The Structure of Semantic Competence: Compositionality as an Innate Constraint of The Faculty of Language*

Guillermo Del Pinal[†]

Department of Philosophy
Columbia University

forthcoming in *Mind & Language*

Abstract

This paper defends the view that the Faculty of Language is compositional, i.e., that it computes the meaning of complex expressions from the meanings of their immediate constituents and their structure. I first argue that compositionality and other competing constraints on the way in which the Faculty of Language computes the meanings of complex expressions should be understood as hypotheses about innate constraints of the Faculty of Language. I then argue that, unlike compositionality, most of the currently available non-compositional constraints predict incorrect patterns of early linguistic development. This supports the view that the Faculty of Language is compositional. More generally, this paper presents a way of framing the compositionality debate, focusing on its implications for language acquisition, that can lead to its eventual resolution, so it will hopefully also interest theorists who disagree with its main conclusion.

Keywords: compositionality; the Faculty of Language; language acquisition; innate constraints. **Words**: 13,066.

[†]Email: ged2102@columbia.edu. Address: Columbia University, Department of Philosophy, 708 Philosophy Hall, MC 4971 1150 Amsterdam Avenue New York, NY 10027

1 Introduction

The human Faculty of Language (FL)—the mental faculty which plays a central role in the acquisition and processing of natural languages—enables us to systematically produce and understand an unbounded number of novel expressions. More precisely, if 'FL_S' is the FL of an arbitrary competent speaker S of some natural language E (English, Spanish, etc.), then:

(PRODUCTIVITY) FL_S can generate *correct* interpretations (relative to E) for complex expressions which S has never encountered before. FL_S has this capacity for indefinitely many distinct complex expressions, generating a distinct meaning for an indefinite number of these expressions.

(SYSTEMATICITY) The generative capacity of FL_S is structured in the following way: if it can generate *correct* interpretations (relative to E) for complex expressions $e_1, \ldots, e_n$, it can generate correct interpretations for all other complex expressions constructed from: (i) constituents of $e_1, \ldots, e_n$ and (ii) syntactic structures employed in any of the complex expressions $e_1, \ldots, e_n$.

For a long time, it was held that the best way to explain PRODUCTIVITY and SYSTEMATICITY (henceforth '$P\&S$')[1] is to assume that FL includes a recursive computational system with a compositional semantics.[2] Roughly, a computational system is 'compositional' if (i) it contains both primitive and syntactically complex symbols, and (ii) it is constrained to determine the semantic properties of its complex symbols only from their structure and the semantic properties of their immediate constituents.

Recently, however, several prominent linguists, philosophers and cognitive scientists have criticized the view that FL is compositional, and especially the argument from $P\&S$.[3] There is some truth to these criticisms. The traditional argument for compositionally is indeed substantially incomplete. No one doubts that assuming that FL is compositional is *one* way of explaining $P\&S$; but there are now other reasonable explanations which assume that FL is *not* compositional.[4] This presents an important challenge. For the assumption that FL is compositional has shaped the way we theorize about basically every central aspect of our linguistic competence.

---

[1] Here we will assume that $P\&S$ are essentially correct. Although widely accepted, esp. in discussions of the non/compositionality of FL, $P\&S$ are not completely uncontroversial (see e.g. Pullum and Scholz [71], [61]). For an extensive defense of $P\&S$, see Del Pinal [56].

[2] The most famous version of this argument was presented by Fodor to defend the compositionality of both natural languages and thought (see Fodor and Pylyshyn [24], Fodor [21], Fodor and Lepore [23]). However, Fodor recently presented some arguments to the effect that natural languages don't seem to be compositional (Fodor [22]).

[3] Some theorists argue on general grounds that our linguistic competence is (probably) not compositional (see Jonsson [39]; Hampton and Jonsson [33]; Johnson [38]; Travis [81]; Lahav [42]). Others argue that although it might be compositional, the general arguments usually taken to establish this are not very persuasive (see Szabo [76], [78]; Dever [18]; Pagin and Westersthl [49]; Baggio et al. [1]).

[4] Jonsson's illuminating [39], which we will discuss in detail in what follows, presents various non-compositional explanations of $P\&S$.

Despite these criticisms, I will argue that assuming that FL is compositional is still the best explanation of $P\&S$. I begin by proposing that we should frame debates about the non/compositionality of FL as debates about the fixed, innate structure of the part of FL which computes the meanings of complex expressions (§2-3). Given this framework, we can determine, for each competing explanation of $P\&S$, its broad empirical consequences for language acquisition and development. I then show that, unlike assuming that FL is compositional, the non-compositional accounts of $P\&S$ entail that, in the course of acquiring a natural language, speakers should go through certain stages of early linguistic development which, it turns out, speakers never seem to go through (§4-5). This strongly suggests that FL is compositional. In the final section I discuss some objections to this argument (§6).

2 The Faculty of Language as a Cognitive Computational System

Following the tradition of computational psychology, we will assume that FL is a language-processing cognitive computational system. This approach has been famously advocated by Chomsky [8], [9], [10]. The view I present below is Chomskian, but in a weak sense that can be welcomed by theorists of different theoretical orientations. This approach revolves around two basic theoretical notions: 'I-languages' and 'languages'. Both terms require some explication.

By a 'language' I mean a set of pairs of acoustic/visual signals and meanings or interpretations which characterizes a natural language. For example, 'English' is a set consisting of certain pairs of acoustic/visual signals and meanings, e.g., (*red*, RED), (*John is happy*, HAPPY (JOHN)), (*red wall*, RED WALL), and so on. A 'language', as I am using the term, is 'extensionally defined' but *not* an external or mind-independent abstract structure of the sort Chomsky [8] argues is of no relevance to the study of FL. Specifically, 'languages' consist of the input/output pairs of representations which 'I-languages' compute, i.e., they specify the main cognitive task which 'I-languages' solve.[5]

Most linguistic theorists agree that, to compute languages, I-languages need to carry-out *at least* three main cognitive tasks: (i) map acoustic/visual signals into expressions (phonetics), (ii) map expressions into syntactic structures (syntax), (iii) map syntactic structures into meanings or interpretations (semantics). An 'I-language' then is a cognitive computational system that can generate phonetic structures, syntactic structures, and semantic structures or interpretations. For example, think of 'I-English' as a cognitive computational system

---

[5]'English', as used here, does not exactly correspond to what we ordinarily mean by 'English'. For example, 'English', as ordinarily used, does not include the expression *the child seems sleeping*. But the meaning of this expression is arguably computed by I-English, so it is part of English, as here defined (see Pietroski and Crain (forthcoming)). Another complication which we can ignore for now is that the semantics does not generate interpretations for all the outputs of syntax. A putative example of this is Chomsky's *colorless green ideas sleep furiously* (but see Camp [7] for rejection of the view that these sorts of sentences don't have a literal interpretation).

that, given certain signals, outputs certain meanings (and vice-versa), thereby computing English. For our purposes, we can remain neutral about the 'nature' of the representations used by I-languages, but in principle this framework can be paired with a Chomskian internalist view (see e.g., Pietroski [53], [54]) or with any of the externalist views more commonly adopted by philosophers and formal semantic theorists (see e.g., Ludlow [44]).[6]

I-languages, as here defined, have three basic properties, crucial to debates about the non/compositionality of FL: (i) they are idiolects, (ii) they have an unbounded generative capacity; and (iii) some of their properties are innate.

I-languages are idiolects. In principle there can be as many distinct I-languages and languages as there are individuals with a FL; or even more, since on the way to its mature and stable state, each FL goes through various developmental stages. However, general cognitive constraints, including some particular to FL, together with general properties of language acquisition environments, reduce the differences between individual I-languages, at least to the extent required for successful communication between speakers of what we informally call a 'linguistic community'.[7] Here we are concerned with general properties of I-languages, so we can ignore variations between the I-languages of members of the same 'linguistic community', when compared at the same stage of development. For example, we will assume that the community of what we informally call 'English speakers' is a linguistically homogenous community, so that 'English' captures (not perfectly) the set of <acoustic/visual signal, meaning> pairs which members of this community use to communicate. Relative to this idealization, I-English is the cognitive computational structure that an arbitrary competent speaker of this homogenous linguistic community uses to compute English.

I-languages have an unbounded generative capacity. I-languages can assign interpretations to an unbounded number of novel expressions, following the patterns specified in $P\&S$. For this reason, theorists studying languages often define or describe (fragments of) them intensionally, rather than by listing their <acoustic/visual signal, meaning> pairs. Different functions-in-intension can define the same set of pairs, the same language, in which case we call them 'extensionally equivalent'. It is sometimes said that if two or more functions-in-intension are extensionally equivalent, the claim that one is the 'correct' one doesn't make sense. This is correct if we assume that the only task of a linguistic theory is to describe or define a language; but incorrect if we assume that part of the task of a linguistic theory is to discover (properties of) I-languages (Chomsky [8]; Evans [20]; Davies [16]). We can discriminate between (at least some) extensionally equivalent models on the grounds that one is, or seems to

---

[6]Various authors now use the term 'I-language' in the sense that I am proposing, i.e., as an individual computational systems which process natural languages, but which are neutral with respect to internal/external individuation. To capture this sense, Ludlow [44] introduced the term 'Ψ-language'. This discussion is indebted to his discussion of the small but important distinction between 'Ψ-languages' and Chomsky's 'I-languages'.

[7]That idiolectical conceptions of language can be wedded to successful accounts of communication has been shown by Larson and Segal [43] and Higginbotham [35].

be, closer to the function actually computed by the I-language's computational processes. Whether a particular piece of psychological data can be used to make this discrimination is usually controversial. Still, there is widespread agreement that relevant evidence can come from data about patterns of language loss, acquisition or revision, and from any neurological data that reveals properties about the computational capacities or structure of the mind/brain.

Some properties of I-languages are innate. FL undergoes development from an initial state prior to exposure to linguistic data, through various intermediate states, to the 'mature' and stable state in which it incorporates I-languages that can fully compute 'natural languages' such as English, Spanish, etc. 'Mature' I-languages consist of certain semantic and syntactic rules and principles, some of which have to be acquired in the course of linguistic development. For example, speakers acquiring I-English have to acquire lexical rules such as $[\![red]\!] = \text{RED}$, and syntactic principles such as that heads precede their complements and that null subject sentences such as *is raining* are not allowed. Other rules and principles, which do not seem to be learned in the course of language development, are more plausibly seen as innate and often fixed properties of FL common to all I-languages. Some candidates for innateness are the syntactic and semantic primitives, the constraints that all syntactic principles are structure-depedent and all syntactic branching is binary.

Precisely which rules and principles are innate, and which of these are unique to FL, is a matter of ongoing debate between nativists who propose a substantial base of innate and language-specific structures (e.g. Berwick et al. [64], Baker [2]), empiricists who propose a minimal base of innate and no language-specific structures (e.g. Elman et al. [19], Pullum and Scholz [60], Perfors et al. [51]), and theorists who defend mixed or intermediate positions (e.g., Xu [85]). But what is not controversial—and what we will assume—is that FL has *some* innate structure, common to all I-languages, except those affected by unusual genetic or developmental conditions. If we drop this assumption, it is impossible to explain how FL can represent and interact with linguistic data to begin to develop into a mature I-language.[8]

To close this brief presentation of our operating conception of FL, I should add that I will *not* assume that FL is an informationally encapsulated cognitive module. This is important because some advocates of compositionality assume

---

[8]There is a heated debate about the learning mechanisms used in language acquisition. Linguists tend to emphasize that acquiring an I-language, as Szabo [77] puts it, 'requires little or no explicit instruction', follows a certain developmental sequence, and 'tends to yield a remarkably uniform level of competence'. This suggests that, for the most part, acquiring an I-language is quite unlike learning social conventions (Szabo [77]) or scientific theories (Chomsky [13]). Unlike knowledge of a scientific theory, speakers are not conscious of—and if prompt cannot state—most of the rules and principles which they 'acquire' as part of their I-languages. When 'acquiring' most rules and principles, speakers do not seem to make the sorts of mistakes they would make if they were constantly testing reasonable but incorrect 'hypotheses' against language data. However, some developmental psychologists disagree with this picture and argue that language acquisition essentially relies on our innate, but domain general, 'science forming' mechanisms (e.g., Gopnik [30]). As will become clear, the argument we present for the compositionally of FL does not *assume* any one of these competing views on the nature of the language learning mechanisms.

that FL is informationally encapsulated, and then appeal to this property to defend its compositionality (see e.g. Borg [6] and Larson and Segal [43]). Critics are justifiably skeptical of this way of defending compositionality (Robbins [68]; Jonsson [39]: chap. 6). I-languages exhibit *some* degree of modularity—they are domain specific, have mandatory operations which are (for the most part) fast, have limited central accessibility, and characteristic patterns of breakdown and development. If to account for these features of I-languages we assume their (almost) total informational encapsulation, we move towards the view that they are likely compositional. For most non-compositional accounts require that, to determine the meanings of complex expressions, the semantics have access to some subset of non-linguistic information. However, as most theorists rightly point out, modularity comes in degrees. Even if we can hold that FL is modular to some non-trivial degree, there is currently no good reason to assume that it is informationally encapsulated to the degree that would be required to make it incompatible with most reasonable non-compositional accounts (Robbins [68]).

