Does the User-Illusion Go? Dennett on Knowing What We Think

Agustín Vicente
Marta Jorba

RESUMEN

El libro de Dennett *From Bacteria to Bach and Back* desarrolla una fascinante historia respecto de cómo nosotros, miembros de la especie *homo sapiens*, nos hemos convertido en criaturas que no sólo actúan por razones, sino que también pueden comprender esas razones. De acuerdo con Dennett, esta capacidad de representarnos nuestras razones depende en gran medida en nuestra capacidad para internalizar intercambios lingüísticos en la forma de habla interna. Por otra parte, Dennett mantiene que el auto-conocimiento es una ilusión. Argumentamos que el punto de vista de Dennett sobre el papel del habla interna entraña en última instancia que existe un límite a los poderes que tienen las apps de interfaz para engañar a sus usuarios. En particular, tanto la fuerza como el contenido de un pensamiento son directamente cognoscibles, tanto si el pensamiento se presenta con el ropaje de una imagen auditiva, como si no. En (al menos) tal caso, mantenemos que no hay ilusión del usuario. La app acierta con lo que hay.

PALABRAS CLAVE: *ilusión del usuario, auto-conocimiento, lenguaje interno, pensamientos desnudos*.

ABSTRACT

Dennett's *From bacteria to Bach and back* develops a fascinating story as to how we *sapiens* became creatures that not only act for reasons but can comprehend such reasons. According to Dennett, this ability of representing our reasons to ourselves, depends mostly on our ability to internalize linguistic exchanges in the form of self-talk or inner speech. On the other hand, Dennett holds that self-knowledge is illusory. We argue that Dennett's view about the role of inner speech ultimately entails that there is a limit to the deceiving powers of user-friendly apps. In particular, both the force and the content of a thought event are directly knowable, whether or not the thought comes in the clothes of auditory imagery. In (at least) such a case, we submit, there is no user-illusion. The app gets things right.

KEYWORDS: *Dennet; User-Illusion; Self-Knowledge; Inner Speech; Bare Thoughts*.

I. OUR MINDS AND THE ROLE OF LANGUAGE

Dennett's *From bacteria to Bach and Back* develops a fascinating story as to how we *sapiens* became creatures that not only act for reasons but can comprehend such reasons. The difference between acting for reasons and comprehending reasons is nicely illustrated when Dennett compares acting in accordance with a practical syllogism and running a practical syllogism in one's own mind [p. 296]. According to Dennett, the key step in the progression from Popperian creatures (creatures that can already test possible behaviours offline) to Gregoryan creatures (creatures that can understand their own reasons), is *language*. The infection of human brains with words (considered by Dennett a special kind of memes), made such transition possible. This is how Dennett's sums up his view:

> The incessant torrent of self-probing and reflection that we engage in during waking life is what permits us, alone, to comprehend our competences and many of the reasons for the way the world is. Thanks to this infestation of culturally evolved symbiont information structures, our brains are empowered to be intelligent designers, of artefacts and of our own lives [p. 370].

All living beings act for reasons. They all have goals that explain their behaviours. However, not all living organisms or parts thereof are equally related to the reasons that move them. Some living structures (Darwinian creatures) have their reasons hard-wired by natural selection; some (Skinnerian creatures) are open to modifying reasons by trial and error; a yet third group of creatures (Popperian creatures) can modify reasons by testing hypotheses offline (by imagining possible consequences of behaviors); and then it is us, Gregoryan creatures capable of meta-representation, i.e., of representing to ourselves the reasons that move us and of acting accordingly[1].

According to Dennett, this ability of having competence with comprehension, i.e. of representing our reasons to ourselves, depends mostly on our ability to internalize linguistic exchanges in the form of self-talk or inner speech[2]. By engaging in internal conversations, asking ourselves questions that we ourselves respond to, we get to think not just about "food, shelter, doors, containers, dangers and the other affordances of daily life but also about *thinking about* food and shelter, and about *thinking about thinking about* food and shelter" [p. 300, Dennett's italics]. In inner speech we produce representations that can be the object of further representations, thus giving us the power to think about what we think about the world around us. Once a creature can do this, it can enter into

what Clark (1998) called "second-order dynamics", gaining a kind of control over its behaviour with unprecedented parallel. We can ask ourselves whether whatever we are going to do next is ok, whether our reasoning is good or not, and even why we think the things we think. Then we can act reflexively. In all these cases, the basis of our self-reflective abilities is that we put our thoughts into words, thereby getting to know what our thoughts/reasons are[3].

