Synthese (DOI: 10.1007/s11229-015-0892-0) manuscript No.
(will be inserted by the editor)

Prototypes as Compositional Components of Concepts

Guillermo Del Pinal

Penultime draft. For published version, see:

http://dx.doi.org/10.1007/s11229-015-0892-0

Abstract The aim of this paper is to reconcile two claims that have long been thought to be incompatible: (a) that we compositionally determine the meaning of complex expressions from the meaning of their parts, and (b) that prototypes are components of the meaning of lexical terms such as *fish*, *red*, and *gun*. Hypotheses (a) and (b) are independently plausible, but most researchers think that reconciling them is a difficult, if not hopeless task. In particular, most linguists and philosophers agree that (a) is not negotiable; so they tend to reject (b). Recently, there have been some attempts to reconcile these claims (Prinz, 2002, 2012; Jönsson and Hampton, 2008; Hampton and Jönsson, 2012; Schurz, 2012), but they all adopt an implausibly weak notion of compositionality. Furthermore, parties to this debate tend to fall into a problematic way of individuating prototypes that is too externalistic. In contrast, I propose that we can reconcile (a) and (b) if we adopt, instead, an internalist and pluralist conception of prototypes and a context-sensitive but strong notion of compositionality. I argue that each of this proposals is independently plausible, and that, when taken together, provide the basis for a satisfactory account of prototype compositionality.

Keywords prototypes · concepts · meaning · compositionality · emergent features

G. Del Pinal
Zentrum für Allgemeine Sprachwissenschaft (ZAS)
Schützenstr. 18
D-10117 Berlin
Tel.: +49-1590-1027848
E-mail: ged2102@columbia.edu

1 Introduction

Theories of lexical concepts have to explain two of our most distinctive cognitive capacities. The first is categorization: we can classify objects and events in such a way that we can make useful inferences about their properties. Scorpions are dangerous. So it is important that, upon facing one, we quickly infer that we should not cuddle it. The second is compositionality: we can combine words into complex phrases to express an unbounded number of novel thoughts. Smith can tell John that *Hegel is a better philosopher than Hume*, and John can understand this sentence even if it expresses an utterly bizarre and completely novel thought.

One of the most influential accounts of lexical concepts is the 'prototype theory'. Roughly, prototypes are structured representations of weighted dimensions and features. For example, the prototype for APPLE includes dimensions such as color, shape and taste, with features for those dimensions such as RED, ROUND and SWEET. The features represented by prototypes are usually typical, diagnostic, or some function of both, relative to their associated category.[1] Prototypes were originally introduced to explain a set of results about lexical concepts and categorization known as 'typicality effects' (Margolis and Laurence, 1999; Murphy, 2002; Machery, 2011):

- When categorizing members of a class, subjects can easily order them according to their degree of typicality (e.g., a sparrow is judged a more typical bird than a penguin).
- Typical members of a class are categorized faster and more reliably than borderline or less typical members (e.g., a cow is judged to be a mammal quicker and more reliably than a Beluga whale).
- Subjects know that most of the features used to represent a category are not necessarily possessed by all members (e.g., there can be male lions that do not have manes).

Many typicality effects were obtained or replicated in experiments which *only* used *linguistic* stimuli.[2] For these reasons, many researchers propose that lexical concepts either are or have prototypes as components.

The problem with this suggestion, however, is that prototypes do not seem to be compositional (Fodor and Lepore, 1996; Fodor, 1999; Murphy, 2002; Connolly et al, 2007; Gleitman et al, 2012). Compositionality says (roughly) that the meaning of a complex expression should be completely determined from the meaning of its parts. This condition seems to be generally violated

[1] A feature x is typical, relative to category A, if the probability is high that an entity has x if it belongs to A, and x is diagnostic if the probability is high that an entity belongs to A if it has x.

[2] For example, if subjects have to decide whether something is a bird, they are faster and more accurate for sparrows than penguins regardless of whether they are shown pictures or words for sparrows and penguins (see Murphy, 2002, Ch. 11). This stimulus independence challenges those who deny that a theory of lexical concepts should do double duty as a theory of linguistic meaning. Most of the data used to argue that concepts have a prototype structure can be used to argue that linguistic meaning has a prototype structure.

by the prototypes associated with complex expressions: e.g., the feature LIVE IN TANKS is arguably part of the prototype of *pet fish*, but it is not part of the prototypes of either *pet* or *fish*. One might dispute this example, but the 'emergent feature' problem it illustrates is general. Several experiments suggest that the processes which determine the prototypes of complex expressions regularly add or subtract features in apparently non-compositional ways (Hampton, 1987; Kunda et al, 1990; Murphy, 1990; Johnson and Keil, 2000). If correct, this means that we cannot take prototypes as conceptual components.

Prototype theorists have addressed this objection (Jönsson and Hampton, 2008; Hampton and Jönsson, 2012; Prinz, 2012). They argue that we can abandon the traditional notion of compositionality and still explain the productivity of language and thought. In its place, they propose a non-compositional constraint that allows for general and context-specific beliefs to enter into the default processes which determines the meaning of phrases. From this perspective, LIVE IN TANKS is part of the prototype of *pet fish* because most subjects believe this from previous encounters with pet fish and they use this belief to determine the meaning of the phrase. I examine and reject this and related non-compositional proposals in §5.

In contrast to previous theorists, I will argue that prototypes are compositional. The argument is based on two independently plausible and mutually reinforcing claims, one about lexical prototypes and the other about compositionality. The first claim is this:

– Lexical prototypes encode information along several dimensions. These include perceptual features but also more abstract dimensions such as functional and genealogical features. In addition, prototypes often encode statistical central tendencies of the corresponding categories (e.g., typicality), but this is not an essential property of prototypes (e.g., they often represent merely salient features).

This account of lexical prototypes is relatively uncontroversial. The problem, as we shall see, is that *in discussions of prototypes and compositionality*, theorists tend to both overemphasize the degree to which prototypes encode central tendencies and to under-appreciate the diversity of the type of information that they encode. The second claim I will defend is this:

– Compositionality allows some degree of context-sensitivity. The effect of context-sensitivity is constrained by the information provided by lexical items. Once a subset of information is selected to represent a lexical item in a language processing task, there is no additional effect of context. In particular, general beliefs do not further affect the combinatorial operations of language.

This account follows recent contextualist proposals, but includes substantial additional constraints. On this view, the processes of phrasal meaning-determination are computationally more tractable than they would be if we assume the radical contextualist or the non-compositional accounts. I will argue that, compared to the weaker alternatives, this account is supported by the available empirical data.[3] Finally, I will show that, given those views about the nature of lexical prototypes and compositionality, we can directly address the 'emergent features' objection to the view that prototypes can be conceptual components (§6-§7).

Given this set up, I should note at the outset that nothing I say is in tension with the assumption that there can be systematic and productive thought that does not involve any linguistic capacities. At the same time, I share the standard assumption that a key desiderata of theories of concepts is that concepts serve as the meanings of words and expressions, at least for common nouns, verbs, and their corresponding modifiers and phrases.[4] Since I will argue that our linguistic competence is compositional, it follows that prototypes can only be plausibly taken as conceptual components if, when taken as components of linguistic meaning, they do not conflict with the compositionality of language. That is what I will try to establish. If successful, this should increase our confidence—or at least open space for the view—that prototypes are conceptual components.

2 Compositional prototype combinatorics

To begin this discussion, we should briefly consider two questions about compositionality: (i) why was compositionality originally thought to be a property of our linguistic competence? and (ii) how does it constrain *prototype* models of phrasal interpretation?

Compositionality, from the perspective adopted here, is a property of the Faculty of Language (FL)—a specialized mental faculty which plays a central role in the acquisition and processing of natural languages. There is much controversy regarding the basic properties of FL, but what is not in doubt is that it is productive and systematic. FL is productive in the sense that it can assign meanings to indefinitely many novel expressions. FL is systematic in the sense that its productivity is structured such that if it can assign meanings to a set of expressions D, it can also assign meanings to other expressions constructed from atomic expressions of D combined using syntactic structures employed in any of the complex expressions of D.[5]

[3] For readers familiar with issues about compositionality and context-sentivity as they arise in recent discussions in Philosophy of Language, I should point out that most contextualist positions, such as Recanati's Truth-conditional pragmatics and Carston's Relevance theory, assume versions of compositionality that are *substantially weaker* than the principle I will defend. The non-compositional views defended in Philosophy of Mind by prototype theorists such as Prinz and Hampton belong to the same family as the radical contextualist views in the Philosophy of Language. These points are elaborated in §5.

[4] This assumption is made in most empirical studies of concepts which use linguistic stimuli. As will become clear below, this assumption allows for a wide variety of views on the relation between the standing meaning of words and the meaning that they take in particular utterances.

[5] As several authors have noted, there are performance and other sorts of limits on the actual productivity and systematicity of language. For our purposes, nothing hangs on this. For further discussion, see Del Pinal (2014) ch. 2-3.

Several prominent authors argued that productivity and systematicity (P&S) can best be explained if we assume that FL is (i) a recursive computational system with (ii) a compositional semantics (Fodor and Pylyshyn, 1988; Fodor, 1999). We will see later that some—esp. prototype theorists—have questioned whether compositionality is really the 'best' explanation of P&S; but we will also see that explaining P&S is *not* the only reason for assuming compositionality. For now, the important point is just that compositionality should be understood as a constraint on the ways in which FL can determine the meanings of complex expressions:

(C) FL determines the meanings of complex expressions from the meanings of their syntactically immediate constituents, the way they are combined, and nothing else besides.