3   Compositionality as a functional constraint of the Faculty of Language

In our framework, to say that FL is compositional is to say that there is a particular constraint on the way in which it generates the meanings of complex expressions: the algorithms which generate semantic interpretations for complex expressions can only use semantic information provided by their immediate constituents and information about their combinatorial structure. This does not tell us, for a particular type of complex expression (e.g., $[_{NP}A\ N]$), what particular algorithm determines its meaning; it only tells us that the algorithm computes a compositional function. We will call general semantic constraints (such as compositionality and other competing constraints) which range over all types of complex expressions, 'meaning-determination constraints' (MDCs). MDCs should be distinguished from particular 'semantic rules' (SRs) which determine the meanings of *particular types* of complex expressions (e.g., $[\![_{NP}A\ N]\!] = f_{NP}([\![A]\!], [\![N]\!]))$. MDCs range over and constraint the general form of particular SRs.

This distinction between MDCs and SRs raises an important question which has been neither sufficiently nor adequately raised in the literature. Should we think of compositionality as a principle that we learn when we acquire some I-language (so that we could have acquired a different MDC)? Or should we think of it more like an innate and fixed property of FL, hence present in all I-languages? The latter option is closer to the way in which I suggest we should understand compositionality and other competing, non-compositional MDCs. Specifically, we should understand MDCs as constraints on what, following Pylyshyn [62], we'll call the 'functional architecture' of the semantics of FL.

We can understand the notion of 'functional architecture' by analogy to the way in which it is used in computer science (Dawson [17]; Pylyshyn [62], [63]).

The functional architecture of a computational system M is the fundamental programming language used to write the algorithms that M computes. This programming language is fundamental in the sense that its primitive operations or functions must be built into the (possibly virtual) machine M. Similarly, the functional architecture of a cognitive computational system C (e.g., an I-language) is something like the basic set of representations and operations available to C. The particular rules and algorithms which can be represented and computed by C are those which can be defined in terms of Cs basic programming language. So if we specify Cs functional architecture we thereby implicitly specify Cs cognitive capacity, i.e., the set of cognitive rules and algorithms which can be represented and processed by C. A 'functional constraint' on C is a way of (partially) specifying Cs functional architecture, hence (implicitly) Cs cognitive capacity.

To further clarify this notion, consider a rule which is clearly *not* a functional constraint, e.g., the lexical semantic rule $[\![red]\!] = $ RED. This rule might be part of some I-languages—e.g., I-English—but it is obviously not a functional constraint. Firstly, particular lexical rules are optional features of I-languages. In certain conditions, FL can acquire the rule $[\![red]\!] = $ RED; but in other conditions, FL can acquire different rules for $[\![red]\!]$, e.g., $[\![red]\!] = $ BLUE, or $[\![red]\!] = $ ANGRY. Secondly, the processes of learning lexical rules such as $[\![red]\!] = $ RED can be usefully understood as a rational learning process in which different hypotheses about the meaning of *red* (e.g., $[\![red]\!] = $ RED, $[\![red]\!] = $ MAROON, $[\![red]\!] = $ DARK ORANGE) can be tested and rejected or accepted. Thirdly, acquired in roughly this way, FL must be capable of explicitly representing the contents of lexical rules.

Functional constraints are fundamentally unlike such *optional* and rationally acquired cognitive rules. Functional constraints specify the *fixed* representational and computational capacity of a cognitive system, i.e., the basic representations and operations used by the system. Hence functional constraints (i) are not acquired via cognitive processes (esp., via processes that can be properly modeled as inferential, or more broadly rational, responses to information), and (ii) we need not assume that they are explicitly represented by cognitive systems. A good example of a functional constraint is the putative informational encapsulation of some modular cognitive systems. A module M is not informationally encapsulated because M learned a rule which specifies that, in its computational operations, M should not use information from other cognitive systems. Rather, M's informational encapsulation is explained by a constraint on its fixed functional architecture: M is implemented in a way that blocks operations of information exchange with other cognitive systems. M's intermodular information restriction is a constraint on M's cognitive capacity; it is not something M can cognitively learn or alter.

MDCs are more like constraints on the exchange of information between some cognitive modules than like optional lexical rules such as $[\![red]\!] = $ RED. At no point in language development does it seem that speakers are trying or have to learn a general rule or principle which, like compositionality, structurally constrains the kind of information which their I-languages can use to deter-

mine the meanings of different types of complex expressions (see §6 below). Furthermore, it does not seem possible to specify a counterfactual acquisition scenario in which speakers would acquire, for cognitive/rational reasons, a different MDC. This suggests that the claim that FL or I-languages satisfy some particular MDC should be understood as a proposal about how to constraint the functional architecture of the semantics of FL.

Taking MDCs as functional constraints ties each competing proposal to a set of characteristic consequences for language acquisition. The reason for this should be clear. To specify the functional architecture of a cognitive system is to implicitly specify the system's cognitive capacity, i.e., the set of cognitive rules and algorithms which the system can represent and process. To hold that FL is constrained by a compositional MDC entails that FL is not *cognitively capable* of instantiating, hence of acquiring, I-languages with non-compositional SRs. In contrast, to hold that FL is constrained by a non-compositional MDC entails that FL is cognitively capable of acquiring I-languages with compositional *and* non-compositional SRs. These differences in the SRs they can 'see' determine the consequences for acquisition of the competing MDCs.

For example, assume that the FL_1 of speaker S_1 has compositional MDC M_1, that the FL_2 of speaker S_2 has non-compositional MDC M_2, and that S_1 and S_2 are beginning the process of acquiring I-language L, compatible with both M_1 and M_2. S_2 has to consider a hypothesis space that includes both compositional and noncompositional SRs. This difference should be manifest in at least slightly different patterns of linguistic development (e.g., in the sorts of mistakes they could make), even if S_1 and S_2 eventually converge at L. Hence even if both M_1 and M_2 can explain $P\&S$ (in the sense that all I-languages compatible with either MDC satisfy $P\&S$) and are compatible with L (in the sense that under certain conditions both speakers could eventually acquire L), we can still prefer one MDC if it predicts patters of development which better fit or explain the course of actual linguistic development. In what follows, I will argue that this is the reason why compositionality is more plausible than the non-compositional MDCs.

4 Compositionality as a MDC

This section presents the notion of compositionality I will defend. The next section present the non-compositional MDCs. To clearly state and compare the competing MDCs, I will use the following terminology:

- A 'lexical rule' is an expression of the form '$[\![x]\!] = m$', where x ranges over particular expressions, e.g., '$[\![dog]\!] = $ DOG' and '$[\![brown\ dog]\!] = $ BROWN DOG'.

- A 'semantic rule' (SR) is an expression of the form '$[\![[_Z X\ Y]]\!] = m$', where '$[_Z X\ Y]$' stands for any arbitrary *type* of syntactic structure (e.g., $[_{NP} A\ N]$), including the most general one, where Z is any branching node with {X, Y} as its immediate constituents.

Compositionality, interpreted as a MDC, amounts to the following constraint:

(CO) If L is an I-language which FL can represent, then:

1. L cannot use lexical rules to determine the meanings of complex expressions.

2. Each SR in L is of the form '$[\![_Z X\ Y]\!] = f_Z([\![X]\!], [\![Y]\!])$', were '$f_Z$' is a humanly computable function defined on the set of meanings.

Condition 1 of CO excludes all I-languages which assign meanings to syntactically complex expressions in a list-like way. To see why this condition should be part of any adequate MDC, including non-compositional ones, consider the consequences for acquisition of dropping it. FLs with MDCs without condition 1 would have to consider, for any particular complex expression, if its meaning is determined through a lexical rule. For example, assume S knows, $[\![brown]\!]$, $[\![dog]\!]$, and $[\![_{NP}A\ N]\!] = f_{NP}([\![A]\!], [\![N]\!])$. If S's FL has a MDC without condition 1, S would still have to consider (without triggering from any special feature of the learning data, e.g., repetition) whether $[\![_{NP}[_A brown][_N dog]]\!]$ is not given by any of a set of lexical rules, which yield not $[\![_{NP}[_A brown][_N dog]]\!] =$ BROWN DOG, but rather $[\![_{NP}[_A brown][_N dog]]\!] =$ ANGRY BROWN DOG or LAME BROWN DOG or any other direct meaning assignment consistent with the learning data. This cognitive 'flexibility' substantially complicates language acquisition and predicts patterns of linguistic development which we never find.[9]

Condition 2 guarantees that all I-languages compatible with CO assign meanings to complex expressions through SRs that have access only to the

---

[9]As will become clear, condition 1 is *not* the point of contention between CO and the non-compositional MDCs; but one might still object to it on the grounds that it seems incompatible with idioms. Explaining idioms is everyone's problem, but some influential recent accounts are consistent with and even support condition 1. Idioms are ambiguous expressions: they have a literal phrasal and a idiomatic meaning. The literal meaning of e.g. *kick the bucket* is KICK THE BUCKET, and its idiomatic meaning is TO DIE. There is substantial evidence that the literal meaning of idioms is automatically processed in parallel with their idiomatic meaning (Tabossi [79], Glucksberg [25]). This suggests, as predicted by condition 1, that I-languages are constrained to determine the literal meaning of complex expressions, even idiomatic ones, via SRs. To explain how the idiomatic meaning of idioms is determined, we have to make a distinction between two types of idioms, based on their syntactic flexibility. Some idioms are syntactically inflexible (except for negation) and behave like words, e.g., *by and large*. There is evidence that the idiomatic meaning of syntactically inflexible idioms is computed directly, as syntactically simple expressions (Glucksberg [25]); so their meaning is determined, consistently with condition 1, in a list-like way via lexical rules. Other idioms are syntactically flexible and behave like phrases, e.g., *spill the beans* can be used as *the terrorist didn't spill a single bean during the interrogation*, or as *John was weak, he spilled all the beans during the interrogation*. There is also evidence that the idiomatic meaning of syntactically flexible idioms is computed in the ordinary compositional way, except that their simple parts are polysemous or ambiguous and, in the idiomatic context, take on the relevant idiomatic meaning (McGlone, Glucksberg, and Cacciari [45]). The idea is that most mature English speakers know, e.g., not only that *spill* means FALL FROM CONTAINER and *beans* means EDIBLE LEGUMES, but also that in some special (idiomatic) contexts they can also mean, respectively, REVEAL and SECRETS. The assumption that parts of the idiomatic phrase correspond to parts of the idiomatic meaning explains why flexible idioms can be internally modified (Nunberg et al. [47]), as in *the investigator spilled some of the beans* or *the suspect quickly spilled all the beans*, with predictable and systematic changes to the meaning of the idiomatic phrase.

(syntactic) mode of composition of expressions and the meanings of their immediate constituents. If we hold CO, conditions 1 and 2 hold in general—i.e., of all I-languages which FL is cognitively capable of instantiating. This ensures that CO provides an adequate structural explanation of $P\&S$.

To further clarify CO let us consider its relation to the syntax and pragmatics interfaces, beginning with the former. CO is compatible with a "strongish" compositional view, in the sense of Jacobson [37]. According to this view, the syntax and the semantics work in tandem: there is no intermediate level (such as the LF of early transformational grammars) that is first built from surface structures—using syntactic operations that have no corresponding semantic operations—and serves as input to the semantics. All the CO compatible solutions to problematic expressions that I discuss later in the paper respect strongish compositionally, but my arguments for CO are compatible with weaker views on Jacobson's scale, e.g., a view according to which there are some syntactic operations (with no corresponding semantic operations) that create LF structures from surface structures. As will become clear, the debate about whether we can hold on to strongish compositionality (which depends on issues like whether we need quantifier/auxiliary raising rules), is independent of the debate between compositional and non-compositional MDCs as I frame it here, and would still arise, mutatis mutandis, even if we hold a non-compositional MDC.

