

# **THOUGHT, LANGUAGE AND THE ARGUMENT FROM EXPLICITNESS**

AGUSTÍN VICENTE AND FERNANDO MARTÍNEZ-MANRIQUE

**Abstract:** This paper deals with the relationship between language and thought, focusing on the question whether language can be a vehicle of thought, as, e.g., Carruthers (1996, 2002) has claimed. We develop and examine a powerful argument - the "argument from explicitness"-, against this cognitive role of language. The premises of the argument are just two: (1) the vehicle of thought has to be explicit, and (2) natural languages are not explicit. We explain what these simple premises mean and why we should believe they are true. Finally, we argue that, even though the argument from explicitness shows that natural language cannot be a vehicle of thought, there is a cognitive function for language.

**Keywords:** compositionality, explicitness, introspection, natural language, underdeterminacy

## **Introduction: language as a vehicle of thought**

For almost two centuries philosophical orthodoxy held that human thought uses natural language as an instrument or vehicle; in other words, that each human being thinks in her own language. This idea, plus the belief that the various languages may differ greatly from each other in many and important respects, led to the thesis of "linguistic relativity", i.e., the thesis that two individuals belonging to two different linguistic communities think, conceive the world and consequently behave in a very different way. This thesis was orthodoxy especially during the heyday of romanticism in the XIXth century, and during the first part of the XXth century, when it crystallized in the

famous “Sapir-Whorf” hypothesis.

Such a position, however, is no longer orthodox or even popular. The cognitive revolution changed the general view on this, and it seems fair to say that there are few authors today who defend that natural language (NL, henceforth) may be such an overarching instrument of thought. There are various reasons for this change. Surely one has to do with the success that cognitive psychology has shown in explaining the behavior of non-linguistic infants and animals. Another is the current conviction that at least some mental activity is needed in order to acquire a language: acquiring syntax requires a wealth of previous cognition; acquiring semantics demands, in addition, a repertoire of previous general concepts (immense, on Fodor’s 1975 view) and some mind-reading capacity (Bloom 2000; Tomasello 2001). The upshot, in any case, is that cognitive psychology has replaced the empiricist theatre-of-sensations view of the mind by a picture where the mind has a considerable innate and independent structure. Instead of a picture where language imposes its conceptual web on mind, mind’s own architectural and conceptual structure enables us to acquire our language.

However, it cannot be denied that our being linguistic creatures is linked to our being particularly intelligent, so that, for instance, babies experience a cognitive explosion as soon as they acquire language. Authors differ in their explanation about this link and, in general, about “what makes us smart” (cf. Spelke 2003). Steven Pinker (1994), for example, believes that language’s role as a means of communication is enough to account for this cognitive difference between linguistic humans and the rest: linguistic exchanges afford an incredible amount of information, both about the world and about the minds of the others. In contrast, Andy Clark (1996), following a Vygotsky-inspired line, holds that NL can help us think, especially by objectifying our own thoughts and making them accessible to consciousness. However, other authors

within mainstream cognitive science defend that NL, once learned, is used as the vehicle, or carrier, of our higher-order thoughts. The position differs from classical Whorfianism in that now NL is not regarded as a necessary conceptual structurer: there is a structured mind before NL is acquired. The architecture of mind would be typically modular, but when NL is acquired, it is used as a vehicle of non-modular and reflexive thinking.

We are going to consider the view that language is used as a carrier or instrument of thought and argue that it is untenable in the light of what we know about NL. As Peter Carruthers is probably the author who has more consistently defended this view, we will begin by reviewing his theses. Even if he has now retreated to a strong Vygotskian line, we think that Carruthers's arguments for his past views are worth considering. Then, we will move on to deal with the central topic of this paper: the argument from explicitness. This argument establishes that our thinking cannot consist in entertaining meanings of NL sentences: NL sentences, being inexplicit about their content, are not suitable vehicles of thought. Last, we will argue for a weak Vygotsky-inspired view on the cognitive use of language that is not affected by the argument from explicitness, and that can explain evidence Carruthers exploits.

### **Language as a vehicle of thought**

For some time Carruthers has consistently contended that (at least) typically human thinking is not realized by tokening Mentalese sentences, but by tokening NL sentences. He is not the only author defending this approach. Daniel Dennett (1996), Derek Bickerton (1995), Steven Mithen (1995), José Luis Bermúdez (2003) or Elizabeth Spelke (2003), to name a few but well-known researchers, have also endorsed variants of this view. As a matter of fact, Carruthers (2002) draws on Mithen and especially on

Spelke in order to strengthen his former (1996) position. What is peculiar to Carruthers is that he has developed two different but complementary arguments purporting to explain a kind of evidence that any account of thought and language should take care of.

The first argument is developed in his (1996). It is a straightforward and compelling argument based on introspective evidence. In a nutshell, it says that our own introspection reveals that we do use language when we think. We can state it in just two premises:

- (a) data of introspection reveal that sometimes we think in a NL, and
- (b) we must begin by taking these data of introspection for what they seem.

What introspection reveals is that we codify linguistically episodic conscious thoughts. But Carruthers goes further and, using considerations of simplicity, argues that the mind also uses NL to codify latent thoughts and unconscious token-thoughts that belong to the same types as those conscious episodic thoughts. Carruthers qualifies his claim, and concedes that NL might not be the only means by which we think conscious thoughts. Some thoughts use a mixed vehicle formed by NL words and images (e.g., 'I want this chair to go there [insert image]'). However, he rejects that conscious thinking uses a specific language of thought, the so-called 'Mentalese' (he also discards that it may use only images).<sup>1</sup>

Later in the paper we are going to dispute proposition (a) above, that is, that introspection reveals that we use NL (or NL-plus-images) as a vehicle of thought. It is also possible to question that we should use introspection as a good source of data, or even that Carruthers is consistent in taking introspective data at face value, for he also troubles to explain data that would look adverse (for instance, quite often we are sure that we are entertaining a thought we cannot express). However, for the time being we will assume that the case from introspection is a good case for the use of language in

thought.