Dennett wants us to think that this use of language is the result of downloading an app (or a set of apps) into our brains. It is an app or set of apps that install a virtual machine that users can exploit to enhance their cognitive powers. As all apps, it incorporates a user interface that generates a user-illusion, akin to the illusions that computer interfaces generate: we may think that we are moving a folder in our computer from one place to another, but what actually happens is nothing like that, as there are no folders and no movement at the hardware level. When we do what we describe as moving the folder from one place to another, we can say that we know what has happened in the manifest image, but have no clue as to what has really happened. According to Dennett, the effect of installing an app such as language in our brains is that we gain a kind of comprehension, but only as users of a user-friendly app. We get to know the reasons that are moving us as such reasons are presented to us. But we need to remember that the way our brain presents reasons to itself is illusory, even if it is the only way that a brain can reach the level of self-management that it reaches. This is the point we will discuss: we want to argue that by using the linguistic app we get to know something about what our brains are doing. In particular, we get to know what the brain is representing. There is no user-illusion concerning thought-contents or, as Dennett puts it, reasons. The representation of thoughts is a faithful representation[4]. Before that, though, we will briefly expand on the general picture presented by Dennett and comment on the user-illusion account.

II. THE USER-ILLUSION VIEW

Dennett argues that the manifest image at large, and in particular the part that concerns ourselves, is a user-illusion. Everything that we regard as part of our conscious life, from experiences (e.g. pain or agency) to emotions and conscious thinking is explainable in terms of a shortcut that the brain has found to interact with itself. This is an idea that one can find in different authors. For instance, Thomas Metzinger (2009) also holds that our manifest image is just the way the super-complex sys-

tem that is the brain represents itself within the model of the world it creates. What seems distinctive of Dennett is the idea that the self-model of the brain develops out of the use of words: the brain starts representing itself in a user-friendly way only after language is acquired and internalized. Thus, immediately before the quote above from page 370, he says:

> Human consciousness is unlike all other varieties of animal consciousness in that it is a product in large part of cultural evolution, which installs a bounty of words and many other thinking tools in our brains, creating thereby a cognitive architecture unlike the "bottom-up" minds of animals. By supplying our minds with systems of representations, this architecture furnishes each of us with a perspective -a user-illusion- from which we have a limited, biased access to the workings of our brains, which we involuntarily misinterpret as a rendering (spread on the external world or on a private screen or stage) of both the world's external properties (colors, aromas, sounds,…) and many of our own internal responses (expectations satisfied, desires identified, etc.) [p. 370].

We won't discuss in detail some of the ideas implied in this view, but we would like to raise a couple of brief remarks regarding the view that consciousness is an illusion. One first concern with this claim is that the referent of the term 'consciousness' is not really further specified: is Dennett just talking about phenomenal consciousness in the sense of what-it-is-likeness, or rather *qualia* (as a necessarily intrinsic property) or access consciousness (as cognitive accessibility), or perhaps experience in general, or the self? The illusion hypothesis seems to include in fact many different mental items: consciousness, inner speech, varieties of self-stimulation and reflection, or the self [pp. 345-346]. It thus seems to be a *general* view on not only consciousness but on the whole mental domain. So it would help to know what notion of 'consciousness' is Dennett operating with in order to better understand his position. Be it as it may, we could say that under one extended use of 'consciousness', claiming it is an illusion involves a conceptual tension. An illusion is something that appears differently from what it really is. The idea is that consciousness, or subjective experience is all a matter of neural and brain circuits (or whatever the NCC are) but they appear to us in the form of thoughts, beliefs, perceptions, etc. Notice, though, that the appearance/reality distinction is not available for a certain concept of experience or consciousness, namely, experience as the *locus* of appearance of things, our way of accessing things. Pain, for instance, is a feeling, and so if you have a pain experience there is no space for claiming that you don't have a pain (you can perhaps be confused about certain aspects of

the pain). If experience *is* this, and is realized in the brain, why should we accept that the realization *is* what experience is and that whatever we think experience or consciousness is, just amounts to an illusion? There seems to be a certain conceptual tension in maintaining that the field of appearance itself –experience– is an illusion: in consciousness (in this sense), what appears to be is what it is (see also Strawson (2010) and Searle (1992) on this point), independently of whether it is reducible or not to neural processes.

Moreover, for Dennett, the mental domain as a whole is the realm of the *manifest image*, and everything in this domain is also part of the user-illusion. However, one could claim that we can also make relevant distinctions within the manifest image that this view doesn't seem able to accommodate: perceptions/illusions, true/false beliefs, and so on. It is not just that the manifest image contains appearances, illusions, etc., but the image seems to leave space for such distinctions (and accounts of them) without appealing to the ultimate scientific image. Given that Dennett recognizes several levels within the scientific image –presumably corresponding to the different sciences– why couldn't we actually acknowledge different levels within the manifest image (accounted for by different disciplines) that are explanatorily relevant even if they do not amount to the bottom neuronal level?

