The following English sentence has never before been produced:

- Having recently exhumed the vicar’s nephew, the stockbroker wistfully contemplated a large wheel of pungent blue cheese.

Nevertheless, any competent speaker will know what it means. What explains our ability to understand sentences we have never before encountered? One natural hypothesis is that those novel sentences are built up out of familiar parts, put together in familiar ways. This hypothesis requires the backing hypothesis that English has a compositional semantic theory:

(Compositionality) A language is compositional if the meaning of each of its complex expressions are derived from the meanings of its simple expressions.¹

The backing hypothesis can seem obviously true. A sentence, after all, just is a collection of words; how could its meaning not be determined by the meanings of its constituent words?

Further examples, however, suggest that English may not have a compositional semantics. Consider the phrase ‘large wheel of pungent blue cheese’ from the above

¹This is only a rough sketch of compositionality; see section 1 for a more careful formulation of the principle.
sentence. If English is compositional, then the meaning of this phrase should be determined by the meanings of its parts. But getting the meaning of ‘large cheese’ (to simplify a bit) out of the meanings of ‘large’ and ‘cheese’ is not straightforward. Appropriate standards of largeness vary depending on the type of object in question. What’s large for a cheese may be quite small for a person.\textsuperscript{2} It’s thus hard to see how the right reading for ‘large cheese’ could be derived by combining independent meanings for ‘large’ and ‘cheese’. Another way of putting the problem: saying that Sveto is a large man cannot amount to saying that he is large and he is a man, because he can be a large man and a sumo wrestler without also being a large sumo wrestler.\textsuperscript{3}

On the other hand, while these examples of \textit{non-intersective adjectives} raise problems for compositionality, those problems are not obviously insuperable. While the meaning of ‘large cheese’ cannot be built from the meanings of ‘large’ and ‘cheese’ in one way – through conjunction, or set intersection – it can be in another way. Suppose that:

\begin{itemize}
  \item $\llbracket \text{cheese} \rrbracket = \{x : x \text{ is a cheese}\}$
  \item $\llbracket \text{large} \rrbracket = f, \ f(X) = \{x \in X : x \text{ is larger than the average size of } X\}$
\end{itemize}

Then the meaning of ‘large cheese’ can be derived from $\llbracket \text{large} \rrbracket$ and $\llbracket \text{cheese} \rrbracket$ through functional application, which yields the set of all cheeses above average in size. The

\textsuperscript{2}The Octuple Gloucester of [59] notwithstanding.

\textsuperscript{3}The reasoning proceeds as follows:

1. Sveto is a large man and a sumo wrestler.
2. Therefore, Sveto is large and a man and a sumo wrestler (by the proposed analysis of ‘large X’).
3. Therefore, Sveto is large and a sumo wrestler and a man (by the commutivity of ‘and’).
4. Therefore, Sveto is large and a sumo wrestler (by ‘and’ elimination).
5. Therefore, Sveto is a large sumo wrestler (again by the proposed analysis of ‘large X’).
same procedure allows Sveto to be a large man without thereby also being a large sumo wrestler. Compositionality is achieved, at the price of a slight complication of the lexical semantics.

Other difficult cases for compositionality abound, as do attempts to deal with those difficulties. The question of whether natural languages have compositional semantics continues to attract considerable interest, as do questions about the reasons for wanting compositionality, the consequences of compositionality, and the very formulation of the principle of compositionality. This overview begins, in section 1, by developing a precise definition of compositionality. In section 2 some technical consequences of that definition are explored. Section 3 examines two compositionally-problematic semantic phenomena, and proposed compositional treatments thereof. Section 4 closes by asking why one might want a compositional meaning theory, and attempting to explain the philosophical significance of compositionality.

1 What is Compositionality?

Compositionality is a tool for limiting what can be relevant to determining the meaning of a complex expression. As such, it represents the simultaneous imposition of two constraints:

(Semantic Closure) Only semantic information can go into the determination of the semantic value of a complex expression.

(Semantic Locality) Only information derived from parts of a complex expression can go into the determination of the semantic value of that expression.

Semantic Closure prevents, for example, the meaning of:

(1) Lois Lane believes that Superman can fly.
from being determined in part by the word ‘Superman’ (its phonetic, morphological, historical-causal\(^4\) properties, etc.), rather than the meaning of the word ‘Superman’ (whatever that meaning turns out to be).\(^5\) Semantic Locality, on the other hand, prevents the meaning of the occurrence of ‘Superman can fly’ in (1) from being a function of (perhaps among other things) the meaning of ‘believes’. Combining Semantic Closure and Semantic Locality yields Compositionality – the requirement that the meaning of a complex expression be determined by the meanings of its parts.

Making this requirement more precise requires clarifying the notion of determination. Two versions of determination dominate discussion of compositionality: the functional analysis and the substitutional analysis.\(^6\)

### 1.1 Compositionality as Functionality

The heart of the functional conception of compositionality is the requirement that the meaning of a complex expression be a function of the meanings of the parts of that syntactic expression and their mode of composition. We give a rather complex implementation of this simple idea. Suppose that, for any language \(L\), we have function \([\cdot]_L\) mapping from \(L\) to some set \(M\) of meanings and a collection of (syntactic formation) \(^4\)See the conception of words defended in [40].

\(^5\)Assuming, of course, that the word ‘Superman’, or its phonetic, morphological, historical-causal, etc. properties are not part of the semantics of ‘Superman’.

\(^6\)See [64] and [63] for an excellent discussion of these two notions of compositionality. Szabo formulates compositionality as a supervenience principle, and then employs Kim’s distinction between weak and strong supervenience (see [44]) to argue for a strengthened conception of compositionality. The current discussion treats compositionality as a weak supervenience principle, and discusses some of Szabo’s arguments for the move to strong supervenience. See [12] for further discussion of Szabo’s use of supervenience.
functions $\delta_n : L^n \mapsto L$. Then we say:

- A meaning theory $\mu$ for language $L$ is compositional relative to (i) a class $L^+$ of extensions of $L$, (ii) a parthood relation $\sqsubseteq_{L_i}$ for each $L_i \in L^+$, where the parthood relations for the extensions in $L^+$ agree on $L$, (iii) a level of structural analysis $\{\delta_n(\theta^1_i, \ldots, \theta^n_i) \mid \epsilon \in L_i\}$ for each $L_i \in L^+$, where the $\theta_i^\epsilon$’s are $\sqsubseteq_{L_i}$-parts of $\epsilon$, and where the levels of structural analysis agree on $L$, and (iv) a class $G^+$ of functions if for each $L_i \in L^+$ there is some function $G_i \in G^+$ such that for all $\epsilon \in L_i$, $[\epsilon]_{L_i} = G(\delta_{\epsilon}^1, [\theta_i^1]_{L_i}, \ldots, [\theta_i^n]_{L_i})$, and the $G_i$’s agree on $L$.

We thus have a four-fold relativized notion of compositionality – a language can be compositional relative to a parthood relation, a level of structural analysis, a collection of possible extensions of the language, and a range of admissible meaning composition functions. Each dimension of relativization responds to a shortcoming of or objection to the simple heart of the functional conception:

1. The parthood relation $\sqsubseteq$ will typically be given by the syntactic theory, but need not be. A syntactic theory may deliver multiple candidate parthood relations, though – do we require compositionality at the level of deep structure or of surface structure? At LF or PF, or some combination of the two, in a minimalist syntax? Relativization to a choice of $\sqsubseteq$ allows the core theory of compositionality to avoid the need to take a stand on such questions.

---

Note that any language is compositional under some parthood relation. Minimally, we want $\sqsubseteq$ not to be reflexive.

---

7In order to maximize generality, no constraint is placed on the set $M$. Thus, for example, in a bilevel Fregean semantic theory compositionality constraints could be imposed separately on each level, by taking $M$ first to be the set of referents and second to be the set of senses. Semantic compositionality thus falls into a genre of mereological reducibility.