Carruthers has lately presented a second argument based on evidence of a different sort. While in his (1996) Carruthers only said that NL had some or other function in thought, now he follows other authors in pointing out a specific function for NL in mind. On the one hand, there are experiments that apparently show that we use NL to integrate outputs coming from different modules (Hermer-Vazquez *et al* 1999; Spelke and Tsivkin 2001a, 2001b; Spelke 2003). For instance, combining geometrical information with color information in a thought like *the object is to the right of the red short wall* becomes more difficult when experimental subjects have to search for the object while they are occupied in a linguistic task. On the other, there is an evolutionary scenario provided by Mithen (1995), in which language appears as a means to let information flow from modules that initially were domain-specific. Carruthers (2002) endorses this general account of NL as an intermodular integrator. In his view, the linguistic module would be ideally located to perform this function because it is an input/output module, i.e., it produces information that is “consumed” by domain-specific central systems, but it also takes information from them.

As we advanced, Carruthers (2006) has now retreated to a different position. Very succinctly, he now takes NL to be a tool for flexible and creative thought, used to broadcast trains of thought globally (i.e., inter-modularly), so as to make them subject of further thinking. More timidly, he speculates on the idea that NL’s combinatorial resources might enable it to be the only candidate for intermodular integration, an idea that seems inspired by what Spelke (2003) holds, i.e., that NL’s syntax makes possible for us to insert the concept RED in a structure that reflects just geometrical information such as ‘to the right of the short wall’. What is important, in any case, is that according to his current position, thinking is not ultimately realized by NL sentences but by

another medium. The reason for this change seems to be that Carruthers has been convinced by something like the argument from explicitness.<sup>2</sup> Now, is this argument so convincing? Let's take a close look at it.

### **The argument from explicitness**

The argument we want to analyze can be given in just two premises:

Premise 1: The instrument of thought must be explicit.

Premise 2: NLs are not explicit.

From this it follows that NL cannot be the instrument of thought.

Both premises are far from self-evident, though we take the second to be more controversial. We will consider them in turn.

#### *Premise (1): the instrument of thought must be explicit*

Pinker (1994) offers support for the first premise. Drawing on ideas of McDermott (1981), his thesis is that the human mind/brain works like a Turing machine, and that no NL can function as a language for the use of a Turing machine. The reason is that Turing machines need a language that is highly regimented, unambiguous and strongly compositional, and that NL is none of that.

Pinker's argument will not convince those that conceive of minds as very different from Turing machines. So we will focus on an argument with a wider scope advanced by Fodor (2001). (See also Levinson 1997, 2003; Carston 2002a). For the most part of the paper, Fodor just assumes that it has been conclusively shown that there must be a language of thought and that this language is compositional. By the end of the paper, however, he advances a new argument for this last claim. Basically, this new argument can be condensed in a sentence: "a thought cannot be inexplicit with respect to its own

content (...) because a thought just *is* its content” (2001, 14). The formulation of his idea does not seem to us very fortunate in this case: a thought is, presumably, the content of a certain sentence in the language of thought. So, saying that a thought is its content is (in the best case) not informative, and does not help to tell thought from meaning, for the same can be said about the propositions expressed by sentences of a natural language: a proposition is its content. So what Fodor may mean is that whatever serves as a vehicle of thought, it cannot be inexplicit about the content of its sentences.

Let us try a formulation of what it is for a language to be explicit. The idea is that its sentences express or codify all the information (no more, no less) that they are intended to express. It is a language that somehow codifies every piece of information necessary to reach the proposition that the sentence is intended to express. So, on a first approximation, a language is explicit if and only if the content of (i.e., the proposition expressed by) its sentences is given only by the semantics of its elements and by its combinatorial syntax (i.e., by compositional rules operating on stable meanings).<sup>3</sup>

Elugardo (2005, 65) points in a similar direction when he reads Fodor as giving two necessary conditions on semantic explicitness for a complex symbol: one that requires each of its syntactic constituents to map onto a semantic constituent, and a converse one that requires each of its semantic constituents to map onto some syntactic constituent. However, it seems that Elugardo misses the element of stability of meaning: the mappings cannot change from one time to another. In our view, this is a key component of Fodor’s position that can be better appreciated if his notion of explicitness is set against the background of his conceptual atomism.

According to Fodor (1998), thoughts are composed by conceptual atoms, which are the semantic or contentful part of the representations postulated by the Representational Theory of Mind. Now, these representations have a stable meaning,

independent of context. The rules that compose the content of wholes are equally stable. Having a thought amounts to this: every time a certain combination of representations is activated, there is a certain thought. So the language of thought is not inexplicit, or it cannot produce inexplicit sentences (in the sense of “explicitness” explained above), because forming a certain sentence (i.e., a combination of representations) means forming a certain thought, which is always the same for that specific sentence. Therefore, the relation between a truth-valuable mental representation and its content is very different from the relation between a NL sentence and its propositional content: having a thought is equivalent to having its content, while in order to obtain a proposition from a NL sentence we need something else, a context.

There is perhaps a more general way to justify Fodor’s claim. Think of an inexplicit language in which the meanings of words or composition rules are not stable (they vary with extra-linguistic context). In order to interpret the sentences produced by such a language one needs a way to put together the information provided by the sentence plus the information provided by the context. If one has to do this by linguistic means, then this second language has to be explicit: otherwise its sentences would have to be interpreted in turn. So the interpretation of an inexplicit language requires an explicit language as the final *locus* of interpretation. As interpretation takes place in thought, it may be concluded that whatever is the (ultimate) language of thought, it has to be explicit. In this same wake, it could be said that a simple phenomenological realization shows that we do not have to interpret our own thoughts: (problems derived from externalism aside) we do not experience thoughts whose meaning is not clear to us.

*Premise (2): NLS are not explicit*

In order to consider the second premise of the argument from explicitness – that NLS are not explicit– some comments about explicitness and compositionality seem appropriate.