C guarantees that if s can determine the meaning of *purple apple* and *red ball*, then s has the competence to determine the meaning of *red apple* and *purple ball*. C doesn't tell us which particular procedure determines the meaning of each type of complex expression—VPs, NPs, and so on. For example, it allows that one type of procedure determines the meaning of Advs combining with Vs and another type of procedure determines the meaning of Adjs combining with Ns. In other words, C specifies a *general* constraint on the ways in which FL determines meanings, in particular, it limits the potential sources of information drawn upon.

As formulated, C does not say anything about what the meaning of words and phrases are. But we can spell-out what C entails for models of compositional prototype combinatorics in terms of two generals constraints:

(Semantic Locality) If α is a complex expression constituted by $\{\gamma, \beta\}$, only features taken from the prototype of γ or the prototype of β can go into the prototype of α.

(Uniform Modification) If α is a complex expression constituted by $\{\gamma, \beta\}$, and of the form $[_Z X\ Y]$, only operations defined for all expressions of type Z can determine the way in which (i) features of γ and β are inherited by α and (ii) the weights on the dimensions of the prototype of α are adjusted.

As before, this leaves open what *particular* operations determine the prototype of different types of expressions, as long as they satisfy Semantic Locality and Uniform Modification—e.g., the operations which determine the meaning of [N N] constructions might be different from those that determine the meaning of [A N] constructions. Following recent debates, we focus on intersective [A N] expressions such as *purple apple*. The crucial property of these expressions is that if x is an [A N], then x is an N.[6]

[6] There is much debate about the basic types of [A N] modifications (Partee, 1998; Morzycki, 2015). Most classifications include intersective, subsective, modal, and privative modifications. In modal and privative modifications—e.g., *former president* and *fake gun*—that x is an [A N] does *not* entail that x is an N. The modal cases do not present any challenges to prototype theory in particular: they shift the time or world of evaluation, and do not affect the content of the head N. The privative cases are problematic for all theories, so

As an example of a compositional prototype model of [A N], consider the Selective Modification Model (SMM) of Smith et al (1988). SMM follows a simple procedure: it selects the dimension of the head N under which the modifier A falls, replaces the feature/s in that dimension with the value of A, and increases the relative weight of the modified dimension. For example, when *purple* modifies *apple*, the color dimension in the apple prototype is selected, its relative weight increased, and its values replaced with the feature PURPLE. This model has two key features:

- Most of the features of the prototype associated with the head N are preserved in the resulting complex prototype for the [A N], and *no* features that do not come from either the A or the N are added to the resulting prototype for the [A N]. This accounts for the fact that it is generally the case that if x is an [A N] then x is an N, while at the same time it respects Semantic Locality.
- When constructing a prototype for *purple apple*, the weight and value of the color dimension of *apple* is changed, but the weights and values of the other dimensions are preserved. Furthermore, the way in which the weights of modified dimensions of the head N are changed can be spelled out in a general way that makes no reference to particular [A N] combinations. This procedure respects Uniform modification.

This procedure for determining the prototype of [A N] expressions is often called 'default to the head stereotype' (DS). Let us clarify the concepts introduced thus far. C is a general constrain on the ways in which FL can determine the meanings of complex expressions. Spelled out for a prototype theory of meaning, C amounts to two constraints: Semantic Locality and Uniform Modification. DS is a *particular* procedure for determining the meaning of [A N] expressions which satisfies Semantic Locality and Uniform Modification.[7]

Most models of prototype combinatorics for [A N] phrases follow DS for two reasons. First, as we said above, for most expressions of the form [A N], it holds that if x is an A N then x is an N. To account for this, we assume that the prototype of the head N is relatively unaltered in the resulting [A N]. For example, whatever happens to the prototype of *apple* when it combines with *purple*, it must still be sufficiently unaltered to entail that if x is a *purple apple* then x is an *apple*. Secondly, by respecting C, DS ensures that our linguistic competence is productive and systematic. This is because

critics of prototype compositionality justly do not appeal to them in their critiques. Still, Del Pinal (2015) argues that privative modifications (e.g., *fake gun*), and certain subsective modifications (e.g., *bad gun*) actually support the view that the meaning of head Ns includes a prototype component, on which the modifiers partly operate. So it is not inaccurate to say that just those modifications that are hard for classical theories are easy for prototype theories, and vice versa. This is why it is particularly important for prototype theories to deal with simple intersective modifications.

[7] To be clear, most models of prototype combinatorics do not stop at the composition stage—see, e.g., Hampton's Composite Prototype Model and Costello and Keane's C^3 Model. What is important is only that they include an initial compositional stage, even if they also try to model some post-linguistic pragmatic modifications.

DS ensures that the procedure for combining prototypes in [A N] structures is uniform (including the readjustment of dimension weights), and that all the features of the complex are derived from the constituents. It follows from these assumptions that if s can determine the prototypes for *purple apple* and *red ball*, s thereby has the competence to determine the prototypes for *purple ball* and *red apple*.

3 The problem: emergent features

The objection to DS and other procedures for determining the prototypes of complex expressions that respect Semantic Locality and Uniform Modification is the following. There are many examples of complex expressions in which general beliefs seem to intrude into the associated prototype, resulting in emergent features (Hampton, 1997; Johnson and Keil, 2000). For example, in some early experiments, participants said that IDEALISTIC is a typical feature of *Harvard-educated carpenter* but not of either *Harvard-educated* or *carpenter*. Unlike *pet fish*, *Harvard-educated carpenter* expresses a concept with whose instances we are likely unacquainted. In these sorts of cases it seems clear that background theories and general beliefs play a role in producing the emergent features.

Indeed, whenever we modify a N with an A that stands for a property that is unusual for the N, the prototype associated with the resulting phrase seems to include emergent features. For example, most of us would likely associate *neon-green carrot* not with a carrot that is neon-green but just like a carrot in all other respects, but with one that lacks the feature EDIBLE or has the feature POISONOUS, although these modifications/features are not directly inherited from *neon-green* (e.g., some neon-green candies are edible). In these sorts of cases, new features are added or removed from dimensions of the prototype of the head N which are not directly modified by the A. This suggests that DS is systematically violated by the apparently non-compositional combinatorial processes which construct prototypes for complex NPs.

For our purposes, a key study of emergent features is Study 1 of Johnson and Keil (2000). This study is important because it focused on novel NPs, making it hard to argue that the emergent features arise because the NPs are, to some degree, idiomatic. They asked participants to generate features for NPs such as *arctic bicycle* and *hospital bicycle*. They also asked other participants to generate features for the modifiers and head nouns, e.g., for *arctic*, *hospital*, and *bicycle*. Most of the features which ranked as typical for the novel NPs were emergent, i.e., were not typical features of the constituents. For example, *arctic bicycle* generated the emergent feature SPIKED TIRES and *hospital bicycle* generated the emergent features METERS and DIALS. These studies seem to confirm the conjectures of philosophers such as Fodor, who argue that typical features of the prototypes associated with NPs do not seem to be derived from the prototypes of the constituents, and hence violate C, and in particular, Semantic Locality.

To be sure, the emergent features objection is not a knock down argument against the view that prototypes are conceptual components. As usual, empirical challenges to theoretical frameworks can be accommodated with sufficient tweaking, but often the repair costs are too high. In this case, we need to determine whether we can address the objection without having to accept ad hoc or implausible claims, in particular concerning compositionality and/or the nature of prototypes. As noted above, some prototype theorists, especially in psychology, responded to the objection by suggesting that we abandon compositionality. But most theorists—and certainly most linguists and philosophers—think that the cost of this move is too high (for now I assume that this is correct, but I will justify this position in §5). I will show that we can address the emergent features objection without such costly moves. The argument has three steps. First, I argue that prototypes do not just encode statistical central tendencies, and that they encode information along various dimensions, including perceptual, functional, and behavioral information (§4). Second, I argue that compositionality allows a type of context-sensitivity at the locus between long-term and working memory that any theory of the lexicon has to accept (§5). Finally, I show that, given those claims, the solution to the emergent features objection is relatively straightforward (§6).

4 Lexical prototypes

This section presents an outline of the basic properties of prototypes. Although the view I present is relatively uncontroversial, it does oppose some problematic tendencies found in discussions of prototypes, especially in debates about compositionality. These are: (i) that prototypes are the sole conceptual components, (ii) that they tend to represent simple perceptual properties, and (iii) that they only encode—or at least aim for—statistical central tendencies of their corresponding categories.

4.1 Prototypes are not the sole conceptual components

Early prototype theorists tried to defend a pure form of the theory according to which concepts are just prototypes. Given this aim, it is natural to think prototypes can only perform their referential/membership function if they encode the actual statistical central tendencies of the category they represent. However, most current prototype theorists do not hold that prototypes are the sole components of lexical concepts, nor do they think that, in most cases, they have an extension/membership determining role (Machery, 2011; Prinz, 2002, 2012; Rosch, 2011; Hampton, 2006). The reasons for this are familiar to most philosophers, so I will be brief. Suppose c is a category. Given suitable scenarios, some entity can be a member of c and not fall under the prototype of c, and vice versa. Most of us believe that typical tigers have distinctive physical features (e.g. stripped coats) and behaviors (e.g., solitary hunters). However, we also believe that some entity can lack these typical features and still be a tiger, and that some entity can have these features and not be a tiger. It follows that, even if we hold that prototypes are conceptual components, we must also hold that they need to be supplanted with other extension-determining components (at least for some categories).[8]

4.2 Prototypes do not just encode perceptual features

Some the earliest and most influential studies in support of prototype theory came from perceptual categorization and learning (Rosch, 1973). When presented with categories of dot patterns made by distortions of a central tendency, or of stick figures in which an average figure is central, subjects learned categories faster if the better examples, particularly the most central members, were presented first, and were faster at categorizing the items which are most similar to the central members. In addition, many categorization studies use static images as proxies for categories. As a result, there is a certain tendency to think that prototype structures represent something like 'a typical exemplar', encoded at a certain level of abstraction. However, this restriction is arbitrary and indefensible. Prototypes also encode more complex information (Murphy, 2002). For example, with respect to animal kinds, we find typicality effects relative to behavioral patterns. With respect to artifact kinds, we find typicality effects relative to uses or functions. Indeed, features related to typical functions are usually produced by subjects who generate feature-lists for artifact kind terms. It follows that the same type of evidence that led us to hold that we encode information about how, say, tigers and chairs typically look, also forces us to hold that we encode more complex information about the typical behavior of tigers and function of chairs. In other words, prototypes are complex, multidimensional structures.