8Note that any language is compositional under some parthood relation. Minimally, we want $\sqsubseteq$ not to be reflexive.
2. A complex expression can typically be given syntactic analysis at varying levels of detail. Suppose:

(2) Some philosopher fears Socrates

receives the rather flat-footed syntactic analysis:

\[ S \]

\[ \begin{array}{c c}
NP & VP \\
D & N & V & NAME \\
\hline
Some & philosopher & fears & Socrates
\end{array} \]

At one level of analysis, (2) is the result of applying a syntactic function \( \delta_1 \) to parts \([NP [D some] [N philosopher]]\) and \([VP [V fears] [NAME Socrates]]\). At another, finer-grained level, it results from applying a different function \( \delta_2 \) to the parts \([D some], [N philosopher], [V fears], \) and \([NAME Socrates]\).\(^9\) By selecting the level of syntactic analysis, we can distinguish between:

(Strong Compositionality) \( L \) is strongly compositional if every expression has a coarsest non-trivial syntactic analysis, and the meaning of every expression is a function of the meanings of its parts and their mode of combination, under that coarsest analysis.\(^10\)

\(^9\)I assume throughout that every expression receives only unambiguous parthood analyses, even if analyses are available at multiple levels of grain. See [55] for an argument that ambiguity blocks the possibility of compositional semantics, [14] and [15] for attempts to show how ambiguity and compositionality can be combined, and [68] for mathematical examination of notions of compositionality designed to take ambiguity into account.

\(^10\)This definition is equivalent to [49]'s notion of strong compositionality.
(Weak Compositionality) **L** is *weakly compositional* if every expression has a finest syntactic analysis, and the meaning of every expression is a function of the meanings of the meanings of its parts and their mode of combination, under that finest analysis.

Suppose, for example, that the semantic value of a definite description were its denotation. Then a strongly compositional theory would be committed to the equivalence of:

(3) In 1983, the president of the United States was named Ronald.

(4) In 1983, the son of George Herbert Walker Bush was named Ronald.

since \[\text{the president of the United States} = \text{the son of George Herbert Walker Bush} = \text{George Walker Bush},\] and since the behaviour of the sentence must be determined by the semantic values of its immediate constituents. A weakly compositional theory, however, could appeal to semantic differences among the components of the definite descriptions (‘president’, ‘George Herbert Walker Bush’, etc.) to ground a semantic difference between (3) and (4).\textsuperscript{11}

A language which is compositional at one level of analysis is compositional under any coarser-grained level of analysis; hence strong compositionality is indeed stronger than weak compositionality.

3. An immediate consequence of a simple functional analysis of compositionality is that if a language contains no synonyms, it is trivially compositional.\textsuperscript{12}

\textsuperscript{11}Since (3) and (4) are clearly not equivalent, strong compositional theories have reason not to use denotation as the semantic value of a definite description.

\textsuperscript{12}See [70] for a brief note of this fact, and [64] for an argument that it shows that the functional approach does not capture the principle of compositionality.
This fact makes functional compositionality appear undesirably weak. Consider the language $\text{Lenglish}$ – a synonym-free fragment of English, with the idiosyncrasy that, given any expression $\epsilon$ occurring in a sentence containing the letter K, $[\epsilon] = \text{T}$. Lenglish should be wildly noncompositional, but it meets the simple functional requirement.

The undesirable result can be avoided through a richer conception of function. The above assumes that $[\cdot]$ is given in extension, but the nature of our semantic competence suggests that meaning functions are in fact given in intension. Semantic mastery of a language extends to the ability to determine what the meanings of complex expressions of the language would be, were they to include (at least some) newly-introduced lexical items of specified meaning. We thus model the desired intensionality by considering a class of extensions of the base language. Lenglish fails of compositionality in natural extensions containing synonyms.\(^\text{13}\)

4. Suppose (following an objection raised by Szabo in [64] and [63], we form a language $L'$ identical to English, except that the meanings of the sentences:

(5) Elephants are gray.

(6) Julius Caesar was murdered on the ides of March

are swapped. If English is (functionally) compositional, so is $L'$, via composition of the English meaning function with a simple permutation function. We

\(^{13}\)Every language will be noncompositional relative to some class of extensions, so the crucial point is what classes of extensions matter to compositionality considerations. Speakers’ semantic mastery provides one natural class of extensions – consisting of extensions in which new lexical items are introduced in existing syntactic categories, with meanings appropriate for those categories, and integrate into complex expressions in accordance with the intensionally-given meaning functions of the core language – but whether this class is the right one to consider depends on why one is interested in compositionality.
might, however, want to call $L'$ noncompositional, perhaps feeling the meaning of ‘Julius Caesar was murdered on the ides of March’ simply cannot be built out of the meanings of ‘elephant’ and ‘gray’ and a few simple syntactic rules. While you may be able to make an omelette from eggs and milk using one recipe (language) and a creme brûlée from the same ingredients using another recipe, there is no recipe for making roast venison from those ingredients – eggs and milk simply don’t combine in that way, no matter what you do with them. Similarly, perhaps there’s just no way to combine large African land mammals and shades of gray to produce a claim about the murder of a Roman general.

We thus require that the function $G$, determining the meaning of a complex expression from the meanings of its parts, be drawn from a class $G^+$ of permissible functions, and relativize the definition of compositionality to a choice of $G^+$. By making $G^+$ small, a quite strong notion of compositionality can be determined. If $G^+$ contains only the technique of functional application, so that $G(<\delta, [\epsilon_1], [\epsilon_2]) > = [\epsilon_1](\epsilon_2)$, where $G \in G^+$, then the corresponding compositionality constraint requires semantics to be done in a Montagovian style. If $G^+$ contains only functions which extend mereological fusion, the result is what Szabo in [63] calls strong parallelism – the view, endorsed by Frege for senses, that the meaning of a complex contains the meanings of its parts as parts.\[14\]

The functional conception of compositionality is frequently connected to the idea of a homomorphism. Assuming that the parthood relation $\subseteq$ reflects a collection $\Delta$ of syntactic constructions, compositionality acts as a requirement that semantic interpretation closely track syntactic form. Suppose that complex expressions in $L$ are built up via applications of two syntactic rules: one of predication of a verb phrase to a subject, and one of adjectival modification of a subject. Then there ought to be two correspond-

\[14\]See, for example, [23], 127.
ing semantic rules: one determining the meaning of a sentence from the meanings of
its component verb phrase and subject, and one determining the meaning of a complex
subject from the meanings of its component adjective and subject. More generally,
the relation between syntax and semantics is given by what [2] called the rule to rule hypothesis:

(R-R) Given $L$ with syntactic rules $\Delta$ and meaning function $\mu$, there is a set
$\Gamma = \{\gamma_\delta : \delta \in \Delta\}$ of functions such that if $\epsilon \in L$ is the result of applying some $\delta$
to some $\epsilon_1, \ldots, \epsilon_n$, then $\mu(\epsilon) = \gamma_\delta(\mu(\epsilon_1), \ldots, \mu(\epsilon_n))$.

Taking the syntax and the semantics of $L$ as algebras (of expressions in one case, of
meanings in the other) generated by operations (syntactic construction rules in the one
case, semantic composition procedures in the other case), the rule-to-rule hypothesis
requires that there be a homomorphism between the two algebras preserving the struc-
ture imposed by the operations of each.\(^{15}\)

Compositionality in the simple functional sense of the previous section and in the
homomorphism sense are trivially equivalent constraints. Suppose $L$ is compositional
in the homomorphism sense. Define $G(<\delta, \mu(\epsilon_1), \ldots, \mu(\epsilon_n)>) = \gamma_\delta(\mu(\epsilon_1), \ldots, \mu(\epsilon_n))$, and
$G$ is a composition function for all of $L$. If, on the other hand, $L$ has a composi-
tion function $G$, then define, for each $\delta \in \Delta$, $\gamma_\delta(\mu(\epsilon_1), \ldots, \mu(\epsilon_n)) = G(<\delta, \mu(\epsilon_1), \ldots, \mu(\epsilon_n)>)$.\(^{16}\)

1.2 Compositionality as Substitutability

Suppose Lex Luthor believes Superman is harmed by Kryptonite, but does not believe
Clark Kent is harmed by Kryptonite. Then if $\llbracket\text{Superman}\rrbracket = \llbracket\text{Clark Kent}\rrbracket$, composi-

\(^{15}\)See [51], [35], [26], [72], [36], and [70] for examples of this approach.