Another puzzling aspect of Dennett's account is the role he seems to attribute to words, or other systems of representation, in the generation of the user-illusion. For it seems that he holds that words can be responsible for making us think that we are experiencing pain. Now, perhaps the experience of pain is a user-illusion –i.e., there is no actual pain anywhere in the real world–, but do words have anything to do with that kind of user-illusion? Wouldn't pain be a user-friendly, direct, way that the brain has to inform itself about the state of the organism it regulates –irrespective of whether the brain has downloaded the linguistic app or not? Why assume that language, or other thinking tools, have to explain all user-illusions? As we say, we will leave issues like these at one side, partly because we are not sure if we have understood Dennett's view on some of these points completely[5].

III. KNOWING CONTENTS

Here we will focus on the issue of *knowing what we think*. Dennett's idea concerning this is that inner speech, that "little voice in the head", is also a user-illusion that the brain employs to keep itself informed about

what it is thinking –whatever thinking really is–. When we (the user end of the brain's relationship with itself) "hear" our own inner voice, we form thoughts like "I am thinking that p", and assume that we know that we are thinking that p. However, what is actually happening in the brain is unknown to us: we are in acquaintance only with a user-friendly representation of what is occurring. In this case, the representation is formed by simulations of sounds or, as Jackendoff (2012) calls them, "pronunciations". Such pronunciations are the digested ways the user receives information about certain cognitive processes that the brain is engaged in. Users interact with such processes via their linguistic app, engaging in self-talk, but at no point do they really know what they are actually doing at the bottom (real) level:

> … there do indeed *seem* to be these items, audible and visible and thinkable, and –setting magic aside– it couldn't seem so unless there were physical tokens in the brain *playing those roles*, but, how, physically, anything in the brain manages to play any of these roles is a topic for future investigation, *not* introspection [p. 185].

Now, the question we want to pursue is: is it true that we do not know *at all* what the brain is doing when we speak to ourselves? In particular, when we say we get to know what we are thinking, i.e. the content of our thoughts, aren't we knowing what our brain is representing –and this, by introspection? If the answer is yes, then we have to conclude that the user-friendly app is not generating an illusion after all –at least with respect to what we think. Moreover, it can be the case that a user-friendly app puts us, users, in direct contact with things that our brains do.

Let us draw an analogy. Suppose that, as Dennett holds, experiences do not correspond to anything that really exist. All these alleged experiences are just intentional objects of beliefs that you have. The user-friendly interface of whatever app the brain has incorporated and is running makes you have such beliefs, but there is nothing that would make your beliefs true. We take it that this means that, e.g., what actually happens when you break your fibula is that neurons in your thalamus and cortex fire (ultimately) in response to the damage, causing two kinds of events: on the one hand, the user is informed of what has happened by means of the experience; on the other hand, a certain behaviour is triggered. Now, the pain experience can be a super-simplified, non-truthful, way to get to know what has happened, but still it can be said to have a content: either representational [Tye (1995)] or imperative [Martínez and Klein (2016)], or even both. This pain experience, presumably a user-

friendly representation, tells the user that the fibula is wrong, that she has to do something with her fibula, or both things at the same time, for instance. In this, the experience plays the *same role* that the underlying reality plays, for neurons can also be said to (ultimately) represent the damage and to trigger the corresponding behaviour. If this is right, then it seems that a certain experience can tell us *something* about what certain physical tokens in the brain represent. That is, the experience does not tell us what *really* goes on in our brains (the neural processes) but it can tell us that whatever goes on in our brains, it represents a certain content.

In the case of inner speech and its role in meta-representation, one could also argue that the pronunciations that users seem to hear in their mind, though misleading them in terms of what it is actually there, still inform users about *the content* of the brain representations. Moreover, it can be argued that users can know what their brains represent by *introspection*. According to Dennett, the linguistic app enables us to "make our competences (somewhat) accessible to users […] who can't know, and don't need to know, the intricate details". So, Dennett seems to agree that thanks to the internalization of speech activities users can get some maybe limited knowledge about their own reasons. Thoughts, however they are formed and vehicled, get translated into strings of phonemes ordered according to the syntax of a natural language, and, as the user accesses such strings, she is able to have some knowledge about what (some of) her brain processes are representing. Thus, according to Dennett himself, the linguistic app seems to be able to (partially, perhaps) reveal the actual nature of our thought-contents. However, we do not see why Dennett seems to have some reservations as to the possibility that the knowledge that we obtain in these cases is exact and why he would deny that users get that kind of knowledge by introspection. Now we will move on to consider the nature of our linguistic thinking and how we know the content of what we think. We will try to show that there is no room for an actual user-illusion when it comes to knowing what we think (or what our brains represent when engaged in propositional thinking).

IV. ON INNER SPEECH AND BARE MEANINGS

IV.1. *Knowing our Inner Speech*

One can distinguish two different types of phenomena in inner speech: on the one hand, we have the pronunciations –the acoustic imagery–; and on the other hand, we have the content of such pronunciations. We think it is difficult to deny that we have introspective access to

the pronunciation side of an inner speech episode. Dennett seems to claim that we lack introspective knowledge even of pronunciations:

> [The extent of your access to your own experience] does not differ much from the access another person can have to those experiences –*your* experiences- if you decide to go public with your account. Your convictions are no doubt reliable but not infallible [p. 351].