4.3 Prototypes do not just encode statistical central tendencies

Theorists often 'define' prototypes as structured representations of features that are typical, diagnostic, or some function of both, relative to their associated categories. Feature x is typical, relative to category A, if the probability is

[8] The view that prototypes are conceptual components can be implemented in various ways. One can be a global or local prototype theorist: i.e., one can hold that all terms include prototypes, or one can hold that, e.g., artifact terms include prototypes but mathematical and some technical terms do not. In this paper, I remain neutral on this issue. This in no way weakens my defence of prototype compositionality. The dialectical situation is as follows. Suppose there are good reasons to hold that prototypes are conceptual components for some class, say, for natural kinds and artifact terms. As mentioned above, many theorists would still resist including prototypes as conceptual components *because* they believe that prototypes are not sufficiently compositional. A clear statement of this position can be found in Fodor and Pylyshyn (2015). Now, my aim is to show that prototypes are sufficiently compositional. If that is correct, we can begin investigating the potential advantages of including prototypes as conceptual components.

high that an entity has x if it belongs to A, and x is diagnostic if the probability is high that an entity belongs to A if it has x. On this view, prototypes are individuated via some statistical central tendency relation to their associated categories. The most prominent central tendency is typicality, but in some cases (e.g., perceptual or magnitude concepts) it can be an average value. Let us call this view 'externalist', since it emphasizes the way in which prototypes encode statistical central tendencies of the external categories they represent.

The externalist view of prototypes is clearly mistaken. As emphasized by Rosch (2011), prototypes which encode central tendencies form an important subclass, but other types of prototypes use features that are e.g. diagnostic or salient, but not typical, of their associated categories. These features are often the extreme values of dimensions. Most people think of the best examples of cities as the largest and most cosmopolitan, rather than average cities. Similar points apply to concepts such as GENIUS, in which our prototype seems to encode the most extreme features associated with brilliance. Consider also concepts for social roles such as PRESIDENT or MOTHER, which tend to encode expectations that can be rather uncommon for actual members of the class. Indeed, even perceptual prototypes such as those for color and geometrical terms might not capture any statistical central tendencies of their corresponding real world categories (e.g., in what sense is fire-engine red either typical or an 'average' of instances of red?). One factor that determines whether a prototype encodes non-typical features is the acquisition route. We often encode salient features of the members we first encountered, and these features can be rather unusual (Hamill et al, 1980). In addition, repetitive and emotionally charged media bias is partly responsible for certain social stereotypes which do not encode central tendencies.

4.4 Prototypes are multi-dimensional and internalist structures

To sum up, prototypes have a 'pluralist' and 'internalist' structure. The pluralism captures the idea that prototypes are not the only conceptual components, and that they encode information along several dimensions.[9] The internalism captures the idea that, although they often encode statistical central tendencies, they also encode features that are merely salient, diagnostic, or normative. The view I am opposing is a kind of simple-minded externalism according to which prototypes represent something like average, abstract and static exemplars of the corresponding categories. Again, I don't think anyone explicitly holds that view of prototypes, but there is a strong tendency to fall

[9] By a 'pluralist' notion of prototypes I do not mean a 'pluralist' theory of concepts, in the sense defended by e.g., Weiskopf (2009). A pluralist theory of prototypes holds that the dimensions along which prototypes are organised are richer than is sometimes assumed: e.g., we not only represent how objects typically look but also how they typically come into being. In contrast, a pluralist theory of concepts holds that, depending on such things as tasks and context, different types of mental structures can serve as the surrogates of concepts. Still, nothing I say here is in tension with pluralism in this other sense.

into it in discussions of compositionality.[10] That slippage is consequential. For to know whether in general we can compose the prototype of a complex phrase from the prototypes of its constituents, we have to individuate the input and output prototypes. In many cases, especially those involving the prototypes of complex expressions, the internalist/pluralist and externalist/simple views suggest different ways of individuating prototypes, which can result in importantly different assumptions about the features of particular prototypes which a combinatorial theory has to account for. Furthermore, once we realise that the way in which we acquire a prototype can determine whether or not it encodes statistical central tendencies, we can accept that the linguistic-compositional way of obtaining prototypes might *not* result in structures that encode central tendencies. If someone made, for the first time ever, neon-green carrots and they happen to be quite healthy, then my linguistically determined prototype for *neon-green carrots* will lack an inductively useful feature of the corresponding category. Of course, after the prototypes for complex NPs are generated by FL, we often immediately begin to modify them—using general beliefs and reasoning—so that they do encode such central tendencies, since this—*sometimes*—increases their usefulness for categorization and induction.

5 Compositionality, context-sensitivity and modulation

The next step in my response to the emergent features objection is to revise the notion of compositionality. Most theorists agree that, although FL is compositional to some degree, it must allow some forms of context sensitivity. The task now is to formulate and defend a context-sensitive version of *C*.

C allows that expressions can have 'characters' as standing meanings, i.e., expressions whose standing meaning is incomplete and only determines a full meaning relative to particular contexts. Although these expression have a fixed standing meaning, their occasion meaning varies across contexts. Expressions with context-sensitive parameters include demonstratives (*that*), indexicals (*I*, *you*), and some Adjs (*small*, *heavy*). When evaluating alleged counter-examples to prototype compositionality, it is important to ask whether the emergent features are due to the presence of open-parameter expressions.

The more interesting challenge to *C* comes from another type of context sensitivity: meaning 'modulations' that cannot plausibly be attributed to the effect of context on open parameters. These are cases in which the meaning of expressions in token utterances is enriched or altered to increase its coherence with the immediate or general context. Consider the following famous examples:

(1) John cut the grass.

(2) John cut the cake.

[10] However, the account of prototypes presented and criticized by Fodor (1999) and Fodor and Pylyshyn (2015) seems quite close to that view.

In most context, we take *cut* in (1) to mean 'cut horizontally' and in (2) to mean 'cut vertically into pieces'. In other contexts, we can also take *cut the grass* to mean 'cut into blocks' and *cut the cake* mean 'cut horizontally'. Based on such examples, various theorists take modulation to be a top-down processes in which the meaning of sub-sentential expressions is altered to make it more 'salient, appropriate or relevant' relative a context and belief set (Recanati, 2010). Call this relevance seeking and meaning-shifting function 'free modulation'. Since modulation seems to affect the intuitive truth-conditions or meanings of utterances, it should be distinguished from post-compositional pragmatic processes such as conversational implicatures.

Contextualists such as Pagin and Pelletier (2007) and Recanati (2010) propose that we reformulate C to allow compositional operations to take the 'modulated' (instead of the standing or occasion) meanings of their immediate constituents. We can formulate this proposal as follows:

(C^{mod}) FL determines the meanings of tokens of complex expressions from the freely modulated meanings of their syntactically immediate constituents, the way they are combined, and nothing else besides.

C^{mod} does *not* allow the combinatorial *operations* that determine the meaning of phrases to be context-sensitive (although some contextualists do allow this). The combinatorial operations are as fixed and compositional as before. For example, if we assume that DS is the compositional procedure for determining the meaning of complex NPs, we can keep that assumption in this context-sensitive framework. What is flexible is the relation between the standing meaning of words and the meaning of their tokens in particular uses.

If we adopt something like C^{mod}, not only can we seemingly deal with cases such as (1) and (2), but also with at least some of the examples of emergent features. Free modulation could access general beliefs which change the default prototype of *carrot* when it appears in the context of *neon-green* (e.g., by adding POISONOUS). Similarly, free modulation could change some of the features of *carpenter* when it appears in the context of *Harvard-educated*. It is not surprising, then, that various prototype theorists have adopted, in response to the emergent features objection, combinatorial principles similar to C^{mod} (e.g., Jönsson and Hampton, 2008; Prinz, 2012).[11] However, I will

[11] These views share the idea that the processes which determine the meaning of complex expressions have access to general beliefs which do not come from the meaning of the parts or from their structure. If we apply these accounts to the view that prototypes are constituents of linguistic meaning, we can formulate the basic idea as follows (cf. Prinz, 2012):

(*Modal C*) As a default, FL determines the meaning of complex expressions from the meaning of their syntactically immediate constituents, the way they are combined, and general beliefs. As a fallback, FL determines the meanings of complex expressions following C.

Modal C says that FL is *capable* of combining prototypes compositionally, but that it does so only when it lacks the relevant general beliefs. Several authors have endorsed the *Modal C* solution to the problem of emergent features (see e.g., Weiskopf, 2009; Robbins, 2002; Schurz, 2012). Strictly, *Modal C* is less constrained than C^{mod}; in practice, however, they allow very similar accounts of emergent features. The reasons I give below to constraint C^{mod} also apply to *Modal C*.

now argue that C^{mod} is too unconstrained, and will propose two constraints that substantially increase its plausibility.

5.1 Constraining the structural locations of modulation

The first constraint regards the level of syntactic structure at which free modulation can apply. As C^{mod} is formulated, free modulation could apply at every level of phrasal construction or branching node. However, we will assume that modulation applies only to lexical items, i.e., to the terminal nodes of syntactic trees. So modulation does not operate on non-terminal nodes such as at the highest node of complex NPs of the [A [A N]] form. This allows a constrained form of meaning modulation which can be described as pre-compositional, in the sense that the effects of the context-sensitive modulation function enter only via the terminal nodes of expressions.