\(^{16}\)See Theorem 4 of [32] for a more precise statement of this result.
tionality has a problem. The two sentences:

(7) Lex Luthor believes Superman is harmed by Kryptonite.

(8) Lex Luthor believes Clark Kent is harmed by Kryptonite.

have different truth values, and hence on almost any semantic theory also have different meanings. But, other than the substitution of ‘Clark Kent’ for ‘Superman’, they are made of the same parts in the same way\(^{17}\), and the two parts ‘Superman’ and ‘Clark Kent’ have the same meaning. Thus the two sentences are made in the same way of parts of the same meaning, and ought, if compositionality holds, to be synonymous. A language in which they are not synonymous, then, cannot be a compositional one.\(^{18}\)

More generally, suppose language \(L\) has two expressions \(\epsilon_1\) and \(\epsilon_2\) with the same meanings, and two complex expressions \(E_1\) and \(E_2\) which differ only in that some occurrences of \(\epsilon_1\) in \(E_1\) have been replaced with occurrences of \(\epsilon_2\) in \(E_2\), and which are such that \(\mu(E_1) \neq \mu(E_2)\). Then \(E_1\) and \(E_2\) are constructed in the same way out of parts with the same meaning, and since they are not synonymous, \(L\) is not compositional. This suggests the following alternative definition of compositionality:

- A meaning theory \(\llbracket \cdot \rrbracket\) for language \(L\) is compositional iff for all complex expressions \(\epsilon_1, \epsilon_2\) such that \(\epsilon_1 = \delta(\eta_1, \ldots, \eta_n)\) and \(\epsilon_2 = \delta(\theta_1, \ldots, \theta_n)\), where for each \(i\)
  \[\llbracket \eta_i \rrbracket = \llbracket \theta_i \rrbracket,\]
  we have
  \[\llbracket \llbracket \epsilon_1 \rrbracket = \llbracket \epsilon_2 \rrbracket.\]

\(^{17}\)This can be denied. Triadic theories of attitude verbs typically take sentences like (7) and (8) to contain additional semantic information (some sort of guise or mode of presentation) which is not syntactically overt, and which differs between (these occurrences of) the two sentences.

\(^{18}\)Partly for this reason that Millians are under pressure to accept that proper names intersubstitute \textit{salva significatio} in propositional attitude contexts.
A language is compositional, that is, if substitution of synonyms always preserves synonymy.  

Like the functional definition of compositionality, the substitutional definition can be relativized to a parthood relation, a level of structural analysis, and a collection of possible extensions of the language. Relativization to a range of admissible meaning functions, however, comes less naturally to the substitutional definition. Abstracting away from the various dimensions of relativization, the functional and substitutional definitions turn out to be equivalent:

**Functionalism implies substitutionalism:** Suppose \( L \) is compositional in the functional sense, and let \( \Sigma(\alpha), \Sigma(\beta) \) differ only by substitution of \( \alpha \) with the synonymous \( \beta \). Some function \( G \) gives the meanings of both \( \Sigma(\alpha) \) and \( \Sigma(\beta) \) from the meanings of their parts and their syntactic structure. But by assumption, each has the same syntactic structure, and each has parts of the same meaning, so \( G \) must assign to each the same meaning. Thus substitution of synonyms preserves synonymy.

**Substitutionalism implies functionalism:** Suppose \( L \) is compositional in the substitutional sense. Suppose \( L \) were not compositional in the functional sense. Then there would be some syntactic rule and some collection of part meanings such that two dif-

---

19 Some care is needed in specifying what counts as the substitution of synonyms. Szabo ([64], 16-18) gives the example of the prima facie synonymous sentences:

(9) Plato was bald.
(10) Baldness was an attribute of Plato.

which do not always preserve meaning when one is replaced for the other, as in:

(11) The philosopher whose most eminent pupil was Plato was bald.
(12) The philosopher whose most eminent pupil was baldness was an attribute of Plato.

If the substitutional conception of compositionality is to be a reasonable one, it must not count (12) as a substitution instance of (11) with (9) swapped for (10). Substitution thus cannot be simple replacement of strings of words, but must respect syntactic structure.
ifferent expressions formed from those part meanings had different meanings. Let $\Sigma(\alpha)$ and $\Sigma(\beta)$ be two such expressions.\(^{20}\) Then $\alpha$ and $\beta$ have the same meaning, but $\Sigma(\alpha)$ and $\Sigma(\beta)$ do not, violating substitutional compositionality. Thus $L$ is compositional in the functional sense.\(^{21}\)

## 2 The Mathematics of Compositionality

With a precise definition of compositionality in hand, we consider the implications of that definition. Before turning to the empirical, in §3 below, we begin with some \textit{a priori} issues. First we examine arguments suggesting that compositionality is a surprisingly weak constraint, one that any meaning theory can meet. Next we turn to an argument in the other direction, one claiming that compositionality is a surprisingly strong constraint – one which (under minimal assumptions) fixes the meanings of all subsentential expressions in a language.

### 2.1 The Weakness of Compositionality

Given the substitutional construal of compositionality, it is a simple matter to construct a noncompositional meaning theory – simply assign the same meaning to two lexical items, and different meanings to two complex expressions differing only via intersubstitution of the now-synonymous lexical items. Despite the ready availability of non-compositional semantic theories, some recent work has claimed that composition-

\(^{20}\)This step illicitly supposes that the two expressions differ only by intersubstitution of a \textit{single pair of} synonymous parts. A more careful approach would either define substitutionalism to allow replacement of arbitrary numbers of synonymous parts, or add additional assumptions in the manner of [32] to guarantee that large-scale substitutions could be achieved by a sequence of single-pair substitutions.

\(^{21}\)See [32], Theorem 4, for more precise versions of each direction of this proof.
ality is a trivial requirement, typically in the sense that all possible languages can be made compositional. Making a language compositional involves changing its meaning function in some way to achieve compositionality; care is needed here to distinguish between adapting an existing meaning theory and giving a wholly new meaning theory (the latter of which can, of course, always be done compositionally). Since compositionality requires that three factors – the meanings of atoms, the meanings of complexes, and the parthood relation between atoms and complexes – be properly aligned, there are correspondingly three ways of making a meaning theory compositional. We set out one of these three ways in detail, and then sketch the thought behind the other two.

1. **Tinkering with the Bottom:** Let \( L \) be an arbitrary language, consisting of a set \( L \) of expressions closed under the syntactic operation of concatenation.\(^{22}\) Suppose \( \llbracket \cdot \rrbracket \) is a meaning function, assigning members of \( L \) to various meanings. \( \llbracket \cdot \rrbracket \) can be wholly arbitrary, and as noncompositional as desired. Nevertheless, Zadrozny, in [72], claims that \( L \) can be given a compositional analysis, and that, more generally:

   We can prove a theorem stating that any semantics can be encoded as a compositional semantics, which means that, essentially, the standard definition of compositionality is formally vacuous. ([72], 329)

How is this possible? Consider a small example using a minimally noncompositional language. Let \( L \) contain atomic elements \( a, b, \) and \( c \), and the following meaning theory:

- \( \llbracket a \rrbracket = 1 \)
- \( \llbracket b \rrbracket = 1 \)

\(^{22}\)I follow [72] in imposing an inessential restriction to languages with only a single syntactic operation.
• \( \|c\| = 2 \)
• \( \|a \sim c\| = 3 \)
• \( \|b \sim c\| = 4 \)

