Clusters:

On the structure of lexical concepts

Agustín VICENTE
Ikerbasque Foundation for Science/ University of the Basque Country
Philosophy Department
Avda. Tolosa 70, 20028 Donostia, Spain
agustin_vicente@ehu.es

ABSTRACT

The paper argues for a decompositionalist account of lexical concepts. In particular, it presents and argues for a cluster decompositionalism, a view that claims that the complexes a token of a word corresponds to on a given occasion are typically built out of a determinate set of basic concepts, most of which are present on most other occasions of use of the word. The first part of the paper discusses some explanatory virtues of decompositionalism in general. The second singles out cluster decompositionalism as the best explanation of the variability of meaning. The third part is devoted to responding to some problems.

Introduction

This paper is devoted to the issue of the structure of lexical concepts, i.e. those concepts which *prima facie* correspond to most of our linguistic simple terms[1]. It is typically assumed that such concepts are of special importance not just because they are lexical, but also because they are the concepts that we typically use in our thinking and in our categorizing of the world. Thus, the paper will deal with those concepts that fulfill the following roles: (a) they are the meanings of the lexical items of a language, (b) they are used in drawing distinctions in the world, (c) they are the constituents (though not necessarily the ultimate constituents) of our thoughts. And the questions to be addressed are: do such concepts have structure? And if they do, what kind of structure do they have?

[1] There is a dispute about whether all linguistic terms express concepts or not. In principle, it can be defended that while most words have conceptual meanings, some of them, such as connectives or quantifiers, do not. This distinction is captured by the classical division between categorematic and syncategorematic terms. In Relevance Theory the distinction is drawn between conceptual and procedural meaning, but Bezuidenhout (2004) calls this distinction into question.

There are three main types of answers to these questions. First, we have the "classical theory", according to which lexical concepts have a definitional structure, a canonical decomposition in terms of a limited set of atomic concepts. The empiricist version of this theory has it that such atomic concepts are observational, whereas most contemporary versions take the set of atomic concepts to be highly abstract, probably innate, concepts such as OBJECT, AGENT, CAUSE, MOTION, CONTAINMENT, SUPPORT, PATH or GOAL. This version of the theory seems to be specially appreciated in linguistics, where it was first developed within generative semantics, drawing on an analogy with the feature account of phonemes. It has been under heavy attack ever since, but it still has well-known defenders such as Steven Pinker (2007). Second, there is the "prototype theory". This theory claims that all concepts, lexical concepts included, have a prototypical structure, by which it is roughly meant that concepts consist in a prototype, which is abstracted from exemplars, situated at their centre and a metric of similarity with respect to it. The origins of the prototype theory trace back to Eleanor Rosch's studies of the 70s' (Rosch and Mervis, 1975, Rosch, 1978) and it is inspired in Wittgenstein's family resemblance metaphor. It is a view that does a good job of accounting for our categorization of the world, and is mainly adhered to by psychologists (see Murphy, 2002). Last, we have Jerry Fodor's (1998) atomism, which holds that lexical concepts do not have any structure at all. Every simple term of a language expresses an atomic concept, which is a mental representation with a certain denotation in the world. The motivation for this theory is basically negative: it is free from the problems that seem to haunt rival accounts. However, it does a good job of explaining two features of concepts which are regarded as crucial: the compositional nature of complex concepts -and, ultimately, of thought- and the fact that we can be mistaken in our categorizations. Besides Fodor, atomism is defended by philosophers such as Stephen Laurence and Eric Margolis (1999) and linguists and pragmaticians like Dan Sperber and Deirdre Wilson (1998), Stephen Levinson (2003) and Robyn Carston (2002).

In what follows, I want to defend a decompositionalist account of lexical concepts which draws heavily on the linguistic version of the classical theory, but which tries to amend its drawbacks. In particular, it is not committed to the existence of definitions for lexical concepts. Rather, it denies that lexical concepts have canonical decompositions. There are several well-known arguments against the definitional structure of lexical

concepts, but most of them highlight the idea that for any definition of any concept that we can think of, we will find cases where the *definiens* is correctly applied while the *definiendum* is not and *vice versa*. One such argument makes use of Wittgenstein's reflections in *Philosophical Investigations* concerning the inexistence of a definition for the meaning of 'game'. Inspired by these reflections, philosophers such as John Searle (1958) purported to amend the classical theory of linguistic meaning by holding that the meaning of a word cannot be captured by a set of necessary and sufficient conditions for its correct application, but rather that the meaning of a word corresponds to a *cluster* of properties which heuristically guide our uses of the word (such that if an important number of the properties apply, the word is taken to be used correctly). What I want to defend is a sort of *cluster theory* for lexical concepts. A simple term in our language does not correspond to a complex concept with a definitional structure: to begin with, the relation between words and concepts is one-to-many. However, the complexes a token of a word corresponds to on a given occasion are typically built out of a determinate set of basic concepts, most of which are present on most other occasions of use of the word[2].

The basic virtue of this decompositionalism is that it can account for the hallmark of linguistic meaning: its stability on the one hand and its variability on the other. However, it has what seem to be several other virtues, which it shares with other, more "classical", decompositionalisms and which have been extensively discussed by other authors. What I will do in what follows is present briefly three of these virtues, which concern the topics of argument realization, impossible words and language and concept acquisition. The discussion of these three issues will hopefully help to see that the decompositionalist program, though still controversial, can offer explanations that go deeper than its rivals into as yet unexplained issues. That is, in this first part I do not purport to show that decompositionalism is right, but only that it gives us a good starting point into some intriguing and complicated phenomena. After that, I will move on to focus on the issue of stability and variability. There I will argue that the kind of decompositionalism I want to defend is the best explanation for what recent contextualism in pragmatics has shown to be the case, i.e. that, in a vast majority of cases, sentential utterances of a given type may and do express different propositions

[2] The idea of lexical concepts having a cluster structure clearly resembles the prototype theory and Jackendoff's preference rule systems (see Jackendoff, 1989). It differs from the former in the kind of features it employs and in its distrust about the centrality of basic level concepts. It departs from the latter in taking it that all –not only some- concepts but basic primitives have a cluster structure.

(have different truth-conditions) in different situations. Finally, I will consider whether some of the typical problems for decompositionalism apply to this version of the theory.

1. Argument realization, impossible words, and language and concept acquisition

1.1. Argument Realization

The issue of argument realization concerns the possible syntactic expressions of the arguments of a verb. For a long time now, it has been proposed (see Beth Levin and Malka Rappaport Hovav's (2005) excellent introduction) that the morphosyntactic expressions of the arguments of a given verb are, at least partly, a projection of the semantics of the verb[3]. There are various theories about how such projection takes place, i.e., about which elements or facets of the meanings of verbs are relevant to the mapping from semantics to syntax. I will briefly introduce one such theory, the *semantic role* theory, which is widely known, so that we can get an idea of what kinds of phenomena need to be explained. Then I will show where it fails. Afterwards, I will go into how decompositionalism tackles the explanatory problems the semantic role theory has, and move on to show that a true decompositionalism can contribute to clarify some other unexplained phenomena. Before I get into it, though, there is an important point to make: these theories, being theories about how semantics projects itself into syntax, are also theories about what needs to be learnt/ known about semantics so that knowledge of syntax is at least facilitated. This is important, because it means that these theories must be constrained by theories and data coming from developmental psychology.

The semantic role theory

The *semantic role* theory holds that each verb has an associated semantic role list which is responsible for its morphosyntactic behaviour. The list draws from a set of semantic roles which include agent, counteragent, object/patient, result, instrument, source, goal and experiencer[4]. Then there are some projection rules based on a hierarchy of roles. For instance, one of these rules states that if the list of semantic roles associated with a verb contains an agent, then the participant with the agent role will occupy the subject

[3] Partly, because it seems that the properties of the arguments of the verb, as well as their fillers, also have a word to say: see Levin and Rappaport Hovav (2005, ch. 7).
[4] This is Fillmore's (1968) original list. Most authors agree that it has to be supplemented.

position. If there is an instrument in the list, then the realizer of the instrument role can also be the subject, but only in the absence of an agent. This explains why these four sentences are acceptable:

(1) John broke the window with a rock;
(2) A rock broke the window;
(3) John hit the fence with a stick;
(4) A stick hit the fence.

However, an object/patient, which is usually selected as a direct object in transitive verbs, can also occupy a subject position in its intransitive counterparts, whereas a place cannot. This accounts for the correctness of (5) and the incorrectness of (6):

(5) The window broke;
(6) *The fence hit.