Why add this structural constraint? Most linguists hold that syntactic and semantic computations work in phases that are sent off for pragmatic interpretation *before* full sentences or clauses are processed by FL. In Minimalist theories, the main phases are vPs and CPs, but due to the 'left edge condition' (Chomsky, 2001b,a), the phases that are sent out for pragmatic processing are more fine grained (Cook and Newson, 2007; Radford, 2004). Theorists who adopt Categorical Grammars also hold that the outputs to pragmatics are sub-sentential phrases such as DPs and VPs (Jacobson, 2012). If interpretation proceeds in such phases which are sent to pragmatic interpretation, there is no reason why we should incorporate into the semantics a function which modulates *both* the inputs and outputs of the compositional operations. The output modulations would be redundant. As I argue below, modulation operations on lexical items are powerful enough account for the relevant data.

5.2 Constraining the operations of modulation

The second constraint regards the kinds of operations on meanings that free modulation can perform. Free modulation is a top-down process which can loosen, enrich, or even create meanings for expressions, depending on such things as context, beliefs, and task demands. Strictly, free modulation is a much more powerful process than is required to deal with meaning shifts such as those in (1) and (2). But *if* we assume a view according to which lexical meanings are informationally impoverished (e.g., an atomic view), then something like free modulation seems required to deal with those sorts of meaning-shifts. However, by providing rich lexical representations, prototype theory opens the possibility of reducing the expressive power of the combinatorial operations. This has several advantages over free modulation.

In the version of prototype theory we adopted, lexical items include rich sets of multi-dimensional information. In particular, lexical items store a larger set of information in long-term memory than is typically activated when the items are tokened in particular language tasks. Consider *book* in (3)-(5):

(3) John is moving to another city. He is worried about how much it might cost to ship his books.

(4) John has to write a piece on the European Union. He is considering which books to review.

(5) John does not know whether to take his favorite book on the trip. It is very interesting, but quite heavy.

The prototype for *book* in each token is somewhat different: in (3) only physical features might be tokened, in (4) only functional features, and in (5) both, since it invokes a trade-off between physical and functional features. All the relevant information is stored in the long-term memory entry for *book*—e.g., the typical physical, functional, and genealogical features of books—but only some of it is used in particular tasks. As a first characterisation, call 'modulation*' the process that selects, *within the options that are lexically available in long-term memory*, the task-relevant aspects of the prototypes.

Modulation* is more constrained than free modulation. The function of modulation* is to activate, for particular tasks, subsets of features available in lexical entries. This coheres well with the manner of cutting examples, since all the candidates are, on a prototype view, arguably lexicalized. Free modulation could be formulated to include similar operations, but it also includes a much more powerful operation, namely, that of taking prototypes as inputs and—via inferential processes which are sensitive to goals, beliefs and contexts—creating new, ad hoc prototypes. It is reasonable to hold that modulation* could be explained on associationist principles, but the additional inferential operations of free modulation are smart and creative in a qualitatively different sense.

Why hold that the combinatorial operations of FL involve only modulation*? For several reasons. (i) There is substantial evidence for and some initial models to explain modulation*. (iii) Free modulation posits additional processes which would make ordinary language processing extremely taxing. (iv) Those processes arguably over-generate unavailable meanings for expressions. Finally, (v) modulation* is sufficiently powerful to deal with prototype compositionally. I discuss (i)-(iv) below, and (v) in §6.

There is substantial evidence for the kind of context-sensitivity of prototypes assumed by modulation*. For example, Roth and Shoben (1983) investigated the ways in which the prototypes of words like *animal* shifts across contexts. Contexts such as *Stacey volunteered to milk the animal whenever she visited the farm*, caused a prototype advantage (e.g., ease of categorization reflected by reading times) to *cow* and *goat* relative to *horse* and *mule*. However, contexts such as *Frank asked her father to let her ride the animal*, caused the opposite prototype effect, favoring *horse* and *mule*. This effect of linguistic context on the prototype and features associated with expressions is also found for basic level terms such as *piano*, *shark*, and *handle* (for an overview, see Barsalou (1987) and Murphy (2002), chpt. 11). For example, Tabossi (1988) used a cross-modal priming paradigm to study whether linguistic context affects the features associated with particular tokens of basic level nouns such as *gold*. The results support the view of selective activation of features assumed by modulation*: e.g., when the importance of colour was made salient, *gold* differentially primed features such as *yellow*, but when malleability was made salient, *gold* did not prime features such as *yellow*.

That prototypes are modulated*, then, is a well-established empirical result. Lexical entries include rich arrays of information, but the subsets of features activated in each token use can vary. Still, none of this involves the generation of new features or prototypes; so there is no need to posit a smart and creative free modulation function. The data only illustrate modulation*, according to which only a subset of the information that is stored in lexical entries is uploaded into working memory in particular tasks. Indeed, in a review of the evidence for the relative instability of lexical prototypes across contexts, Barsalou (1987) defends what essentially amounts to modulation*. He distinguishes between the information associated with words in long term memory, and the token prototypes extracted to working memory to perform particular tasks. The prototypes extracted to working memory are informationally pruned, relative to their corresponding long-term memory databases. The process of selective activation is influenced by various contextual factors—such as preceding discourse, goals, and relevant task constrains—but is basically associationist. In other words, modulation* can be naturally modelled as a byproduct of the general associationist processes that underlie the extraction of information from long-term to working-memory, and hence is not an ad hoc stipulation used to explain the nuances of prototype compositionality.[12]

Crucially, the processes of information selection from long-term to working memory cannot be used to explain the more paradigmatically inferential operations of free modulation. To allow modulation on lexical and phrasal prototypes once they are in working memory, as required by free modulation, we would have to posit additional top-down operations on working memory units for which we have no independent evidence. On this view, everyday language processing would involve: (i) holding the meanings of multiple lexical items and phrases in mind, (ii) comparing them with each other, general beliefs, and the context, and (iii) performing coherence raising operations on them which are responsive to other single items and phrases, also held in mind and operated upon—and all this prior to and in between compositional operations. On this view, simultaneous operations performed on multiple items held in

[12] Modulation* processes fall squarely in what two system theorists of reasoning call the 'associative system' or 'system 1' (Sloman, 2002; Morewedge and Kahneman, 2010). On the standard view, the operations of system 1 are based on associations, not abstract rules, and operate on prototypes in an automatic and unconscious manner. These are also characteristics of pre-compositional modulations*. In the *cut* examples, subjects assume a particular manner of cutting, without being aware of the processing behind it. Indeed, this explains why the effects of modulations* are taken as part of the intuitive truth conditional content or meaning of utterances. In contrast, the additional relevance seeking, inferential operations of free modulation fall squarely in what two system theorists call the "rule-based system" or "system 2". Rule-based processes are sensitive to goals and beliefs, and tend to be conscious and relatively effortful. Insofar as the general distinction between associative and rule-based systems has psychological reality, this further justifies distinguishing between modulation* and free modulation, and taking modulation* as the basic pre-compositional process.

working memory are a general feature of ordinary language processing. However, it is well-known that even simple independent operations on two or more items in working memory can be extremely taxing (Kahneman, 1973, 2011). In particular, such operations are taxing in a way that does not seem to correspond to the usual easy flow of everyday language processing. To be sure, it is empirically possible that FL is a module with special working memory properties which are fine tuned to perform this type of multiple operations and calibrations between the meanings of terms and phrases. Still, this is a strong empirical commitment that proponents of free modulation rarely mention but would eventually have to defend.

There is, however, a more serious problem with free modulation: it arguably over-generates meaning shifts that are not available. Consider the following example, due to Asher (2011):

(6) a. Mary stopped the apple.
 b. Mary stopped eating the apple.

In most contexts, the meaning expressed by (6-b) would be the most relevant or salient interpretation of (6-a). However, that meaning is unavailable. If a relevance seeking free modulation function could apply to certain key items at any level of interpretation, that readings would, it seems, be easily obtained. Even adding a priming context does not result in the desired modulations. Consider:

(7) John was busy, but is now ready to go for lunch.
 a. He finally stopped the garden.
 b. He finally stopped mowing the garden.

In this case, *stopped* in (7-a) cannot mean 'stopped mowing', as in (7-b), although this would be the easiest and most relevant modulation in the context. Consider another example:

(8) John and Mary want to hang the paintings.
 a. John began the nails.
 b. John began hammering the nails.

In this case, *began* in (8-a) cannot mean 'began hammering', as in (8-b), although this would result in a relevant modulation and could be achieved by a simple enrichment of *began*.[13] Examples like this suggest that adopting free modulation to explain cases such as (1) and (2) was an over-reaction. The ways of cutting we need to invoke to explain the context effects in (1) and (2) all involve typical ways in which things can be cut; so all we need is to appeal

[13] These examples suggest that something like free modulation cannot be affecting the intuitive truth conditions or meanings assigned to expressions. This should not be confused with claims about limits on what utterances can communicate in context via traditional pragmatic means such as Gricean conversational implicatures. Naturally, if we have to determine what someone who used, say (7-a) meant to communicate in that context, we might very well say that it is something like (7-b).

to modulation* to explain how the typical manner of cutting represented in each particular task is selected as a function of context.

To sum up, there are several reasons why we should replace free modulation with modulation*. Modulation* is a byproduct of the way in which information from long-term memory is extracted and used in particular working memory tasks. It is an interface property of FL that every model of language processing has to eventually incorporate. In contrast, free modulation postulates additional operations, intrinsic to FL, which face explanatory and descriptive problems: its 'calibration' operations would impose severe demands on working memory and they arguably over-generate unavailable meanings. To be clear, each of these problems is open to debate.[14] The point, however, is that if we can model prototype compositionality using only modulation*—as I argue in the next sections—then there is no need to assume, merely to respond to the emergent features objection, that language processing generally involves free modulations.[15]

5.3 Compositionality and modulation*

In light of the previous discussion, we will constrain C^{mod} as follows:

(C^{mod*}) FL determines the meanings of tokens of complex expressions from the meanings of their syntactically immediate constituents, the way they are combined, and nothing else besides. If the syntactically immediate constituents are terminal/non-branching nodes, these meanings may be modulated*.