Another important issue at the syntax interface concerns the relation between syntactic rules and types of SRs. Some Montague-style theories use particular phrase-structure rules such as $S \rightarrow NP\ VP$ and pair them with construction-specific compositional SR such as (1):

(1) $[\![_S NP\ VP]\!] = f_S([\![NP]\!], [\![VP]\!])$

where 'f_S' is a function which given $[\![NP]\!]$ and $[\![VP]\!]$, outputs $[\![S]\!]$. As stated, CO is compatible with those views; but we will make a stronger assumption, namely, that CO requires general, not construction-specific, SRs. So we will assume that, in the formulation of CO above, Z stands for any branching node with $\{X, Y\}$ as its immediate constituents. This is in any case how we would have to interpret CO if it is paired with a syntactic theory, such as Minimalism, that does *not* have category-specific phrase-structure rules. A famous theory along these lines is the type-driven theory presented in Heim and Kratzer [34]. Type-driven theories do not require category-specific syntactic phrase-structure rules. Heim and Kratzer assume that the syntax delivers to the semantics bare-phrase structures. Translated into our terminology, this means that the semantics sees only the most general type of syntactic structure, a branching node and its immediate constituents. Given this assumption about what the syntax delivers to the semantics, CO entails that the SRs have to be general, ranging over all (types of) complex expressions (including NPs, VPs, Ss, etc.). An example of a general SRs is Functional Application (FA):

(FA) If α is a branching node, $\{\beta, \gamma\}$ is the set of α's daughters, and $[\![\beta]\!]$ is a function whose domain contains $[\![\gamma]\!]$, then $[\![\alpha]\!] = [\![\beta]\!]([\![\gamma]\!])$

10

However, CO is compatible with various accounts of how the compositional operations work, i.e., of the nature of the general SRs. According to recent Neo-Davidsonian accounts, composition is a uniform operation such as predicate conjunction over monadic concepts (Pietroski [55]). Other linguists, closer to Montague's original framework, use rules such as FA, predicate modification, and various type-shifting rules (Jacobson [37], Heim and Kratzer [34]). The important point, for our purposes, is that we could state the whole dialectic between compositional and non-compositional MDCs by assuming either view.[10]

Consider now the interface with pragmatics. CO is compatible with at least two kinds of context-sensitivity. Firstly, CO allows the meaning of some, most, or all lexical items to be characters. We can represent this by saying that, for any expression e:

- $[\![e]\!]_c = f_e(c)$

where f_e is the character of e and $f_e(c)$ is the occasion meaning of e in c. If e has no free parameters, then for all c's, $f_e(c) = m$, where m is the standing meaning of e. Secondly, CO allows SRs to take the modulated (instead of the standing or occasion) meanings of the immediate constituents of complex expressions. Following Recanati [67], we can represent the modulated meanings of an expression e, $[\![e]\!]_{M,c}$, as follows:

- $[\![e]\!]_{M,c} = mod(e,c)([\![e]\!]_c)$

mod takes as an argument an expression e and context c in which e occurs and returns as value the modulation function $f_{M,e}$, which takes $[\![e]\!]_c$ and returns the meaning that is salient/relevant/appropriate for e in c.

There are two main ways of implementing the context-sensitive mod function to get general SRs that determine the meanings of complex expressions in terms of the modulated instead of the standing or occasion meanings of their constituents. On an unconstrained view, mod is generalized to apply at every level of interpretation. On a constrained view, which is the one we will adopt here, mod applies only on lexical items. To illustrate, let us implement this constrained version of mod in a type-driven framework. Focusing again on FA, our interpretation should be formulated as follows (assume for brevity that all non-branching nodes are terminal nodes):

(TN_M) If α is a terminal node, then $[\![\alpha]\!]_{M,c} = mod(\alpha,c)([\![\alpha]\!]_c)$, where $[\![\alpha]\!]_c$ is specified in the lexicon.

(FA_M) If α is a branching node, $\{\beta,\gamma\}$ is the set of α's daughters, and $[\![\beta]\!]_{M,c_1}$ is a function whose domain contains $[\![\gamma]\!]_{M,c_2}$, then $[\![\alpha]\!]_{M,c} = [\![\beta]\!]_{M,c_1}([\![\gamma]\!]_{M,c_2})$

---

[10]This is not to deny, of course, that the outcome of this debate affects the interpretation of CO. For example, neo-Davidsonians often avoid type-shifting rules by positing covert syntactic elements (see e.g. Pietroski's account of proper names in [55]); so assuming that composition is predicate conjunction might entail that we abandon strongish compositionality.

On this account, mod does not operate on the outputs of FA_M (or other rules for interpreting the meaning of complex expressions), but only on terminal nodes/lexical items. This allows a constrained form of meaning modulation. Since the compositional step (i.e., the determination of the meaning of complex expressions), in this sort of framework, corresponds to the FA_M rule, we can say that meaning modulation is pre-compositional.

Pragmatic processes also modify the outputs of the semantics, but there is no good reason to model post-compositional pragmatic processes as part of FL. Many linguists now think that syntactic/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 [11, 12]), the phases that are sent out for pragmatic processing are more fine grained (Cook and Newson [14], Radford [65]). Theorists who adopt Categorical Grammars also usually assume that the outputs to pragmatics are sub-sentential phrases (Jacobson [37]). If interpretation proceeds in such phases, which are inputs to (primary) pragmatic processes, then there is no reason why we should incorporate into the semantics a generalized version of mod, i.e., a function which modulates both the inputs and outputs of the compositional operations. Such output modulations would be redundant; indeed, in actual case studies (as in the CO compatible accounts we present below), most of the modulation operations operate on lexical items.

To conclude, let me emphasize the most important consequences, for this discussion, of interpreting CO as excluding construction-specific SRs. Firstly, assuming that FL can only represent general SRs has the advantage of substantially diminishing the amount of rules which we have to assume speakers either acquire or innately possess. Secondly, that assumption also coheres nicely with an important current trend in generative linguistics, namely, to generalize or eliminate phrase-structure rules (see Chomsky [8]; Heim and Kratzer [34]). But most importantly, assuming that FL cannot represent construction-specific SRs has distinctive empirical consequences for language acquisition and development. For example, Heim and Kratzer's type-driven theory entails that speakers need not acquire or innately posses a construction-specific SR like (1) for each type of syntactic construction. If speakers know, about a complex expression, (i) the meanings of its parts and (ii) its structure, this theory predicts that they should have the linguistic competence to adequately determine its meaning. In other words, once they know (i) and (ii), there is no space for speakers to make a mistake that leads to an incorrect understanding of a complex expression. Non-compositional theories make different predictions about the sorts of mistakes speakers can make. For as we will now see, each non-compositional MDC has to assume that speakers can acquire construction-specific SRs. This entails that, despite knowing (i) and (ii), speakers could, early in development, systematically assign incorrect meanings to tokens of certain types of complex expressions, for they could assign an incorrect construction-specific SR to any type of complex expression.

5 CO vs Non-compositional MDCs

The claim we will defend is that assuming a MDC approximately like CO is currently the best explanation of $P\&S$. CO seems correct insofar as it requires that the meaning of complex expressions be determined via SRs. But we might suspect that CO is too restrictive insofar as it requires that SRs determine the meaning of complex expressions *only* from information derived from their immediate constituents and their structure.

Szabo [78] eloquently elaborates this suspicion. The fact that S, an arbitrary competent speaker of English, understands some token of a novel complex expression e shows only that the information necessary to determine e's meaning is available to S in the context and information state in which S processed e. Part of the information S has in this state is information about e's structure and about the meaning of its constituents. But as Szabo reminds us, S also has access to other information—e.g., other linguistic information, general features of the context and certain general beliefs. This information may partly determine the meaning of tokens of e and other complex expressions, in which case structural and constituent information is not generally sufficient to determine the meaning of complex expressions. Since the opposite is assumed by CO, we might conclude that we should replace CO with a 'weaker' MDC compatible with the possibility that there is a set of non-constituent-derived but generally available information which partly determines the meaning of certain types of complex expressions.

This suspicion against CO derives most of its initial plausibility from its generality. To show this, we will now examine *particular* proposals for types of information that could, via non-compositional SRs, partly determine the meaning of certain types of complex expressions. The non-compositional proposals most commonly presented appeal to certain types of (i) contextual information and (ii) general beliefs. From our perspective, the often overlooked point to note is that to allow information of type (i)-(ii) to partly determine, via non-compositional SRs, the meanings of certain types of complex expressions, we have to assume, in each case, that the MDC of FL weakens condition 2 of CO to allow the desired non-compositional SRs.[11] However, in doing that each non-compositional MDC is also made compatible with many other (unintended) SRs. As a result, unlike CO, each non-compositional MDC predicts patterns of early linguistic development that seem obviously incorrect.

5.1 Non-compositional MDCs which use contextual information

The first proposal we will consider is to replace CO with a MDC which allows non-constituent contextual information to partly determine the meaning

---

[11]There are various reasons why this is often overlooked. One is that critics often focus only on a particular type of complex expression, and on SRs for that type of expression, and fail to consider the *general* consequences of adopting a non-compositional MDC that is weak enough to permit the particular SR they are considering.

of (certain types of) complex expressions. The motivation for adopting this non-compositional MDC is that there seem to be complex expressions with 'unarticulated semantic constituents': their meaning is determined by the meanings of their parts, their structure, and certain contextual information which is not the meaning of any of its constituents.

The paradigmatic examples of expressions with 'unarticulated constituents' are simple 'meteorological expressions' like:

(2)   It is raining.

In most contexts, tokens of (2) seem to express the proposition that it is raining at a certain time and place. One can hold that the relevant time is signified by the tense of the auxiliary verb, so that the time is represented at LF. But it seems that no constituent of (2) signifies or indexically encodes the relevant place. At the same time, it seems that competent speakers who understand the parts of (2) can understand (2), i.e., it seems that competent speakers can productively and systematically arrive at these interpretations. This suggests that: (i) the proposition expressed by (2) includes information of a location, and (ii) this information is not determined by either the structure or the constituents of (2). According to this account, then, the meaning of simple meteorological expressions like (2) is not determined compositionally.

Note that unlike the unarticulated constituency account just sketched, most accounts of the meaning of meteorological expressions respect CO. For example, Borg [6] argues that at LF (2) has a time but not a location variable; but she denies that the proposition literally expressed by (2) has a location specification. Recanati [66] defends an account similar to Borg's: (2) can be used to assert a proposition that is indefinite with respect to location, which suggests that location definite uses of (2) involve primary pragmatic enrichments.

Since there are CO compatible accounts of simple meteorological expressions, why replace CO with a MDC that allows non-compositional accounts of meteorological expressions? The issue in this discussion is not whether there are accounts of meteorological expressions compatible with CO; the issue is whether we should take CO as the MDC of FL. We might question this if adopting CO *forces us*, a priori, to dismiss otherwise plausible accounts of the meaning of certain types of complex expressions, such as the unarticulated constituency account of simple meteorological expressions. It *seems* preferable to adopt a MDC which allows both compositional and non-compositional SRs.

Basically for these reasons, Jonsson [39] argues that we could replace CO with LOC, which we here reformulate as a MDC:

(LOC) If L is an I-language which FL can represent, then:

1. L cannot use lexical rules to determine the meanings of complex expressions.

2. Each SR in L is of form (a) or (b):
   (a) '$[\![_Z X\ Y]\!] = f_Z([\![X]\!], [\![Y]\!])$', where '$f_Z$' stands for a humanly computable function defined on the set of meanings

(b) '$[[_Z X\ Y]] = f_{Z'}([\![X]\!], [\![Y]\!], g)$' where '$g$' stands for a location function (functions from contexts to places) and '$f_{Z'}$' stands for a humanly computable function defined on the set of meanings and location functions.

LOC is weaker than CO in the sense that condition 2 allows, on the right-hand side of SRs, reference to location functions which are not the meaning of a constituent of the complex expressions whose meaning they determine.[12] This opens space for unarticulated constituency accounts of simple meteorological expressions, via construction-specific SRs. To illustrate, assume that (2) has the following simple structure:[13]

(2) $[_S[_E\ It][_{VP}[_{Aux}\ is][_V\ raining]]]$

The unarticulated location function could be introduced, via a construction-specific non-compositional SR, at the level of the S or VP. For our purposes this choice does not matter, but assume it is introduced at the level of the VP:

(3) $[\![[_{VP} Aux\ V]]\!] = f_{VP'}([\![Aux]\!], [\![V]\!], g_1)$

Rule (3) contains the function g_1, a location function which is not the meaning of any of the constituents of the left-hand side of the rule.

As Jonsson [39] argues, adopting LOC does not affect the explanation of $P\&S$. Firstly, LOC, like CO, prohibits lexical rules to determine the meaning of complex expressions. Secondly, the sorts of unarticulated meanings which LOC allows—i.e., location functions from contexts to places—are constituted by information which speakers generally have access to, and there is no reason to deny that FL can access this kind of contextual information. For these reasons, we might be tempted to conclude that the non-compositional LOC is a better choice of MDC than CO.

However, LOC entails that speakers have to face certain choices in language acquisition that, judging from the general patterns of early linguistic development, speakers never seem to face.

To see that, note, first, that if we assume LOC, then to generate English (using non-compositional SRs) speakers would have to acquire an I-language with construction-specific SRs such as (3), and not one with general SRs such as FA. For only in the case of *some types* of complex expressions—e.g., meteorological expressions—is it plausible to assume that unarticulated location functions partly determine their meanings. For example, the meanings of most NPs of the form $[_{NP} A\ N]$—*black cat*, *angry cow*, *pretty dolphin*, etc.—do

---

[12]As stated, LOC is logically weaker than CO: every I-language compatible with CO is compatible with LOC but not vice-versa. However, I do not emphasize this because we can impose additional constraints on LOC (some of which we will discuss below) which entail that there is no logical strength ordering between LOC and CO. For our purposes what is crucial is only that LOC has to weaken condition 2 of CO to allow the desired type of non-compositional SRs.