In his (2001) Fodor seems to regard compositionality and explicitness as extensionally equivalent, so that a language is explicit if and only if its semantics are compositional (meanings of wholes are functions of meanings of their parts). However, this identification of compositionality and explicitness introduces a misleading element in the debate, for the discussion may well revolve around whether NLS are compositional or not. This is the case of Elugardo's account. He concentrates most of his criticism of Fodor's argument on showing how a NL can be compositional without satisfying the necessary conditions for explicitness stated above (Elugardo 2005, 66-75). We think this is a false lead for our discussion because while explicitness entails compositionality, compositionality does not entail explicitness. So showing that NL is not compositional would be a good way to conclude that it is inexplicit, and hence unable to be the carrier of thought.<sup>4</sup> But showing that NL is compositional has no bearing on whether it is explicit and capable of that function.

In fact, it is always possible to hit on a compositionality function for any language (Hodges 2001). Does this mean that any language can be shown to be explicit? It does not. In order to be explicit, a language must consistently produce sentences that are not ambiguous, polysemous, or, in general, open to numerous interpretations. So its sentences must code, by themselves, and barring vagueness, determinate contents. Not all languages that can be described as compositional happen to be explicit. The compositionality function may well throw an inexplicit, underdetermined meaning. To take one simple example, according to Sainsbury (2002), English might be regarded as

an orthodox compositional language, though it would be scarcely explicit, given the number of ambiguous terms that it has. The meaning of sentences would be given by composition rules operating on the meanings of parts, but, as a relevant number of these are underdetermined, the meaning of the whole would be equally underdetermined, waiting for contextual information to resolve the underdeterminacy. Let us now move to see what reasons there may be for considering NLs inexplicit.

### **Unarticulated constituents and pragmatic enrichment**

Many authors working on pragmatics (Sperber and Wilson 1986/1995; Recanati 2001, 2003; Carston 2002a) claim that the truth-conditional meaning of most utterances is not given only by their semantics; rather, the “real” or “intuitive” truth conditions of such utterances are “enriched” or “modulated” using contextual information. Here are some typical examples where enrichment takes place:

- (1) It is raining.
- (2) All the students have gone on strike.
- (3) The chair is broken.
- (4) John’s car is empty.
- (5) Peter is small.

Some very probable truth conditional meanings of these utterances are the following:

(1) it is raining [here] (in the place of the utterance), (2) all the students [of the university/school of the speaker] have gone on strike, (3) the [specific] chair [about which both speaker and hearer know well] is broken, (4) [the car owned by] John is empty, (5) Peter is [a] small [man]. However, these meanings do not coincide with the alleged codified meaning: (1) means that it is raining (place unspecified), (2) that all students (no domain restriction) have gone on strike, (3) that there is only one chair and

that such a chair is broken, etc. In all these cases, then, there are some syntactically unarticulated constituents whose semantics enrich the alleged codified meaning.

As the codified meanings of (1-5) are unusual, it can be argued that such meanings cannot be the meanings of the expression-types. The truth conditions of those sentences would coincide with their typical (pragmatically) enriched meanings, rather than with their alleged codified meanings. In other words, the truth conditional meaning of the sentences (1-5) is not semantically given and obtained by composition. There are pragmatic processes at work in order to get the “intuitive truth conditions” of these sentences (cf. Recanati 2002b, 2003).<sup>5</sup>

The defender of the explicitness of NL has two kinds of responses available. First, she can hold that the meaning of (1-5) is the codified meaning, and that the “intuitive meanings” are just the result of Gricean inferences. The classical Gricean-style response here would be that there is just one kind of meaning, given by semantic composition. The so-called “intuitive meaning” is just the end-point of a process that starts with the appreciation that the meaning of the sentence does not fit what (according to the maxims of communication) is expected from the speaker. This seems to be the line taken by some “minimalists”, such as Emma Borg (2004).<sup>6</sup> A divergent Gricean view, such as Sperber and Wilson’s, would instead recognize two types of meanings: (1) literal meaning, resulting from decodification; (2). enriched meaning (the explicature, in their terminology), obtained through the Principle of Relevance, which may involve the construction of new concepts on-line. (In turn, enriched meanings would serve as input to further processes of inference guided by the Principle of Relevance to obtain “what is implicated”). The difference between these two Gricean schools lies in the treatment of the “enriched meaning”: the classical view takes it as “speaker’s meaning”, while in relevance theory “enriched meaning” is usually equated with ‘what is said’. Both views,

however, agree in defending a notion of literal meaning for at least one kind of meaning.

The second response available to the defender of the explicitness of NL is to grant that the meaning of sentences such as (1-5) coincides with the so-called “enriched meaning”, but to contend that this meaning is obtained by decodification. For instance, even though the syntactic surface of (2) contains no mark pointing to a domain restriction, there is such a mark, though hidden (Stanley and Szabo 2000; Stanley 2002). This response may look rather *ad hoc* but it has been defended by means of an attractive argument, called ‘the argument from binding’. Briefly, what this argument says is that if a variable can be bound, then such a variable exists, even if it appears nowhere in the syntactic surface of the sentence. So by introducing quantifiers we may “exhume” variables that would otherwise be invisible. For instance, take (2) ‘All students have gone on strike’. The sentence contains a free variable behaving like an indexical that restricts the domain of the universal quantifier. This variable becomes evident when we form

(2’) In every department all their students have gone on strike.

To see just another example, in (5) ‘Peter is small’ there is a hidden indexical pointing to a reference class. This is revealed when we take

(5’) There is a class of beings such that Peter is small for such a class.

At first blush, this second kind of response is best suited for the defender of NL-explicitness, since it preserves the isomorphism between syntax and semantics. Every semantic constituent of the proposition would be represented in syntax, thus ensuring explicitness (though only *prima facie*: see later on). Indeed Elugardo (2005, 72-5) appeals to analogous considerations when he argues that ellipses are not conclusive for the presence of unarticulated constituents, and hence for non-explicitness.

But the argument from binding is not exempt from problems. One is that it forces to introduce too many hidden variables. Take this example:

(6) Your eyes are green.

Presumably, this sentence means (under any notion of meaning) that the eyes of the hearer are green, period. However, it is possible to bind a variable pointing to a determinate green color by proceeding just in the same way as the preceding examples:

(6') There is a shade of green such that your eyes are green of such shade.

So the argument from binding seems to exhume not just invisible variables but also inexistent ones.