We agree that reports based on introspection are corrigible, but we disagree with the idea that they are corrigible because they are in any way accessible from the second-person point of view. The second-person can question a first report about what went on in your mind, and such questioning can have the effect that you revise your report and change it. For instance, suppose someone, A, reports that she just told herself "I wish I could be flying now to Hawaii". B, her friend, asks, "are you sure you told yourself exactly those words? Maybe you did not pronounce the initial 'I'?". She can also query for details: "Could you tell me about the pitch and prosody of your saying? Did it sound like an optimistic wish or more like a desperate one?". A may be puzzled by B's questions, and it may even occur that she cannot answer some of them in that moment. However, B is in a much worse condition. If the questions are answerable, they are answerable only by A. A can get back to her saying and realize that she didn't pronounce the 'I' –she just said "Wish I could be flying now to Hawaii", but it can also happen that she is unable to respond to that question in this occasion. This certainly shows that this mechanism is not infallible and might be subject to all sorts of errors. However, she can train and improve her introspective abilities. An example of how this can be done is Russell Hurlburt's descriptive experience sampling (DES) method. The method involves giving subjects a beeper to carry during their daily activities. The beeper beeps at several moments of the day and the subjects are asked to write down what is going on in their minds in that precise moment. These "brute" annotations are then commented in a series of interviews with researchers and carefully examined in a way such that subjects become "better" at introspecting or telling about their inner activities, potentially removing presuppositions and previous preconceptions[6].

When introspection is understood as some kind of observation that can be always improved (also with external help), we find no reason for which the kind of introspective access we will have to the pronunciation part of inner speech should be denied. Casual introspection is clearly fallible, but the way to gain a better knowledge of the features of our expe-

rience is by introspecting better. That is, introspection is the method of knowing the features of our inner speech experiences (and of experiences in general). Perhaps Dennett could agree with this view, since it denies that casual introspection will tell us everything about a certain experience, which many times seems to be the precise point Dennett also wants to reject. In fact, it is interesting to note how close the DES method might be to the "spirit" of Dennett's proposed method of "heterophenomenology":

> Collaborating with other investigators on the study of your own consciousness (adopting, if you like, the "second-person point of view") is the way to take consciousness, as a phenomenon, as seriously as it can be taken [p. 351].

However, Hurlburt and colleagues take the results of the DES method to be information about our *real* experiences obtained by introspection, i.e., experiences only accessible to the subject. We do not see how the method can be interpreted as showing that "[the extent of your access to your own experience] does not differ much from the access another person can have to those experiences".

We also disagree with Dennett (as well as with Carruthers and others) with respect to the issue of knowing the *content* part of our inner speech episodes. We want to argue that we can know the content of our inner speech episodes also by introspection. That is: an educated observation of what goes on in our minds can tell us what our brains are representing. That is, in principle, we can get full introspective access to the content of our brain representations. It can even be argued that being wrong about the content of an inner speech episode is more difficult than being wrong about the features of the vehicle. In overt speech at least one usually knows what she herself meant, although she may be wrong about what she actually said –or how she actually said what she said.

In his book, as well as in Huebner and Dennett (2009), Dennett endorses an interpretive account of inner speech. This account, also endorsed by Carruthers (2018), inter alia, has it that we interpret our own inner speech in the same way that we interpret someone else's overt speech. In particular, the model holds that we use the mind-reading system to guess what the speaker intended to convey. In Carruthers' model, inner speech interpretation is facilitated by having easy access to the contextual features that sometimes make other-talk interpretation difficult. Even a fragmentary, subsentential inner speech utterance, such as 'the seminar!' is easily understood as expressing the reminder that there is a seminar in ten minutes because the speaker had that particular seminar in

mind right before she spoke. So the contextual elements are active and accessible, which means that there is no much room to get it wrong. Once the message has been interpreted, the content is bound to the inner speech event and broadcasted together with its auditory features. The result is that we "hear" the pronunciation *and* the content. However, according to Carruthers, the content is not experienced. Only the perceptual part of the inner speech episode is experienced; the content part is just accessed, that is, it is information available to the direct rational control of thought and action [Block (1995); see Jorba & Vicente (2014), for a critique of this view].

Note that even in Carruthers' interpretive model it is possible to hold that, in normal conditions, we do know the content of our inner speech episodes. However, apparently at least, we do not know them by introspection, but by using our mind-reading system of interpretation. In contrast with this, Dennett claims that "our access to our own thinking, and specially to the causation and dynamics of its subpersonal parts, is really no better than our access to our digestive processes" [p. 346]. He adds that our access is in fact to the *result* of the neuronal activity, which is where all the causal interaction happens. Although we agree that we don't have access to the subpersonal process by which thinking is neurologically realized, we would resist the idea that what seems to us in inner speech and thinking is an illusion, precisely because we can know the content of our inner speech (and thinking, as we will see) via introspection.