[13]This structure is obviously not the one that would be assigned by a serious syntactic theory. For a more realistic structure, see footnote 14. However, none of the points I will make depend on the particular structure assigned to simple meteorological expressions.

not include an unarticulated constituent that, given a context, determines a location. The same is true of most Ss, e.g., *John is thinking, Einstein's idea is fantastic, Empiricism is dead*, etc. Indeed, even construction-specific SRs like (3) would need to be reformulated in terms of more fine-grained syntactic categories, for (3) incorrectly assigns a location specification to *all* tensed VPs with the syntactic structure $[_{VP}\ Aux\ V]$, e.g., *is thinking* and *is happy*.[14] So if we accept LOC, speakers would have to either acquire or innately posses (fine-grained) construction-specific SRs.

Consider the first option, that construction-specific SRs (a fortiori, phrase-structure rules) are innate. If we take this option, we would have to attribute *substantially* more innate knowledge to speakers than if we adopted CO, while gaining no descriptive coverage. In any case, this option is empirically implausible. Assuming a syntax that uses phrase-structure rules, there is substantial cross-linguistic evidence that at least some of these rules have to be acquired (Roeper [69]). For example, compounds are recursive in Germanic languages but not in Romance languages. Possessives are recursive in English but not in German. Prenominal As are recursive in English but not in French, and the opposite holds for post-nominal As. There are plenty of other examples like this. If phrase-structure rules have to be acquired, then the construction-specific SRs for such rules cannot plausibly be innate. So assuming that construction-specific SRs are innate is not a viable option for defenders of LOC.

The other option is to assume that construction-specific SRs are learned or acquired. In itself, this is not a problem. Assume that S knows (i) that '$\gamma\ \beta$' is an expression of the form $[_{VP}Aux\ V]$, and (ii) $[\![\gamma\ \beta]\!]$, $[\![\gamma]\!]$, and $[\![\beta]\!]$. For S to acquire a rule like (3) from (i) and (ii), we have to assume that S is able to use information (i) and (ii) to test hypotheses like (6) and (7):

(6) $[\![\gamma\ \beta]\!] = f_{VP'}([\![\gamma]\!], [\![\beta]\!], g_1)$

---

[14]This point does not depend on assuming that simple meteorological expressions have that structure: regardless of the particular syntactic structure we assign to simple meteorological expressions, there are other expressions which are syntactically identical *at the level of structure where the non-compositional rule applies*, but which do not have a location specification. For example, assume that simple meteorological expressions have the syntactic structure they are assigned in most P&P syntactic theories: $[TP\ [[PRN]\ [\bar{T}\ [T\ VP]]]]$ (see Radford [65]). The non-constituent contextual function g could be introduced at the level of the complex TP, via an SR for $[_{TP}PRN\ \bar{T}]$, or at the level of the complex $\bar{T}$, via an SR for $[_{\bar{T}}T\ VP]$. As a result, expressions like *he is thinking, he is depressed, it is said that life is short, it is sad to think about suffering* would be incorrectly assigned a location specification.

To avoid this problem of implausibly over-saturating all sorts of complex expressions with unarticulated location specifications, the SRs acquired by LOC constrained FLs have to use construction-specific SRs with fine-grained syntactic categories. For example, to give an unarticulated constituency account of expressions like *it is raining* in a way that doesn't over-saturate with location specifications the meanings of similarly structured sentences, we can include a new category for a subset of verbs, call it V_e. Assume V_e includes verbs of physical events like *snows* and *rains*, but not verbs of mental events or states like *loves* and *thinks*. A model which includes fine-grained category V_e/VP_e can include fine-grained construction-specific SRs like (4)-(5):

(4) $[\![[_{VP_e}Aux\ V_e]]\!] = f_{VP'}([\![V_e]\!], [\![Aux]\!], g_1)$

(5) $[\![[_{\bar{T}_e}\ T\ VP_e]]\!] = f_{\bar{T}_e}([\![T]\!], [\![VP_e]\!], g_1)$

(7) $[\![\gamma \ \beta]\!] = f_{VP}([\![\gamma]\!], [\![\beta]\!])$

If (7) fits the data, S assumes that $[\![\gamma \ \beta]\!]$ is determined from its parts using compositional rule (8):

(8) $[\![_{VP}Aux \ V]\!] = f_{VP}([\![Aux]\!], [\![V]\!])$

If (6) fits the data, S assumes that $[\![\gamma \ \beta]\!]$ is determined from its parts using non-compositional rule (3), here repeated:

(3) $[\![_{VP}Aux \ V]\!] = f_{VP'}([\![Aux]\!], [\![V]\!], g_1)$

Suppose that S concludes that hypothesis (6) fits the data better than hypothesis (7), then S can generalize to all structures of the form $[_{VP}Aux \ V]$ and so acquire non-compositional SR (3).

Note, however, that if we replace CO with LOC analogous learning procedures enter each case of acquiring a construction-specific SR. For unlike CO, LOC allows the *general* possibility that construction-specific non-compositional SRs determine the meaning of *every syntactic type* of complex expression. Assuming only LOC, each time S acquires a construction-specific SR, S would be open to consider compositional and non-compositional alternatives. For example, assume that (9)-(10) are correct SRs, in the sense that they output the correct meaning-assignments, relative to English. En route to acquiring (9)-(10), S is cognitively 'open' to consider a hypotheses space like (9*)-(10*), where each g_i stands for a location function:

(9) $[\![_{S}NP \ VP]\!] = f_S([\![NP]\!], [\![VP]\!])$

(10) $[\![_{NP}AP \ NP]\!] = f_{NP}([\![AP]\!], [\![NP]\!])$

(9*) $[\![_{S}NP \ VP]\!] = f_S([\![NP]\!], [\![VP]\!])$ or $f_{S'}([\![NP]\!], [\![VP]\!], g_1/g_2/.../g_n)$

(10*) $[\![_{NP}AP \ NP]\!] = f_{NP}([\![AP]\!], [\![NP]\!])$ or $f_{NP'}([\![AP]\!], [\![NP]\!], g_1/g_2/.../g_n)$

This entails that, at some early stage in the acquisition of I-English, S *could* make mistakes like (11)-(13), even if eventually S ends up acquiring the correct SRs (9)-(10) (in the examples below, assume that S knows the relevant syntax and the correct meanings for the lexical items, relative to I-English).

(11) $[\![_{S_{ms}}NP \ VP_{ms}]\!] = f_{S'}([\![NP]\!], [\![VP_{ms}]\!], g_1/.../g_n)$, where '$VP_{ms}$' is a subcategory of VPs headed by Vs of mental states. So S takes $[\![John \ is \ happy]\!]$ to be JOHN IS HAPPY HERE/CLOSE TO HOME/EVERYWHERE/... (note that LOC allows location functions other than g_1), $[\![Mary \ is \ sad]\!]$ to be MARY IS SAD HERE/ CLOSE TO HOME/EVERYWHERE/... and so on.

(12) $[\![_{S_{mt}}NP \ VP_{mt}]\!] = f_{S'}([\![NP]\!], [\![VP_{mt}]\!], g_1/.../g_n)$, where '$VP_{mt}$' is a subcategory of VPs headed by Vs of mental traits. So S takes $[\![John \ is \ silly]\!]$ to be JOHN IS SILLY HERE/CLOSE TO HOME/EVERYWHERE/..., $[\![Mary \ is \ courageous]\!]$ to be MARY IS COURAGEOUS HERE/CLOSE TO HOME/EVERYWHERE/..., and so on.

17

(13) $[\![_{NP_c}AP_c \ NP)]\!] = f_{NP'}([\![AP_c]\!], [\![NP]\!], g_{10}/.../g_n)$, where '$AP_c$' is a subcategory of APs headed by color As. So S takes $[\![gray \ shark]\!]$ to be GRAY SHARK WHEN UNDER WATER/WHEN I SEE IT/WHEN OUTSIDE WATER/..., $[\![red \ fish]\!]$ to be RED FISH WHEN UNDER WATER/WHEN I SEE IT/WHEN OUTSIDE WATER/..., and so on.

The meaning assignments in (11)-(13) are consistent with the learning-data encountered by most speakers during early language acquisition. For example, S can reasonably assume that color As, when combined with common Ns, result in color attributions that have location restrictions, resulting in cases like (13). The non-obvious mistakes in (13) would take time to correct, since the location-restricted assertion-conditions of each color A and common N complex expression is a commonly used subset of their (mature English) non-location restricted assertion-conditions. However, speakers do not make mistakes about the meaning of color A and common N compounds analogous to those presented in (13). Similarly, S can reasonably assume that expressions which attribute mental states or traits are location restricted, like (11)-(12). The non-obvious mistakes in (11)-(12) would also take time to correct, since the location-restricted assertion-conditions of each expression attributing a mental state is a commonly used subset of their (English) non-location restricted assertion conditions. However, speakers do not make mistakes about the meanings of expressions attributing mental states or traits analogous to (11)-(12).

That speakers, even early in development, never go through states like (11)-(13) suggests that they never consider a hypothesis space like (9*)-(10*), i.e., a hypothesis space which includes, for each type of complex expression, a set of possible non-compositional SRs, each involving some unarticulated location function. However, this is the hypothesis space that would be open to speakers if their MDC was LOC. So either we reject LOC, or it is a mystery why speakers acquiring I-English never adopt or even consider 'reasonable but mistaken' construction-specific non-compositional SRs like those presented in (11)-(13).

To sum-up, although LOC does not undermine the explanation of $P\&S$, it seems to entail false predictions about patterns of early linguistic development. LOC entails that language acquisition partly consists in acquiring phrase-structure rules and construction-specific SRs. Each time speakers acquire a construction-specific SR, they would be open to consider a whole set of competing non-compositional SRs. Given the type of information encountered in early language acquisition scenarios, the fact that speakers do not adopt, at least temporarily, some reasonable but 'incorrect' non-compositional SRs is left completely unexplained. In addition, replacing CO with LOC buys our linguistic theories no additional descriptive coverage. For these reasons, CO is a more plausible MDC than the weaker, non-compositional LOC.

5.2 Non-compositional MDCs which use general beliefs

Another recent and more influential proposal is to replace CO with a MDC that allows general beliefs to partly determine the meanings I-languages assign to

(certain) types of complex expressions (Prinz [59], Jonsson and Hampton [33]). One reason to propose such a MDC is that there are complex expressions that seem to have semantic features neither present in nor determined by the semantic features of their constituents. For example, consider complex NPs such as *black cat* and *brown cow*. For some speakers, $[\![black\ cat]\!]$ seems to include the information that their presence brings bad luck; and for others $[\![brown\ cow]\!]$ seems to include the information that brown cows produce bad milk. To linguistically account for the 'free-enrichment' of complex NPs, some theorists propose that we allow general beliefs about the extension of complex NPs—usually called 'extensional feedback' beliefs—to partly determine their meaning, even if such beliefs are not part of the meaning of their constituents (Murphy [46], Prinz [58, 59], Hampton [31], Jonsson and Hampton [33]). There are other types of 'free-enrichments' of complex expressions, including NPs not plausibly modeled as computations involving the use of extensional feedback. For now we will focus on the extensional feedback class.

There are accounts of 'free-enrichment' complex NPs compatible with CO. One proposal assumes that the meaning of tokens of common Ns is enriched on-line to include some (relevant) encyclopedic information (Barsalou [3]; Carston and Wilson [83]). According to this view, in some contexts, $[\![cow]\!]$ is enriched to include information like 'produces bad milk if brown', $[\![cat]\!]$ to include information like 'brings bad luck if black', etc. There is no a priori reason to reject the claim that the meaning of some lexical items can be enriched online to include such information. Another proposal is to hold that the free-enrichments of the meanings of complex NPs are post-linguistic pragmatic enrichments, even if they are often sub-personal and automatic. The intuitive meanings of many kinds of expressions are affected by pragmatic enrichments, which are often sub-personal and automatic (Recanati [66]; Glucksberg [26]). There is no a priori reason to reject the claim that such post-compositional processes account for the 'intuitive' meanings of free-enrichment complex NPs.

Since there are CO compatible accounts of free enrichment complex NPs, why replace CO with a MDC that allows non-compositional accounts of the meanings of these complex NPs? Again, the issue in this discussion is *not* whether there are CO compatible accounts of free-enrichment complex NPs; the issue is whether we should take CO as the MDC of FL. We might question this if adopting CO *forces* us, a priori, to dismiss apparently reasonable accounts of the meaning of certain types of complex expressions, such as the extensional feedback account of free-enrichment complex NPs.

To allow non-compositional accounts of free-enrichment complex NPs, Jonsson [39] proposes GEN, which we here reformulate as a MDC:

(GEN) If L is an I-language which FL can represent, then:

1. L cannot use lexical rules to determine the meanings of complex expressions.

2. Each SR in L is of form (a) or (b):

(a) '$[[_Z X\ Y]] = f_Z([\![X]\!], [\![Y]\!])$', where '$f_Z$' stands for a humanly computable function defined on the set of meanings

(b) '$[[_Z X\ Y]] = f_{Z+}([\![X]\!], [\![Y]\!], b)$' where '$b$' stands for a set of general beliefs and 'f_{Z+}' stands for a humanly computable function defined on the set of meanings and general beliefs.