In addition, the biggest problem for this account as a defense of the explicitness of language is that there is no way to fill in all those hidden indexicals automatically. Rather, one has to look into the “wide context” (see below) and do some pragmatics in order to decide what value a given indexical is taking in a certain context (King and Stanley 2004). This is best revealed if one looks at enrichment as a case of underdeterminacy –which we are about to do– rather than if one regards it as ellipsis, which is Elugardo’s focus in his discussion of Fodor’s explicitness argument (understandably, given that the examples mainly offered by Fodor fall into that case). As we will show, even though one cannot construe a good argument against explicitness from the facts of pragmatic enrichment, the consequences of underdeterminacy are lethal. So a defender of the explicitness of NL had better not look for help in Stanley’s position.

Now, suppose that both responses –the Gricean response and the hidden-variable response–were to fail, and consequently that the “real meaning” of (1-5), and of most NL sentences, were not explicit. Still, it is possible to wonder what this may imply for the question about the use of NL in cognition.

Let us distinguish two senses of 'explicit language', a strong and a weak one. In the strong sense, a language L is explicit if and only for every sentence S of L, the proposition expressed by S is determinate and obtained solely by semantic and syntactic information (i.e., without the involvement of pragmatic processes).<sup>7</sup> In the weak sense, a language L is explicit if and only if for every intended determinate proposition, it can produce a sentence S that expresses such a proposition by means of its syntax and semantics alone. If L is not explicit in the strong sense, then pragmatic enrichment is needed to obtain the truth conditions of at least some sentences. However, it may be the case that L is not explicit in the strong sense but it is explicit in the weak sense. In this case, for any sentence S of L that needs pragmatic enrichment to express a particular proposition there is another sentence S' in L that expresses semantically (without pragmatic enrichment) the same proposition. If L is not explicit even in the weak sense then L does not have the resources to express semantically what it expresses with the help of pragmatic enrichment. It seems that a natural languagelike English is explicit in the weak sense, for it can produce sentences that express semantically what sentences such as (1-5) express by means of pragmatic enrichment. In other words, it can generate sentences such as (1a) 'it is raining here', (2a) 'all the students of the department have gone on strike', etc.

It may be the case that NLs are not explicit because they function primarily as an instrument of communication (McDermott 1981; Pinker 1994). In order to communicate our thoughts we exploit contextual factors as much as we can because we have limited time to say what we want to say.. Thus we omit all that information that can be easily gathered from context, and our audience expects us to do so. However, this explains why most of the sentences *that we utter* are inexplicit, but it does not follow that this is the case for *all possible* sentential utterances. Therefore it is not a good explanation for

why natural languages *should* be inexplicit, and it seems that NLS are in principle capable of producing explicit sentences<sup>8</sup>. Hence, for all that has been discussed up to now, the argument from explicitness might be blocked at premise (2). So in the next section we will consider a final problem for the defender of explicitness: the problem posed by underdetermined terms.

### **Underdeterminacy<sup>9</sup>**

Take example (4) above, i.e., ‘John’s car is empty’. What is the meaning of (4)? As Recanati (2001) explains, there is no definite answer without additional information about the context of utterance: it may mean (and it is an open disjunction) that the car owned by John is empty, or that the car driven by John is empty, or that John’s favorite car is empty, etc. The genitive expresses that there is a relation R between John and a car but it is not specified what sort of relation R is, that is, we cannot give a determinate interpretation for this sentence until we have used some pragmatic means to decide what the genitive stands for. So it seems that the existence of underdetermined terms poses a further, and harder, problem for explicitness.

Before we develop this idea, we want to note that here it is possible to find a departure between the demands of explicitness and the demands of compositionality. We think the presence of a genitive in a sentence renders it inexplicit about its content. Does this mean that its meaning is not compositionally given? Recanati himself (2002b) suggests a negative answer. According to him, the meaning of a sentence is compositional if it is “grammatically driven”, or in other words, if the pragmatic effects that affect its meaning are prompted by elements of its syntax. In this respect, the genitive behaves like a free indexical variable, so that the assignment of semantic values to it is grammatically driven: the grammar of the sentence provides slots that prompt

you to look for a semantic value. When the slot has been filled so as to obtain a truth-conditional structure, then we say that the sentence is *saturated*.

Thus, there is an important difference between a sentence like (4) and sentences such as (1) 'it is raining'. In (4) the genitive ['s] provides explicitly a slot that asks for saturation, whereas the element ['here'] that enriches the meaning of (1) (in its typical interpretation) is not given, neither explicitly nor implicitly (i.e., in terms of a "hidden variable"). As Recanati puts it, ['here'] is a genuine unarticulated constituent, whose presence is optional:<sup>10</sup> there is a truth-conditional sense of (1) that does not include a hidden locative. In contrast, it is not optional to give a particular interpretation to the constituent ['s] in (4): without this interpretation we do not have a proposition to begin with. In a nutshell, what happens with underdetermined terms like ['s] is not that different from what happens with indexicals: in both cases there are grammatical instructions that say "look at the context in order to assign a value to this expression". The dissimilarity lies in that indexicals<sup>11</sup> tell you to look at the narrow context (place of utterance, speaker, time...), whereas the genitive requires a wider context, including just about everything. But the principle of compositionality is not altered by this peculiarity, as it is by the presence of unarticulated constituents.<sup>12</sup>

Now, if we maintain that the claim of explicitness requires that the proposition expressed cannot depend on obviously nonsemantic contextual information, the presence of underdeterminacy, however it is understood, precludes the possibility of explicitness. For it means that the sentence by itself cannot be the "carrier" of content. We take it that this is the relevant point in order to decide whether a certain language can be a vehicle of thought. If the arguments above for premise (1) (the vehicle of thought must be explicit) are sound, what we need is a language whose meanings and rules are stable and require no interpretation. Hence, we may conclude that the

phenomenon of underdeterminacy, even if it just consisted in the presence of free variables, does pose a problem to the use of NL in thought.