To develop our point we have to go into some details about how inner speech is produced, at least according to many contemporary theorists. The explanation is that first we form a certain high–level, subconscious intention to express a content that can be more or less structured. We then go through a series of steps, which can include a certain, maybe high, degree of back and forth [Hickok and Poeppel (2007)], where we flesh out that high-level, programmatic intention, giving it form in terms of semantic and syntactic structures, phonological representations, and finally articulatory instructions, or motor commands, to produce sounds. This is the part where inner speech does not differ from outer speech. However, unlike in overt outer speech, right after the motor command is issued, it is inhibited. There is time, however, for the monitoring system to receive a copy of the motor command and to predict what will happen when the motor command is realized (what signals the body is going to receive). According to many proponents of this theory [e.g., Jeannerod (2006); Carruthers (2011); Guillot et al. (2012); Perrone-Bertolotti et al. (2014)], when something like this happens, i.e., when in general a

motor command is inhibited, the result is imagery: if the motor command concerns producing speech sounds, the result is the phenomenology of inner speech, i.e., our inner voice. Plausibly, the phenomenology of inner speech is generated by making conscious the prediction that does not match any actual output. In general, simulating is making a prediction conscious [Jeannerod (2006)].

This theory draws on the corollary discharge model of perception proposed by Helmholtz (1860), and extended by von Holst and Mittelstaedt (1950) and Sperry (1950) to deal with motor acts. The model, in general terms, works in the following way: whenever a motor command is issued, the brain predicts, based on an efference copy (or corollary discharge) and the work of some 'forward models', what proprioceptive and sensory feedback will ensue. This prediction is used, together with a prediction of how the motor command will be executed, to monitor the execution of the motor instructions. Both predictions are used to detect errors in execution. The prediction of incoming signals is used to detect errors, through comparison with the actual sensory signals received. If the prediction matches the signal, the action is self-ascribed –maybe generating the 'sense of agency' [Feinberg (1978); Frith (1992), (2012)]. When the monitoring system does not work properly, subjects may end up experiencing their actions as alien, as in passivity phenomena such as control delusion, auditory verbal hallucinations, and thought insertion [Feinberg (1978); Frith (1992)].

Now, this is a theory of how the inner voice is produced, and it focuses only on motor commands and predictions related to sensory signals. However, predictions, given their monitoring role [Hickok (2012)], are plausibly issued at different levels [Pacherie (2008)]. In the case of speech, some influential psycholinguists argue that the production system issues predictions at many levels: at least concerning semantics, syntax and phonology [Pickering and Garrod (2013); Garrod et al. (2013); Gambi and Pickering (2016)]; or at the 'formulator' (semantics and syntax) and the articulator (motoric) levels [Hartsuiker (2014)]. Pickering and colleagues argue that typically, overt speech comparisons for error detection do not use the auditory input of the subject's own voice, but work on the basis of the predictions issued at the different levels. Error detection and repair is too quick to rely on comprehension of the external input (i.e., overt speech). On their account, semantic errors, such as selecting a meaning or semantic representation that does not correspond to the intended meaning, are detected and corrected very quickly (in about 300 ms). This suggests that the error is spotted and the correction

made on the basis of the predicted meaning [for a full development of this view, see Vicente and Jorba (2017)].

We want to highlight at this point that if a prediction of the sound of a speech act is the inner voice, a prediction of the full semantics of a speech act will be the meaning/content of such a speech act. That is, a prediction of the full semantics of a speech act is nothing other than what we identify as a propositional content. It is plausible to believe that our typical inner speech is not just a string of auditory representations or an inner voice; rather, it has sound-like properties *and* meaning. How could this be? Carruthers' view (see above) is that we bind the content of what we have said in our inner voice after interpreting our inner utterance via the comprehension system. Another possibility, however, is that our inner speech already 'comes' with a meaning. As we have said, when we form the intention to express a certain content and we refine that intention to the point of executing the motor commands that ultimately realize it, we issue predictions not just of how the utterance is going to sound, but also of what it is going to mean. Now, if instead of executing the motor command, we inhibit it, it may be that what enters our consciousness is not just an acoustic image, but an acoustic image with a meaning, given that we have issued both kinds of predictions. There is no reason why only the prediction related to the motor command should be conscious (that is, unless one is independently committed to the view that only sensory states are conscious). The way we experience inner speech, in contrast, is very often as if we were thinking in words, that is, as if our inner voice were the vehicle of our thoughts. This part of the phenomenology of inner speech can be explained if the inner voice comes with its meaning attached; in our terms, if the prediction of the acoustics is accompanied by a prediction of the meaning.

Now, if this is the case, and if, as just mentioned, one does not adopt the view, defended by, inter alia, Jackendoff (2012), Prinz (2011) and Carruthers (2015), that only sensory elements can be experienced, there is ample room to defend that we can have introspective access to the content of our inner speech episodes. To the extent that our inner speech episodes express the contents of some of our brain states, then it follows that we can know what such states represent by introspection.