GEN is weaker than CO in the sense that it allows, to partly determine the meaning of complex expressions, general beliefs which are not part of the meaning of their constituents.[15] This opens space for non-compositional accounts of free-enrichment complex NPs, including extensional feedback accounts, via non-compositional SRs:

(14) $\quad [[_{NP} A\ N]] = f_{ex}(f_{NP}([\![A]\!], [\![N]\!]), b)$

(14) refers to the extensional feedback belief set b, which is not the meaning of any of the constituents of the complex expressions whose meaning it partly determines. According to (14), the meaning of complex NPs is a function f_{ex} from the value of the ordinary compositional function f_{NP} for NPs, which applies to the meaning of its constituents, and from the extensional-feedback belief set b. To illustrate f_{ex}, assume that b stands for S's extensional feedback, which includes the belief that black cats bring bad luck:

(15) $\quad f_{ex}(f_{NP}([\![black]\!], [\![cat]\!]), b) = $ BLACK CAT AND BRINGS BAD LUCK

What f_{ex} does is to incorporate into the meaning of complex NPs, as compositionally determined, whatever beliefs about the NP there are in the extensional feedback belief set.

As Jonsson [39] argues, adopting GEN does not seem to affect the explanation of $P\&S$. Firstly, GEN, like CO, prohibits lexical rules for complex expressions. Secondly, the beliefs which GEN allows to partly determine the meaning of free-enrichment complex NPs are beliefs which speakers have access to, at least at the personal level. If we assume that FL also has access to these general beliefs—i.e., if we assume that FL is *not* an informationally encapsulated module—I-languages constrained by GEN would satisfy $P\&S$.

Why should we stick with CO, which blocks the possibility that speakers can acquire I-languages with some non-compositional SRs like (14) to deal with extensional feedback complex NPs, when we can, apparently without violating $P\&S$, adopt a weaker non-compositional MDC which allows that possibility? The problem faced by GEN is quite similar to the problem faced by LOC.

Note, first, that SRs like (14) are actually problematic, relative to English. If S's I-language has (14), it follows that, if α is an expression of the form $[_{NP} A\ N]$, its meaning is partly determined by *all* of S's extensional feedback beliefs about

---

[15]As stated, GEN is logically weaker than CO: all I-languages compatible with CO are also compatible with GEN, but not vice-versa. However, what I said with respect to LOC also applies in this case: what is crucial, for our purposes, is only that GEN has to weaken condition 2 of CO to allow for the desired type of non-compositional SRs. This is compatible with there being some additional constraints on GEN (some of which we discuss below) that entail that there is no strict logical strength ordering between GEN and CO.

the referent of α. For example, if S believes not only that black cats bring bad luck, but also that cats are felines and felines are never underwater, that cats are animals and that animals are robots controlled by Martians, then:

(16) $f_{ex}(f_{NP}(\llbracket black \rrbracket, \llbracket cat \rrbracket), b) =$ BLACK CAT AND BRINGS BAD LUCK AND IS NEVER UNDER WATER AND IS A ROBOT CONTROLLED BY MARTIANS

However, whatever peculiar beliefs about cats, black cats, black animals, colored animals, etc., S has, not *all* of them are part of $\llbracket black\ cats \rrbracket$, as determined by S's I-language. Some theorists have defended the idea that beliefs about the referent of an arbitrary complex expression can, in principle, affect the meaning speakers assign to it; but with the exception of radical holists, no one defends the view that *all* such beliefs affect the meaning speakers assign to it, as (14) entails.[16] Most theorists who assume that extensional feedback beliefs about the entity denoted by a complex expression partly determine its meanings implicitly assume that only *some* of those beliefs play such a role. Of course, GEN is *compatible* with constrained versions of (14), such as (14+):

(14+) $\llbracket [_{NP}A\ N] \rrbracket = f_{ex}(f_{NP}(\llbracket A \rrbracket, \llbracket N \rrbracket), b^+)$

(14+) is just like (14) except that the extensional belief set b^+ is a subset of the extensional belief set b. What subset? One option, assuming that we can represent something like degrees of belief, is that b^+ only includes extensional feedback beliefs that pass a threshold. In this way, not *all* of S's extensional feedback beliefs about e.g. black cats would be included in the meaning which S's I-language assigns to *black cat*, but only the ones that S 'really' believes.

Still, although rule (14+) might work for cases like *black cat* and *brown cow*, it is incorrect for English. For example, take *gray shark* and *fierce lion*, which have the form $[_{NP}A\ N]$, so that (14+) applies to both of them. Suppose again that at some early stage of linguistic development, S 'really' believes that lions are only fierce in their territory, and that color attributions to fish are restricted to the way they look when underwater. In this case, rule (14+) would have the result that $\llbracket fierce\ lion \rrbracket$ is something like FIERCE LION IN HIS TERRITORY, and that $\llbracket gray\ shark \rrbracket$ is something like GRAY SHARK WHEN UNDERWATER. In short, just as in the case of LOC, if we adopt GEN, we have to assume that speakers acquire construction-specific SRs which can use fine-grained syntactic categories, such as the following:

(17) $\llbracket [_{NP_a}A_c\ N_t] \rrbracket = f_{ex}(f_{NP_a}(\llbracket A_c \rrbracket, \llbracket N_t \rrbracket), b^+)$

---

[16]The most famous holist about linguistic meaning is probably Block [4]. However, I agree with Block [5] that holistic inferential role theories of meaning are compositional. The reason for this is simple. According to these theories, the meaning of an expression is given by all of its inferential roles. Hence extensional feedback beliefs such as that black cats bring bad luck are part of $\llbracket cat \rrbracket$. From this perspective, to hold that extensional feedback beliefs are incorporated into the meanings of tokens of complex expressions via non-compositional SRs would be entirely superfluous, since they are already part of the meaning of the constituents. For further discussion, see Szabo's [75] response to Fodor and Lepore's [23] claim that total inferential roles are not compositional.

'NP_a' is a category of NPs formed out of color As and terrestrial animate beings. In this case, (17) applies to *brown cow* and *black cat*, but not to *fierce lion* and *gray shark*.

However, if we adopt GEN, S could, early in linguistic development, test 'mistaken' construction-specific SRs like (14) and (14+) for at least some types of complex expressions, in cases in which they output reasonable but incorrect meaning assignments. Consider the following examples (assume that S knows the relevant syntax and the correct meanings for the lexical items, relative to I-English):

(18) $[\![_{S_m} NP\ VP_m]\!] = f_{S_m}(f_S([\![NP]\!], [\![VP_m]\!]), b^+)$, where '$b^+$' stands for the set of highly weighted extensional feedback beliefs, and 'S_m' is a category of Ss formed by a NP and a VP_m which predicates some mental trait. In this case, if S assumes that attributions of mental traits are restricted to certain locations, S would take $[\![John\ is\ silly]\!]$ to be JOHN IS SILLY IN HIS HOUSE (or some other reasonable location restriction), and so on.

(19) $[\![_{NP_m} A_m\ N_{an}]\!] = f_{NP_m}(f_{NP}([\![A_m]\!], [\![N_{an}]\!]), b^+)$, where '$NP_m$' is a category of NPs formed by mental trait As and Ns that stand for animate beings. In this case, if S believes that certain mental traits of animate objects are restricted to certain locations, S would take $[\![fierce\ lion]\!]$ to be something like FIERCE LION WHEN IN HIS TERRITORY (or some other location restriction), $[\![silly\ student]\!]$ to be SILLY STUDENT WHEN IN HIS SCHOOL (or some other reasonable location restriction), and so on.

(20) $[\![_{NP_p} A_m\ N_{an}]\!] = f_{NP_p}(f_{NP}([\![A_m]\!], [\![N_{an}]\!]), b^+)$, where '$NP_p$' is a category of NPs formed by physical trait As and Ns that stand for animate beings. In this case, if S believes that certain physical traits of animate objects are strongly correlated with certain mental traits, S could take $[\![strong\ cat]\!]$ to be something like STRONG AND MEAN CAT, $[\![strong\ student]\!]$ to be something like STRONG AND BULLYING STUDENT, and so on.

Again, we do not find patterns of early linguistic development in which speakers test reasonable but incorrect SRs such as (18)-(20). This suggests that FL is not as unconstrained as it would be if GEN was its MDC. Note, in addition, that GEN faces basically the same problems faced by LOC. For if assuming GEN, speakers would have to test and eliminate reasonable but incorrect construction-specific SRs which have the same output as those we considered in (11)-(13), which we used to criticize LOC. To illustrate this, consider (12) again:

(12) $[\![_{S_{mt}} NP\ VP_{mt}]\!] = f_{S'}([\![NP]\!], [\![VP_{mt}]\!], g_1/.../g_n)$, where '$VP_{mt}$' is a subcategory of VPs headed by Vs of mental traits. So S takes $[\![John\ is\ silly]\!]$ to be JOHN IS SILLY HERE/CLOSE TO HOME/EVERYWHERE/..., $[\![Mary\ is\ courageous]\!]$ to be MARY IS COURAGEOUS HERE/CLOSE TO HOME/EVERYWHERE/..., and so on.

GEN does not allow SRs like the one used in (12), but it allows SRs that have, under similar conditions, very similar outputs, as is illustrated by (18) and (19). As we argued before, if the MDC of FL doesn't exclude these options, it would be 'reasonable', early in linguistic development, for S to *believe* that assertions using certain types of complex expressions have location restrictions, as in (11)-(13), even if these beliefs are eventually abandoned.

So GEN inherits most of the problems faced by LOC, and introduces some of its own. If we hold that GEN is the MDC, we have to explain how speakers acquire rules closer to (17) than to (14) or (14+), how they acquire something like (9)-(10) from an initial hypotheses space that is much wider than (9*)-(10*), and so on. But even if that can be explained, the crucial point is that we should find, early in linguistic development, mistakes that reveal the use of construction-specific SRs like (18)-(20), which result in various kinds of reasonable but incorrect free-enrichments of certain types of complex expressions.

This objection to GEN is important because *many* critics of compositionality hold that what we should infer from $P\&S$ is only that FL respects some weak constraint along the lines of GEN—a constraint which allows the meaning of complex expressions to be compositionally determined, but also allows the meaning of some types of complex expressions to be partly determined by general beliefs (Murphy [46], Prinz [58,59], Hampton and Jonsson [33]). But if we change our perspective and take MDCs not as convenient methodological assumptions but as empirical hypotheses about the functional architecture of FL, the problematic consequences for language acquisition of weak constraints such as GEN are intuitively easy to see.

Suppose that, using any general learning strategies at your disposal, you are given the task of acquiring the semantics (S_T) of a target I-language (I_T). To do this, you are given subsets of the language (L_T) generated by I_T and some hints about what S_T *cannot* be like. Suppose that the only 'hint' about S_T—the only hint about the form of the SRs—that you are given is that it satisfies something like GEN. This amounts to the following hint: for a complex expression e of any type generated by I_T, $[\![e]\!]$ is determined by the meanings and structure of its immediate parts, and possibly *any* other general beliefs which are consistent with the data. Given only this hint, you would begin the task of acquiring the SRs of I_T with a substantially unconstrained hypothesis-space. Even if you eventually acquire S_T, you would have to consider and reject many reasonable but 'incorrect' construction-specific SRs such as (18)-(20). You would have to learn, as you encounter more L_T data, that only some construction-specific SRs that involve extensional feedback and/or some subsets of general beliefs are correct for I_T.

This process of testing and rejecting reasonable construction-specific SRs is not a process which we ever, at least systematically, observe in actual patterns of early linguistic development, which suggests that normal speakers do not begin the acquisition of target I-languages with a MDC as unconstrained as GEN.

It is natural to think that we can, if not avoid, at least weaken the force of this objection by defending instead a constrained version of GEN. For example, consider a MDC, call it 'GEN^+', which allows non-compositional SRs but only

23

of the extensional feedback type. We need not state GEN^+ in detail, since it is obvious how to do this. GEN^+ might seem like an ad hoc MDC, proposed merely to reject CO. However, extensional feedback beliefs have a special place amongst our general beliefs, e.g., they are explicitly stored in memory, are for the most part easily retrievable, and can be incorporated into meaning assignments without having to appeal to complicated inferential computations. So it is not implausible to suggest that FL has selective access to extensional feedback beliefs, but not to other types of general beliefs. Another option is a MDC, call it GEN^*, which allows non-compositional SRs that involve only highly-weighted beliefs. Highly-weighted beliefs also have a special place amongst our beliefs. There is no known reason to deny that FL has selective access only to highly weighted beliefs, but not to other types of general beliefs. However, note that the construction-specific non-compositional SRs used in (18)-(20) are compatible with both GEN^+ and GEN^*, since they refer only to sets of highly-weighted extensional feedback beliefs. So the same sort of general objection raised against LOC and GEN can be raised against GEN^+ and GEN^*.