Nevertheless, it is possible to think that this new problem is on a par with the problem of pragmatic enrichment: it only shows that the meaning of some of our utterances is not given by their semantics alone, but it does not show that NL is not explicit in the weak sense. Just as we can go from ‘it is raining’ to ‘it is raining here’, so can we go from ‘John’s car is empty’ to ‘the car owned by John is empty’. It seems that for every underdetermined utterance our NL can produce a determined explicit one. So it seems that NL is explicit in the weak sense introduced above.

Now, it is not obvious whether underdeterminacy is remediable, or its remedy comes at too high a price. Recanati (2001), for instance, says that the presence of underdetermined terms is *generalized*. This does not look *prima facie* unlikely: from ‘John’s car is empty’ we can go to ‘the car owned by John is empty’, but what is the meaning of ‘empty’? It demands determinacy. Perhaps we can express it as ‘the car owned by John has nobody inside’. Now, what is the meaning of ‘inside’ here? We do not want to commit ourselves to the claim that this process of specification will never come to an end, but it looks as though it will stop (if it does) in a long and probably unmanageable sentence.

Carston (2002a) provides powerful considerations in support of the generalization thesis, and against the idea that this specification process may stop somewhere. According to her, both referential and predicative terms are underdetermined. For a start, proper names should be treated as indexicals whose rule is not “pure” (it roughly says “this name refers to the contextually prominent bearer of it”). Even the reference of complete definite descriptions depends on the context, where the context in question can be the actual world, a possible world or even a belief world. The work of linguists such

as Culicover and Jackendoff (2004) and even Chomsky (2000) give also some support to this underdeterminacy thesis applied to referential terms. Culicover and Jackendoff persuasively argue that any proper name can be referring either to a person/object (its bearer) or to some iconic representation of it, such as a picture or a statue. For instance, the sentence 'Ringo is the Beatle that I like the most' can mean either that the actual drummer of the Beatles is the musician (or person) that I like the most, or that Ringo's statue at Mme. Tussaud's is the statue of a Beatle that I like the most.<sup>13</sup> Chomsky (2000) offers an analogous discussion of the contribution of 'London' to 'London has moved to the South Bank'

As for predicates, examples such as 'the kettle is black', 'the table was covered with butter' and 'Hugo is a sailor' (Travis 1985; Carston 2002a), where 'black', 'covered with butter' and 'sailor' may mean a variety of things, suggest that the semantic value of predicates is also typically underdetermined. Recanati (2001) goes much in the same line when he says that even the *adj. + noun* construction, such as 'red pen', is semantically underdetermined. Travis (1996, 2000) has made this idea quite plausible: in a famous example he shows that even a simple sentence such as 'those leaves are green' does not have determinate truth-conditions. 'Those leaves are green' may be true of springtime leaves, but also, in some contexts, of leaves which are dead but painted green. In conclusion, the phenomenon of underdeterminacy seems widespread enough to support premise (2), and so it becomes a real problem for the use of NL as a vehicle of thought.

Note finally that the situation does not change much in this respect when we turn to consider mixed vehicles like those Carruthers proposes. The use of images is a good way to fill incomplete sentences like 'I want this chair to go there'. By inserting an image first after 'this chair' and then after 'there', we may get an explicit sentence.

However, not all underdetermined sentences are determinable in this way. For example 'John's car is empty' is doubtfully completable by means of an image.

So we want to conclude that underdeterminacy precludes NL from being a suitable vehicle of thought. However, there is a final position we should address. For a number of authors, the so-called 'minimalists', claim that there is no underdeterminacy problem, and that sentences such as 'that pen is red' express a perfectly complete thought (see especially Cappelen and Lepore, 2005).<sup>14</sup> According to minimalists, any well-formed sentence of a NL expresses, just by itself, a proposition, that is, a well-formed sentence does not need any contextual completion of the kind suggested by most pragmatics in order to express a thought. If this were so, then our argument for the inexplicitness of NL would be wrong, and sentences, contrary to what we have said, could be the carriers of thoughts.

However, we have serious doubts about this minimalist proposal. On Cappelen and Lepore's view, a sentence such as 'John is ready' has complete truth-conditions, namely, 'John is ready' is true if and only if John is ready. The question, then, is: do we entertain general thoughts such as JOHN IS READY?<sup>15</sup> We think we do not. As far as introspection goes, we would say that our thoughts are much more complete than this. When thinking about John's readiness, we always think that he is ready for this or that (his lecture, his having a baby, whatever). But also and more important, if thoughts are to be responsible for our actions, they had better be more explicit than these putative minimal propositions. For, what would we do after tokening the very general thought that John is ready, punk?

This problem is common to all minimalist proposals, seen in the light of our current interests. For instance, Borg (2004) suggests that minimal propositions are usually existentially quantified ones. Thus, 'John is ready' would express the thought

that there is something such that John is ready for it. Following this line, we could account for the underdeterminacy cases by quantifying over senses in which things can be. For example, 'that pen is red' might express the minimal proposition that there is a sense of being red such that a salient thing that is a kind of pen is red in that sense. Now, we do think that these propositions can be entertained. Think of overhearing a conversation where 'that pen is red' is uttered. In the absence of contextual information, the overhearer may well just think that some pen is red in some of the various senses in which a pen can be said to be red. And the same would go for an utterance of 'John's car is fast': an overhearer could well form the thought that someone called 'John', possibly a male adult, is in some particular relation to something which is a kind of car, something which is also fast, presumably in relation to other cars of the same type. An overhearer is left just with semantic information, and it is our claim that the semantics of 'John's car is fast' gives you only this kind of very general information. However, these are not the kinds of propositions that we typically entertain. And they are not the kinds of thoughts that cause our actions. So if NL sentences encode only these sorts of very general thoughts, NL is not the instrument of our thought.

### **An alternative cognitive use for language**

We are going to change gears now. We hope to have made clear that Carruthers's past views are representative of a position that has to confront some deep problems. Now we want to bring in a straightforward criticism to such a position while at the same time we begin to present our positive view on the matter.