Thus fails the substitution test, and is not compositional. Zadrozny, however, gives us a procedure for producing a compositional meaning function \( \mu \) “which agrees with the function \( \langle \cdot \rangle \)” ([72], 330). Zadrozny’s \( \mu \) function satisfies the following two constraints:

- For all \( \epsilon, \mu(\epsilon)(\epsilon) = \|\epsilon\| \)
- For all \( \epsilon, \eta, \mu(\epsilon \sim \eta) = \mu(\epsilon)(\mu(\eta)) \)

The first constraint gives the promised agreement with \( \langle \cdot \rangle \), while the second gives compositionality. Finding \( \mu \) is then a matter of solving a system of equations:

- \( \mu(a) = \langle a, 1 \rangle, \langle \mu(c), \mu(a \sim c) \rangle \)
- \( \mu(b) = \langle b, 1 \rangle, \langle \mu(c), \mu(b \sim c) \rangle \)
- \( \mu(c) = \langle c, 2 \rangle \)
- \( \mu(a \sim c) = \langle a \sim c, 3 \rangle \)
- \( \mu(b \sim c) = \langle b \sim c, 4 \rangle \)

The two unsolved terms resolve to:

- \( \mu(a) = \langle a, 1 \rangle, \langle \langle c, 2 \rangle, \langle a \sim c, 3 \rangle \rangle \)
- \( \mu(b) = \langle b, 1 \rangle, \langle \langle c, 2 \rangle, \langle b \sim c, 4 \rangle \rangle \)

The resulting \( \mu \) meets both the agreement and the compositionality constraints.

More generally, \( \mu \) is constructed from \( \langle \cdot \rangle \) by solving the set of simultaneous equations given by \( \mu(\epsilon) = \langle \epsilon, \|\epsilon\| \rangle \cup \{ \langle \mu(\eta), \mu(\epsilon \sim \eta) : \epsilon \sim \eta \in \mathbb{L} \} \).\(^{23}\)

\(^{23}\)A solution is always available in the set theory AFA, with an anti-foundation axiom (see, e.g., [1] for formal details on AFA). The use of AFA is essential for Zadrozny’s result – the system of equations lack a
The compositional \( \mu \) function matches the noncompositional \( m \) function in that \( \mu \), when applied to an expression \( \epsilon \) produces not the \( \llbracket \cdot \rrbracket \)-meaning of \( \epsilon \), but another function which, when \( \mu \) is applied to \( \epsilon \), produces \( \llbracket \epsilon \rrbracket \). Thus, in the above example, \( \mu(\epsilon) \neq \llbracket \epsilon \rrbracket \) for every expression in the language. The resulting sense of “agreement” is thin, and fails to preserve synonymies – \( \llbracket a \rrbracket = \llbracket b \rrbracket \), but \( \mu(a) \neq \mu(b) \). In fact, the construction of \( \mu \) guarantees that no distinct terms of a language \( L \) have the same \( \mu \)-meaning. But if \( \mu \) allows no synonymies, then it is trivially compositional by way of satisfying the substitution test. Once this point is seen, it becomes obvious that there are many ways of building a \( \mu \) which (a) is compositional and (b) allows extraction of \( \llbracket \cdot \rrbracket \)-meanings. For example:

- Define \( \mu(\epsilon) = \{ \langle \eta, \llbracket \eta \rrbracket \rangle : \eta \in L \} \) for all \( \epsilon \in L \). Then the \( \mu \)-meaning of each term in the language encodes the \( \llbracket \cdot \rrbracket \)-meanings of every term and compositionality is trivial.

- Define \( \mu(\epsilon) = \langle \llbracket \epsilon \rrbracket, \epsilon \rangle \). Encoding of \( \llbracket \cdot \rrbracket \)-meanings is obvious, and compositionality is guaranteed because no expressions are \( \mu \)-synonymous.\(^{24}\)

If part meanings can be set however one pleases (and requiring that some pretheoretic part meanings be somehow or other encoded in the theoretically dictated part meanings represents no significant deviation from perfect freedom), then achieving compositionality is no trick. Apparently noncompositional behaviour at higher levels can be finessed by ‘value loading’ lexical items with information about the desired semantic behaviour of complexes formed out of them.\(^{25}\) If,

\(^{24}\) See [11] for more details on alternative ways of producing Zadrozy’s result.

\(^{25}\) Thus, for example, if even a single lexical item is left wholly unconstrained in its meaning, then a semantic theory can be given such that all sentences which have that lexical item as a part behave compositionally,
however, the meaning function must meet significant constraints on the meanings of the atoms, then this route to cheap compositionality is blocked.²⁶,²⁷

2. Tinkering with the Top: Consider a syntax \( L \), complete with atomic and complex expressions and a parthood relation between the two. Suppose you are given meanings for all of \( L \)'s atomic expressions, but are told that the meanings of the complex expressions are entirely up to you, and then are asked to give a compositional meaning theory for all of \( L \). Under these conditions, your task is trivial. Since compositionality requires properly aligning part meanings and complex meanings via parthood structure, and since one variable in this equation has been left unconstrained, solutions abound. For example, you could assign the same meaning to every complex expression, guaranteeing satisfaction of the substitution principle. Given complete freedom to set the meanings of complexes, those meanings can always be made functions of any meanings of atomic expressions. Complete freedom to tinker at the top, then, gives a clear case in which compositionality places no constraint on meanings — as before, compositionality is a tug-of-war between the meanings of parts and the meanings of complexes, and if one side lets go the rope, the war is easily won. The formal point is straightforward, but it also threatens to be uninteresting. One cannot, in just by having that item carry case-by-case instructions on what each sentence is to mean. Value loading, of coarse, need not be so crude, and is not always inappropriate in a semantic theory. The proposal in the introduction for treating 'large' as a function from \( \overline{N} \) extensions to extensions is a more refined version of value-loading, with 'large' carrying upward instructions on how to interact with different modified nouns.

²⁶One obvious example: if the meaning function is required to respect pretheoretic judgements of synonymy, then compositionality is impossible if the distribution of such pretheoretic judgements violates the substitution principle.

²⁷See [11], [36] [43], and [70] for additional discussion of Zadrozny.
general, construct a useful semantic theory via utter indifference to the meanings of complex expressions.

Because pre-theoretic views on the meanings of sentences are typically more robust than those on the meanings of individual words, tinkering with the top is not a common strategy for achieving compositionality. Arguably Horwich’s claim that ‘the compositionality of meaning places no constraint at all on how the meaning properties of words are constituted’ ([33], 154) rests on such tinkering, backed by a minimalist view of sentence meaning evidenced in comments such as ‘once one has worked out how a certain sentence is constructed from primitive elements, and provided one knows the meanings of those elements, then, automatically and without further ado, one qualifies as understanding the sentence’ ([33], 155). The more common interpretation, however, is that Horwich fails to account for the role of the variable mode of composition in determining the meaning of the sentence.²⁸

3. **Tinkering in Between**: Let \( L \) be a language with expressions \( L \) and a fixed meaning function \( \mu \). If the parthood relation \( \sqsubseteq \) is subject to no pretheoretic constraints, then it can always be chosen so as to make \( L \) compositional.²⁹ Most trivially, this can be done by setting \( \sqsubseteq \) to be the identity relation on \( L \). Alternatively, \( \sqsubseteq \) can be chosen in any way such that no two expressions have the same parts. Pretheoretic judgements about parthood can be partially respected: suppose that associated with each expression \( \epsilon \) are two sets \( \sqsubseteq^+\epsilon \) and \( \sqsubseteq^-\epsilon \) of expressions which, pretheoretically, are and are not (respectively) parts of \( \epsilon \). If, for all \( \epsilon \in L \),

---

²⁸For more detailed criticisms of Horwich’s views on trivial compositionality, see [20] and [25].

²⁹See [36] and [35]. Janssen proves the stronger result that if the parthood relation is recursively enumerable and the meaning theory \( m \) is computable, then the composition function is also computable.
then the pretheoretically-given constraints on parthood can always be extended to a parthood relation $\sqsubseteq$ making the language compositional. If $L$ is allowed to contain an arbitrary set of expressions whose $\sqsubseteq$-behaviour is unconstrained by pretheoretic judgements (corresponding, perhaps, to theoretical entities of syntax), then $\sqsubseteq$ can always be chosen to make $L$ compositional.