This is a sketchy and simplified statement of the semantic role theory. However, it is enough to get a rough idea of what it looks like and to understand what its weaknesses are, which is what interests us at this point. The theory has several problems that have to do with predictive failures that I will keep aside (see Levin and Rappaport Hovav, 2005, for a detailed critique). More interesting are its explanatory debts. Levin and Rappaport Hovav (2005) list three explanatory failures: (i) The theory does not explain why we find some combinations of semantic roles and not others; for instance, it seems that no verb can be associated with a list such as “place, experiencer, patient”: why is this so? (ii) The theory says nothing about the number of roles that can be associated with a given verb; however, the lists tend to be restricted to three roles: why is there such regularity? (iii) The theory does not explain why generalizations, such as “agents take subject position” hold, and, in general, it does not explain what backs the hierarchy of roles proposed.

Decompositionalism: outline of the theory

Decompositionalist theories can provide a more explanatory account of these issues. According to decompositionalism, the syntactic behaviour of verbs depends on their

lexical decomposition. For instance, all causative verbs can be said to have a decomposition of the form "[x CAUSE [y BECOME <z>]". Thus, the structure of 'sink' in 'the captain sunk the boat' is [x CAUSE [y BECOME <SUNK>], with 'the captain' saturating the *x*, and 'the boat' occupying the place of the *y*[5]. The structure of 'splash', in turn, would be [x CAUSE [y BECOME <SPLASHED>], etc. Now, this makes clear that causatives are transitive verbs that demand two arguments: a causer and a theme or patient. It also explains why agents are expressed as subjects and why there is a hierarchy within semantic roles in the way they occupy subject positions: the hierarchy is imposed by the structure of the type of event denoted by the verb; thus, in a causal event such as a sinking event, the agent appears as the subject because it is the first, non-embedded, argument in the event structure, which in turn seems to be a projection of causal order[6]. This approach is not free from problems, though: on the one hand, it seems to provide no reason why the subject of 'want' is the person who wants instead of the wanted object or proposition: nothing in the event structure denoted by 'want' establishes that; on the other, it arguably should predict that the subject of 'see' is the object seen instead of the experiencer, given that it is the causer of the seeing event.

This very general view of argument realization has to be improved, though, for two main reasons. First, whatever its merits as a linguistic theory, it is disappointing as a decompositionalist theory. It may well explain why causatives take at least two arguments, but it seems to be content with too many primitives, since for every causative that is decomposed we have an intransitive that functions as a primitive. Such primitives are called 'roots' in the theory, and they are not usually further analyzed. Second, the theory, as has been presented, cannot explain why certain alternations hold and others do not. For instance, (7), (8) and (9) are acceptable, but (10) is not (see Pinker, 2007):

(7) John loaded the van with bricks;
(8) John loaded bricks into the van;
(9) John filled the van with bricks;
(10) *John filled bricks into the van.

[5] In what follows, I will mainly restrict the discussion to causatives.
[6] I am here lumping together two different theories of thematic hierarchies: the event structure account and the causal order account. However, their differences are not important for the purposes of the paper.

In principle, ‘load’ and ‘fill’ denote events of the same structure (CAUSE [y BECOME <LOADED/ FILLED>). Why then can ‘load’ alternate and ‘fill’ not? Levin and Rapaport Hovav (1998) propose that the meaning of the content-locative construction (8) is that A causes B to go to C, whereas the meaning of the container-locative construction (7) and (9) is that A causes C to change state. Now, why is it that ‘fill’ cannot express that something has been caused to go somewhere?

Refinement

Drawing on Levin (1993), the answer Pinker (2007) gives to these questions can solve the other concern expressed, i.e. the one related to massive primitivism. In order to explain alternations, be they locative, dative or of some other comparable kind, we have to look at the meaning of the root in the event structure. An adequate decomposition of it will tell us why some verbs enter into certain alternations and others do not. Thus, he says, ‘fill’ “is all about the state of the container… It is apathetic about how the container became full. By specifying the change of state of a container, it is compatible with a construction that is about a state-change… Because it says nothing about a cause or a manner of motion of the contents, it isn’t compatible with a construction that is all about motion”. On the other hand, ‘load’ “simultaneously specifies something about how a content has been moved *and* how the container has changed (p. 50; his italics)”. His final proposal is that this kind of information can be captured by a composition of concepts drawn from a limited list of basics. It is certainly difficult to come up with a list of basics, but Pinker offers a tentative beginning:

> (i) a cast of basic concepts: event, state, thing, place, property, manner; (ii) a set of relationships that enmesh these concepts with one another: acting, going, being, having; (iii) a taxonomy of entities: human vs. nonhuman, animate vs. inanimate, object vs. stuff, individual vs. collection, flexible vs. rigid, one-dimensional vs. two-dimensional vs. three-dimensional; (iv) a system of spatial concepts to define places and paths, like the meanings of ‘on’, ‘at’, ‘in’, ‘to’, and ‘under’; (v) a time line that orders events, and that distinguishes instantaneous points, bounded intervals, and indefinite regions; (vi) a family of causal relationships: causing,

letting, enabling, preventing, impeding, encouraging; (vii) the concept of a goal, and the distinction between means and ends (p. 81)[7].

According to Pinker, this basic repertoire is the main tool we have in order to account for argument alternations in particular and argument realization in general. That is, the ingredients that give us the meaning of the root verb, the ones responsible for its argument alternations, are the ones listed above. Predictions and explanations of alternations are impossible without correct decompositions of root verbs into these atoms. The list of basics is the main tool, but it is not the end of the story: argument alternations are not universal. Thus, we find once and again that the counterpart in language L' of a verb *v* of L refuses to enter into the locative alternation, while the alternation for *v* is admissible in L. According to Pinker, this is due to the particular conventions of a linguistic community. Different linguistics communities form different "microclasses" of verbs with similar meanings (decompositions) and similar morphosyntactic behaviour. The metric of similarity is conventionally established, but the point is that we have to know the meaning of a given verb to form a hypothesis about its behaviour[8].

This decompositional account of argument realization and argument alternations has some problems (see Levin and Rappaport Hovav, 2005) and perhaps does not emerge as a simple, elegant, theory at the end of the day. Yet, the point to highlight is that it seems to be better than any of its rivals, so that it seems to be a (or rather, *the*) research line worth pursuing.

1.2. Impossible Words

[7] Pinker's thesis is that by explaining argument realization in this way we thereby discover the basic conceptual units of our thought. This is so because the primitives that account for argument realization are the first concepts to arise in our minds. For, it has to be remembered, the theory does not purport to explain just why certain verbs have certain grammatical expressions, but also why learning a grammar is facilitated by knowing the meaning of words. So, the units we use in our decompositions have to be empirically adequate in two ways: (i) they have to be able to explain the phenomena related to argument realization, and (ii) they have to respect what developmental psychology has to say about the first concepts to arise in our thought.

[8] I have some doubts about the adequacy of this general account. In particular, I do not know how it can explain the behaviour of verbs that stand for basics. For instance, how does it account for the argument realization of 'cause'? It cannot be that 'cause' has the argument structure that it has in virtue of having a certain decomposition, because it is a primitive. Then, does the argument realization of some verbs have one kind of explanation, and that of others another? It could be said that the knowledge of the argument structure of 'cause' simply falls out from having the CAUSE concept. But that response opens the door for the atomist to generalize that account to the rest of verbs. Probably what has to be said is that, given that we learn the verb 'cause' long after other causatives have been acquired, we make use of our knowledge of the morphosyntactic behaviour of the latter.

The issue of impossible words is related to all this. We don't find a verb such as 'blik', which enters into constructions such as 'the desk blikked Mary', meaning that Mary broke the desk (see Johnson, 2004). It is not just that we do not find any such verb, but that it looks like its existence is impossible. Now, why is this? The decompositionalist has an answer: the event structure of MARY BROKE THE DESK is MARY [CAUSED [THE DESK BECOME <BROKEN>]. Now, if this is so, then 'Mary', not 'the desk', has to occupy the subject position in the sentence, for it denotes the unembedded argument, which in turn is the causal origin of the causal event, while 'the desk' denotes the patient. This kind of explanation is unavailable to an atomist, who denies that verb meanings have any kind of structure.

Fodor and Lepore (1999, 2005) have responded to this argument in the following way: it is true that 'blik' is an impossible word in English. However, the atomist and the decompositionalist are on an equal footing concerning the explanation of such impossibility. For the only thing that the decompositionalist could explain is why 'blik' is not a possible word assuming its meaning is derived from others, i.e. assuming it has a decompositional structure, i.e. assuming it is not a primitive. However, BLIK could be an underived notion. The decompositionalist cannot exclude this, according to Fodor and Lepore, since she lacks an explanation of why BLIK is not, or could not, be a primitive. Thus, the atomist and the decompositionalist have the same problem, namely, that they cannot explain why 'blik' is not a word in English, except by telling us that we lack the concept BLIK.