Let us take the argument from introspection that Carruthers presents. A premise of this argument is that introspection shows that we use NL as a vehicle of thought. Now, one may wonder whether we do have such introspective data. Suppose NLS were

explicit and so were capable of being the vehicle of our thought. Then, we should introspectively experience complete univocal sentences. However, except in the case of reading and writing, we do not entertain anything close to complete sentences. Moreover, it seems that the linguistic items that we experience are especially fragmentary, i.e., even more economical than the sentential or quasi-sentential items of spoken language. We think this is something a simple reflection on our phenomenology will reveal: when we feel that we are thinking in language most times we do not discover entire sentences but two or three words that sum up a complete proposition. In fact, many times we have the feeling that we are “speaking to ourselves”, that is, that we are in a situation where linguistic economy can be maximized because speaker and audience have a maximal mutual knowledge (why one speaks at all in such a case is something that we will try to explain below). Internal monologues in the novel of the XIXth century were composed by entire sentences, but the novel of the XXth century, with its fragmentary internal monologues, is more realistic in this sense.

To sum up, introspection does not reveal us the kinds of NL sentences that are capable of expressing complete, exact thoughts, even if these sentences existed.<sup>16</sup> Introspection reveals fragmentary, patently inexplicit pieces of language that cannot fulfill the requirements of Carruthers’s view. Plato said that when thinking the soul speaks to itself. Our position comes close to this: when using language, the subject is “speaking to herself” and by doing this she helps her own thinking. So we do give a cognitive use to NL, not as a vehicle of thought but as a tool that we use for thinking better.

Andy Clark (1997, 1998) has developed a related view about what we do with NL in cognition. Drawing our attention to some habitual uses of language, he brings forward a cognitive role for language that does not turn it into the instrument of thought,

but only into a facilitator (or, as he likes to say, a “scaffolding”) for thought.<sup>17</sup> Quite often, in writing or speaking, one discovers herself having thoughts that otherwise she would not have. By “objectifying” the thoughts and contemplating them, be they written or spoken, one has a different access to them, which seems to provide a different stimulus. According to Clark, although language’s proper function is to communicate our thoughts its capacity to objectify thoughts endows it with a derived function as an aid to reflect, revise and, in general, facilitate our access to our thinking. In Clark’s words, language gives birth to a “second-order cognitive dynamics”, i.e., a return to, and revision of, our thoughts. This way, language becomes an external tool that we use in order to gain knowledge about our own mental life. This also explains why there is a correlation between having a language and being capable of deliberation.

However, there is a more interesting sense in which language is a cognitive tool. There are external tools that are internalized. For instance, we make multiplications with the external help of pencil and paper, but we also do it internalizing that external help and imagining numbers written on a piece of paper. Following a Vygotskian line, Clark suggests that language can be so internalized. As the cognitive uses that we make of language as an external tool become cognitive uses of part of our mental life, it can be regarded as internal to the mind. In general, they are uses that are involved in hard or resource-consuming cognitive tasks, such as multiplication or conscious deliberation, so it is unsurprising that they resort to an external aid.

The picture that results may explain what we can call the ‘real data of introspection’: the fragmentary pieces that we experience are parts of an “inner talk” that helps us to think. Those linguistic items are not “carrying thoughts”: they are just provoking them. And you do not need an especially long sentence to provoke a thought in someone who knows pretty well what you are intending to say. In addition, we think

that the experimental evidence by Spelke and Tsivkin –mentioned above– about the difficulty to integrate information without language does not falsify the hypothesis either. For instance, searching for an object when the relevant information comes from different sources –geometry and color– may be one of those difficult cognitive tasks that the use of NL makes easier. So it is to be expected that blocking the linguistic system will have an effect on the degree of success for that task.

## **Conclusions**

We have argued that the argument from explicitness poses a real problem for the view that minds use natural language as the instrument for thought. Drawing on considerations from philosophy of mind and from recent developments in pragmatics, we have contended that both premises of the argument –that the instrument of thought must be explicit, and that NL is not explicit– are actually plausible. Thoughts demand a kind of completeness and stability of meaning that NL sentences –being remarkably underdetermined– cannot provide.

On the other hand, the real data of introspection speak in favor of a use of NL in thought as if we spoke to ourselves. Two questions spring to mind: (1) if speaking to oneself can be described as a case in which the mind of speaker and hearer is reciprocally transparent, why does one speak at all, if that, or something similar, were indeed what one does?, and (2) if we use Mentalese for thinking, why would we use language too? The answer to both these questions is the same: we use language as a cognitive dynamist, to “contemplate” our own thoughts. By so doing we have an easier access to certain thoughts and trains of them, we can fix our attention on them and we bring them back to memory much more easily. We have argued that this position explains, and coheres with, the real data of introspection, as well as with empirical data

that allegedly support the cognitive view of language. Moreover, this position has no problem with the argument from explicitness: an incomplete or ambiguous expression can help us to think insofar as we can interpret it. We are able to do this if (1) we know enough about the mental states of the speaker (as we do), and (2) the language that we use to interpret the inexplicit utterances is free of ambiguities (as Mentalese is).<sup>18</sup>

*Philosophy Department*

*University of Valladolid*

*Pza. del Campus s/n*

*47011 Valladolid*

*Spain*

*agusvic@fyl.uva.es*

*Department of Philosophy*

*University of Granada*

*Campus Cartuja, Edif. Psicología*

*18071 Granada*

*Spain*

*fnmanriq@ugr.es*

## **References**

Bach, K. 1994. "Conversational Implicature." *Mind and Language*, 9: 124-62.

Bermúdez, J.L. 2003. *Thinking Without Words*. Oxford: Oxford University Press.

Bickerton, D. 1995. *Language and Human Behavior*. Seattle: University of Washington Press.