6 Objections and Open Issues

All the MDCs we examined can account for $P\&S$, but the non-compositional ones predict incorrect patterns of early linguistic development. This strongly suggests that CO is the most plausible MDC currently on the table. Let us now consider some objections to this argument.

Objection 1: Is CO as descriptively adequate as the non-compositional MDCs?

Assume that a MDC satisfies the condition of 'descriptive adequacy' relative to L if an I-language compatible with it generates L. One might question the claim that, in terms of descriptive adequacy relative to English and other natural languages, CO and the non-compositional MDCs are on equal footing. We did show, for each type of complex expression which motivated the introduction of a non-compositional MDC (meteorological expressions for LOC and extensional feedback complex NPs for the versions of GEN), that there are plausible accounts of how their meaning is determined compatible with CO. However, these types of expressions are only a subset of the types of expressions often considered problematic for compositionality, which include conditionals, genitives, nominal compounds, etc. If some of these types of expressions can only be given a non-compositional account, then relative to descriptive adequacy CO is in worse shape than some of the non-compositional MDCs. Indeed, several critics have argued that the view that FL is compositional is empirically false because the meanings of various types of complex expressions don't seem to be compositionally determined (see e.g. Fodor [22], Lahav [42] and Hampton [32])

This is a reasonable worry, but as is clearly illustrated in most recent surveys on this issue, it has been substantially addressed by the collective effort of

theorists who have proposed various compositional accounts for each of the problematic constructions.[17] These compositional accounts generally use the tools we used to show that there are CO compatible accounts of the meaning of simple meteorological expressions and extensional feedback complex NPs: lexical context-sensitivity, primary pragmatic processes, including meaning modulation, and lexical entries which in occasions of used are informationally enriched. The moral we should draw from this is that, as things currently stand, there are no types of expressions that can be taken as *direct* empirical counter-examples to CO.[18] This is why to resolve debates about the MDC of FL we need to move beyond descriptive adequacy.

To be clear, a consequence of adopting CO is that some data and intuitions about the meaning of certain expressions have to be dealt with at the level of pragmatics, hence not by appealing only to the workings of FL. With some important differences, most theorists accept this. For example, Borg [6]and other minimalists propose that we substantially 'clean' the data coming from meaning intuitions; while Recanati [66] and other contextualists try to account for a wider range of our pre-theoretical meaning intuitions. In discussions of competing MDCs what is important, when one of the accounts deals with some pre-theoretic intuition only in conjunction with non-linguistic cognitive processes, is that such decisions follow in a principled way from the assumed division between semantics and pragmatics, a division which everyone, including non-compositionalists, have to accept.

Objection 2: Why doesn't CO over-generate meanings in ways that parallel the non-compositional MDCs?

Assume CO is descriptively adequate in the sense just specified—and especifically that I-languages compatible with it can account for meteorological expressions and extensional feedback complex NPs. Why, if a CO constrained FL can account for that phenomena, can't it also over-generate meanings in a way that parallels the reasonable but unattested mistakes allowed by FLs constrained by non-compositional MDCs such as LOC and GEN?

Consider a way in which CO could allow mistakes that *seem* to mirror the sorts of mistakes which we used to object to non-compositional MDCs. Here is an obvious candidate: *if* for some speaker S, $[\![cat]\!]$ = CAT AND MEAN IF STRONG—i.e., if it includes such conditional information—then $[\![strong\ cat]\!]$ would mean STRONG AND MEAN CAT. We can model this as a result of such conditional information being included in S's lexical entry for *cat*, or of its arising from particular meaning modulations (enrichments) in certain occasions of use of

---

[17]Jonsson [39]: ch. 5 presents an up-to-date review of compositional accounts of many *problematic* linguistic constructions. See also Dever [18], Szabo [74], [76], Recanati [67], Pagin [48], [49], and Partee [50].

[18]At the end of a survey of 'problem cases' for compositionality, Jonsson concludes—echoing other theorists—that 'whether semantic theories in the end should be compositional...cannot be settled by attempting to provide examples that cannot be handled by a compositional (explicated in terms of CO) account since there does not seem to be any such cases' (Jonsson [39]: ch. 5).

cat. Similar observations can be used to generate incorrect locative restrictions in cases such as *gray cat*. Suppose such (incorrect) conditional information could, in certain circumstances, be part of the occasion meaning of lexical items; you might then suspect that a CO constrained FL could give rise to the same incorrect patterns we invoked against the non-compositional MDCs.

In response, note that the mistake when S determines, under CO, $[\![strong\ cat]\!]$ can be traced to an overly enriched standing or occasion meaning for *cat*, hence does not systematically affect the meaning-assignments to other expressions, including complex expressions that do not have *cat* as a constituent. So mistakes which can be traced to lexical items (which are allowed by CO and by the non-compositional MDCs) are quite different from those allowed only by the non-compositional MDCs. Consider (19) above, which involves the SR "$[\![[_{NP_m}A_m\ N_{an}]]\!] = f_{NP_m}(f_{NP}([\![A_m]\!],[\![N_{an}]\!]),b^+)$", where '$NP_m$' is a category of NPs formed by mental trait As and Ns for animate beings. In (19), S is testing a non-compositional construction-specific SR, allowed by GEN, which systematically assigns to expressions of the form $[_{NP_m}A_m\ N_{an}]$—such as *fierce lion, courageous soldier, angry pit-bull, loving father*—a meaning which includes an incorrect (but reasonable) locative restriction. The mistake in (19) is not due to a particular lexical item, and affects a much wider range of expressions than the subset of expressions of which that item is a constituent. The mistake is due to S's 'trying out' an incorrect non-compositional SR for expressions of the form $[_{NP_m}A_m\ N_{an}]$. The objections which I raise against LOC, GEN, and its variants are of that form. They apply even under the assumption that S assigns the correct meanings to all lexical items.

In short, it is not true that the relevant patterns of mistakes that can occur assuming LOC or GEN can also occur assuming CO, which blocks speakers from considering construction-specific SRs. The reason why CO, which allows meaning modulation of lexical items, does not allow patterns of mistakes analogous to (11)-(13) and (18)-(19) is due to the case-by-case nature of pragmatic modulations: whether S takes an utterance of *gray shark* to mean GRAY SHARK WHEN UNDERWATER or *strong cat* to mean STRONG AND MEAN CAT, under this account, depends on particular features of the contexts of utterances. In other words, it is a case-by-case decision which does not systematically affect S's literal meaning assignments to expressions of the same form in other contexts, for S's FL, constrained by CO, cannot acquire construction-specific non-compositonal SRs. In contrast, if S, constrained by one of the non-compositional MDCs, is trying out a construction-specific non-compositional SR, then the inclusion of the locative restrictions or general beliefs would result automatically—i.e., from the automatic processing of FL—and would range over all expressions (including novel ones) of the form over which the adopted SR ranges.[19]

Objection 3: Aren't we ignoring crucial trade-offs between CO and the non-compositional MDCs?

One might argue that the previous argument for CO depends on ignoring the full set of trade-offs between the competing MDCs. Take the arguments against LOC and GEN. We said that we can account for the meaning of simple meteorological expressions and extensional feedback complex NPs in ways that are consistent with CO. But to do that we have to assume that there is a distinction between the standing and the occasion meaning of expressions, that modulation functions can further modify the occasion meaning of lexical items, and that some of these modifications consist of enriching their meaning (see section 4). However, except for dealing with an over generation worry in Objection 2, we didn't carefully consider other costs of using those tools.

The key to address this worry is to note that, as far as we know, these are tools which we need to incorporate into any plausible linguistic model, including those that adopt a non-compositional MDC, to account for expressions other than simple meteorological expressions or extensional feedback complex NPs. This is obvious in the case of having to appeal, as Borg, Recanati, and Glucksberg do, to (primary) pragmatic effects to account for the intuitive meaning of tokens of certain types of complex expressions. Furthermore, assuming that lexical items are semantically non-atomic and include rich arrays of information is also becoming popular in accounts of genitives, possessives, privative As and adverbial modifications (see Vikener and Jensen [82], Coulson and Fauconnier [15], and Wunderlich [84]). Finally, (although here we did not appeal to this tool) assuming that some lexical items, which are not obvious indexicals, have context-sensitive parameters is common in recent accounts of words like $tall$, $flat$, $green$, $home$, $faraway$ etc. (see Szabo [74], Segal and Rothschild [70], Kennedy and McNally [41], and Recanati [67]: ch. 3). Non-compositional MDCs also have to account for these sorts of expressions, and to do so have, in many cases, to use these same tools, for the non-compositional SRs which are allowed by each non-compositional MDC do not provide a general way to deal with all or most of these expressions. So whatever the cost of introducing these tools, it is also incurred by models which adopt non-compositional MDCs. In addition, the non-compositional MDCs, but not CO, also introduce other technical tools—non-compositional, construction specific SRs—which have problematic consequences for language acquisition.

Objection 4: Why not take CO as just a default or cognitive bias?

An interesting response to the claim that CO is more plausible than any of the non-compositional MDCs is to propose a middle-ground. The idea is that we interpret CO as a default or cognitive bias of FL, such that under certain conditions language learners can drop this default and acquire non-compositional construction-specific SRs. There are various ways of implementing this proposal. For example, we can formulate a 'mixed' MDC, call it '$CO(LOC)$', that

has CO as a default and LOC as a secondary constraint on SRs. Following the same recipe, we can construct a MDC, call it '$CO(GEN)$', that has CO as a default and GEN as a secondary constraint on SRs. To fully determine if mixed MDCs with a compositional default are more plausible than CO would require an extensive discussion; here I will only briefly explain why it seems unlikely.

The challenge for mixed MDCs is to specify the properties of the learning data that would trigger the use of the non-default part of the constraint, i.e., the search for a non-compositional SR. The triggering conditions would have to include data that can be reliably used to infer that a compositional SR does not output the correct meanings for tokens of some type of complex expression, e.g., for tokens of simple meteorological expressions or extensional feedback complex NPs. For concreteness, let us focus on $CO(LOC)$ (the points I make apply equally well to $CO(GEN)$). The point of proposing that we replace CO with $CO(LOC)$ is to allow speakers to acquire construction-specific non-compositional SRs that assign location restrictions to meteorological expressions, while keeping a general bias for compositional SRs for at least most other types of complex expressions. So the triggering conditions have to meet two requirements: (i) token assertions of meteorological expressions trigger the search for a non-compositional SR, and (ii) token assertions of complex expressions that do not have a location restriction do not trigger the search for a non-compositional SR (which would explain why we don't observe patterns of mistakes like those in (11)-(13)). What could the triggering conditions be?

One might think that the search for a non-compositional SR should be triggered when S concludes that the meanings of complex expressions of some type seem to involve location restrictions, and that these meanings do not match the compositionally determined (default) meanings, which do *not* involve location restrictions. Things are not so simple, however, for data about the meanings of complex expressions, which comes mostly from linguistic interchanges, is usually noisy and plagued with mismatches between the compositionally determined and the asserted content of tokens of complex expressions. Speakers have tools to deal with these mismatches without having to revise the relevant default compositional SRs (see Objection 3 above). For example, in the case of assertions of meteorological expressions, speakers could assume that the location restrictions are due to primary pragmatic effects, such as the free-enrichment of the meanings of tokens of lexical items such as *rains* and *snows*. The question then is this: under what conditions should language learners resolve the mismatches between the asserted and the compositionally determined meanings of tokens of complex expressions of some type without using any of these available tools, but by revising instead the relevant construction-specific SRs?

This question reveals the problem with $CO(LOC)$ and other mixed MDCs with a compositional bias: simply put, all the ways of specifying the triggering conditions either render the mixed MDCs superfluous, or entail incorrect patterns of linguistic development similar to those entailed by pure non-compositional MDCs. On the one hand, if the conditions that trigger the search for a construction-specific, non-compositional SR are too demanding, this search would never begin. This is because, as we just said, most cases of mismatches

between the asserted and the compositionally determined meanings of tokens of complex expressions of some type (including those involving meteorological expressions) can be resolved *without* revising the default compositional SRs, e.g., by revising instead the relevant lexical entries or factoring-in systematic primary pragmatic effects. In this case, replacing CO with a more complex mixed MDC such as $CO(LOC)$ would be an entirely superfluous theoretical move. On the other hand, if we weaken the conditions that trigger the search for a non-compositional SR, this search would be triggered by mismatches (involving location restrictions) between the asserted and the compositionally determined meanings of tokens of expressions of various types, in addition to meteorological expressions. In this case, we would expect to find some speakers that test 'incorrect' SRs such as those in (11)-(13). For if we assume these weak triggering conditions, why would speakers resolve location restriction mismatches by revising the relevant SR in the case of meteorological expressions but *not* in the case of other types of expressions such as those in (11)-(13)? However, as we argued before, there is not much, if any, testing of non-compositional construction-specific SRs going on in early language acquisition, contrary to what would be predicted by mixed MDCs such as $CO(LOC)$ when paired with weak conditions that trigger the search for non-compositional SRs.