Compositionality marks the convergence of three components of a semantic theory: the assignment of meaning to the parts, the assignment of meaning to the complexes, and the parthood relation between parts and complexes. It is thus only as stringent a requirement as the prior constraints on these three components are robust. Should any of the three be arbitrary (i.e., available for free choice by the theory constructor), then the satisfaction of compositionality is no task. If the intended use of the principle of compositionality is as a tool to choose among a range of otherwise acceptable semantic theories, then that range of theories needs to be dictated by rich views, from whatever source, on what sentences mean, what words mean, and what the grammar of the language is.

### 2.2 The Strength of Compositionality

Having seen the weakness of the compositionality constraint in the ways in which it can be trivially satisfied, we now examine the strength of that constraint by setting out a surprising consequence of it due Hodges ([32]).

Suppose we have a set $L$ of expressions, and two meaning functions $\mu_1$ and $\mu_2$ whose domains are (possibly

\[ [\eta : \zeta L_{\eta} \equiv \zeta L_{\eta} \wedge \zeta L_{\eta} \equiv \zeta L_{\eta} \mid \leq [\mu \mid \zeta L_{\eta} \cup \zeta L_{\eta} ] \]

30 The following discussion is a simplified presentation of Hodges’ results. The main point of simplification is that I ride roughshod throughout over Hodges’ distinction between *compositional* and *1-compositional*. The two are equivalent under the assumption that the meaning functions meet a condition Hodges calls being *Husserlian*. I thus simply presuppose (rather unjustly, given Hodges’ care in highlighting the importance of the condition) that all meaning functions are involved are Husserlian.
improper) subsets of \(L\). \(\mu_1\) and \(\mu_2\) can thus be thought of as specifying two languages whose expressions are drawn from, but may not exhaust, \(L\). Then \(\mu_2\) is a Fregean cover of \(\mu_1\) if the following two conditions are met:

1. If \(\mu_2(\epsilon) = \mu_2(\tau)\) then \(\Sigma(\epsilon)\) is in the domain of \(\mu_1\) if \(\Sigma(\tau)\) is, and if both are, then \(\mu_1(\Sigma(\epsilon)) = \mu_1(\Sigma(\tau))\). (\(\mu_2\)-synonyms are \(\mu_1\)-intersubstitutable; hence \(\mu_1\) is ‘almost’ compositional; \(\mu_1\)-meanings of complexes are functions of \(\mu_2\)-meanings of parts.)

2. If \(\mu_2(\epsilon) \neq \mu_2(\tau)\), then either:
   - There is some \(\Sigma\) such that \(\Sigma(\epsilon)\) is \(\mu_1\)-meaningful and \(\Sigma(\tau)\) is not, or vice versa.
   or:
   - There is some \(\Sigma\) such that \(\Sigma(\epsilon)\) and \(\Sigma(\tau)\) are both \(\mu_1\)-meaningful but \(\mu_1(\Sigma(\epsilon)) \neq \mu_1(\Sigma(\tau))\)

Terms differ in \(\mu_2\)-meaning only if complex terms which differ only by substitution of the starting terms themselves differ in \(\mu_1\)-behaviour; this is intended to be a version of Frege’s Context Principle:

never to ask for the meaning of a word in isolation, but only in the context of a proposition. ([22], x)

\[31\]

\[31\] The context principle is frequently taken to be in tension with, the principle of compositionality, because it suggests that meanings of wholes are prior to meanings of parts, and thereby threatens to make nonsense of the compositional idea that meanings of wholes are built out of meanings of parts. See [13], 3-5 for one influential attempt to reconcile Frege’s apparent commitment to both context and compositionality, and [56] and [37] for more recent discussions.
A Fregean cover of a meaning function $\mu$ is another meaning function which assigns two expressions different meanings when and only when intersubstitution instances of those two expressions differ in $\mu$-meaning. Suppose, for example, that $\mu$ assigns meanings to all of the sentences in $\mathbb{L}$. A Fregean cover of $\mu$ whose domain included subsentential expressions of $\mathbb{L}$ would treat subsentential meanings as driven by the need to account for the sentential data, assigning different subsentential meanings only when there was a corresponding difference in sentential behaviour.

Fregean covers are always compositional. Suppose $\mu_2$ is a Fregean cover of $\mu_1$. Then if $\mu_2$ violates the substitution principle, $\mu_1(\epsilon) = \mu_2(\tau)$ and $\mu_2(\Sigma(\epsilon)) \neq \mu_2(\Sigma(\tau))$ for some $\epsilon$, $\tau$, and $\Sigma$. But then, since $\mu_2$ is a Fregean cover, there must be some larger context $\Omega$ such that $\mu_1(\Omega(\Sigma(\epsilon))) \neq \mu_1(\Omega(\Sigma(\tau)))$ (or one of the two is not in the domain of $\mu_1$). But terms synonymous under a Fregean cover must be intersubstitutable according to the covered meaning function, so this is a contradiction. Thus $\mu_2$ is compositional.

Every meaning function has a Fregean cover. Let $\mu_1$ be an arbitrary meaning function on a subset $D_1$ of language $\mathbb{L}$. Call $\epsilon$ and $\tau$ co-categorical if $\Sigma(\epsilon)$ is $\mu_1$-meaningful iff $\Sigma(\tau)$ is. Then define $\mu_2$ as follows:

- $\mu_2(\epsilon) = \{ \tau : \tau$ and $\epsilon$ are co-categorical for $\mu_1$ and $\mu_1(\Sigma(\epsilon)) = \mu_1(\Sigma(\tau))$ whenever $\Sigma(\epsilon)$ is in $D_1 \}$

Then $\mu_2$ is a Fregean cover for $\mu_1$:

- Suppose $\mu_2(\epsilon) = \mu_2(\tau)$, and $\Sigma(\epsilon)$ is in $D_1$. $\epsilon \in \mu_2(\epsilon) = \mu_2(\tau)$, so $\epsilon$ and $\tau$ are co-categorical. Thus $\Sigma(\tau)$ is in $D_1$. By construction, $\mu_1(\Sigma(\epsilon)) = \mu_1(\Sigma(\tau))$.

- Suppose $\mu_2(\epsilon) \neq \mu_2(\tau)$. Then there is some $\eta \in \mu_2(\epsilon)$ but not $\eta \in \mu_2(\tau)$ (or vice versa). Then either $\eta$ and $\tau$ are not co-categorical for $\mu_1$, or there is some $\Sigma$ such that $\mu_1(\Sigma(\eta)) \neq \mu_1(\Sigma(\tau))$.

  - If the former, then $\epsilon$ and $\tau$ are not co-categorical for $\mu_1$ and there is some $\Sigma$ such that $\Sigma(\epsilon)$ and $\Sigma(\tau)$ are not both $\mu_1$-meaningful.
If the latter, then since $\mu_1(\Sigma(\eta)) = \mu_1(\Sigma(\epsilon))$, we have $\mu_1(\Sigma(\epsilon)) \neq \mu_1(\Sigma(\tau))$.

In either case, $\mu_2$ meets the requirements for a Fregean cover.