- Bloom, P. 2000. *How Children Learn the Meanings of Words*. Cambridge, Mass.: MIT Press.
- Borg, E. 2004. *Minimal Semantics*. Oxford: Oxford University Press.
- Cappelen, H. and E. Lepore. 2005. *Insensitive Semantics: A Defense of Semantic Minimalism and Speech Act Pluralism*. Oxford: Blackwell.
- Carruthers, P. 1996. *Language, Thought, and Consciousness: An Essay in Philosophical Psychology*. Cambridge: Cambridge University Press.
- . 2002. "The Cognitive Functions of Language." *Behavioral and Brain Sciences*, 25 (6): 657-726.
- . 2006. *The Architecture of Mind: Massive Modularity and the Flexibility of Thought*. Oxford: Oxford University Press.
- Carston, R. 2002a. *Thoughts and Utterances*. London: Blackwell Publishing.
- . 2002b. "Linguistic Meaning, Communicated Meaning and Cognitive Pragmatics". *Mind and Language* 17: 127-48.
- Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.
- Clark, A. 1997. *Being There*. Cambridge, Mass.: MIT Press.
- . 1998. "Magic Words: How Language Augments Human Computation." In *Language and Thought: Interdisciplinary Themes*, edited by P. Carruthers and J. Boucher, 162-83. Cambridge: Cambridge University Press.
- Culicover, P. and R. Jackendoff. 2004. *Simpler Syntax*. Oxford: Oxford University Press.
- Dennett, D. 1996. *Kinds of Minds*. New York: Basic Books.

- Elugardo, R. 2005. "Fodor's Inexplicitness Argument". In *The Compositionality of Meaning and Content. Volume I: Foundational Issues*, edited by M. Werning, E. Machery and G. Schurz, 59-85. Frankfurt: Ontos-Verlag.
- Fodor, J. 1975. *The Language of Thought*. Cambridge, Mass.: Harvard University Press.
- . 1998. *Concepts*. New York: Oxford University Press.
- . 2001. "Language, Thought and Compositionality." *Mind and Language*, 16: 1-15.
- Hermer-Vazquez, L., E. Spelke, and A. S. Katsnelson. 1999. "Sources of Flexibility in Human Cognition: Dual-task Studies of Space and Language." *Cognitive Psychology* 39: 3-36.
- Hodges, W. 2001. "Formal Features of Compositionality." *Journal of Logic, Language and Information*, 10: 7-28.
- Jackendoff, R. 1997. *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.
- King, J. and J. Stanley. 2004. "Semantics, Pragmatics and the Role of Semantic Content." In *Semantics vs. Pragmatics*, edited by Z. G. Szabo, 111-64. Oxford: Oxford University Press.
- Levinson, S. 1997. "From Outer to Inner Space: Linguistic Categories and Non-Linguistic Thinking." In *Language and Conceptualization*, edited by E. Pederson and J. Nuyts, 13-45. Cambridge: Cambridge University Press.
- . 2003. "Language and Mind: Let's Get the Issues Straight." In *Language in Mind*, edited by D. Gentner and S. Goldin-Meadow, 25-46. Cambridge, Mass.: MIT Press..
- McDermott, D. 1981. "Artificial Intelligence Meets Natural Stupidity." In *Mind Design*, edited by J. Haugeland, 143-60. Cambridge, Mass.: MIT Press.

- Mithen, S. 1995. *The Prehistory of Mind*. London: Thames and Hudson.
- Pagin, P. 2005. "Compositionality and Context." In *Contextualism in Philosophy*, edited by G. Preyer and G. Peter, 303-48. Oxford: Oxford University Press.
- Pinker, S. 1994. *The Language Instinct*. New York: Morrow.
- Recanati, F. 2001. What Is Said. *Synthese*, 128: 75-91.
- . 2002a. "Does Linguistic Communication Rest on Inference?". *Mind and Language*, 17: 105-127.
- . 2002b. "Unarticulated Constituents." *Linguistics and Philosophy*, 25: 299-345.
- . 2004. *Literal Meaning*. Cambridge: Cambridge University Press.
- Sainsbury, M. 2002. "Two Ways to Smoke a Cigarette." In *Meaning and Representation*, edited by E. Borg, 94-114. Oxford: Blackwell.
- Spelke, E. 2003. "What Makes Us Smart? Core Knowledge and Natural Language." In *Language in Mind*, edited by D. Gentner and S. Goldin-Meadow, 277-311. Cambridge, Mass.: MIT Press.
- Spelke, E. and S. Tsivkin. 2001a. "Initial Knowledge and Conceptual Change: Space and Number." In *Language Acquisition and Conceptual Development*, edited by M. Bowerman and S. Levinson, 70-97. Cambridge: Cambridge University Press.
- . 2001b. "Language and Number: A Bilingual Training Study." *Cognition*, 78: 45-88.
- Sperber, D. and D. Wilson. 1986. *Relevance: Communication and Cognition*. (2<sup>nd</sup> edition, 1995). Oxford: Blackwell.
- Stanley, J. 2002. "Making It Articulated." *Mind and Language*, 17: 149-169.
- Stanley, J. and Z. G. Szabo. 2000. "On Quantifier Domain Restriction." *Mind and Language*, 15: 219-261.

- Tomasello, M. 2001. "Perceiving Intentions and Learning Words in the Second Year of Life." In *Language Acquisition and Conceptual Development*, edited by M. Bowerman, M. and S. Levinson, 132-58. Cambridge: Cambridge University Press.
- Travis, C. 1985. "On What Is Strictly Speaking True." *Canadian Journal of Philosophy*, 15:187-229.
- . 1996. "Meaning's Role in Truth." *Mind*, 105: 451-66.
- . 2000. *Unshadowed Thought: Representation in Thought and Language*. Cambridge, Mass.: Harvard University Press.
- Vicente, A. and Martínez Manrique, F. 2005. "Semantic Underdetermination and the Cognitive Uses of Language". *Mind and Language*, 20: 537-558.

---

<sup>1</sup>A word about this: It is in the context of discussing this last point (i.e., whether images can be another vehicle of thought or not) where Carruthers concedes that some thoughts may use a mixed vehicle. He tries to show that images alone cannot carry thoughts (they are non-propositional), but concedes that they may be involved in some thoughts. Even imagistic thoughts are language-related: "in many of those cases where a thought is carried by an image of an object, an embedding in a linguistic context may be necessary to confer on that image a determinate content", (1996, 253). For this reason, we take it that Carruthers's cautious phrasing of his theses as "conscious thinking involves [instead of 'uses only'] public language" is directed to convince the reader that his role for language covers thoughts that include images, and not to deny that some thoughts are entirely linguistic. This is a direct consequence of his introspection argument: he takes for granted that some thoughts appear consciously as completely linguistic. What he is at pains to show is that even those thoughts that do not appear as linguistic (i.e.,

---

imagistic thoughts) have often a linguistic component. Be it as it may, we take it that the arguments to be considered apply equally (or with minor variations -see later) to these two theses: (a) conscious (or whatever) thinking involves NL, and (b) conscious (or whatever) thinking uses NL as its only vehicle.