IV.2. *Unsymbolized Thinking*

The model that we have sketched enables us to go a step further. Dennett grants that there can be wordless thinking, what he calls "bare meanings or thoughts" [p. 184]. He also asks whether we could have this

kind of thinking "if we didn't have the neural systems, the "mental discipline", trained up in the course of learning our mother tongue". We wonder how we could get to know what we think without words if not by introspection. But our point here is another. Our model can explain how this kind of thinking is possible. Suppose that instead of inhibiting our intention to express some thought content at the level of speech commands, we inhibited it at a previous level, say, at the level where we have given form to our intention in terms of semantics and syntax, but not yet in terms of phonology. In such a case, we would still have issued a prediction concerning the meaning of the utterance we were intending to make. What would happen next? According to the model, the prediction would be made conscious, presenting to you the content of the utterance you did not make, which would come in a linguistic format, i.e., with the structure and semantics of a sentence of the language you use to talk to yourself. This would be an episode of what Hurlburt [Hurlburt and Akhert (2008)] calls 'unsymbolized thinking': an episode involving a thought with a propositional content but with no sensory accompaniment. It would also be an episode of what Vygotsky (1987) took to exemplify ultra-condensed inner speech, a kind of inner speech where all wording has disappeared. Thus, in our construal, Dennett's wordless thinking would have as contents precisely the contents that a subject intends to express once they have been structured according to that subject's language. In this, we completely agree with his idea that such thinking is enabled also by the linguistic app [Vicente and Martínez-Manrique (2016); Vicente and Jorba (2017)].

Hurlburt and collaborators hold that we engage very often in this kind of wordless thinking. Using their DES method, they discovered that many subjects report experiencing determinate propositional thoughts that are not accompanied by any kind of sensory phenomenology. If we believe that unsymbolized thinking is possible –and Dennett himself seems to believe it–, then wouldn't this be a case where a user-friendly app turns out to not deceive the user in any way? We would be knowing what we think, we would be knowing it by introspection, and we would be also knowing what our brain states are representing. We would not know *how* our brain states represent what they do, but we would be acquainted with the content of the brain representations. In this case, therefore, the user-friendly app does not generate a user-illusion: what you introspect is what there is.

V. Coda

We take it that apps generate user-illusions when they make the user believe that there are things that are not there. According to Daniel Wegner (2002), for instance, there is nothing like conscious willing: our behaviour is caused by mental states we are largely ignorant about but the user believes that the cause is her conscious will. However, the experience that we associate with consciously willing is just the way users represent certain patterns of causes and effects. So, the users are deceived, as they believe that things occur because they consciously willed them to occur. According to Dennett, qualia are also user-illusions: we believe that our minds are hosting a certain experience when in fact all that is occurring is that some neurons of our brain are firing in a certain way. It is like when we believe that we are eliminating a folder in our computer screen by moving an icon to the recycle bin icon. We believe that things are in a way that they are not.

However, there seems to be a limit to the deceiving powers of user-friendly apps. If you experience a pain in your kidney, you are informed that something in your kidney is going wrong, and you are also told to do whatever you can do to improve the situation. The representational content of your pain experience gets it right, and its imperative content gets it right as well –it is an adequate command. But perhaps the representational and imperative contents of a pain experience are not transparent to us. However, when it comes to thinking, both the force and the content of the thought event are directly knowable, whether or not the thought comes in the clothes of auditory imagery. In (at least) such a case, we submit, there is no user-illusion. The app gets things right.

Ikerbasque and University of
the Basque Country (UPV/EHU)
Department of Linguistics and
Basque Studies
Paseo de la Universidad 5
01006 Vitoria-Gasteiz, Spain
E-mail: agustin.vicente@ehu.eus

University of the Basque
Country (UPV/EHU)
Department of Linguistics and
and Basque Studies
Paseo de la Universidad 5
01006 Vitoria-Gasteiz, Spain
E-mail: marta.jorba@ehu.eus

NOTES

¹ There seem to be some meta-representational capacities in some animals. At least, the discussion is open: see Kornell et al. (2007); Smith et al. (2008).

² This is a fairly common thesis. To date, the most direct empirical support for a link between language and metacognition is a very recent paper by Langland-Hassan et al. (2017).

³ Following what we think is Dennett's usage, we will not distinguish here between reasons and thoughts or mental states in general.

⁴ We are not sure to what extent Dennett would disagree with us. At some points, he seems to hold that we can actually know reasons. At some other points, he seems to argue that everything we think we know, including reasons, is illusory knowledge, which can be more or less approximate. In any event, our contribution can be seen as a way of fleshing out the idea, whether or not endorsed by Dennett, that we can know reasons. We will also argue that we can know reasons by acquaintance –we doubt that this last part would be endorsed by Dennett.