However, I think Fodor and Lepore's argument is feeble. The decompositionalist can safely assume that if BLIK is going to be a concept of our conceptual repertoire, it has to be derived. It is true that, despite Pinker's attempt, we probably do not know which are the basic elements of our conceptual machinery. However, we know that we have to try to do with (a) concepts that tend to arise early in development and universally, and (b) concepts that can be theoretically fruitful, by which I mean concepts that can enter into compositions and explain the structure of other concepts. None of these constraints are fulfilled by BLIK, so it seems safe to exclude it from the list of basics. And if it is not a basic notion, it must be derived, if 'blik' is going to be a possible word at all. Now, as it cannot be derived, it is explained why it is not a possible word.

That is, the decompositionalist has a perfectly sound explanation of why impossible words are indeed impossible. In a nutshell: the meaning of an impossible

word is not a primitive of our conceptual repertoire. Since it is not a primitive, it must have a decomposition. However, once we have the decomposition of the putative concept the word corresponds to, we see that the word cannot have such a meaning and at the same time behave in the way it is supposed to behave.

1.3. Language and concept acquisition

Now, for the topic of concept and language acquisition. Pinker (2007) presents his decompositionalist proposal as an attempt at explaining how children can form their hypotheses about the morphosyntactic behaviour of verbs. In fact, all theories about argument realization that share the idea that argument expression is a projection of the semantic properties of verbs are explanations of how children need to know the meaning of verbs before they can form any hypothesis about their morphosyntactic argumental expression. The idea, in a nutshell, is that once they can guess what the decomposition associated with a verb is, they can start projecting their semantic competence into syntactic structures. This solves some problems about how syntax starts to be learnt.

On the other hand, decompositionalist theories have always been a good start on the issue of concept acquisition. As is well known, Fodor (1975) posed a problem for any theory of concept learning: since to learn something consists in forming and confirming hypotheses, concept learning is impossible, for the only way to form a hypothesis about what a concept consists in is by having it beforehand. This problem still haunts theories of conceptual development nowadays (see, e.g. Carey, 2001). Many psychologists reject Fodor's model of learning, but it is certainly difficult to find an adequate substitute[9]. The point, in any case, is that Fodor's problem concerns only simple atomic concepts. Complex concepts can be learnt, at least in the sense that the complex can be a new mental entity acquired through hypotheses-testing and not be available without it. The compounds of the complex may all be innate; yet, the complex

[9] Apart from inductive and probabilistic models of learning, there are some interesting recent proposals. Thus, Carey and Sarnecka (2005), along with other developmental psychologists concerned with number concepts, suggest that concepts can be learnt by a bootstrapping strategy in which language plays a major role. Dewar and Xu's (2009) impressive result that mere labelling seems to generate categories is another approach worth taking into account (although Xu seems to be committed to there being innate concepts). Weiskopf (2007) has recently proposed a model for concept acquisition (which resembles what Herder had to say about the origins of thought) that resorts to internal coining and Margolis (1998) developed the idea of "sustaining mechanisms" for the acquisition of atomic concepts. So, perhaps the correct way to put my point, as an anonymous referee suggests, is in conditional form: if concept learning implies hypothesis-testing, then decompositionalism fares better than atomism.

as such may be learnt. Thus, the narrower the list of atomic concepts is, the more concepts can be acquired by means of a learning process. So the decompositionalist is in a better position than the atomist if we want to preserve the idea that most concepts arise in the mind as a result of learning. Now, there seems to be a growing consensus on what concepts are the first to arise in a child's mind. Both nativists (such as Spelke and Carey) and anti-nativists (such as Mandler) coincide in citing the dichotomy between OBJECT and AGENT (and the related notion of GOAL), the CAUSE relation and a number of spatial notions such as PATH, MOTION, CONTAINMENT and SUPPORT[10]. These sole concepts cannot be the only building blocks of the rest of the concepts, and perhaps they admit further decompositions (e.g. AGENT can be said to decompose in terms of SELF-MOTION, CHANGE-OF-STATE-CAUSED-AT-A-DISTANCE and some other concepts). And most probably, treating lexical concepts as compounds of basic concepts will not solve the issue of concept acquisition completely, for it seems that at least some concepts, such as number concepts (see Carey and Sarnecka, 2005) cannot be acquired by forming complexes. The point, however, is that if we have a stock of concepts which include these and similar abstract notions and we are able to provide decompositions of lexical habitual concepts using them, then we are in a good position to simplify the issue of concept acquisition.

Fodor himself acknowledged this much when in his (1975) he proposed a decompositionalist account of concepts. He repudiated it when he became convinced that concepts do not have canonical decompositions (i.e. definitions). Now, if we are able to do without definitions while remaining decompositionalists, which is what I propose here, then this is an explanatory virtue of decompositionalism to be taken into account.

2. Cluster decompositions

The explanations of argument realization, impossible words and concept acquisition are possibly the three main strengths of any version of the linguists' account of decompositionalism, the classical theory included. What I want to do in this section is show that a cluster decompositionalism fares significantly better than the classical

[10] Mandler (2004) is very explicit and convincing about the stages in the mind's conceptual development. According to her, the first concepts to arise are these very general and abstract notions. Basic-level concepts arise only much later, probably as compounds of these abstract notions. The fact that children acquire basic-level words first is due simply to their being the words they hear more often.

theory in explaining some other facts, in particular, facts about meaning. The classical theory, as well as atomism, is committed to words' having univocal meanings. I will argue that this commitment is unfounded: the relation between words and concepts (or concept structures) is one-to-many, and what concept a token of a word expresses depends on the linguistic and extra-linguistic context. The first and second subsections are devoted to rejecting the classical theory. The first develops some typical problems for the definitional account and gives a first step towards a cluster solution. The second discusses some amendments to the classical view, which try to account for the variability of word meanings with the sole resource of canonical decompositions. It will be argued that this account can explain only a very restricted kind of phenomena. The last subsection develops the contextualist thesis that utterance meaning is semantically underdetermined. This thesis implies that what concept or conceptual structure a token of a word expresses depends both on the linguistic and on the extralinguistic context. Some contextualists grant this much, but remain atomist. I try to show then that a decompositionalist account is a simpler account of the phenomenon of semantic underdetermination. In conclusion, I try to convince the reader that a cluster decompositionalism is a more natural and successful way to explain the variability of word meanings than either the classical theory or atomism.

2.1. The problem of definitions

As explained above, the classical view has two important problems: one is the "problem of ignorance and error" (Laurence and Margolis, 1999), in principle shared by all non-denotationalist theories; the other is the inexistence of definitions. The problem of ignorance and error has its origins in Kripke's and Putnam's observations that the meaning of a proper name or a natural kind term cannot have a definitional structure, since any definition we can think of can fail to capture the bearer of the name or the natural kind. That is, definitions can be wrong routes to categories.

However, the most devastating problem for the classical view is the absence of definitions. Philosophy can be thought as the search for definitions of a restricted but at the same time wide range of terms. Now, we can safely say that no proposed definition has been shown to be free from counterexamples, and thus that we have quite strong inductive evidence against the view that words can be defined in terms of other words. This problem is indeed as damaging as it is taken to be. However, it may be good to briefly explore, first of all, one way the defender of the classical view may react to this

kind of inductive arguments. Basically, the strategy consists in claiming that what has been identified is a problem of words, not of concepts. One thing is a definition of words in terms of other words and another a definition of concepts in terms of other concepts. A problem for the former will be a problem for the latter only if concepts can be captured by single words, e.g. if the meaning of 'cause' is CAUSE[11]. However, we can drop this assumption. We can say that concepts are ineffable, so that CAUSE is an abstract concept which does not correspond to the variable meaning of 'cause'. If this move is legitimate, then showing that, e.g. 'knowledge' cannot be defined in terms of 'justified', 'true' and 'belief' would not show that KNOWLEDGE does not have a canonical decomposition in terms of JUSTIFIED, TRUE and BELIEF[12].

However, the recourse to ineffability, which might be of some use in other debates will not save the defender of the classical view[13]. Even if basic concepts are ineffable, definitions of concepts in terms of other concepts can fail just as definitions of words in terms of other words do fail. Suppose we say that KILL can be defined as CAUSE TO DIE. Now, as Fodor (1981) claims, one can cause someone to die without killing her. This means that one can apply the complex CAUSE TO DIE without applying the concept KILL. Nothing here has to do with the relation between words and concepts.