Objection 5: Why can't MDCs be learned?

We argued that we should understand MDCs as innate constraints on the functional architecture of FL, specifically, as innate 'over-hypotheses' on SRs, i.e., constraints on the general form of the SRs which FL is cognitively capable of representing. But are we really forced to hold that MDCs are innate?

Learning from experience requires some innate constraints. In particular, FL must have some innate constraints which help speakers learn the semantics of target I-languages. However, maybe the innate constraint on FL is more abstract than particular MDCs such as CO, LOC, or GEN, and is more like an 'over-over-hypothesis' which constrains possible MDCs in a way analogous to the way in which MDCs constrain possible SRs. This more abstract constraint, call it '$O(MDC)$', could constraint possible MDCs so that each allows only SRs that, for any type of complex syntactic structure $[_Z X\ Y]$, determine $[\![[_Z X\ Y]]\!]$ as a function of the meaning of its immediate constituents {X, Y}, and *possibly* something else. Although weak, $O(MDC)$ does restrict the set of possible MDCs. Speakers would still have to learn which of CO, LOC, GEN, etc., is correct for the target I-language. If constrained only by something like $O(MDC)$, then early in linguistic development S could consider, among others, SRs such as:

(21) $[\![[_Z X\ Y]]\!] = f_Z([\![X]\!], [\![Y]\!])$

(22) $[\![[_Z X\ Y]]\!] = f_{Z'}(f_Z([\![X]\!], [\![Y]\!]), b)$

(23) $[\![[_Z X\ Y]]\!] = f_{Z''}(f_Z([\![X]\!], [\![Y]\!]), g)$

where 'b' stands for a set of extensional feedback and 'g' for a location function. According to this view, when testing hypotheses like (21)-(23) against the data (i.e., subsets of the language), S not only selects the best fitting SR, but at the same time, and prior to acquiring other particular SRs, S also selects the best fitting MDC. For example, if S determines that (21) generates the correct meaning assignments for tokens with the form $[_Z X\ Y]$, this in turn suggests to S that the MDC is probably closer to CO than to LOC or GEN—assuming that S applies a learning mechanism which selects the logically strongest and simplest MDC, among those allowed by $O(MDC)$, which is consistent with the selected SR.

This sketch of how MDCs could be learned is especially interesting because in other cognitive domains similar learning processes—of acquiring over-hypotheses of the sort usually assumed to be innate by nativists—have been modeled using Hierarchical Bayesian Models (HBMs) (Kemp et al. [40]). To be clear, the reason we suggested that MDCs are likely innate is *not* that, given developmentally plausible subsets of the target language and any power-ful domain-general learning mechanism, it is otherwise impossible to explain how someone could acquire a target I-language such as I-English; the reason is that there seems to be no learning of the relevant sort going on in actual language acquisition—i.e., no learning of MDCs and even of construction-specific SRs. Of course, this does not conclusively show that although MDCs could in principle be learned they in fact are innate. For without considering specific HBMs, it is very hard, if not impossible, to determine whether the sorts of patterns Bayesian learners would have to go through to acquire MDCs—which given certain ways of modeling the problem, could be quite minimal—are consistent with the general patterns of early linguistic development. As far as I know, no HBMs that can acquire particular MDCs have been tested, although in principle such models can certainly be constructed.[20] Still, even if we can construct cognitively plausible HBMs that can acquire MDCs and so seriously consider the view that MDCs are acquired early in language acquisition, HBMs have some general properties which suggest that this result would be consistent with the main claim defended here—namely, that CO is the MDC on I-languages.

We can illustrate this last point by considering in broad outline a hypothetical HBM ('HBM$_L$') that, constrained by an $O(MDC)$, acquires a particular MDC in the processes of acquiring the semantics (i.e., the SRs) of I-English.

---

[20]There are two models that might be thought to bear on this issue. The first models the cultural evolution of natural languages, and shows that 'compositional' languages would be selected over 'non-compositional' languages (Kirby [72]). This model is not relevant to our problem—whether a Bayesian learner, given developmentally plausible bits of the language, would acquire CO over the non-compositional MDCs—because the 'non-compositional' languages considered by the model are implausible extremes, e.g., they do not satisfy any of the non-compositional MDCs considered here (they are even more unconstrained). The second model acquires a 'compositional semantics' (Tenenbaum et al. [52]). However, the main task faced by this model is only to pair simple expressions with lambda-types. The model is set-up to calculate the meaning of all complex expressions in the same general way, composition-ally via FA. So this model's functional architecture instantiates CO—it can't even represent non-compositional SRs.

We know something about the general conditions that would have to be met by 'HBM$_L$'. The process of acquiring a target I-language requires MDCs of some sort; so HBM$_L$ has to select a MDC early in the process of acquiring the full set of SRs, even if this selection is later revised. This is not a problem. As we just said, HBMs can be set-up to acquire the relevant over-hypotheses before they acquire most of the specific lower-level hypotheses. Now, assume that, given some subset E_1 of English, HBM$_L$ selects the MDC from the set of CO, LOC and GEN with the highest conditional probability. HBM$_L$ computes that as a function of the prior probability of each MDC—$P(CO)$, $P(LOC)$ and $P(GEN)$— and of the likelihood of E_1 given each MDC—$P(E_1|CO)$, $P(E_1|LOC)$ and $P(E_1|GEN)$. HBMs assign the highest prior probability to the simplest over-hypotheses. CO is the simplest MDC, since it has the fewest free parameters (for the same reason, compositional SRs are simpler than non-compositional SRs), so $P(CO) > P(LOC)$ and $P(CO) > P(GEN)$. The likelihoods partly depend on what is included in E_1. We are modeling the earliest stages in the acquisition of I-English, when speakers are beginning to learn how to determine the meaning of very simple complex expressions, so E_1 includes very simple complex expressions such as *red ball*, *green apple*, and *daddy away*, each paired with a representation of a stereotypical exemplar or situation (Pinker [57]). For this reason, it is safe to hold that the meanings of tokens of complex expressions included in E_1 can be generated using a compositional SR, such as (21) above. Now, compositional SRs such as (21) are compatible with CO, LOC and GEN. But since LOC and GEN also generate other SRs, e.g., (22) and (23) respectively, $P(E_1|CO) > P(E_1|LOC)$ and $P(E_1|CO) > P(E_1|GEN)$. It follows that:

(24) $P(CO|E_1) > P(LOC|E_1)$ and $P(CO|E_1) > P(GEN|E_1)$

such that:

(25) $P(CO|E_1)/P(LOC|E_1) > P(CO)/P(LOC)$ and
      $P(CO|E_1)/P(GEN|E_1) > P(CO)/P(GEN)$

Informally, (25) tells us that, although HBM$_L$ is initially biased to favor the simpler CO over the non-compositional MDCs, HBM$_L$ favors CO even more strongly after processing E_1. Furthermore, once there is a strong initial bias for CO, this selection would likely remain stable through the rest of the process of acquiring I-English, essentially for the reasons given in response to Objection 4. So even if we hold that MDCs are acquired in early language acquisition, it seems quite likely—given that HBMs are currently our best idea of how this process could work—that the MDC selected by language learners would be more like CO than like any of the non-compositional MDCs.

Open Issue: Implications for Constructionist Approaches

Constructionist approaches to language have been gaining popularity (Smith [73], Hoffman and Trousdale [36]). A full discussion of the implications of our

argument for CO for Constructionist approaches is outside the scope of this essay. Still, the issue merits some preliminary discussion, in part because it might seem that Constructionist approaches are, on the one hand, in tension with some of the basic assumptions we made about FL, and, on the other, undermined by the arguments against non-compositional MDCs. However, I will briefly explain why the implications are more nuanced and interesting.

What sets apart Constructionist approaches from mainstream Generative approaches is their emphasis on phrasal constructions in language acquisition. Like traditional lexical items, phrasal constructions are *learned* pairings of form and function (for an overview, see Goldberg [28]). Different types of phrasal constructions are associated with different types of functions. In our terminology, this means that phrasal constructions are associated with particular construction-specific SRs. According to Constructionist views, for speakers to determine the meaning of a complex expression of a certain phrasal type, it is not enough that they know its form and the meanings of its parts; they must also know which construction-specific SR is associated with the phrasal type. This entails that Constructionists implicitly reject the strong version of CO according to which SRs must be general.

We should not conclude from this that our approach and Constructionist approaches are incommensurable.We can straightforwardly frame a version of the compositionality debate in a Constructionist framework. But to do that we have to make an important modification: the compositional MDC would have to be a weak version of CO which allows construction-specific SRs, but only compositional ones. The competing non-compositinal MDCs could still be formulated basically like LOC, GEN, and its variants, since these allow learned pairings of phrasal types with particular SRs. What distinguishes the non-compositional MDCs from weak CO is that only the former allow non-compositional SRs.[21]

Once framed in this way we can see that the criticism based on unattested patterns of mistakes such as (11)-(13) and (18)-(20) would apply straightforwardly to a Constructionist approach that assumes one of the non-compositional MDCs, but not to one that assumes a weak version of CO. For although weak CO allows construction-specific SRs, it does not allow non-compositional construction-specific SRs, which blocks cases like (11)-(13) and (18)-(20).[22]

---

[21]Some of the literature that contrasts Constructionist with mainstream Generative approaches tend to characterize the former as non-compositional (see e.g. Smith [73]: p. 380). What I think they usually mean when they make those remarks is that Constructionists cannot accept that FL represents only general SRs, hence they cannot say that, in all cases, the meaning of a complex expressions is determined by the meaning of the parts and their structure. But clearly Constructionists can in principle accept a compositional view, as long as this view allows construction-specific compositional SRs. In this case the position can be expressed by saying that the meaning of a complex expressions is determined by the parts, their structure, and the function associated with that structure.

[22]This is not to deny that our approach is more in tune with mainstream Generative approaches. For example, some of the considerations we presented in favor of compositionality tend to support a compositional MDC which only allows general SRs. What I am arguing here is that, if for other reasons we favor a Constructionist approach, we can frame a version of the compositionality debate within this approach. Once we do that, we can see that our previous

Non-compositionalist Constructionists cannot respond that the phrasal form-function pairings are innate, since they are committed to these pairings being learned (Goldberg [27, 28], Tomasello [80]). In addition, our argument didn't make any strong assumptions about the learning mechanism responsible for acquiring SRs (e.g., they can be domain general), so there is no obvious assumption there to reject (for discussion see footnote 8). Finally, although most Constructionist are anti-nativists (Goldberg [27, 29], Tomasello [80]), this does not force them to deny that, unlike particular SRs, MDCs are plausibly innate; but if they do deny that, then the remarks made in Objection 5 above—regarding why Bayesian learners would tend to acquire a compositional over-hypothesis—directly apply to this case.

In short, our argument for CO is not in tension with all Constructionist approaches. Accepting a Constructionist approach does require replacing a strong version of CO with a weaker version. Aside from that, none of the further assumptions we made about FL or language acquisition to defend compositional MDCs are inconsistent with Constructionist approaches per se. But once properly modified, the argument for CO supports the view that we should adopt compositional over non-compositional Constructionists view.

7   Conclusion

Debates about the non/compositionality of FL seem to reach a standstill when we acknowledge, with recent critics, that there are non-compositional MDCs that can account for $P\&S$. To resolve this standstill, we first argued that we should frame these debates as debates about which MDC is the most plausible functional constraint on the semantics of FL, specifically, on the allowed forms of SRs which FL can represent. We then saw that each non-compositional MDC involves a weakening of CO, the point of which is to make FL compatible with some adequate non-compositional (construction-specific) SRs. However, theorists generally fail to notice that each weakening also makes FL compatible with many other incorrect SRs. As a result, if FL was constrained by these non-compositional MDCs, speakers would, in the course of early linguistic development, have to test and reject at least some reasonable but incorrect construction-specific non-compositional SRs. This predicts patters of early linguistic development which actual speakers never seem to go through. In contrast and more consistent with actual linguistic development, CO predicts patterns of development that do not involve any testing of reasonable but incorrect construction-specific SRs. We also considered some seemingly plausible additional constraints on the non-compositional MDCs which seek to constrain the search space of construction-specific SRs during acquisition. The two main proposals are that construction-specific SRs are innate and that non-compositional MDCs are the non-default options of complex MDCs with compositional defaults. We showed that none of these moves saved the non-compositional MDCs without giving rise to other unacceptable problems. This strongly suggests that

argument favors weak compositional over non-compositional Constructionist approaches.

33

the MDC of FL is closer to CO than to any of the non-compositional MDCs currently on offer. Admittedly, this argument for CO is not a wholly general argument that ranges over all conceivable non-compositional MDCs combined with all conceivable additional constraints. This is why it is important that the particular non-compositional MDCs and additional constraints which we examined are currently the most plausible, motivated and popular. Perhaps more importantly, this overall approach—which focuses the implications of MDCs on patterns of language acquisition via their constraints on the SRs which FL can represent—can be used to evaluate future proposals for MDCs.