A Fregean cover for a given $\mu$, however, need bear little similarity to $\mu$. Suppose $\mu$ is a meaning function defined on a subset of the closure of $a, b, c$, and $d$ under concatenation, fully characterized by:

\[
\begin{align*}
\mu(a \sim b) &= 1 & \mu(a \sim d) &= 2 & \mu(c \sim b) &= 2 \\
\mu(c \sim d) &= 1 & \mu((a \sim b) \sim b) &= 3 & \mu((a \sim b) \sim d) &= 3 \\
\mu((a \sim d) \sim b) &= 4 & \mu((a \sim d) \sim d) &= 5 & \mu((c \sim b) \sim b) &= 4 \\
\mu((c \sim b) \sim d) &= 5 & \mu((c \sim d) \sim b) &= 3 & \mu((c \sim d) \sim d) &= 5
\end{align*}
\]

Constructing a Fregean cover for $\mu$ defined on the closure of $\mu$'s domain under the part-hood relation yields:

\[
\begin{align*}
\mu_1(a) &= \{a\} & \mu_1(b) &= \{b\} \\
\mu_1(c) &= \{c\} & \mu_1(d) &= \{d\} \\
\mu_1(a \sim b) &= \{a \sim b\} & \mu_1(a \sim d) &= \{a \sim d, c \sim b\} \\
\mu_1(c \sim b) &= \{a \sim d, c \sim b\} & \mu_1(c \sim d) &= \{c \sim d\} \\
\mu_1((a \sim b) \sim b) &= \{(a \sim b) \sim b, (a \sim b) \sim d, (c \sim d) \sim b\} \\
\mu_1((a \sim b) \sim d) &= \{(a \sim b) \sim b, (a \sim b) \sim d, (c \sim d) \sim b\} \\
\mu_1((a \sim d) \sim b) &= \{(a \sim d) \sim b, (c \sim b) \sim b\} \\
\mu_1((a \sim d) \sim d) &= \{(a \sim d) \sim d, (c \sim b) \sim d, (c \sim d) \sim d\} \\
\mu_1((c \sim b) \sim b) &= \{(a \sim d) \sim b, (c \sim b) \sim b\} \\
\mu_1((c \sim b) \sim d) &= \{(a \sim d) \sim d, (c \sim b) \sim d, (c \sim d) \sim d\} \\
\mu_1((c \sim d) \sim b) &= \{(a \sim b) \sim b, (a \sim b) \sim d, (c \sim d) \sim b\} \\
\mu_1((c \sim d) \sim d) &= \{(a \sim d) \sim d, (c \sim b) \sim d, (c \sim d) \sim d\}
\end{align*}
\]

The crucial point is the pattern of synonymies, so any meaning function (on the requisite domain) which makes synonymous the groups $a \sim d$ and $c \sim b; (a \sim b) \sim b,$ $\ldots$ and $(c \sim d) \sim b$; $a \sim d$ and $c \sim b; (a \sim d) \sim b$ and $(c \sim b) \sim b; (a \sim d)$ and $(c \sim d)$, $(c \sim b) \sim d$, and $(c \sim d) \sim d$ is a Fregean cover of $\mu$. But the synonymies of the Fregean cover needn’t be the same as those of the covered function: $\mu_1$ denies the
synonymy that $\mu$ asserts between $a \sim b$ and $c \sim d$.

However, if $\mu$ is compositional, then the relation between it and its Fregean covers is more intimate. We then have Hodges’:

(Theorem 2) Suppose $\mu_1$ is a meaning function on some subset $D_1$ of $\mathbb{L}$. Then $\mu_1$ is compositional if and only if there is a Fregean cover $\mu_2$ on

$$D_2 = \{ e \in \mathbb{L} : \exists \eta \in D_1 \ e \sqsubseteq \eta \}$$

such that for all $\epsilon \in D_1, \mu_2(\epsilon) = \mu_1(\epsilon)$.

**Proof**: Left to right: Let $\mu$ be an arbitrary Fregean cover of $\mu_1$. Suppose $\epsilon$ and $\tau$ are $\mu_1$-synonymous. Since $\mu_1$ is compositional, $\Sigma(\epsilon)$ and $\Sigma(\tau)$ are $\mu_1$-synonymous for all $\Sigma$, and any Fregean cover must also treat $\epsilon$ and $\tau$ as synonymous. If $\epsilon$ and $\tau$ are not $\mu_1$-synonymous, then $\mu$ cannot treat them as synonyms either, given the second clause of the definition of Fregean cover, with $\Sigma$ the null context.\(^{32}\) Thus $\mu$ must have exactly the same synonymies as $\mu_1$ on $D_1$. Now define:

$$\mu_2(\epsilon) = \begin{cases} 
\mu_1(\epsilon) & \epsilon \in D_1 \\
\mu(\epsilon) & \text{otherwise}
\end{cases}$$

$\mu_2$ is then a Fregean cover of $\mu_1$ agreeing with $\mu_1$ on $D_1$. Right to left: Since $\mu_2$ is a Fregean cover, it is compositional. Hence it satisfies the substitution constraint. Since $\mu_1 \subseteq \mu_2 \mu_1$ also satisfies the substitution constraint, and is compositional. $\therefore$

Suppose a semanticist seeks a meaning theory for language $\mathbf{L}$. Perhaps he is given meanings for all of the sentences of $\mathbf{L}$, and needs an assignment of meanings to individual lexical items compatible with the given sentential meanings. General worries about the underdetermination of theory by data might have led us to suspect that there would be many ways for the semanticist to complete his task. However, Theorem 2 shows that if $\mathbf{L}$ is compositional on the sentential level\(^{33}\) and the semanticist wants the

\(^{32}\)A Fregean cover can only break, not introduce synonymies.

\(^{33}\)This is a non-trivial requirement if sentences can contain sentences as proper parts.
lexical meanings to be well-fitted to the sentential meanings (in the sense of satisfying the Context Principle), then there is, up to equivalence, only a single meaning function available, and a compositional one. The puzzle of semantics has an almost unique solution.\footnote{See \cite{67} for an application of Hodges’ result along these lines.}

On reflection, however, Hodges’ result is less surprising, and hence also less significant, than it might at first seem. A commitment to constructing meaning theories in accord with the Context Principle carries with it a commitment to distributing meanings among expressions in exactly one pattern – not making so few distinctions in meaning as to violate compositionality, and not making so many distinctions in meaning as to differentiate expressions which contribute in the same way to complex expressions. Inevitably, then, meaning theories constructed in accord with the Context Principle are unique up to equivalence. If the starting fragment is compositional, then it fits into the semantic agenda dictated by the Context Principle, so the final pattern of synonymies contains the starting pattern as a fragment. But our goals in semantic theorizing may outstrip the commitments of the Context Principle. Suppose, for example, that, in attempting to extract ontological commitments from our semantic practice, we come to wonder whether the lexical item ‘gavagai’ should be understood as meaning \textit{rabbit} or \textit{undetached rabbit part}. Given meanings for all of the sentences, adherence to the Context Principle dictates what ‘gavagai’ is synonymous with, but does not tell us what it means. The mere synonymy information does nothing to settle the ontology acquired via the use of the term ‘gavagai’. Clearly, many conceptions of semantic theory will regard this only as the barest of starts toward construction of an adequate theory.
3 Some Problem Cases for Compositionality

While the results of §2.1 show that when semantic theories are sufficiently unconstrained, compositionality can be cheaply obtained, the question remains whether a *satisfactory* semantic theory for a natural language, one properly responsive to natural constraints on semantic and syntactic facts, can be given compositional form. In this section we consider two problem cases for the construction of compositional semantics, examining the data which resist a compositional treatment and then considering ways of overcoming that resistance. The goal is not to settle the question of whether natural languages have compositional semantics, or even the smaller questions of whether the particular phenomena discussed here have a compositional semantics, but rather to see how questions of compositionality influence semantic theorizing.\(^{35}\)

3.1 What the Hell

The Problem: Compare the following two sentences:

(13) Who bought that book?

(14) Who the hell bought that book?\(^{36}\)

\(^{35}\)The literature abounds in problem cases for compositionality and treatments thereof. In addition to the issues addressed below, see among many others discussion of compositionality and: independence-friendly logics in [28], [29], [30], [31], [32]; prototype theory in [16] (ch. 5), [19], [18], [39], [53], [62]; idioms in [42], [41], [52], [71]; ‘unless’ in [27], [54], [66]; propositional attitude cases in innumerable places, but especially [60] (ch. 4), [61] (ch. 8), [46], [8]; ‘any’ and other negative polarity items in [48], [47], [4], [50], [45]; anaphora in [38], [24], and [34]. [36] also contains an overview of several compositionally-problematic semantic phenomena.