The same holds for the much debated concept PAINT (see Fodor, 1981, Pinker, 2007). Pinker (2007: 102) proposes a detailed decomposition of such a concept: x

[11] Some decompositionalists (see e.g. Fodor's quote of Jackendoff's: Fodor, 1998; 52, Johnson, 2004) in effect hold that CAUSE is not the meaning of 'cause' but is an ineffable concept. Thus, it seems open for them to say further that the problem of definitions is not a problem that affects concepts. I am not sure whether this quote of Jackendoff's (1989) expresses precisely this idea: "[T]his observation shows only that if there are principles of lexical conceptual composition, they are not entirely identical with the principles of phrasal lexical composition".

[12] In a number of places, Fodor (e.g. 1998 and Fodor and Lepore, 2005) has opposed this move on the grounds that it comes down to a rejection of the idea that language expresses thought. In effect, it seems that if we say that concepts are ineffable, then the contents of our thoughts must be ineffable too. However, I think it is possible to hold both that concepts are ineffable and that language expresses thought. The point is to clarify what is meant by 'ineffability'. That concepts are ineffable does not have to mean that they cannot be expressed by linguistic means at all. It may just mean that concepts cannot be captured by single words alone. If we go along with recent contextualism, we can say that words express concepts only in a (linguistic and extralinguistic) context. If this is so, concepts cannot be captured by words or even by words-in-sentences, for a change in the context of utterance of a sentence can bring about a change in the thought it expresses. Now, this shows that neither sentences nor words have univocal meanings. However, it does not imply that thoughts cannot be expressed. They can, but they are expressed by utterances, not by sentences. So the issue of ineffability does not have to do with whether or not a basic concept can be expressed by linguistic means –it can, in the context of an utterance- but with whether or not we have the means to produce a sentence which captures it as part of its proper/encoded meaning.

[13] The recourse to the ineffability of basic concepts seems to provide a way out to some regress problems raised by Fodor (1998). I deal with these in the last section. I can advance that I do not think that regarding basic concepts as ineffable helps in any way.

paints y, according to him, if and only if x is an agent, and x covers intentionally the surface of y with paint and does so not as a means to cover the surface of z with paint, but as an end in itself. However, there are cases where the concept PAINT is applied and its putative decomposition is not, be it because the decomposition is too detailed or because it is too general. For instance, (11) and (12) are cases where the complex associated with 'paint' must be *prima facie* simpler than Pinker's definition, since we drop at least the agency and intentionality features:

(11) The machine is painting the cars with the wrong colour;

(12) After the Exxon Valdez disaster, the sea painted the beaches black.

Also, the *coda* related to the distinction between means and ends can be relaxed. Pinker uses it to explain why we do not think that the brush we are using to paint a wall counts as painted, even if its surface is intentionally covered with paint. But now suppose I am painting the wall by attaching to it cloths that I have coloured before. Will I be wrong if I say (or think) (13)?

(13) First I paint these cloths; then I paint the wall with them.

On the other hand, (14) and (15) show that 'paint' can stand for an even more detailed complex of concepts:

(14) I want you to paint the house;

(15) I want you to paint the car.

To paint a house is, by default, to paint its inside surface, and a house can be taken as painted even if not all of its surfaces are painted. To paint a car, on the other hand, is to paint its exterior surface, and a car is taken as painted if it is painted all over.

Last, suppose you are a professional painter. Your assistant gives only a first coat of paint to the walls and then leaves. She has covered the surface of the walls with paint. However, when you arrive and see what she's done, you utter:

(16) That's *not* painting a house!

Thus, PAINT is not applied homogeneously: it matters whether you are thinking about machines or humans and brushes or cloths, that is, inanimate or animate entities and habitual or non-habitual painting instruments. It also matters whether you are thinking of houses, boxes or cars. The upshot is that there is no such a thing as *the* PAINT concept, but a variety of compounds that can be expressed by tokens of the word

‘paint’ and tend to share some of the concepts in Pinker’s compound. For instance, COVER A SURFACE WITH PAINT (which we know cannot be the definition of PAINT) seems to be present in all the uses of ‘paint’ that we have looked at. But usually ‘paint’ points to a more complex concept. On many occasions, it points to Pinker’s compound or something similar, but in many other cases it points either to an even more detailed complex or to a less stringent characterization. And as has been said above, all this involves concepts directly, not words.

2.2. Co-compositions

The strategy

The example of “painting a house” vs “painting a car” illustrates one of the sources of the variability of the correspondence between words and concepts, namely the linguistic context. Some authors refer to this variability as co-composition (Pustejovsky, 1995, Jackendoff, 2002). According to them, in linguistic compounds, the conceptual information of lexical entries may determine the conceptual expression of others. Pustejovsky (1995) holds that such co-compositions are generative and can be predicted from the knowledge of the canonical decompositions of the concepts involved. This idea can be seen as an amendment of the classical theory: the conceptual structure a word stands for is not always its canonical decomposition, but what conceptual structure a certain word-token expresses can be calculated on the basis of its canonical decomposition and the canonical decompositions of the words that form part of the sentence it appears in. So, in an important sense, words do have definitions. It can predicted, from what ‘jump’ and ‘run’ canonically mean, that ‘he jumped for three hours’ means that he jumped *repeatedly* for three hours while ‘he ran for three hours’ does not include any reference to repetition. In a similar vein, it can be said that ‘paint’ in ‘paint a house’ means what it means on account of what ‘paint’ and ‘house’ canonically mean. However, this way of rescuing the classical theory is doomed to failure. Let’s see why:

Pustejovsky (1995) has famously applied the co-composition idea to the study of various cases of polysemous nouns: a lexical entry for a noun, on his view, is associated with a structured conceptual representation which informs about its origins, its purpose, its constitution and its distinguishing properties such as shape, colour, etc. Depending on the linguistic context, one or more of these features comes into focus. Thus, it is explained why ‘book’ means different things in these different sentences:

(17) I am starting the book.

(18) The book is on the shelf.

(19) The book has sold 1,000 copies.

(20) Joan has started writing her second book.

In (17) what may be meant is that I have started reading the book. The feature of 'book' that is highlighted is its function. In (18) we are referring to the book as a physical object, whereas (19) refers to the type of physical objects. Finally, (20) combines two different senses associated with book: its textual information sense and its physical object sense. Although Pustejovsky represents what he calls the *qualia* structure non-linearly, I take it that all this variability could be captured by a single (rough) decomposition, say: a book is a physical object of a rectangular shape formed by two covers and a number of usually written pages whose contents derive from an author's thoughts and whose purpose is to be read by other people. Now, when this structure enters into a process of composition with other concepts, then some parts come into the fore while others remain inhibited.

The problems

However, Pustejovsky's account cannot be right. It is true that the whole decomposition is not there in all uses of 'book'. We call 'books' (or apply the concept of book) to a number of things to which only a part of the decomposition applies. I may (correctly) call 'a book' something that appears on the web and will never be converted into a physical object or I may refer as a book to something that cannot be read because it is full of non-sensical chains of letters. In the first case, there is no trace of the part of the decomposition about the constitutive and formal features of books; in the second, one of the things that are lacking is the information having to do with purpose. Thus, what 'book' contributes to the thought expressed by an utterance does not depend only on co-composition processes.

On the other hand, there are cases where the co-composition forces us to go beyond the *qualia* structure that Pustejovsky proposes. For instance, when a noun is modified by the adjective 'real', the meaning of the noun may be affected in such a way that it goes beyond its putative canonical decomposition. Thus, if I say

(20) That's not a real book,

I may simply mean that what I am referring to is not interesting enough to deserve to be called 'a book'. But that books are interesting is nowhere in the canonical decomposition of 'book'.

Last, note that a sentence such as (17) is underdetermined between several readings: I may be starting reading the book, I may be starting writing it, or I may be starting manufacturing it. That is, Pustejovsky's conceptual structures, all by themselves, do not determine which meaning a certain word has in a certain sentence[14].

2.3. Semantic Underdeterminacy

The phenomenon

This brings us to the thesis of semantic underdeterminacy (henceforth SU). As I see it, this thesis is the best reason we have to believe in a cluster decompositionalism, for it is possible to mount a direct attack on the view that words stand for concepts (be they complexes or atoms) in a one-to-one way. We have been approaching the denial of univocal correspondences between words and concepts by means of examples. What I want to show in this subsection is that we can construe an argument for this denial, and for the positive thesis that lexical concepts have a cluster structure.