References

[1] Giosue Baggio, Michel Van Limbalgen, and Peter Hagoort. The processing consequences of compositionality. In Markus Werning Wolfram Hinzen and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 32, pages 656–672. Oxford University Press, New York, 2012.

[2] Mark C. Baker. *The Atoms of Language*. Basic Books, New York, 2001.

[3] Laurence Barsalou. Ad hoc categories. *Memory & Cognition*, 11:211–27, 1983.

[4] Ned Block. Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, 10(1):615–78, 1986.

[5] Ned Block. Holism, hyper-analyticity and hyper-compositionality. *Philosophical Issues*, 3:37–72, 1993.

[6] Emma Borg. *Minimal Semantics*. Oxford University Press, Oxford, 2004.

[7] Elizabeth Camp. On the generality constraint. *Philosophical Quaterly*, 54:209–231, Aug 2004.

[8] Noam Chomsky. *Knowledge of Language: Its Nature, Origin, and Use*. Praeger Publishers, Connecticut, 1986.

[9] Noam Chomsky. Language and nature. *Mind*, 104, 1995.

[10] Noam Chomsky. *New Horizons in the Study of Mind Language*. Oxford University Press, Oxford, 2000.

[11] Noam Chomsky. Beyond explanatory adequacy. Technical report, MIT, Cambridge, MA, 2001.

[12] Noam Chomsky. Derivation by phase. In M. Kenstowitz, editor, *Ken Hale: A Life in Language*, pages 1–52. MIT Press, Cambridge, MA, 2001.

[13] Noam Chomsky. Reply to gopnik. In Louise M. Antony and Norbert Hornstein, editors, *Chomsky and his Critics*, chapter 11, pages 316–325. Blackwell Publishing Ltd, 2003.

[14] V. J. Cook and Mark Newson. *Chomsky's Universal Grammar: An Introduction.* Blackwell Publishing Ltd, Oxford, 2007.

[15] Seana Coulson and Gilles Fauconnier. Fake guns and stone lions: Conceptual blending and privative adjectives. In B. Fox, D. Jurafsky, and L. Michaels, editors, *Cognition and Function in Language.* CSLI, Palo Alto, CA, 1999.

[16] Martin Davies. Tacit knowledge and semantic theory: can a five per cent difference matter? *Mind*, 96:441–462, 1987.

[17] Michael Dawson. *Understanding Cognitive Science.* Blackwell Publishers Ltd, Oxford, 2001.

[18] Josh Dever. Compositionality. In Ernest Lepore and Barry Smith, editors, *Handbook of Philosophy of Language.* Oxford University Press, Oxford, 2006.

[19] J. Elman, E. Bates, M. Johnson, and A. Karmiloff-Smith. *Rethinking Innatness: A Connectionist Perspective on Development.* MIT Press, Cambridge, MA, 1996.

[20] Gareth Evans. *Collected Papers.* Oxford University Press, Oxford, 1985.

[21] Jerry Fodor. *Concepts: Where Cognitive Science Went Wrong.* Oxford University Press, Oxford, 1998.

[22] Jerry Fodor. Language, thought and compositionality. *Mind & Language*, 16:1–15, February 2001.

[23] Jerry Fodor and Ernest Lepore. *The Compositionality Papers.* Oxford University Press, Oxford, 2002.

[24] Jerry Fodor and Zenon Pylyshyn. Connectionism and cognitive architecture: A critical analysis. Technical report, Rutgers Center for Cognitive Science, 1988.

[25] Sam Glucksberg. *Understanding Figurative Language: From Metaphors to Idioms.* Oxford University Press, New York, 2001.

[26] Sam Glucksberg. The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, 7(2):92–96, 2003.

[27] Adele E. Goldberg. *Constructions at Work: The Nature of Generalizations in Language.* Oxford University Press, Oxford, 2006.

[28] Adele E. Goldberg. Constructionist approaches. In Thomas Hoffman and Graeme Trousdale, editors, *The Oxford Handbook of Construction Grammar*, chapter 2. Oxford University Press, Oxford, 2013.

[29] Adele E. Goldberg. Explanation and constructions. *Mind & Language*, forthcoming.

[30] Alison Gopnik. The theory theory as an alternative to the innatness hypothesis. In Louise M. Antony and Norbert Hornstein, editors, *Chomsky and his critics*, chapter 10, pages 238–255. Blackwell Publishing Ltd, Oxford, 2003.

[31] JA Hampton. Emergent attributes in combined concepts. *Creative thought: An investigation of conceptual structures and processes*, pages 83–110, 1997.

[32] James A. Hampton. Conceptual combination and fuzzy logic. In Radim Belohlavek and George J. Klir, editors, *Concepts and Fuzzy Logic*, chapter 8. The MIT Press, Cambridge, MA, 2011.

[33] James A. Hampton and Martin L. Jonsson. Typicality and compositionality: The logic of combining vague concepts. In Marcus Werning, Wolfram Hinzen, and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 18, pages 385–402. Oxford University Press, New York, 2012.

[34] Irene Heim and Angelina Kratzer. *Semantics in Generative Grammar*. Blackwell Publishers Ltd, Oxford, 1998.

[35] James Higginbotham. Languages and idiolects: their language and ours. In Ernest Lepore and Barry Smith, editors, *The Oxford Handbook of the Philosophy of Language*. Oxford University Press, Oxford.

[36] Thomas Hoffman and Graeme Trousdale, editors. *The Oxford Handbook of Construction Grammar*. Oxford University Press, Oxford, 2013.

[37] Pauline Jacobson. Direct compositionality. In Wolfram Hinzen Marcus Werning and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 5, pages 109–28. Oxford University Press, Oxford, 2012.

[38] Kent Johnson. On the systematicity of language and thought. *Journal of Philosophy*, 101(3):111–139, 2004.

[39] Marrtin L. Jonsson. *On Compositionality: Doubs on the Structural Path to Meaning*. PhD thesis, Lund University, 2008.

[40] Perfors A. Kemp, C. and J. B. Tenenbaum. Learning overhypotheses. In *Proceedings of the Twenty-Eighth Annual Conference of the Cognitive Science Society.*, 2006.

[41] C. Kennedy and B. Levin. Measure of change: The adjectival core of degree achievements. In Louis McNally and Christopher Kennedy, editors, *Adjectives and Adverbs: Syntax, Semantics, and Discourse*. Oxford University Press, 2008.

[42] Ran Lahav. Against compositionality: the case of adjectives. *Philosophical Studies*, 57:261–279, 1989.

[43] Richard Larson and Gabriel Segal. *Knowledge of Meaning*. The MIT Press, Cambridge, 1995.

[44] Peter Ludlow. *The Philosophy of Generative Grammar*. Oxford University Press, Oxford, 2011.

[45] M.S. McGlone, S. Glucksberg, and C. Cacciari. Semantic productivity and idiom comprehension. *Discourse Processes*, 17:167–190, 1994.

[46] Gregory L. Murphy. *The big book of concepts*. The MIT Press, Cambridge, Massachusetts, 2002.

[47] Geoffrey Nunberg, Ivan A. Sag, and Thomas Wasow. Idioms. *Language*, 70(3):491–538, 1994.

[48] Peter Pagin and Jeffrey Pelletier. Content, context, and composition. *Context-Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics*, page 25, 2007.

[49] Peter Pagin and Dag Westerstahl. Compositionality. October 2008.

[50] Barbara Partee. *Compositionality in Formal Semantics*. Blackwell Publishers Ltd, Oxford, 2004.

[51] Amy Perfors, Joshua B. Tenenbaum, and Terry Regier. Poverty of stimulus? a rational approach. *Cognition*, 118:306–338, 2011.

[52] S. T. Piantadosi, N. D. Goodman, B. A. Ellis, and J. B. Tenenbaum. A bayesian model of the acquisition of compositional semantics. In *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*, 2008.

[53] Paul Pietroski. The character of natural language semantics. In A. Barber, editor, *Epistemology of Language*. Oxford University Press, 2003.

[54] Paul Pietroski. Minimalist meaning, internalist interpretation. *BIOLINGUISTICS*, 2(4), 2008.

[55] Paul Pietroski. Semantic monadicity with conceptual polyadicity. In Marcus Werning, Wolfram Hinzen, and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 6, pages 129–148. Oxford University Press, Oxford, 2012.

[56] Guillermo Del Pinal. *The Architecture of the Faculty of Language: Compositional Operations and Complex Lexical Representations*. PhD thesis, Columbia University in the City of New York, New York, forthcoming 2013.

[57] Steven Pinker. Language acquisition. In *An Inivitation to Cognitive Science*, volume 1, chapter 6. The MIT Press, 1995.

[58] Jesse Prinz. *Furnishing the mind: concepts and their perceptual basis*. MIT Press, Cambridge, Massachusetts, 2002.

[59] Jesse Prinz. Regaining composure: a defense of prototype compositionality. In Markus Werning, Wolfram Hinzen, and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 21. Oxford University Press, Oxford, 2012.

[60] Geoffrey K. Pullum and Barbara C. Scholz. Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 19(9-50), 2002.

[61] Geoffrey K. Pullum and Barbara C. Scholz. Recursion and the infinitude claim. In Harry van der Hulst, editor, *Recursion and Human Language*, chapter 5, pages 113–139. De Gruyter Mouton, Germany, 2010.

[62] Zenon Pylyshyn. *Computation and Cognition: Towards a Foundation for Cognitive Science*. MIT Press, Cambridge, 1986.

[63] Zenon Pylyshyn. The role of cognitive architecture in theories of cognition. *Architectures for intelligence: the twenty-second Carnegie Mellon Symposium on Cognition*, 1991.

[64] B. Yankama N. Chomsky R. Berwick, P. Pietroski. Poverty of the stimulus revisited. *Cognitive Science*, 35:1207–1242, 2011.

[65] Andrew Radford. *Minimalist Syntax: Exploring the Structure of English*. Cambridge University Press, Cambridge, 2004.

[66] Francois Recanati. *Literal Meaning*. Cambridge University Press, Cambridge, 2004.

[67] Francois Recanati. *Truth-Conditional Pragmatics*. Oxford University Press, Oxford, 2010.

[68] Phillip Robbins. Minimal semantics and modularity. In G. Preyer and G. Peter, editors, *Context Sensitivity and Semantic Minimalism: New Essays on Semantics and Pragmatics*. Oxford University Press, 2007.

[69] Tom Roeper. The acquisition of recursion:how formalism articulates the child's path. *Biolinguistics*, 5(57-82), 2011.

[70] Daniel Rothschild and Gabriel Segal. Indexical predicates. *Mind and Language*, 24:467–493, 2009.

[71] Barbara C. Scholz and Geoffrey K. Pullum. Systematicity and natural language syntax. *Croation Journal of Philosophy*, 21:375–402, 2007.

[72] Kenny Smith and Simon Kirby. Compositionality and linguistic evolution. In Wolfram Hinzen Markus Werning and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 25. Oxford University Press, Oxford, 2012.

[73] Neil Smith. Introduction: Syntax symposium. *Mind & Language*, 28:377–391, September 2013.

[74] Zoltan Szabó. Adjectives in context. In R. Harnish and I. Kenesei, editors, *Perspectives on Semantics, Pragmatics, and Discourse*, pages 119–46. John Benjamins, Amsterdam, 2001.

[75] Zoltan Szabó. Review of the compositionality papers. *Mind*, 113:340–344, 2004.

[76] Zoltan Szabó. Compositionality. *Stanford Encyclopedia of Philosophy*, 2007.

[77] Zoltan Szabó. Structure and conventions. *Philosophical Studies*, 137(3), 2008.

[78] Zoltan Szabó. The case for compositionality. In Markus Werning, Wolfram Hinzen, and Edouard Machery, editors, *The Oxford Handbook of Compositionality*, chapter 3. Oxford University Press, Oxford, 2012.

[79] P. Tabossi and F. Zaron. The activation of idiomatic meaning. In Schenk Everaert, Van den Linden and Schreuder., editors, *Idioms: Processing, Structure, and Interpretation*, pages 145–163. Hillsdale, New Jersey, 1995.

[80] Michael Tomasello. *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, MA, 2003.

[81] Charles Travis. On constraints on generality. *Proceedings of the Aristotelian Society*, 44:165–188, 1994.

[82] Carl Vikner and Per Anker Jensen. A semantic analysis of the english genitive. interacion of lexical and formal semantics. *Studia Linguistica*, 56(2):191–226, 2002.

[83] Deidere Wilson and Robyn Carston. A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In Noel Burton-Roberts, editor, *Pragmatics*, pages 230–259. Palgrave, London, 2007.

[84] Dieter Wunderlich. Lexical decomposition in grammar. In Marcus Werning, Wolfram Hinzen, and Edouard Machery, editors, *Oxford Handbook of Compositionality*, chapter 14. Oxford University Press, Oxford, 2012.

[85] Fei Xu. Rational statistical inference and cognitive development. In Stephen Laurence Peter Carruthers and Stephen Stich, editors, *The innate mind: Foundations and the future*, volume 3, chapter 10, pages 199–215. Oxford University Press, New York, 2007.