⁵ Regarding the latter point, it seems that for Dennett having an experience requires having a previous judgment that one is having that experience: "It is like something to be you *because* you have been able to tell what's like to be you!" [p. 344]. So, if language is required in order to have that judgement, then language would be necessary to have experiences.

⁶ For a discussion on the method itself, see Hurlburt and Schwitzgebel (2007), and for further confirmation of the result of the method based on fMRI scans, see Kühn et al (2014).

REFERENCES

BLOCK, N. (1995), "On a Confusion about a Function of Consciousness"; *Behavioral and Brain Sciences* 18, pp. 227-247.

CARRUTHERS, P. (2011), *The Opacity of Mind: An Integrative Theory of Self-knowledge*; New York: Oxford University Press.

— (2015), *The Centered Mind: What the Science of Working Memory Shows Us About the Nature of Human Thought*; Oxford University Press.

— (2018), "The Causes and Contents of Inner Speech"; in P. Langland-Hassan & A. Vicente (eds.). *Inner Speech: New Voices*; Oxford University 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, pp.162-183.

DENNETT, D. (2017), *From Bacteria to Bach and Back: The Evolution of Minds*: Penguin Books, Limited.

FEINBERG, I. (1978), "Efference Copy and Corollary Discharge: Implications for Thinking and its Disorders"; *Schizophrenia Bulletin* 4, pp. 636-640.

FRITH, C. (1992), *The Cognitive Neuropsychology of Schizophrenia*; Hover: Lawrence Erlbaum Associates.

— (2012), "Explaining Delusions of Control: The Comparator Model 20 Years On"; *Consciousness and Cognition* 21, pp. 52-54.

GAMBI, C. and PICKERING, M. (2016), "Predicting and Imagining Language"; *Language; Cognition and Neuroscience*, 31 (1), pp. 60-72.

GARROD, S., GAMBI, C. and PICKERING, M. (2013), "Prediction at All Levels: Forward Model Predictions Can Enhance Comprehension"; *Language, Cognition and Neuroscience*, 29:1, pp. 46-48,

GUILLOT, A., DI RIENZO, F., MACINTYRE, T., MORAN, A. and COLLET C. (2012), "Imagining Is Not Doing but Involves Specific Motor Commands: A Review of Experimental Data Related to Motor Inhibition"; *Frontiers in Human Neuroscience,* 6, 247.

HARTSUIKER, R. (2014), "Monitoring and Control of the Production System; in M. Goldrick, V., Ferreira, and M. Miozzo (eds.); *The Oxford Handbook of Language Production*; Oxford: Oxford University Press, pp. 417-436

HELMHOLTZ, H. V. (1860), *Treatise on Physiological Optics*; New York: Dover.

HICKOK, G. and POEPPEL, D. (2007), "The Cortical Organization of Speech Processing"; *Nature Reviews Neuroscience* 8: pp. 393-402.

HICKOK, G. (2012), "Computational Neuroanatomy of Speech Production"; *Nature Reviews Neuroscience,* 13(2), pp. 135-145.

HUEBNER, B. and DENNETT, D. (2009), "Banishing 'I' and 'We' from Accounts of Metacognition"; *Behavioral and Brain Sciences* 32, pp. 148-149.

HURLBURT, R. T. and AKHTER, S. A. (2008), "Unsymbolized Thinking"; in *Consciousness and Cognition* 17, pp.1364-1374.

HURLBURT, R. T. and SCHWITZGEBEL, E. (2007), "Describing Inner Experience? Proponent Meets Skeptic"; Cambridge, MA: MIT Press.

JACKENDOFF, R. (2012), *A User's Guide to Thought and Meaning*; New York: Oxford University Press.

JEANNEROD, M. (2006), *Motor Cognition: What Actions Tell the Self*; Oxford: Oxford University Press.

JORBA, M. and VICENTE, A. (2014), "Cognitive Phenomenology, Access to Contents, and Inner Speech"; *Journal of Consciousness Studies* 21 (9-10), pp. 74-99.

KORNELL, N., SON, L., and TERRACE, H. (2007), "Transfer of Metacognitive Skills and Hintseeking in Monkeys"; *Psychological Science*, 18, pp. 64-71.

KÜHN, S., FERNYHOUGH, C., ALDERSON-DAY, B. and HURLBURT, R. T. (2014), "Inner Experience in the Scanner: Can High Fidelity Apprehensions of Inner Experience Be Integrated with fMRI?"; *Frontiers in Psychology* 5, pp.1-8.

LANGLAND-HASSAN, P., GAUKER, C., RICHARDSON, M. J., DIETZ, A., and FARIES, F. R. (2017), "Metacognitive Deficits in Categorization Tasks in a Population with Impaired Inner Speech"; *Acta Psychologica* 181, pp. 62-74

MARTÍNEZ, M. and KLEIN, C. (2016), "Pain Signals Are Predominantly Imperative"; in *Biology & Philosophy,* 31; pp. 283-298.