The corresponding revisions to our account of prototype combinatorics are straightforward:

(Semantic Locality) If α is a complex expression constituted by $\{\gamma, \beta\}$, only features derived from the modulated* prototype of γ or the modulated* prototype of β can go into the prototype of α.

[14] In particular, the claim that free modulation over-generates meanings has been intensively debated by philosophers and linguistics. Some of the most influential arguments against free modulation can be found in the papers re-printed in Stanley (2007). For a direct response to Stanley's main arguments and examples see Hall (2008), among others. I think some of the best and yet somewhat ignored examples against free modulation are presented in Asher (2011) and ves (1996). An important (and ultimately critical) empirical investigation of whether contextualist theories with free modulation such as Recanati's successfully predict which meaning shifts are licensed is reported in Rabagliati et al (2011). In the end, even Recanati (2010) accepts that it is likely that, as stated, free modulation is too unconstrained, and welcomes suggestions for how to constrain it (see his Introduction).

[15] Prototype theorists should welcome this strategy. Although most have adopted unconstrained combinatorial principles such as C^{mod}, they usually did so as a response to the emergent features objection. If my proposal for reconciling compositionality with prototype theory is successful, theorists such as Prinz, Hampton and Jönsson need not adopt such radically unconstrained principles, nor depend—to defend their preferred theory of meaning—on the eventual resolution of the debate regarding over-generation and free modulation.

(Uniform Modification) If α is a complex expression constituted by $\{\gamma, \beta\}$, and of the form $[_Z X\ Y]$, only operations defined for all expressions of type Z can determine the way in which (i) features of γ and β are inherited by α and (ii) the weights on the dimensions of the prototype of α are re-adjusted.

According to Semantic Locality, the features of the prototype of a particular complex expression e in context c are derived from the modulated* prototypes of the lexical constituents of e in c. According to Uniform Modification, the combinatorial operations themselves are not context-sensitive. Since we are dealing with intersective NPs, we continue to assume that the process of prototype composition follows DS (default to the head prototype/stereotype). According to this version of DS, all the features of the modulated* head N prototype which are not directly affected by the modulated* modifier, are inherited into the NP.

6 Emergent features reconsidered

I have now presented the two main components of framework that we will use to deal with prototype combinatorics: (i) an internalist and pluralist account of prototypes which emphasizes that the features associated with simple and complex expressions often do not encode statistical central tendencies, and (ii) C^{mod*} and the corresponding context-sensitive versions of Semantic Locality and Uniform Modification (and since we are dealing with intersective NPs, this amounts to a context-sensitive version of DS). Call this framework 'CCIP' (for Context-sensitive, Compositional, and Internalist Prototypes). The task now is to use CCIP to address the emergent features objection to the compositionality of prototypes. I will illustrate the steps that, according to CCIP, determine the prototypes of complex expressions and show that CCIP can deal with the usual examples of emergent features.

Consider again the infamous *pet fish*. Strictly, the prototype for *fish* is quite complex, but let us assume for simplicity that a trout is indeed a good exemplar of the prototype associated with the standing meaning of *fish*. The prototype associated with the standing meaning of *pet* likely includes some typical pet functions, even if in some occasions we also use more static perceptual features associated with typical pets. Given those choices, what is a reasonable candidate for a compositionally determined prototype for *pet fish*? If we assume a context-*insensitive* version of DS, then this prototype would seem to include perceptual information about some rather ordinary fish, and functional information about the typical functions associated with pets. A good exemplar of this would be a trout that keeps one company or that lives in one's house. As critics point out, this result seems problematic.

However, the context-sensitive version of DS adopted by CCIP entails that when processing *pet fish* the prototypes of both *pet* and *fish* are modulated*. Recall some of the previous results. In some contexts, the prototype for a token of *animal* can be closer to a representation of a cow than of a horse, and this order shifts in other contexts. Also depending on the context, the prototype of *gold* can include its perceptual features, its social features, its intrinsic physical properties, or combinations of those dimensions. Terms such as *pet* and *fish* should be context-sensitive in similar ways. In the context of talking about fish, the prototypical functions associated with *pet* involve more aesthetic uses than companionship uses. Similarly, in the context of talking about pets, the perceptual features associated with *fish* will be closer to those typical of smaller ornamental fish than to those typical of wild fish.

Still, even after modulation* of the prototypes associated with the constituents of *pet fish*, the linguistically determined prototype, assuming a process of combination that respects DS, is unlikely to encode some prominent central tendencies of the category of pet fish. So the features of the linguistically determined *pet fish* prototype—namely, that they are something like smallish ornamental fish—are unlikely to fully correspond to the features we associate, after some reflection, with *pet fish*. From the perspective of CCIP, there is nothing problematic about this. Our intuitions when given time to think about the 'typical' features of *pet fish* are about the representation associated with *pet fish* after the prototype assigned by FL has been modified based on pragmatic considerations and using general beliefs. Since we have extensional knowledge about pet fish—e.g., their typical small size, bright colours, living conditions, and decorative uses—we can use that information to modify the perceptual and functional features of the linguistically determined prototype. This subsequently modified prototype associated with *pet fish* is closer to encoding the central tendencies of the category of pet fish.

Similar points apply to *arctic bicycle*, *hospital bicycle*, and other examples of emergent features that have actually been obtained experimentally (Johnson and Keil, 2000; Kunda et al, 1990). Considered in isolation, the typical bicycle is something like an ordinary street bicycle that can be used for recreational travel in cement or pavement. However, we also know that there are bicycles specialised for various sorts of terrains, and various sorts of uses, including racing and indoor stationary activities. In the context of talking about arctic or snow vehicles, the prototype of *bicycle* is modulated* in response to certain general facts about the need for special traction. Still, it is likely that the prototype assigned by FL to *arctic bicycle* is something like an ordinary bicycle with a special function, say that of being usable in snow and ice. This is just an informal way of saying that the combinatorial processes follow DS. This prototype might well fail to encode some important central tendencies of the category of arctic bicycles. For example, maybe most arctic bicycles do not look like ordinary bicycles: maybe they have special pedaling systems and two slides with a manual propeller instead of wheels.

To generate a more useful concept, the linguistically determined prototype for *arctic bicycle* has to be pragmatically modified. In this case, most of us do not have any relevant extensional beliefs, since acquaintance with exemplars of arctic bicycles is rare. However, we have plenty of relevant general beliefs, e.g., that vehicles designed for snow-use need specially powerful traction systems and tend to have special types of tires. We also know something about why this is the case. We can use these general beliefs to modify the prototype associated with *arctic bicycle* so that it includes emergent features such as SPIKED TIRES. This process is comparatively slow and deliberate. Indeed, it is usually unmentioned—but crucial for this discussion—that studies such as Johnson and Keil (2000) which ask subjects to generate feature lists for novel NPs, do not impose any substantial time constraints on the task. Most subjects take several minutes to generate the feature lists that include the emergent features. From the perspective of CCIP, this slow reasoning tasks transcend the linguistic operations of FL (for further discussion, see §7.2 below).

In short, if we adopt CCIP the cases of emergent features found in the literature do not threaten prototype compositionality. According to CCIP, there are two sources of emergent features. First, the prototypes associated with lexical terms are often modulated* to select the subset of the lexically available information that is relevant in the linguistic and discourse context. These shifts—which can be quite subtle—may result in the addition or elimination of features relative to the default prototypes of lexical terms. Secondly, the compositionally determined prototypes for complex expressions are often taken as inputs by general pragmatic processes which modify them so that they better encode features relevant to the current task (this often requires aiming for statistical central tendencies). These two sources of emergent features are compatible with the compositionality of prototypes.

7 CCIP and recent experimental results

I have argued that each component of CCIP—the account of prototypes and the revised notion of compositionality—is independently plausible. I have also argued that, if we adopt CCIP, the phenomenon of emergent features is reconciled with the compositionality of prototypes. On this view, some emergent features are due to modulation*, and others are due to post-compositional pragmatic processes. This last claim is, of course, open to direct empirical investigation, but I take the previous argument to show that it should, at this point, be our default assumption. Still, it is possible that FL does compute (perhaps via some non-compositional route different from free modulation) some of the reasoning-based emergent features which we attributed to post-linguistic processes. To determine whether there is any evidence of this, I will examine two key experiments—one presented against and one in favour—of prototype compositionality. I will argue that CCIP coheres well with the results, and end this section by briefly describing what kind of experiment we could use to directly test CCIP in the future.

7.1 The modifier effect

Connolly, Fodor, Gleitman, and Gleitman (2007) present an argument against the compositionality of prototypes based on an experiment that seems to show that violation of DS is more widespread than suggested by the classic examples of emergent features. Even in cases were there are no obvious extensional (cf. pet fish living in tanks) or general beliefs (cf. Harvard-educated carpenters being idealistic), they argue that subjects do not respect DS when constructing prototypes for [A N] expressions. Subjects were presented with sets of four types of sentences, and asked to score each sentence on a scale of 1-10 for its likely truth, with 10 = 'highly likely' and 1 = 'highly unlikely' (the numbers indicate average scores for sentences of each type):

(A) *Ducks have webbed feet* (8.3). A sentence S with an unmodified familiar N as subject and a predicate that is typical of instances of the subject.

(B) *Quacking ducks have webbed feet* (7.7). Modifies S by adding a modifier to the subject N that is a typical feature of the N.

(C) *Baby ducks have webbed feet* (7). Modifies S by adding a modifier to the subject N that is a non-typical feature (but neither bizarre nor contradictory) of the N.