According to the SU thesis, proposed by recent contextualists, such as Recanati (2004) and defenders of Relevance Theory (RT), the semantic information of a sentence fails to provide definite truth-conditions or to express a determinate proposition, even after saturating all indexical expressions. We can capture this idea by saying that, the problem of indexicality aside, sentences do not carry truth-valuable contents. Usually contextualists argue for SU by means of examples such as[15]:

(21) The house is adequate for *three* persons (at least three/exactly three/at most three?) (Atlas, 2005)

---

[14] According to Pustejovsky (17) exemplifies "type coercion": a verb such as 'start' requires an event as a complement. However, in (17) it is followed by an NP. So coercion applies in order to change the complement NP into an event. But the point is that the event in question is left unspecified by the entry for 'book', with the result that we have an underdetermined sentence. This example, however, does not show –as the others do- that the basics of utterance interpretation cannot be canonical decompositions: it only shows that interpretation has to resort to the extralinguistic context. So the example is included here in order to link this subsection with the next.

[15] I put in italics the term which *prima facie* (see below for a revision of this appearance in the case of (23)) seems to be the source of the underdeterminacy, and in brackets the possible meanings of it. Perhaps the last example needs some explanation: it exemplifies what according to Culicover and Jackendoff (2004) is a rule of our language, namely that an NP can stand either for its referent or for a physical representation of it.

(22) John*'s* car is empty (the car he owns/the car he drives/his favourite car…?) (Recanati, 2001)

(23) The ink is *blue* (blue in its external aspect/ blue in its writing?) (Travis, 2000)

(24) *Ringo* is the Beatle that I like the most (Ringo the drummer/ a statue of Ringo in Mm. Tusseaud's/ a picture of Ringo/ a picture of the statue?) (Culicover and Jackendoff, 2004).

The examples can be produced without limit, as the above discussion about 'paint' or 'book' has made clear[16]. Now, what the examples seem to show in any case is that we need more than purely semantic knowledge (i.e. knowledge of lexical meanings plus knowledge of rules of composition) to extract what we take to be the meaning of an utterance or to assign truth-conditions to it. What we are said to need is contextual information. *Prima facie*, this means that either the meaning of one or more lexical unit is semantically underdetermined or that composition rules are. I think this *prima facie* idea is basically right. Now, in the former case, which I take it is the most usual, we are led to conclude that the underdetermined terms lack univocal meanings. If, as it seems, examples of SU can indeed be produced without limit, we have to end up concluding that the existence of univocal meanings –be they atoms or definitions- is a myth. On the other hand, I will suggest that decompositionalism is also better suited to explain the cases of SU whose source is the underdeterminacy of a composition rule.

A note on the dialectics of this subsection: its purpose is to show that a cluster decompositionalism can account for the phenomenon of SU better than its rivals, be they atomists or definitionalists. However, it will be mostly devoted to arguing against atomism. The reason is that some of the most prominent contextualists have strong atomist tendencies.

Some distinctions

Now, under the label of SU fall quite different phenomena. Some of them are explainable by a cluster decompositionalist account, but others are neutral about the decompositionalism/atomism issue. For instance, in (22) the genitive is said to behave

[16] For a defence of the generalization of SU, see Carston (2002) and Vicente and Martínez Manrique (2005, 2008).

like a free variable[17]. Its lexical information must restrict the range of values it can take, but the point is that such values could well be atomistic dyadic properties, such that one could well say that the genitive tells you to look for an atomic concept within the *relations* category.

Travis (1996) puts forward another type of example where what seems to be going on is that the variability of the truth-conditions associated with a sentence does not depend on meaning, but on the circumstances of evaluation. Thus, suppose that while looking at a squash game you utter:

(25) The ball is round.

This assertion can be taken as true even if the ball in that moment is hitting the wall and is not round. But it can also be false precisely for that reason. Now, 'round' seems to have the same meaning in both cases. The difference in the truth-values depends solely on the circumstances of evaluation.

However, there are many other cases of SU which seem to constitute grist for the cluster decompositionalist's mill. As I have said, it looks as though the underdeterminacy of sentences with respect to propositions is a symptom, whose usual causes are the underdeterminacy of lexical components and/or of the rules of composition. In either case, and I insist, within a restricted set of phenomena, cluster decompositionalism fares better than atomism.

From SU to cluster decompositionalism

Let's start with the underdeterminacy of rules. Travis (2000) discusses example (23) above. Here we have a simple sentence whose truth-conditional meaning is not fixed. The ink may be blue in its writing or it may be blue in its external aspect (though it writes black, let's say). The atomist can say that 'blue' means BLUE and that 'ink' means INK, but that BLUE is modulated into BLUE WRITING or into BLUE LOOKING, or something along those lines (this could be Recanati's approach), as a result of pragmatic processes. Now, the decompositionalist can offer a simpler and more conservative account. 'Ink' has a non-canonical variable decomposition which typically includes the

[17] See Recanati (2001). Cognitive linguists such as John R. Taylor (2003), however, disagree. On their view, the various relations the genitive may stand for are but metaphorical extensions of the possession relation. I also have some problems with the idea that the genitive behaves like a free variable: the approach has to say something else about how linguistic information constrains the range of values the genitive can take –and this is probably not an easy task.

concepts of DENSE LIQUID STUFF USED FOR WRITING/MAKING MEANINGFUL MARKS ON PAPER. Then we have a rule of composition which states that predication consists in assigning a property concept to an entity concept. However, this is too vague a rule, given that, on the decompositionalist approach, entity concepts are typically multi-faceted. Thus, the rule tells us that BLUE has to be applied to INK, but this can be understood in several ways: BLUE can be applied to the function associated with INK or it can be applied to the information about ink's constitutive properties. So here we have why the meaning of a sentence such as (23) is underdetermined: the rule of composition is unable to fix its meaning in the absence of extralinguistic information.

Next we have the underdeterminacy of the lexical units that form part of a sentence. As a start we can take Sperber and Wilson's (1998) widely known example of the following dialogue:

(26) Peter: Let's go to the cinema

Mary: I'm tired.

Now, Mary means that she's tired in a particular way: she's tired in such a way that she can't cope with going to the cinema. What a cluster decompositionalist can say is that she's expressing a complex concept that picks out some of the elements that form part of the family of TIRED concepts. However, this is not the way Relevance theorists prefer to go. What Sperber and Wilson (1998) hold is that Mary's use of 'tired' encodes the atomic concept TIRED which, in the context, and guided by the relevance principle, prompts the construction of another –non-lexicalized- atomic *ad hoc* concept TIRED*. This is done mainly by looking at the encyclopaedic information associated with TIRED.

As I see it, there is a basic problem with this approach, namely, that it conflicts with one of the basic tenets of contextualism: the view that there are no eternal sentences, i.e. sentences that *carry*, and not merely *express*, thoughts[18]. Or to put it in other words, the approach compromises the contextualist claim that the basic units of truth-conditional meaning are utterances, not sentences, given that sentences are semantically underdetermined and are thus unable to carry thoughts. A way to see it is this:

[18] That a sentence carries a thought means that it expresses it by semantic means alone. Sentences can express thoughts without carrying them, as when they are used to communicate a thought by resorting to contextual information.

(i) Contextualism entails the ineffability of thoughts in the sense that thoughts cannot be captured by sentences alone: only utterances "vehicle" thoughts: sentences, given SU, cannot.

(ii) According to RT's approach many sentences would carry thoughts. Thus, 'Mary is tired' would carry the thought MARY IS TIRED. That is, if there are univocal meanings for words, be they atoms or definitions, then many sentences must be capable of encoding thoughts.

(iii) Even thoughts that are supposed to be truly ineffable are in fact encodable by sentences. As explained above, RT holds that Mary's utterance of 'I am tired' expresses an *ad hoc* concept which, as Sperber and Wilson claim, is not lexicalized. This view on *ad hoc* concepts, however, can be disputed. It's not very clear how *ad hoc* concepts are construed or reached on this thoroughly atomist account, while we can have a simple intelligible explanation if *ad hoc* concepts are complexes made up of atomic concepts[19]. However, endorsing the latter approach entails not only decompositionalism about *ad hoc* concepts, but also dropping ineffability.

Take a simple example such as 'Sally is an angel', uttered with the metaphoric or loose meaning that Sally is kind and good. Relevance theorists hold that what is expressed by this utterance is SALLY IS AN ANGEL*, and that ANGEL* is a truly ineffable *ad hoc* concept to which it is associated new encyclopaedic information[20]. Yet, the new concept seems to be given by the composition of KIND, GOOD and probably some other concepts, while its encyclopaedic information equally consists of the similar composition of the encyclopaedic informations pertaining to 'kind', 'good', etc. Of course this turns ANGEL* into an effable, lexicalisable concept. It may be difficult, perhaps impossible, to give an adequate characterization of the concept. Surely it is not exhausted by 'kind and good'. But, as a matter of metaphysical fact, we can say that there is such a characterization since the concept ANGEL* is made up of encoded lexicalized concepts. In conclusion, thoughts some of whose components are *ad hoc* concepts are as effable as those constituted by encoded concepts[21].