METZINGER, T. (2009), *The Ego Tunnel: The Science of Mind and the Myth of the Self*; NY: Basic Books.

PACHERIE, E. (2008), "The Phenomenology of Action: a Conceptual Framework"; *Cognition* 107, pp. 179-217.

PERRONE-BERTOLOTTI, M., RAPIN, L., LACHAUX, J.-P., BACIU, M., and LŒVEN-BRUCK, H. (2014); "What Is that Little Voice Inside My Head? Inner Speech Phenomenology, Its Role in Cognitive Performance, and Its Relation to Self-Monitoring"; in *Behavioural Brain Research,* 261, pp. 220–239.

PICKERING, M. and GARROD, S. (2013), "An Integrated Theory of Language Production and Comprehension"; *Behavioral and Brain Sciences* 36, pp. 329-347.

PRINZ, J. (2011), "The Sensory Basis of Cognitive Phenomenology", in Bayne, T. and Montague, M. (eds.) (2011); in *Cognitive Phenomenology*. Oxford: Oxford University Press, pp. 174-196.

SEARLE, J. (1992), *The Rediscovery of the Mind*; Cambridge, Mass., The MIT Press.

SMITH, J. D., BERAN, M. J., COUCHMAN, J. J., and COUTINHO, M. V. C. (2008), "The Comparative Study of Metacognition: Sharper Paradigms, Safer Inferences"; *Psychonomic Bulletin & Review,* 15, pp. 679-691.

SPERRY, R. W. (1950), "Neural Basis of the Spontaneous Optokinetic Response Produced by Visual Inversion"; in *Journal of Comparative and Physiological Psychology* 43, pp. 482-489.

STRAWSON, G. (2010), *Mental Reality* (2nd edition); Cambridge, London: MIT Press.

TYE, M. (1995), "A Representational Theory of Pains and their Phenomenal Character"; *Philosophical Perspectives*, 9, pp. 223-239.

VICENTE, A. and JORBA, M. (2017), "The Linguistic Determination of Conscious Thought Content"; *Noûs*. DOI: 10.1111/nous.1223

VICENTE, A and MARTÍNEZ-MANRIQUE, F. (2016), "The Nature of Unsymbolized Thinking"; *Philosophical Explorations*, 19 (2), pp. 173-187.

VON HOLST, E. and MITTELSTAEDT, H. (1950), 'Das Reafferenzprinzip'; *Die Naturwissenschaften* 20, pp. 464-476.

VYGOTSKY, L. S. (1987), *Thinking and Speech*; in *The Collected Works of L. S. Vygotsky* (Vol. 1), New York: Plenum (Original work published 1934).

WEGNER, D (2002); *The Illusion of Conscious Will*; Cambridge, Mass., The MIT Press.

OXFORD
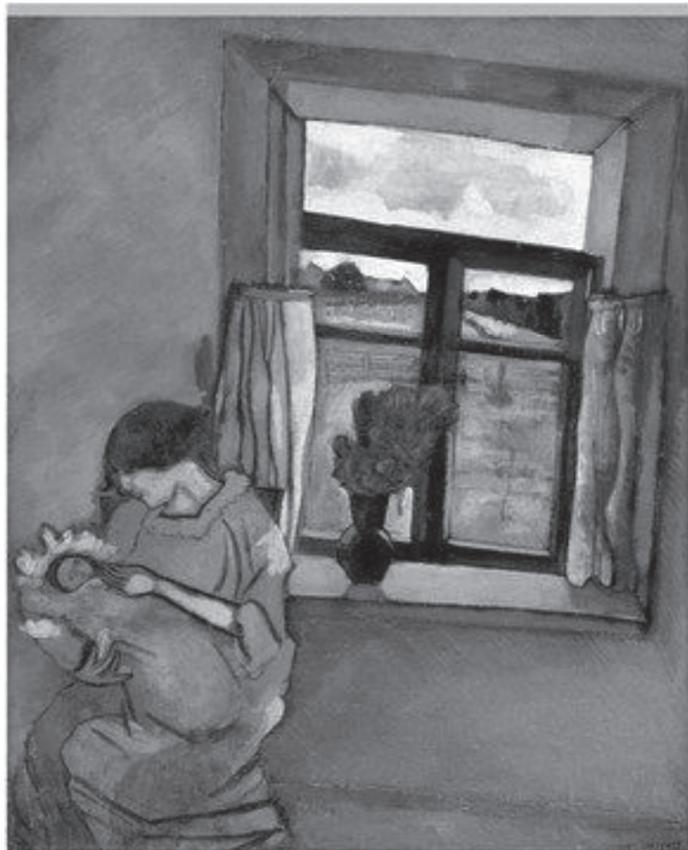
WHY WORRY ABOUT
FUTURE GENERATIONS?

Samuel Scheffler