(D) *Peruvian baby ducks have webbed feet* (6.5). Modifies S by adding an additional modifier to the one in condition C.

According to Connolly et al, models which accept DS predict that the likelihood of truth scores in each of these conditions should be roughly equivalent. For if you assume that features of the head N which are not directly implicated in the modification should be inherited, then *baby ducks* should be judged to be just as likely to have webbed feet as *baby Peruvian ducks*, since in neither case are means of locomotion directly implicated in the modification. That prediction, however, is incorrect. For each set of sentences, there was a systematic deviation away from the likelihood of truth scores obtained in the baseline condition A. Sentences of type B had reliably lower scores than those of type A, although the modifiers in B stood for typical features of the head Ns. Furthermore, the introduction of one, in C, and two, in D, non-typical modifiers resulted in progressively lower scores. These results supposedly show that subjects do not follow DS when determining the prototypes of complex NPs[16]

The question is: given CCIP, what results should we expect in these experiments? Arguably, something very close to the actual results.

Consider first the aspects of this pattern which are easily explained: the pattern of confidence decrease from type A to B/C to D cases. As Connolly et al acknowledge, there is a relevant Gricean explanation. When the subject head Ns are modified, readers might assume that the modifications are significant,

[16] The results could signal a violation of DS for two reasons. First, they could signal a violation of Semantic Locality: if when computing the prototype of *baby ducks* or *baby Peruvian ducks* some emergent feature is added regarding their type of feet, this would explain why subjects are less confident that *baby ducks* and *baby Peruvian ducks* have webbed feet compared to *ducks*. Second, the results could also signal a violation of Uniform Modification: if the way in which weights are readjusted in the case of *baby ducks* is different to the case of *quacking ducks*, this would explain why subjects' drop in confidence with respect to webbed feel is different in each case. In either case, DS would be violated.

i.e., the act of modification implicates an interpretation of the head N that deviates from the unmodified case. Furthermore, models of prototype combinatorics that follow DS pull in the same direction (Prinz, 2012). For example, Hampton (1997)'s feature pooling model assumes that, when prototypes are combined, weights for dimensions and features are readjusted systematically. Since in [A N] combinations the weight of the dimension of the N which takes the modifier is increased, there is a decrease in the relative weight assigned to other dimensions and features.

Weight readjustment can explain the pattern of confidence decrease from type A to B/C to D cases, but not the decrease reliably shown between B and C cases (e.g., between *quacking ducks* vs. *baby ducks have webbed feet*). For to satisfy Uniform Modification, these weight readjustments have to be computed in a uniform way, and since the NPs in B and C cases have the same syntactic structure and number of modifiers, the (small) decrease in confidence when comparing one typical (type B) and one non-typical but non-conflicting modifier (type C) has to be explained in some other way.

According to CCIP, pragmatic modulation can operate on sub-sentential phrases. So responses to whether the predicative VP *have webbed feet* is likely true of the subject NP are sensitive to the prototype of the subject NP after it has already been modified by general beliefs to ensure that, among other things, it represents a coherent concept and, in some cases, one that encodes statistical central tendencies. What general beliefs could account for the difference between the B and C cases? Consider again the sample stimuli for B and C, where brackets indicate phases sent to pragmatic interpretation:

(B) [Quacking ducks] [have webbed feet] (7.7)

(C) [Baby ducks] [have webbed feet] (7)

The key difference between B and C cases is that, since the C cases involve non-typical modifiers, the subject NPs seem to stand for 'atypical' subclasses of the head N. Being atypical, they are less likely than normal members (subjects of B cases) to have the typical properties (in all cases the predicates stand for typical features of the subject head). There is a natural way of modelling this within CCIP. Subjects follow DS and compositionally compute prototypes in both the B and C cases. This means that the predicated features, being typical of the subject head N, are in all cases inherited into the prototype for the subject NP. However, participants are slightly less confident in the C cases that the resulting prototypes encode statistical central tendencies. So even if having webbed feet is part of the *baby duck* prototype, subjects are comparatively more careful in assuming their typicality.[17] In short, CCIP is perfectly compatible with the 'modifier effect' of Connolly et al (2007).

7.2 Time-sensitive experiment

Time-sensitive paradigms are the most useful tools to investigate to what degree FL is compositional. Although existing experiments do not decisively support or undermine CCIP, it is worth discussing some of them in detail. To appreciate the potential of these paradigms, let us first briefly discuss the main limits of the original experiments of emergent features.

I have argued that, if taken as evidence against the compositionality of prototypes, the main shortcoming of the classic emergent features experiments—e.g., Kunda et al (1990) and Johnson and Keil (2000)—is that they do not control for the time-course of the production of such features. In the basic paradigm, subjects could reason, without significant time limits, about the sorts of features which different classes could have. Recall some examples obtained by Johnson and Keil: *arctic bicycles* had the emergent feature SPIKED TIRES, and *hospital bicycles* had the emergent features METERS & DIALS. In a study which followed the feature production task, Johnson and Keil tried to tap into the reasoning processes behind the generation of emergent features by asking subjects to fill reasoning schemes such as (R):

(R) Since the ARCTIC ____________________
and since BICYCLES ____________________
then arctic bicycles have spiked tires.

Note that (R) ends with the emergent feature obtained in the first study. In this case, most participants reasoned as follows:

(R) Since the ARCTIC *is covered with ice and snow*
and since BICYCLES *need traction*
then arctic bicycles have spiked tires.

In the reasoning scheme for *hospital bicycles*, most participants said that since hospitals *need to monitor patient's health* and since bicycles *require physical*

[17] This explanation of the modifier effect is compatible with the account provided by Gagne and Spalding (2014). They explain the modifier effect as a result of subcategorisation, in conjunction with the belief that subcategories might lack some of the properties of categories. In the explanation above, we also appealed to that belief. Gagne and Spalding (2014) argue that this belief is not based on extensional knowledge, but is more like a meta-belief about the general relation between categories and subcategories. I agree with this, but would add that subjects do seem to be sensitive to the difference between typical and atypical modifiers, otherwise the decrease in confidence between the B and C cases could not be explained. Gagne and Spalding (2014) might be somewhat sceptical about this additional point. The reason is that they present an experiment which introduces a new condition, call it type E, in which the modifier of the NP is an unknown word, e.g., *blika ducks have webbed feet.* The result is that E cases show a decrease that is stronger than B cases. However, we can easily reconcile these results with our interpretation: subjects treat unknown modifiers as atypical modifiers (clearly a reasonable assumption); hence, E cases are treated as C cases.

exertion, hospital bicycles have meters and dials. As these examples illustrate, the reasoning behind the generation of emergent features for novel NPs can be quite subtle. In some cases, subjects might use simple heuristics, but usually subjects took several minutes to generate these sorts of emergent features. Since introducing free modulation to account for these processes within the compositional operations of FL gives rise to several problems (see §5.2), it is reasonable to assume, with CCIP, that reasoning-based emergent features are computed after the initial combinatorial operations of FL.

Now, the conjecture that reasoning-based emergent features are not computed by FL is open to direct empirical (dis)confirmation. The best way to test this prediction of CCIP in the future is to perform time-sensitive studies of the main types of emergent features. Unfortunately, existing studies do not explicitly control for the types of emergent features which they use. A look at their materials suggests that, unlike Johnson and Keil (2000), they overwhelmingly focus on the types of extension or memory based features which, according to CCIP, are due to modulation*.[18] Still, I will discuss one time-sensitive experiment, conducted by McElree, Murphy, and Ochoa (2006). This experiment provides some limited support to one aspect of CCIP, and it can be fruitfully used to clarify how future experiments could be designed.

McElree et al (2006) used a speed-accuracy trade-off paradigm. Participants were presented with sentences consisting of a complex NP as subject and a predicated property. Their task was to verify whether the predicated property was true (or false) of the subject. They were required to respond at six time windows, and the dependent measure was response accuracy at each time window. Consider an example:

(9) a. Water pistols have triggers. ['true N and NP']
b. Water pistols are harmless. ['true only NP']

In the (a) cases, the predicated property is true of both the head noun, and the noun phrase ('true N and NP'). In the (b) cases, the predicated property is an emergent feature, i.e., is true only of the noun phrase ('true only NP'). In the experiment, the noun-phrase appeared on a screen for 600 msec. Then the predicate appeared. Then either 300, 500, 700, 900, 1,500 or 3,000 msec after the presentation of the predicate a tone was heard that signalled to participants that they had to verify ('yes' or 'no') if the predicate applied to the noun-phrase. Participants were trained to respond within 100 to 300 msec after hearing the tone. The results were as follows. By 2 sec after the

[18] In §8.1 I discuss how we can operationalise the distinction between extensional feedback or memory based and reasoning based features. For now, I will assume that there is an intuitive distinction. To see this, consider some representative examples of the phrases and emergent features used in the time-sensitive studies we will examine: *peeled bananas* and WHITE, *peeled apples* and WHITE, *water guns* and HARMLESS, *boiled celery* and SOFT. These are all categories with which most participants are likely acquainted, unlike the examples used by Johnson and Keil, which focus on novel or surprising categories such as *arctic bicycle*. In addition, note that in the memory based examples it is very hard to reason causally to explain why the emergent feature is typical of the class, without appealing to direct knowledge that the class has in fact that feature.

presentation of the predicate subjects were extremely accurate at verifying the predicates, in both conditions. Based on the accuracy patterns at the earlier time-windows, the experimenters were able to determine that the emergent features in the 'true only NP' condition were available around 84 msec later than were the noun and phrase features in the 'true N and NP' condition.