---

[19] RT's notion of *ad hoc* concepts draws from Barsalou's work. Yet, Barsalou seems to think of *ad hoc* categories as constituted by features, i.e. as non-atomic (see e.g. Barsalou, 2002)

[20] A word about ineffability: that ANGEL* is ineffable does not mean that it cannot be expressed by the word 'angel', when 'angel' is uttered in a given linguistic and extralinguistic context. What it means is that ANGEL* is not expressed, or captured, by any term, simple or complex, of our language alone.

[21] See Vicente and Martínez Manrique (2010) for further developments of this argument. In particular, we try to show that the reasons Relevance theorists may have for having *ad hoc* concepts atomic are not well founded. If this is right, then the default view on *ad hoc* concepts should be the one here developed, and RT would end up committed to minimalism.

Now, from the last two premises it seems safe to conclude:

(iv) Most thoughts can be encoded by sentences, i.e. they can be carried by sentences alone,

which enters into contradiction with claim (i) above.

On the other hand, any view that postulates the existence of univocal meanings (encoded meanings, in RT's terms) for words has to face a further problem related to their somewhat ghostly existence. Given that words may have a variety of meanings, how are we to decide which one is *its* meaning?[22] And also: what role do these alleged literal meanings play? Take Recanati's (2004) proposal: Recanati holds that there are univocal/encoded meanings for words, but that there are no (psychologically realistic) minimal propositions because the process of composition of meanings in order to form a thought is realized only after a number of pragmatic processes –what he calls "primary pragmatic processes"- have taken place. That is, what are composed are not encoded meanings but meanings that have already been affected by the context in one way or another. However, the picture may work just as well without postulating the existence of encoded meanings. There seems to be no need of having intermediate stations in the process of interpretation, especially if these are entities that we cannot identify in a principled way. In conclusion, I think that the theory of lexical concepts that best suits the phenomenon of SU is a cluster decompositionalist account. On this account, what pragmatics does for us is basically select what part of the cluster of concepts associated with a given lexical entry has to be active in the recovery of the thought expressed.

Conclusion

I take it that the kind of cases of SU here discussed shows how widespread polysemy can be taken to be. Now, polysemy has always been a problem for atomism, to the extent that Fodor (1998) claims that "there is no such thing as polysemy" (1998: 53). It is certainly difficult to decide when a word is polysemous. Zwicky and Saddock (1975) proposed their famous *zeugma* criterion (which Fodor and Lepore (2002), on the other

[22] This is a pressing problem also for a position which held that there are univocal encoded meanings, that these are definitions, and that all departures from these alleged univocal meanings are modulations of definitions or constructions of new *ad hoc* definitions (see Recanati, forthcoming). Which one, of all the possible meanings a term can have, is its proper definition?

hand, use to identify homonyms). According to it, a word is polysemous if two juxtaposed appearances of it create a *zeugmatic* effect, like in 'I played the guitar and a game of chess'. However, a lot of juxtapositions of intuitively polysemous words do not create *zeugmas*: for instance, there is no *zeugma* in 'I wrote the book, and now it is in every bookstore', even though 'book' is prototypically polysemous (see Cruse, 2000). And as Cruse shows, zeugmatic effects are context-sensitive[23]. But however difficult may be to detect polysemy and to be clear about what it consists in, the fact is that the combination of stability and variability in word meanings is a widespread phenomenon of language. On the other hand, it seems that a cluster decompositionalism is especially well-suited to account for it.

However, any kind of decompositionalism has to face a number of problems, most of them put forward by Fodor and his followers. I will briefly comment on two of them, probably the most well-known, and dedicate more space to a third, which I think is particularly pressing for the account here developed.

3. Problems

3.1. The problem of ignorance and error

Both the definitional theory and the prototype theory are difficult to reconcile with Kripke's and Putnam's views that the denotation of natural kind terms is independent from our web of beliefs about them. We may believe that water is liquid, has no colour or taste and fills rivers, lakes and oceans. However, if we were to use such beliefs to identify water, we may well be mistaken, for we may pick up twin-water instead of water. Neither the classical theory nor the prototype theory can account for the possibility of being wrong. The classical theory holds that the way we categorize the world is by comparing its entities with our definitions. So if we find some stuff that fits our definition of water, then that stuff must be water. The only way to be wrong is by applying the definition wrongly; that is, the classical theory makes no room for being erred if the definition is correctly applied. The propotypical theory, on the other hand, has it that we categorize on account of similarity with respect to the prototype. So if we find something that looks much like the prototypical watery stuff, then it must be water. Again, there is no room for being wrong if what is there really looks like the prototype.

[23] His example is the following pair. The first sentence is not zeugmatic, but the second is:
(a) The arm of the ocean resembles that of a giant;
(b) The giant put his arm into that of the ocean.

Kripke and Putnam claim that the way to individuate a natural kind is by looking at its hidden essence. According to Gelman's (2003) empirical research, that's what we do: we are sensitive to the essential properties of natural kinds, and we tend to think that apparent properties are an effect of having a certain hidden essence. Thus, if we are presented with a tiger that undergoes several changes so that it no longer looks like a tiger, we still think it is a tiger, as long as we believe that its hidden essence –its genetic endowment, say- has not been altered. In principle, this is highly problematic for any decompositionalist theory which claims that prototypes, definitions, or whatever, are made out of apparent property concepts. Now, there is hope for a decompositionalist to amend her theory. The idea that a decompositionalist theory has to work only with apparent property concepts seems misplaced. If we stick to the definitional theory, there is no reason why a definition of, say, TIGER, must be the well-known definition FEROCIOUS FELINE WITH BLACK AND YELLOW STRIPES. A definition of tiger may include the fact that tigers have a certain hidden essence responsible for their apparent properties. Thus, TIGER may be defined as FELINE WITH A CERTAIN, PARTICULAR, HIDDEN ESSENCE SHARED ONLY BY ITS CONSPECIFICS, WHICH IS FEROCIOUS, etc. If we admit the possibility of such definitions, then we can explain Gelman's findings and be immune to the typical counterexamples of the sort "what if we encounter something that looks like a tiger but find out that it belongs to another species?"

In a nutshell, I don't think that this problem is as devastating for a definitional theory of concepts as it is sometimes thought to be. The decompositionalism I have been arguing for is not threatened by this problem either. I have presented it as a cluster theory according to which there is not a canonical decomposition for a certain concept, but a number of atomic concepts, most of which are applied most of the times we apply the concept in question. In the case of natural kind concepts, the theory must be polished up slightly: As long as a concept behaves as a natural kind concept, that part of the decomposition that makes reference to its having a certain essence is unmovable.

2.2. Compositionality

Compositionality is a requirement on the nature of thoughts and complex concepts. Such a requirement states that any complex conceptual construction, be it truth-valuable or not, must depend solely on their constituents and on determinate rules of composition. Complexes must not have emergent properties. Two of the theories of concepts that have been presented in this paper comply straightforwardly with the

compositionality requirement, namely, the Fodorean view and the classical approach. The Fodorean view, being denotational, can explain how PETFISH, which is individuated by the property *being a petfish*, is constructed out of PET, whose denotation is *pet*, and FISH, whose denotation is *fish*: a petfish is something that is both a pet and a fish. The classical, definitional, theory meets the requirement in a different way: the concept PETFISH has a definitional structure, and the definition is constructed out of the definitions of PET and of FISH by juxtaposition (say, if PET is ANIMAL HELD AT HOME TOWARDS WHICH ONE FEELS SOMETHING LIKE LOVE, and FISH is AQUATIC ANIMAL WITH GILLS AND SCALES, PETFISH will be ANIMAL HELD AT HOME…WHICH IS AQUATIC, HAS GILLS AND SCALES).

However, the requirement of compositionality is supposed to be a hard problem for a prototype theory of concepts. As Fodor (1998) remarks the prototype of a petfish does not stem from the prototype of a pet and the prototype of a fish. That is, if what individuates the concept of fish and what individuates the concept of pet are their respective prototypical structures, centred around a prototype in each case, the prototypical structure of PETFISH should be predictable and explainable from them. Yet, it is not. Rather, the prototypical structure of PETFISH has largely emergent properties.