\(^{36}\)These two examples are drawn from [10]. The behaviour of ‘the hell’ phrases was first noted in [58].
The two are roughly synonymous. While the addition of ‘the hell’ alters the rhetorical impact of (14) (perhaps encouraging the conversational implicature that it is surprising that the book was bought), the core semantic value of each is a request for information about the identity of a book buyer. Whatever the semantic contribution of ‘the hell’, it must be compatible with the close semantic proximity of (13) and (14).

However, ‘who’ and ‘who the hell’, or, more generally, ‘wh-’ and ‘wh- the hell’ expressions, diverge in meaning in other contexts. Thus:

- The minimal variant of adding a modal auxiliary causes a difference to emerge:

(15) Who would buy that book?

(16) Who the hell would buy that book?

The first of these is most naturally read as a request for information, but the second is most naturally, and perhaps obligatorily, read as an indirect assertion that nobody would buy that book.38

- When the original examples are embedded in an indirect question, a difference in grammaticality emerges:

  – I know who bought that book.

  – *I know who the hell bought that book.

‘Who the hell’ phrases are grammatical only in negative contexts, whether overt:

37The first three of these examples are drawn from [10]; the last draws from [10] and [58].

38[10] claims that only the indirect assertion reading of (16) is available, but I find the data less univocal. The pressure toward the indirect assertion reading, in my judgement, increases with the strength of the attached vulgarity. In order to keep this volume suitable for a family audience, I have used ‘the hell’ throughout, but the reader is encouraged to substitute as his imagination allows.
– I don’t know who the hell bought that book.

in the antecedent of a conditional:

– If anyone knows who the hell bought that book, please tell me.

or in the scope of so-called *adversative attitude verbs*:39

– John refused to tell me who the hell bought that book.

• ‘The hell’ blocks certain scope readings of sentences with multiple quantifiers. Thus:

– What did everyone buy for Max?

is ambiguous between a reading on which ‘everyone’ takes wide scope, and people make separate purchases for Max, and a reading on which ‘everyone’ takes narrow scope, and there is some one thing bought by everyone for Max. However:

– What the hell did everyone buy for Max?

allows only the second of these two readings.

• ‘Wh- the hell’ phrases, unlike normal ‘wh-’ phrases, cannot enter into anaphoric attachments. Thus:

– Someone, walked in the park, but I don’t know who.

is acceptable, but:

– *Someone, walked in the park, but I don’t know who the hell.*

is not. Similarly, ‘which’ phrases, which require an anaphoric link to a contextually provided range of salient objects, do not allow ‘the hell’ modification:

---

39See [47].
Which the hell book did you read that in?

The puzzle for compositional semantics is to show how ‘the hell’ can systematically contribute to the meanings of larger expressions in a way that allows its impact to be minimal, if anything at all, in (14), but much greater in the other cases set out above.\footnote{A compositional semantics need account for the failures of \textit{grammaticality} such as ‘I know who the hell bought that book’ and ‘Which the hell book did you read that in’ only if the syntax is not thoroughly autonomous. It is tempting to think that grammatical failures due to failures of anaphoric linkage, at least, have a semantic explanation.}

A Solution: A simple ‘who’ question can have its interpretation influenced by linking the range of admissible answers to a contextually-provided domain. Thus consider the following dialogue:

(17) A: Various friends of mine voted for each of the different presidential candidates in the 2000 election.

B: Really? Who voted for David McReynolds?

B’s question is not answered by specifying an arbitrary McReynolds voter (and does not require listing \textit{all} such voters); rather, it calls for a (or all) McReynolds voters \textit{among A’s friends}. In another context, however, ‘Who voted for David McReynolds?’ can receive an unlinked reading, in which it calls for the total list of McReynolds voters.

Suppose the semantic function of ‘the hell’ is to require that the range of admissible answers to a wh-question include \textit{novel answers} – ones not already provided as possible by contextual linkages of the sort just discussed. When a wh-question is an unlinked one, as on one natural reading of (13), adding ‘the hell’ has no effect, because when unlinked, all answers are novel. But when the wh-question is a linked one, adding ‘the hell’ has a semantic impact. Thus consider:

\footnote{The following solution is a simplified and modified version of the proposal of [10]. Any shortcomings of it are due to the present alterations.}
Various friends of mine voted for each of the different presidential candidates in the 2000 election.

B: Really? Who the hell voted for David McReynolds?

This dialogue, unlike the first, creates the implicature that B expects all of A’s friends not to have voted for McReynolds. If the effect of adding ‘the hell’ is to insist on the admissibility of novel answers (here, people other than those B counts as A’s friends), this new implicature is to be expected. The various effects of ‘the hell’ noted above now fall out:

- ‘Wh- the hell’ phrases refuse anaphoric linkage because that linkage dictates the range over which the wh- phrase ranges, which contradicts the novelty requirement imposed by ‘the hell’.42 ‘Which’ phrases, which always require anaphoric/contextual linkage, can thus never combine with ‘the hell’.

- A question of the form ‘Who would buy that book?’ takes as answer pairs of people and possible situations.43 Given the broad total range of possible situations, such a question is typically linked to a contextually-provided range of admissible situations.44 Adding ‘the hell’ to form ‘Who the hell would buy that book?’ requires the admissibility of novel answers, and thus defeats any contextually-provided restriction on admissible situations. But once all possible situations are provided, the question becomes trivialized: anyone would, in some situation, buy the book. The asking of trivial questions, though, is pragmatically proscribed.

42 Although note the acceptability of:

- Someone walked in the park, but I don’t know who the hell it was.

43 Thus: ‘Albert, if it has a chapter on direct reference’, ‘Louisa, if autographed copies are available’, etc.

44 Thus ruling out answers such as ‘Brian, if we threaten to kidnap his dog if he doesn’t’.

29
and an alternative communicative explanation is favoured, such as the explanation that the speaker is emphasizing the remoteness of any situation in which the book is bought.

- The requirement of novelty imposed by ‘the hell’ is impossible to fulfill when the ‘wh- the hell’ phrase is simply imbedded in an operator of positive epistemic commitment. To say that I know who the hell bought the book is to undermine, by my knowledge, the requisite novelty of the admissible book buyers. Similarly an epistemically positive operator in the antecedent of a conditional, such as the earlier:

  – If anyone knows who the hell bought the book, please tell me.

creates no conflict with the novelty requirement, since the function of the antecedent is to entertain hypothetical situations. Operators of negative epistemic commitment, such as ‘refused to tell’, for similar reasons allow ‘the hell’ modification. The novelty requirement thus explains the distributional facts noted above.

- The novelty requirement makes ‘wh- the hell’ phrases negative polarity items, where various sorts of negation license the introduction of novelties. Suppose that negative polarity items are subject to:

  (Immediate Scope Constraint) A negative polarity item can appear only in the immediate scope of its licensing negative item.45

Consider again:

  – What the hell did everyone buy for Max?

45 See [50] for formulation and defense of the Immediate Scope Constraint.
and assume that the licensing item is the marker of interrogative force. If ‘everyone’ is raised to give it scope over ‘what the hell’, it intervenes between ‘what the hell’ and its licenser, violating the Immediate Scope Constraint. The unavailability of a reading wide-scoped for ‘everyone’ is thus explained.

### 3.2 Many Scandinavians

**The Problem:** Fourteen Scandinavians have won the Nobel prize in literature. Since there have been only 99 Nobel laureates in literature, and since Scandinavians are only about 0.5% of the world’s population, the following claim looks acceptable:

(19) Many Scandinavians have won the Nobel prize in Literature.

On reflection, however, the acceptability of (19) is puzzling. Fourteen, after all, is not by most natural standards *many*. Consider the oddity of:

(20) Many Scandinavians have emigrated to the United States. Fourteen, in fact.