The results suggest that extension or memory based emergent features (those used in the 'true only NP' condition) have almost the same computational profile as paradigmatically compositional features (those used in the 'true N and NP' condition). This provides some limited support for CCIP, according to which the computational profile (hence measures of time/effort) of these sorts of emergent features should be closer to that of paradigmatically compositional features (since bot are a result of FL) than to that of paradigmatic post-compositional pragmatic processes.[19] At the same time, CCIP also entails that, compared to memory based features, reasoning based emergent features for novel NPs—such as those in Johnson and Keil's cases—involve more complex post-linguistic computations, hence should show significantly different measures of processing time/effort. Since no time-sensitive studies use these sorts of emergent features, this conjecture is still open to direct empirical investigation.

Still, how we *could* test that conjecture should at this point be clear: compare the time-course of memory vs reasoning based emergent features. According to CCIP, the latter type of emergent features should show a significant processing delay. To test this, we can use time-sensitive behavioural paradigms, or neuroscientific tools with good temporal resolution, such as EEG. To illustrate: the event related potential N400 inversely correlates with semantic expectations, and can be used as an electrophysiological measure of semantic priming (Kutas and Federmeier, 2011).[20] As stimuli, we can use sentences that predicate reasoning-based emergent features of subject NPs (e.g., *Arctic bicycles have spiked tires*), sentences that predicate memory based emergent features of subject NPs (e.g., *Peeled bananas are white*), and, as baselines, sentences that predicate features typical of *both* the head Ns and the complex NPs (e.g., *Arctic bicycles have two tires*). If we get a significantly larger N400 when processing the predicates that stand for reasoning-based features relative to the other cases, this would suggest that such emergent features are *not* computed during online processing, as predicted by CCIP.

[19] A study which used a cross-modal lexical priming paradigm provides converging evidence for this conclusion (Swinney et al, 2007). Specifically, it showed that complex NPs prime memory based emergent features by the onset of the word that came right after the head N of the NP. The stimuli used by Swinney et al (2007) can be accessed at: http://lcnl.ucsd.edu/publications.html.

[20] Famously, sentences which predicate typical features of the subject such as *Dutch trains are yellow* generate less N400 activity at the predicate than sentences such as *Dutch trains are white/sour*, which do not predicate typical features.

8 Objections

The construction of models of prototype combinatorics is an active and challenging area of research. I have presented a broad framework that, if adopted by particular models, insulates them against a general and very influential objection to the entire enterprise. To close this discussion, let us consider two concerns which, I think, some philosophers would be tempted to raise. The first focuses on the view of the semantics/pragmatics distinction which follows from CCIP. The second challenges the idea that we need to reconcile compositionality with prototype theories of meaning.

8.1 CCIP and the semantics/pragmatics distinction

According to CCIP, although modulation* accounts for some emergent features, other types of emergent features—paradigmatically, the reasoning-based features studied by Johnson and Keil (2000)—arise due to inferential, post-linguistic pragmatic processes. Since CCIP invokes a version of the semantics/pragmatics distinction, one might reasonably ask: why not just accept the traditional Gricean semantics/ pragmatics distinction, pair it with a non-context-sensitive version of compositionality and DS, and hold that all emergent features are due to post-linguistic inferential processes? If we adopt this suggestion, many compositionally determined phrasal prototypes would fail to encode central tendencies, but CCIP also adopts a view of prototypes according to which this is not, by itself, a serious problem; so why not stick to the traditional distinction and simply hold that *all* emergent features are due to inferential, post-linguistic processes?

There are both explanatory and descriptive reasons why we cannot adopt the traditional semantics/pragmatics distinction. First, I argued that modulation*, and the constrained context sensitive version of compositionality that follows from it, is not just a theoretical assumption made to explain how prototypes could be compositional. Rather, modulation* is a basic property of cognitive systems which have lexical items with rich informational contents, only a subset of which can be profitably used in each instance of a working memory task. Furthermore, modulation* entails that many emergent features will arise during the first stage of purely compositional, linguistic processes, and explains why they are part of the intuitive literal meaning of expressions. Paradigmatic examples of these are memory based emergent features. The time-sensitive experiments show that memory based features have a computational profile that is almost indistinguishable from paradigmatic compositional features. In addition, priming studies show that complex NPs prime memory-based emergent features very rapidly—specifically, before full clausal phrases are computed (Swinney et al, 2007). If we take those results seriously, it is simply no longer an option to hold, as implied by the traditional view, that these emergent features are computed by post-linguistic processes which take the meaning of clauses or full sentences as inputs.

To be clear, I have also emphasised aspects of CCIP which are still open to direct empirical confirmation. In particular, that reasoning-based emergent features are computed by relatively slow, inferential processes is a reasonable default assumption that should be directly tested. In §7.2 I suggested one way in which it could be tested. Now, both that suggestion, and much of what I have said here, assumes that we can usefully distinguish between memory and reasoning based emergent features—and though this is something that one can reasonably doubt, I think we can be optimistic.

Any operationalisation of that distinction will have to admit ambiguous cases. For example, the feature LIVE IN TANKS for *pet fish* is likely based on extensional feedback, but maybe one could also generate it via reasoning. However, all we need to directly test CCIP is a method that allows us to reliably pick out a set of paradigmatic examples of memory and of reasoning based emergent features. The intuitive difference is clear enough: most members of the relevant community are (directly or indirectly) acquainted with exemplars of the first kind, but not with exemplars of the second kind. So we should focus on categories with whose members and relevant emergent features we are likely acquainted in the extensional cases, and on clearly novel categories in the reasoning based cases.[21] To control for the reliability of our intuitions, we can construct, for each NP and emergent feature pair, reasoning schemes such as (R) above. Filling out the reasoning schemes should be reliably easier for paradigmatic reasoning-based features, and harder or even impossible for purely extension based emergent features. So we can ask subjects to fill out such schemes, rate each task as natural/easy or hard/weird, and then use those ratings to classify the NP and emergent feature pairs.[22]

8.2 Why account for prototype combinatorics at all?

Some theorists have argued that we can avoid the whole conundrum of coming up with a combinatorial model of prototypes. The idea goes back to a debate about compositionality and conceptual role theories. Fodor argued that conceptual roles are not compositional: the inference from *brown cow* to *dangerous* might be part of the conceptual role of *brown cow*, even if the inference to *dan-*

[21] For example, in their study of emergent features in combinations involving social concepts, Kunda et al (1990) found that how novel or surprising subjects found the class denoted by a combination, was a good predictor of whether other subjects would give causal explanations for their emergent features.

[22] For example, in a pilot study conducted to test this proposal, we used reasoning schemes such as (R) and (C): (R) “Since the ARCTIC ________________ , and since BICYCLES ________________ , then arctic bicycles have spiked tires” and (C) “Since PEELED things ________________ , and since BANANAS ________________ , then peeled bananas are white”. (R) uses a paradigmatic NP plus reasoning-based emergent feature pair, and (C) uses a paradigmatic NP plus memory based emergent feature pair. After each reasoning scheme, we asked subjects to rate the task on a 1 (easy/natural) to 7 (hard/weird) likert scale. NPs plus reasoning based feature pairs consistently rank as significantly easier/more natural than NPs plus memory based feature pairs. So there is some initial evidence that we can use tasks such as this to classify emergent features.

gerous is not part of the conceptual role of either *brown* or *cow*. In response, Horwich (1998) and Recanati (2006) argue that Fodor's argument unjustifiably assumes the 'Uniformity Assumption', i.e., that what is constitutive of lexical concepts—say, prototypes or conceptual roles—should also be constitutive of complex concepts. In particular, Recanati argues that once conceptual roles determine the reference for lexical terms, all we need to say about complex terms is that their reference is determined compositionally: to know the meaning of *brown cow* is just to know that it is that concept whose extension is the intersection of the brown things and the cows. This move might seem attractive to dual-content theorists. For if we hold that prototypes are one component of meaning, the other being reference determining cores, we can also hold that although simple concepts have prototypes, complex concepts do not: prototypes do not compose, only reference-determining conceptual components. Margolis and Laurence (1999) proposed essentially this response to save dual-content prototype theories from the emergent features challenge.

However, if we reject the Uniformity Assumption prototype theory would lose much of its explanatory power, and the claim that prototypes are conceptual components would become an unwarranted stipulation. To illustrate, one reason why prototype theory is becoming increasingly attractive to linguists and philosophers is that there are modifiers that seem to operate on the prototypes of nouns and verbs (Taylor, 2009; Sassoon, 2011; Leslie, 2015; Del Pinal, 2015). These accounts are only available if we assume that prototypes are *compositional* components of meaning. Take modifiers such as *typical* and *perfect*. Although a gun might not have been made to shoot, a *typical* gun ought to have been made to shoot. Although there are guns that might be lousy shooting instruments, *perfect* guns shoot well. Modifiers such as *typical* and *perfect* seem to operate on subcomponents of the prototypes of the head Ns. Other modifiers that seem to access the prototypes of Ns are privative Adjs such as *fake* and *counterfeit* (Del Pinal, 2015), and certain modifiers of social-role concepts such as *real* and *true* (Leslie, 2015). Although friends can betray, *real* friends don't betray, the suggestion being that in complex expressions such as *real friends*, aspects of the prototypes associated with the head Ns (e.g., some ideal) can be accessed by the modifier. In all these cases, prototype modifiers can operate on complex NPs such as *fake plastic gun*, *real childhood friend*, and *typical faithful priest*. So these accounts presuppose that, at least in general, phrases such as *plastic gun*, *childhood friend*, etc., also have prototypes, as required by the Uniformity Assumption.

These promising accounts are only available if we hold that, in general, complex expressions also have prototypes. Prototype theory would lose much of its explanatory power if, as a 'solution' to the emergent features challenge, we drop the Uniformity Assumption. So we should try to meet the challenge directly, even if we accept a dual-content theory. CCIP allows theorists to take advantage of the compositional role of prototypes to freely explore how this might help deal with certain linguistic constructions, including some that have previously been taken as problem cases for compositionality.