The requirement of compositionality and how it should be understood is the centre of an interesting discussion (see, e.g. Robbins 2002). However, my focus will be on whether the decompositionalism I have been arguing for meets the compositionality requirement as it is usually understood. In particular, the question to address is whether abandoning the definitional account of concepts puts the cluster theorist in a predicament similar to that of the prototype theory. Such a question could receive a straightforward answer: there is no problem of compositionality, given that the only difference between a cluster decompositionalism and a classical theory is that there are no definitions. This apparently implies only that the composition may vary from token to token of a complex concept, due to the variability of the decompositions of its constituent concepts. But, as for the rest, if PET in circumstance *c* is composed of concepts C and C’ and FISH is composed by C’’ and C’’’, then PETFISH will consist in the juxtaposition of C, C’, C’’ and C’’’.

However, this response is not satisfying. For what PET contributes to the complex PETFISH surely depends on the fact that it is part of a complex. That is, what PETFISH consists in probably does not solely depend on what PET consists in and what FISH

consists in. In a way it does, but in another it does not: the structure of PETFISH on a given occasion depends on the fact that PET and FISH have been put together and adjusted for the purposes of forming a compound. That is, on the present account, the decomposition of a certain concept is sensitive to influences of the context of its tokening: as occurs with the meaning of linguistic items, one of the contextual factors to be taken into account is the concept's being embedded within a larger whole. So, probably the cluster of concepts associated with 'petfish', or with any other non-monomorphemic expression, does not stem straightforwardly from the clusters associated with its monomorphemic constituents. What concept a certain monomorphemic item of a language stands for when it is combined with another depends on its being so combined. However, I do not think this is so problematic. Even if, on every occasion, PET and FISH are mutually adjusted in order to form PETFISH, the resulting compound has no emergent properties: every concept that forms part of it comes from one of the other of its compounding complexes PET and FISH[24].

3.3. Regress

The problem of regress finds its more detailed exposition in Fodor (1998: 51ff). Jackendoff's decompositionalism (see Jackendoff, 1992) is commonly regarded as a promising explanation of the phenomenon of polysemy (see, e.g. Levin and Rappaport, 2005) as it is exemplified by his discussion of the variability of the meaning of 'keep'. What Jackendoff holds about 'keep' is, first, that 'keep' is a polysemous expression, since it has different but related meanings in utterances such as 'I keep the car in the garage', 'I keep the crowd happy' and 'I keep that promise', and, second, that this polysemy can be explained by resorting to a common conceptual nucleus, namely [CAUSE [ENDURANCE OF STATE]], and the way it is determined by the semantic field it belongs to: LOCATION, EMOTION, COMMITMENT, etc. As a matter of fact, one can understand that what Jackendoff proposes is that the KEEP concept stands for the structure [CAUSE [ENDURANCE OF STATE (X)]], with X acting as a free variable that has to be saturated depending on what is kept.

Now, the problem of regress is the following: Decompositionalism says that there is something common to the different senses of 'keep', and that this

[24] This probable won't satisfy Fodor's demands, according to which "the compositionality thesis says that complex representations inherit their content from simple ones, not *viceversa*" (Fodor, 2003: 96). For a response, see Recanati (2009)

commonality is mentally represented in the corresponding conceptual structures. Suppose that part of what is common is, in effect, the concept CAUSE. The trouble is that there are different ways of causing involved in the different uses of 'keep'. So we are using a concept that, just as the meaning of 'keep', varies according to the other elements with which it is associated. The decompositionalist faces a dilemma. If she tries to solve the problem by resort to different concepts, $CAUSE_1$, $CAUSE_2$, $CAUSE_3$, etc, that underlie each different use, then she misses what was allegedly common to those uses. And if she opts for decomposing CAUSE into further elements, then she faces a regress: any concept that we find at the end of the decompositional chain will be as variable as CAUSE is.

One way to understand this problem is this: if CAUSE is invariable, and the meaning of 'keep' is not, then we can have a reason to consider CAUSE primitive with respect to KEEP. However, it is plausible to suppose that 'cause' is as polysemous as 'keep'. So in the absence of a difference in the behaviour of the words, it seems we should treat both concepts on an equal footing. If we decide to hold CAUSE as primitive, despite the variability of the meaning of 'cause', then so we may do with KEEP.

It may be conceded that 'cause' is, in effect, a polysemous word, while sticking to the idea that there is an atomic and invariable CAUSE concept. The price to be paid is regarding CAUSE as an ineffable concept[25]. This may or may not be regarded as problematic in itself, but I take it that it is preferable to follow another thread in response to Fodor's challenge with regards to the analysis of KEEP. If the decompositionalist concedes that 'cause' (though not CAUSE) is polysemous in constructions such as 'cause the endurance of a state of location', then she should provide an account of its polysemy which accords with the model sketched for other polysemies. However, such a model resorts to the internal structure of concepts, while CAUSE is said to be unstructured[26]. So it seems that the decompositionalist has no *ad hoc* explanation of the polysemy of 'cause'. If this is right, then she had better deny

[25] Fodor (1998; 52) quotes the following words of Jackendoff (p.c): "I'm not claiming that keep is univocal, nor that cause is. Rather, the semantic field feature varies among fields, the rest remaining constant. And the rest is abstract and cannot be expressed linguistically, because you have to choose a field feature to say anything". The problem with this move is that it leaves us without any idea of what CAUSE is.

[26] Besides, it seems that what 'cause' stands for in the "keep constructions" is not a complex only partly constituted by the CAUSE concept. See below.

that ‘cause’ is polysemous in these constructions. That is, Fodor presents the decompositionalist with a dilemma: given that ‘cause’ is polysemous in the “keep constructions”, either there are various CAUSE concepts or CAUSE is not atomic after all. What I propose that the decompositionalist ought to do in response is deny that ‘cause’ is polysemous and thus avoid the dilemma. This means that it must be open for the decompositionalist to say that ‘cause’ means CAUSE invariably here. Now, is this possible? Fodor holds that the same reasons that make us think that ‘keep’ is polysemous should make us regard ‘cause’ as polysemous. If we say that there are different senses of ‘keep’ involved in ‘keep the car in the garage’ and ‘keep the crowd happy’ because what we have to do in order to keep cars in garages and to keep crowds happy are sensibly different things, then we should say the same about ‘cause’ in ‘cause the car to stay in the garage’ and ‘cause the crowd to remain happy’. However, it is possible to claim that all instances of causing are covered by a very general and abstract concept which covers a wide range of events. There is some evidence that such an abstract notion of causing may be innate (see Saxe and Carey, 2006). But even if it is abstracted from more particular notions, to do with physical contact, such as was proposed by Alan Michotte (1963), the fact is that there seems to be such a notion of causality.

Now, a counterargument suggests itself: why would we not say the same about the KEEP concept? Why, if we avail ourselves of this kind of approach for CAUSE, should we not extend it to all the other lexical concepts and hold that they are general abstract notions which cover a wide range of somewhat dissimilar entities in the world? In the end, this is what Fodor holds: there is no polysemy; the variability of meanings of words is only in the eye of the decompositionalist observer. I think this is a pressing problem for the decompositionalist once she takes the route here presented. She can go back to the list of explanatory virtues of decompositionalism and insist on the problems of atomism, but another convincing argument not to go the way the atomist proposes has to do with principles of parsimony. The reason why it is not advisable to extend what has been said about CAUSE to KEEP, or to many other concepts, is that we can explain the latter in terms of the former. That is, there is no need to postulate the existence of an atomic concept KEEP: we have a theory that can do with fewer primitives.

Some psychologists (see especially Mandler, 2004) argue that the first concepts to arise in the infant’s mind are highly abstract notions related to movement,

causation and spatial scenarios. Armed with these concepts, the child is able to construe less abstract notions such as those of animate being or physical object. Then, little by little her mind gets more concrete and finally arrives at the so-called basic level concepts. It seems possible to explain all this process of specification as a process of complex concept building. The idea that I want to put forward is that this constitutes a simple explanation, and that it should be abandoned only in the presence of a more explanatory account. For all the reasons given above, atomism is still not such a theory.

This, however, should not imply that the words that we use to capture conceptual primitives are univocal in all their occurrences. I have restricted my discussion to the role of 'cause' in the "keep constructions", where the CAUSE concept does not seem to be modified by any other concept. In those cases 'cause' stands for CAUSE, but there may be cases where 'cause' stands not for CAUSE but for a complex of which CAUSE forms part. As a way to make this point clear, let me take another putative primitive: MOTION. Suppose that, in effect, MOTION is one of our primitives. It can be seen, however, that 'move' does not have a univocal meaning. Imagine that you are a dance teacher and the group you are teaching is doing really badly: they cannot follow the rhythm and are unable to synchronize their steps. Then you utter:

(27) That's not moving!