<sup>2</sup> See chapter 3. There Carruthers only claims that it has been shown by Sperber and others that NL is multiply ambiguous so the content of its sentences cannot be propositional. In our opinion, the argument has to be developed further in order to make it sound convincing, and this is what we mean to do in this paper.

<sup>3</sup> We will modify this formulation later on.

<sup>4</sup> This is the path we follow in Vicente and Martínez-Manrique (2005), where we discuss two different ways of construing the argument from semantic underdeterminacy to the non-compositionality of natural language.

<sup>5</sup> As Carston (2002b, 132) points out, Recanati's intuitive truth-conditions, and his correlated notion of 'what is said', denote very much the same sort of entity as Bach's 'implicature' (Bach 1994) or Sperber and Wilson's 'explicature' (Sperber and Wilson 1986/95), even though there are some major differences in their wider semantic/pragmatic frameworks.

<sup>6</sup> For a characterization of and an attack on minimalism, see Recanati (2001). For a defense see Borg (2004), Cappelen and Lepore (2005). It is not clear to us, however, whether Cappelen and Lepore would endorse this model of linguistic processing. It is true that they defend the psychological reality of minimal propositions, but we think that not in the sense that they are entertained before enriched ones, i.e., as necessary steps in the process of interpretation.

<sup>7</sup> One may relax this condition, and make it gradable, establishing that L is explicit iff *most* of its sentences have their truth conditions given just by its semantics. This allows

---

for a language that is largely explicit but with small "islands" of non-explicit sentences.

We take it that Recanati refers to this second, gradable sense when he argues that the majority of NL sentences are semantically underdeterminate.

<sup>8</sup> Fodor's (2001) argument for the inexplicitness of NL uses examples of sentences that are pragmatically enriched, so it fails to show that NL is not explicit in the weak sense. Consequently, it falls short of proving that NL cannot be the vehicle of thought: it might be the case that the NL sentences deployed in thought are explicit, not the inexplicit sentences that we actually utter.

<sup>9</sup> Sometimes the label of 'underdeterminacy' is used to cover all cases of inexplicitness, including both "genuine" underdeterminacy and examples of pragmatic enrichment in general (see, e.g. Carston, 2002a). For our purposes, it is important to distinguish the inexplicitness caused by the presence of underdetermined terms (the dangerous kind) from that whose origin is pragmatic enrichment (the remediable kind).

<sup>10</sup> The notion of 'optional' for Recanati is not so straightforward as it seems at first sight. We think that the following is the most plausible reading of his views: The elements added by free enrichment, such as ['here'] in (1), are optional in the sense that they are not necessary in order to obtain a truth-conditional rendition of the sentence, but they are not optional in order to obtain the *intuitive* truth-conditions. In other words, there is a sense in which the context *mandates* the addition of a particular element, on penalty of misunderstanding in a gross manner what the speaker intends to say. This contrasts with the process of obtaining an implicature, which is much more optional: the hearer may disavow the implication 'you ought to take an umbrella' that may be inferred from an utterance of (1).

---

<sup>11</sup> And not even all of them, only the so-called *pure* indexicals (if there are any).

Demonstratives, for instance, also require the wide context to be given a suitable assignment.

<sup>12</sup> See Pagin (2005) for an extended discussion on compositionality and unarticulated constituents. Pagin's paper is also devoted to the discussion of Fodor's (2001) argument against the use of language in thinking. However, Pagin focuses on what Fodor has to say about compositionality of language and thought and on the examples he does use of provision of unarticulated constituents. Here we are pulling from the explicitness thread, with the conviction that it is explicitness and not compositionality what is the real issue in these matters.

<sup>13</sup> The underdeterminacy in 'Ringo is the Beatle I like the most' can be solved by means of an explication: 'the statue of Ringo is the statue of a Beatle I like the most'. This should clarify the possible ambiguity, but does it? Well, it clarifies *that* ambiguity, but this new sentence happens to be ambiguous itself, for we could be speaking either about the statue of Ringo or about a picture of it (we took pictures of all the statues and the picture of Ringo's statue happens to be the best). Culicover and Jackendoff sum up the lessons of these examples in the following "statue rule" (2004, 371):

"The syntactic structure NP may correspond to the semantic/conceptual structure PHYSICAL REPRESENTATION (X) where X is the ordinary interpretation of NP".

<sup>14</sup> We thank an anonymous referee for pressing this point.

<sup>15</sup> This is not a question about whether we introspect or entertain phonological representations of the sentence 'John is ready'. We do not claim that we don't. We do introspect tokens of that kind of sentences, as well as more fragmentary discourse. The question is about whether when we do so we are also entertaining the minimal thought that these kinds of sentences are said to express.

---

<sup>16</sup> This also holds for NL-plus-images sentences.

<sup>17</sup> Jackendoff (1997) suggests a similar role for language, from assumptions having to do with his theory of consciousness. The moral of the story is pretty similar to the one told in this section, given that the cognitive role of language arises from its being at the interface between mind and world. Bermúdez (2003) endorses Clark's overall view, while at the same time holding that NL is the vehicle of the typically human thought. Carruthers (2006) comes close to Clark's position, and tries to integrate it within a modularist view of the mind, where NL would be a means to globally broadcast our thoughts.

<sup>18</sup> This paper is thoroughly collaborative. Order of authorship is arbitrary. The paper has benefited from discussions held at the University of Granada and Barcelona University (Logos group). We are especially grateful to an anonymous referee of *Metaphilosophy*. (S)he has done much to improve our first draft. Work for this paper has been funded by Research Projects HUM2005-03211/FISO, HUM2005-07539-C02-00 and HUM2005-07358.