9 Conclusion

The idea that linguistic meaning includes prototypes was originally defended by Putnam (1997) alongside his influential views on the directly referential component. However, the widespread acceptance of the objection that prototypes are not compositional is largely responsible for obliterating the interest among semanticists and philosophers in developing the prototype component of Putnam's original theory of meaning. As a result, most of the theories of meaning explored and developed by semanticists and philosophers in the last couple of decades have been excessively unresponsive to advances in the cognitive psychology of concepts. The recent efforts by prototype theorists to bridge this gap and address the emergent features objection by adopting radically unconstrained combinatorial principles, or by dropping compositionality altogether, have proven to be unconvincing. In contrast, CCIP allows theorists to keep a principle that, for various good reasons, most are unwilling to drop—i.e., some form of strict compositionality. At the same time, CCIP opens space for the guiltless exploration of theories of linguistic meaning which are more closely informed by recent advances in the empirical study of concepts.

Acknowledgements For helpful comments and discussions of earlier drafts of this paper, I am grateful to Luca Barlassina, Akeel Bilgrami, Brian H. Kim, Karen S. Lewis, Eleonore Neufeld, Daniel Rothschild, and Achille Varzi. I am also extremely grateful to two anonymous referees whose excellent comments led to substantial improvements of the paper. This work was supported by the Alexander von Humboldt Foundation and by BMBG Grant No. 01UG1411.

References

(1996) Lexical limits on the influence of context, Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, Lawrence Erlbaum Associates Publishers

Asher N (2011) Lexical Meaning in Context: A Web of Words. Cambridge University Press, Cambridge, UK

Barsalou LW (1987) The instability of graded structure: Implications for the nature of concepts. In: Neisser (ed) Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization, Cambridge University Press, Cambridge, pp 101–140

Chomsky N (2001a) Beyond explanatory adequacy. Tech. rep., MIT, Cambridge, MA

Chomsky N (2001b) Derivation by phase. In: Kenstowitz M (ed) Ken Hale: A Life in Language, MIT Press, Cambridge, MA, pp 1–52

Connolly AC, Fodor J, Gleitman L, Gleitman H (2007) Why stereotypes don't even make good defaults. Cognition 103(1):1–22

Cook VJ, Newson M (2007) Chomsky's Universal Grammar: An Introduction. Blackwell Publishing Ltd, Oxford

Del Pinal G (2014) The architecture of our semantic compentence: Compositional operations and complex lexical representations. PhD thesis, Columbia University, New York

Del Pinal G (2015) Dual content semantics, privative adjectives and dynamic compositionality. Semantics and Pragmatics 8(6):1–56

Fodor J (1999) Concepts: where cognitive science went wrong. Oxford University Press, New York

Fodor J, Lepore E (1996) The pet fish and the red herring: why concepts still can't be prototypes. Cognition 58(2):329–43

Fodor J, Pylyshyn Z (1988) Connectionism and cognitive architecture: A critical analysis. Tech. rep., Rutgers Center for Cognitive Science

Fodor JA, Pylyshyn ZW (2015) Minds without Meanings. The MIT Press, Cambridge, MA

Gagne CL, Spalding TL (2014) Subcategorization, not uncertainty, drives the modification effect. Language, Cognition and Neuroscience 29(10):1283–1294

Gleitman LR, Connolly AC, Armstrong SL (2012) Can prototype representations support composition and decompositions? In: Werning M, Hinzen W, Machery E (eds) The Oxford Handbook of Compositionality, Oxford University Press, Oxford, chap 20

Hall A (2008) Free enrichment or hidden indexicals? Memory & Language 23:426–56

Hamill R, Nisbett RE, Wilson TD (1980) Insensitivity to sample bias: generalizing from atypical cases. Journal of Personality and Social Psychology 39(4):578–589

Hampton J (1997) Conceptual combination. In: Lamberts K, Shanks D (eds) Knowledge, Concepts, and Categories, The MIT Press, Cambridge, MA, pp 133–60

Hampton J (2006) Concepts as prototypes. In: Ross BH (ed) The psychology of learning and motivation: advances in research and theory, vol 46, Elsevier, pp 79–113

Hampton JA (1987) Inheritance of attributes in natural concept conjunctions. Memory & Language 16:579–591

Hampton JA, Jönsson ML (2012) Typicality and compositionality: the logic of combining vague concepts. In: Werning M, Hinzen W, Machery E (eds) The Oxford Handbook of Compositionality, Oxford University Press, Oxford, chap 18

Horwich P (1998) Concept constitution. Philosophical Issues 9:15–19

Jacobson P (2012) Direct compositionality. In: Marcus Werning WH, Machery E (eds) The Oxford Handbook of Compositionality, Oxford University Press, Oxford, chap 5, pp 109–28

Johnson C, Keil F (2000) Explanatory understanding and conceptual combination. In: Keil F, Wilson R (eds) Explanation and Cognition, The MIT Press, Cambridge, pp 327–359

Jönsson ML, Hampton JA (2008) On prototypes as defaults (comments on connolly, fodor, gleitman and gleitman 2007). Cognition 106:913–23

Kahneman D (1973) Attention and Effort. Prentice-Hall, Inc, Englewood Cliffs, New Jersey

Kahneman D (2011) Thinking Fast, Thinking Slow. Farrar, Straus, and Giroux, New York

Kunda Z, Miller DT, Claire T (1990) Combining social concepts: The role of causal reasoning. Cognitive Science 14:551–577

Kutas M, Federmeier KD (2011) Thirty years and counting: finding meaning in the n400 component of the event-related brain potential. Annual Review of Psychology 62:621–47

Leslie SJ (2015) "Hilllary Clinton is the only man in the Obama administration": Dual character concepts, generics and gender. Analytic Philosophy 56(2):111–141

Machery E (2011) Concepts: A tutorial. In: Belohlavek R, Klir GJ (eds) Concepts and Fuzzy Logic, The MIT Press, Cambridge, Massachusetts, chap 2

Margolis E, Laurence S (1999) Concepts and cognitive science. In: Margolis E, Laurence S (eds) Concepts: core readings, The MIT Press, Cambridge, Massachusetts, chap 1, pp 3–82

McElree B, Murphy GL, Ochoa T (2006) Time course of retrieving conceptual information: a speed-accuracy trade-off study. Psychonomic Bulletin & Review 13(5):848–853

Morewedge CK, Kahneman D (2010) Associative processes in intuitive judgment. Trends in Cognitive Sciences 14(10):435–40

Morzycki M (2015) Modification, forthcoming edn. Cambridge University Press, Cambridge, UK

Murphy GL (1990) Noun phrase interpretation and conceptual combination. Journal of Memory and Language 29:259–288

Murphy GL (2002) The big book of concepts, The MIT Press, Cambridge, Massachusetts

Pagin P, Pelletier J (2007) Content, context, and composition. Context-Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics p 25

Partee B (1998) Lexical semantics and compositionality. In: Gleitman L, Liberman M (eds) An Invitation to Cognitive Science: Volume 1, vol 1, The MIT Press

Prinz J (2012) Regaining composure: a defense of prototype compositionality. In: Werning M, Hinzen W, Machery E (eds) The Oxford Handbook of Compositionality, Oxford University Press, Oxford, chap 21

Prinz JJ (2002) Furnishing the mind: concepts and their perceptual basis, MIT Press, Cambridge, Massachusetts

Putnam H (1997) Is semantics posible? In: Mind, Language and Reality: Philosophical Papers, vol 2, Cambridge University Press, Cambridge

Rabagliati H, Marcus GF, Pylkkanen L (2011) Rules, radical pragmatics and restrictions on regular polysemy. Journal of Semantics 28(4):485–512

Radford A (2004) Minimalist Syntax: Exploring the Structure of English. Cambridge University Press, Cambridge

Recanati F (2006) Indexical concepts. In: Garcia-Carpintero M, Macia J (eds) Two-Dimensional Semantics, Claredon Press, Oxford
Recanati F (2010) Truth-conditional pragmatics. Oxford University Press, Oxford
Robbins P (2002) How to blunt the sword of compositionality. Nous 36:313–334
Rosch E (1973) Natural categories. Cognitive Psychology 4:532–547
Rosch EH (2011) Slow lettuce: Categories, concepts, fuzzy sets, and logical deduction. In: Belohlavek R, Klir GJ (eds) Concepts and Fuzzy Logic, The MIT Press, Cambridge, Massachusetts, chap 4
Roth EM, Shoben EJ (1983) The effect of context on the structure of categories. Cognitive Psychology 15:346–378
Sassoon GW (2011) Adjectival vs. nominal categorization processes: The rule vs. similarity hypothesis. Belgium Journal of Linguistics 24:104–147
Schurz G (2012) Prototypes and their composition from an evolutionary point of view. In: Werning M, Hinzen W, Machery E (eds) The Oxford Handbook of Compositionality, Oxford University Press, Oxford, chap 26
Sloman S (2002) Two systems of reasoning. In: Gilovich T, Griffin D, Kahneman D (eds) Heuristics and Biases: The Psychology of Intuitive Judgment, Cambridge University Press, Cambridge, UK, chap 22
Smith EE, Medin DL, Rips LJ, Keane M (1988) Combining prototypes: a selective modification model. Cognitive Science 12:485–527
Stanley J (2007) Language in Context. Oxford University Press, Oxford, UK
Swinney D, Love T, Walenski M, Smith EE (2007) Conceptual combination during sentence processing. Psychological Science 18(5):397–400
Tabossi P (1988) Effects of context on the immediate interpretaion of unambiguous nouns. Journal of Experimental Psychology: Learning, Memory, and Cognition 14(1):153–162
Taylor JR (2009) Linguistic Categorization. Oxford University Press, Oxford
Weiskopf D (2009) Atomism, pluralism and conceptual content. Philosophy and Phenomenological Research LXXIX(1):131–164