Clearly, what you mean by 'moving' is not the putative MOTION primitive. Now, with respect to the "keep constructions" I have said that the decompositionalist must accept that 'cause' is univocal in its meaning given her lack of resources to explain its hypothetical polysemy. Are we now in the same setting regarding 'moving'? I do not think so. For the decompositionalist can say that 'moving' in (27) stands for a complex concept partly constituted by our primitive, say, RYTHMIC SYNCHRONIZED MOTION. The words that we use to refer to the conceptual primitives can be used to express other concepts. When they are so used, they typically stand for complexes which include a primitive. There is nothing in this explanation that is not continuous with the position developed through this paper.

4. Conclusions

I have argued for an account of lexical concepts which claims that lexical concepts have a cluster structure. There are some antecedents of this view: prototype theory is vaguely one, but the account which is more similar to the one here proposed is Jackendoff's preference rule system. The structure of the argument has been the following: First I have presented three explanatory virtues of decompositionalism *vis à vis* atomism. These, however, are virtues of the "linguistic" version of decompositionalism in general, which means that they are also virtues of the classical theory. So in the second section I have tried to single out the cluster theory by showing that it is the best explanation of the variability of meaning. This means that it is a better explanation than the classical theory, but also that it fares better than atomism. As some recent contextualists are atomist, I have paid special attention to their views, thus showing that the explanation of the variability of meaning is another field where decompositionalism does better than atomism. Finally, I have tried to give a response to three pressing problems for the cluster theory[*].

REFERENCES

ATLAS, J. 2005, *Logic, Meaning and Conversation*. Oxford University Press.

BARSALOU, L. 2003, "Situated Simulation in the Human Cognitive System", *Language and Cognitive Processes*, 18: 513-562.

BEZUIDENHOUT, A. 2004, "Procedural Meaning and the Semantics/Pragmatics Interface", in: C. Bianchi, ed., *The Semantics/Pragmatics Distinction*, CSLI Publications.

CAREY, S. 2001, "Whorf vs. continuity theorists: Bringing data to bear on the debate", in: M. Bowerman, and S. Levinson, eds., *Language Acquisition and Conceptual Development.* Cambridge: Cambridge University Press.

CAREY, S. and SARNECKA, W. 2005, "The development of human conceptual representations: A case study", in: Y. Munakata and M. Johnson, eds., *Processes of Change in Brain and Cognitive Development*. Oxford University Press

[*] I especially thank Fernando Martínez Manrique, Mark Jary, Begoña Vicente, and Juan Barba and the rest of the group at Valladolid for their help. An ancestor of this paper was presented at the Linguistics Dpt. of UCL. I am very grateful to Robyn Carston and Deirdre Wilson for their comments and kindness. I also thank the work of two anonymous referees, who helped me to improve on an earlier version of the paper, and Elisabetta Lalumera for inviting me to submit my work for this special issue. Funding for this research was provided by the Spanish Government (project FFI2008-06421-C02-02) and the Junta de Castilla y León (VA077A07).

CARSTON, R. 2002, *Thoughts and Utterances*. London: Blackwell.

CRUSE, D. A. 2000, "Aspects of the Micro-structure of Word Meanings", in Y. Ravin and C. Leacock, eds., *Polysemy: Computational and Theoretical Approaches*. Oxford University Press.

CULICOVER, P. and JACKENDOFF, R. 2004, *Simpler Syntax*. Oxford University Press.

DEWAR K. & XU, F. 2009, "Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants", *Psychological Science*, 20, 252-257

FILLMORE, C. J. 1968, "The Case for Case", in: E. Bach and R. T. Harms, eds., *Universals in Linguistic Theory*, New York: Holt, Rinehart and Winston.

FODOR, J. 1975, *The Language of Thought*, Cambridge, MA: Harvard University Press.

FODOR, J. 1981, "The Present Status of the Innateness Controversy", in J. Fodor, *Representations*, Cambridge, MA: MIT Press.

FODOR, J. 1998, *Concepts*, New York: Oxford University Press.

FODOR, J. 2003, *Hume Variations*, Oxford: Oxford University Press.

FODOR, J. and LEPORE, E. 1999, "Impossible Words?", *Linguistic Inquiry* 30: 445-453.

FODOR, J. and LEPORE, E. 2002 "The emptiness of the lexicon: Reflections on Pustejovsky", in J. Fodor and E. Lepore *The Compositionality Papers.* Oxford: Oxford University Press.

FODOR, J. and LEPORE, E. 2005, 'Impossible Words: Reply to Johnson', *Mind and Language*, 20: 353–356.

GELMAN, S. 2003, *The Essential Child: Origins of Essentialism in Everyday Thought*, Oxford University Press.

JACKENDOFF, R. 1989, "What is a Concept, that a Person May Grasp It?", *Mind and Language*, 4: 68-102, reprinted in E. Margolis and S. Laurence (1999).

JACKENDOFF. R. 1992, *Languages of the Mind*. Cambridge, MA: MIT Press.

JACKENDOFF, R. 2002, *Foundations of Language*. Oxford: Oxford University Press.

JOHNSON, K. 2004, "From Impossible Words to Conceptual Structure: The Role of Structure and Processes in the Lexicon", *Mind and Language* 19: 334 – 358.

LEVIN, B. 1993, *English Verbs Classes and Alternations: A Preliminary Investigation*, Chicago: University of Chicago Press.

LEVIN, B. and RAPPAPORT HOVAV, M. 2005, *Argument Realization*, New York: Cambridge University Press.

LEVINSON, S. 2003, "Language and mind: Let's get the issues straight", in: D. Gentner and S. Goldin-Meadow, eds., *Language in Mind.* Cambridge, MA: MIT Press.

MANDLER, J. M. 2004, *The Foundations of Mind*, Oxford University Press.

MARGOLIS, E. 1998, "How to Acquire a Concept", *Mind and Language* 13:347-369

MARGOLIS, E. and LAURENCE, S. 1999, *Concepts: Core Readings*. Cambridge, MA: MIT Press.

MICHOTTE, A. E. 1963, *The perception of causality*, London: Methuen. (Originally published in 1946).

MURPHY, G. L. 2002, *The Big Book of Concepts*, Cambridge, MA: MIT Press.

PINKER, S. 2007, *The Stuff of Thought*, Penguin Books.

PUSTEJOVSKY, J. 1995, *The Generative Lexicon*, Cambridge, MA: MIT Press.

RECANATI, F. 2001, "What is Said", *Synthese*, 128, 75-91.

RECANATI, F. 2004, *Literal Meaning*, Cambridge: Cambridge University Press.

RECANATI, F. 2009, "Compositionality, Flexibility, and Context-dependence", in: W. Hinzen, E. Machery and M. Werning, eds., *The Oxford Handbook of Compositionality*.

ROBBINS, P. 2002, "How to Blunt the Sword of Compositionality", *Nous* 36: 313-334.

ROSCH, E. 1978, "Principles of Categorization", in Rosch and Lloyd (eds.) *Cognition and Categorization*. Hillsdale, NJ: Erlbaum

ROSCH, E. and MERVIS, C. B. 1975, "Family Resemblances: Studies in the Internal Structure of Categories", *Cognitive Psychology*, 7: 573-605.

SAXE, R. and CAREY, S. 2006, "The Perception of Causality in Infancy", *Acta Psychologica*, 123: 144-165.

SEARLE, J. 1958: "Proper Names", *Mind*, 67: 166-173.

SPERBER, D. and WILSON, D. 1998, "The mapping between the mental and the public lexicon", in: P. Carruthers & J. Boucher, eds., *Language and Thought: Interdisciplinary Themes*, Cambridge, MA: MIT Press.

TAYLOR, J. R. 2003, *Linguistic Categorization*, Oxford: Oxford University Press.

TRAVIS, C. 1996, "Meaning's Role in Truth", *Mind*, 105: 451-466.

TRAVIS, C. 2000, *Unshadowed Thought*, Cambridge, MA: Harvard University Press.

VICENTE, A. and MARTÍNEZ MANRIQUE, F. 2005, "Semantic Underdetermination and the Cognitive Uses of Language", *Mind and Language* 20: 537-558.

VICENTE, A. and MARTÍNEZ MANRIQUE, F. 2008, "Thought, Language and the Argument From Explicitness", *Metaphilosophy* 39: 381-401.

VICENTE, A. and MARTÍNEZ MANRIQUE, F., "Relevance Theory's Atomistic Commitments", in: B. Soria and E. Romero, eds., *Explicit Communication: Essays on Robyn Carston's Pragmatics*. London: Palgrave.

WEISKOPF, D. 2007, "The Origin of Concepts", *Philosophical Studies*, 140: 359-384.

ZWICKY, A. and SADOCK, J. 1975, "Ambiguity tests and how to fail them", in: J. Kimball, ed., *Syntax and Semantics* 4. New York: Academic Press.