The acceptability of (19) seems to derive from having fourteen qualify not as *many Scandinavians*, but rather as *many winners of the Nobel prize in literature*, and thus from reading (19) as:

(21) Many winners of the Nobel prize in literature have been Scandinavian.

A similar apparent ‘swapping of positions’ can be found in sentences with adverbs of quantification:

46Bjørnstjerne Bjørnson, Selma Lagerlöf, Verner von Heidenstam, Karl Gjellerup, Henrik Pontoppidan, Knut Hamsun, Sigrid Undset, Erik Karlfeldt, Frans Eemil Sillanpää, Johannes Jensen, Pär Lagerkvist, Nelly Sachs, Eyvind Johnson, and Harry Martinson. This observation and the subsequent problematic sentence (19) are both due to [69] in 1985. The intervening years have not been kind to the recognition of Scandinavian literature, and the number remains 14.

47See [9] and [7]
(22) Scandinavians often win the Nobel prize in literature. (The Nobel prize in literature is often won by Scandinavians.)

and in sentences with generics:\__48__:

(23) Scandinavians win the Nobel prize in literature, but Americans win the Nobel prize in economics. (The Nobel prize in literature is won (generically) by Scandinavians, but the Nobel prize in economics is won (generically) by Americans.)

Some other determiners, such as ‘few’ and ‘several’, exhibit similar behaviour. These cases look like violations of the semantic locality component of compositionality: ‘many’ in imposing its cardinality constraint requires semantic interaction with the syntactically distant verb phrase:\__49__

\__A Solution\__: We concentrate on addressing the problem in its ‘many’ form, deferring integration of these suggestions with theories of adverbs of quantification and generics. In some cases, the semantic role of ‘many’ in a sentence of the form ‘Many X’s Y’ is merely to require that the number of X’s that Y is above some minimum threshold cardinality $\kappa$.\__50__ This sort of ‘many’ can be given a straightforward compositional semantics:

\__48\See [6] and [7]\__

\__49\The apparent position swapping can be taken as suggesting that ‘many’ is symmetric, in the sense that ‘Many X’s Y’ and ‘Many Y’s X’ are equivalent. From the perspective of compositionality, however, the observation that ‘many’ is symmetric (if correct) merely restates the problem – how can ‘many’ be symmetric, given that symmetry requires giving equal semantic footing to the immediate complement of ‘many’ and a syntactically distant verb phrase?\__

\__50\The required number of Y-ing X’s will surely be vague, but the idealization to a specific $\kappa$ is harmless here.\__
• \[\text{many} = f : \wp(D) \mapsto \wp(\wp(D)), \quad f(X) = \{Y : |X \cap Y| \geq \kappa\}\]
• \[\llbracket_{DP} \text{many } X \rrbracket_{VP} Y\rrbracket = \text{true if } \llbracket Y \rrbracket \in \llbracket\text{many}\rrbracket(\llbracket X \rrbracket)\]

‘Many’ applied to a noun phrase thus yields a collection of sets, each of which contains many of the satisfiers of the noun phrase. If any of those sets is the extension of the verb phrase, the sentence is true. ‘Many’ is symmetric on this ‘cardinality’ semantics, in the sense that ‘Many X’s Y’ is equivalent to ‘Many Y’s X’, which would allow for position swapping without alteration of truth value, but no plausible value for \(\kappa\) accounts for the truth of (19).

In other cases, though, the impact of ‘many’ seems more subtle than a simple cardinality constraint. Compare the following:

(24) Many philosophers of mathematics have read Russell and Whitehead’s *Principia Mathematica*.

(25) Many Brazilians have read Russell and Whitehead’s *Principia Mathematica*.

Suppose that among the 2000 philosophers of mathematics, 800 have read the *Principia*, and that among the 180 million Brazilians, again 800 have read the *Principia*. Then (24) looks true and (25) false, which is impossible if ‘many’ simply imposes a cardinality constraint. This suggests the following alternative reading of ‘many’:

• \[\text{many} = f : \wp(D) \mapsto \wp(\wp(D)), \quad f(X) = \{Y : |X \cap Y| \geq \rho|X|\}\]
• \[\llbracket_{DP} \text{many } X \rrbracket_{VP} Y\rrbracket = \text{true if } \llbracket Y \rrbracket \in \llbracket\text{many}\rrbracket(\llbracket X \rrbracket)\]

where \(\rho\) sets the threshold percentage for manyness. Symmetry is now lost, since the complement noun phrase to ‘many’ has the privileged role of providing the number of objects a percentage of which must satisfy the verb phrase. This second, ‘proportionate’ reading thus also gives the wrong analysis of (19).

On both the cardinality and the proportionate analyses, if the number of X’s who Y is the same as the number of Z’s who W, then many X’s Y if and only if many Z’s W. However, some examples fail to fit this pattern:
(26) Many penguins live in Antarctica.

(27) Many penguins live in my bedroom.

Twelve penguins in a bedroom suffices for many, but will hardly do for a whole continent. Here ‘many’ means something like more than one would expect, where expectations are set (in part) by the verb phrase. This ‘expectation’ semantics can be roughly characterized by requiring that X’s Y at a rate greater than the general rate of Y-ing:

- \[ \text{many} = f : \varphi(D) \mapsto \varphi(\varphi(D)), \quad f(X) = \{ Y : |X \cap Y| \geq \frac{|X||D|}{|D|} \} \]
- \[ \text{true iff } Y \in \text{many}(X) \]

This is a crude measure of expectation, but it suffices for a start. The resulting semantics is symmetric, like the cardinality semantics, and gives the desired result for (19), since Nobel laureates in literature are represented among Scandinavians at a higher rate than that at which they appear in the general population.

The ‘expectation’ semantics uses ‘value loading’ to get the right truth conditions for (19) in a compositional manner, by granting ‘many’ a parameterized sensitivity to the verb phrase interpretation which is ‘passed up’ until semantic composition meets that part of the sentence. However, it fails to explain why the ‘position swapped’ reading of (19) seems preferable to the straight reading, given that it makes the two equivalent. It also yields undesirable results in closely related cases. Nobel laureates in literature appear in the general population at a rate of about one in every 60 million. Thus St. Lucia, with its population of some 200,000 and a single Nobel laureate in literature, dramatically exceeds the expectation threshold. But the claim:

(28) Many St. Lucians have won the Nobel prize in literature.

\[ ^{51} \text{I idealize here by assuming all Nobel laureates in literature are currently alive.} \]

\[ ^{52} \text{Derek Walcott} \]
seems false. One laureate out of the 99 is too few, no matter how few the St. Lucians are. This result is predicted if the ‘proportionate’ semantics is applied to the position swapped:

(29) Many winners of the Nobel prize in literature have been from St. Lucia.

But the position swapping is then again a problem for compositionality. Two possible moves at this point:

• Combine the ‘proportionate’ semantics with a syntactic story swapping the argument positions of ‘many’ sentences prior to semantic analysis. Thus (19), at the level of semantic analysis, would be ‘Many winners of the Nobel prize in literature have been Scandinavian’, which would then combine with the ‘proportionate’ semantics to yield to desired result. However, note that quantification over empty classes is typically pragmatically disfavoured; hence the peculiarity of:

(30) Many Freedonians have won the Nobel prize in literature.

But as [7] observes, a sentence like:

(31) Many Scandinavians have won the Nobel prize in silly walks.

seems simply false, rather than pragmatically disfavoured, despite the fact that the syntactic swapping story would make the class quantified over the empty class of Nobel laureates in silly walks.

• Give a ‘reverse proportionate’ semantics, by altering the class a minimal percentage of which needs to behave as required:

$$\|\text{many}\| = f : \wp(D) \mapsto \wp(\wp(D)), \quad f(X) = \{Y : |X \cap Y| \geq \rho|Y|\}$$

(19), under this approach, requires a certain minimal percentage of the Nobel laureates in literature to be Scandinavian. Note that this approach allows (28) to be false. However, we must now explain why
'many’ allows both proportionate and reverse proportionate semantics, while ‘most’ allows only the proportionate semantics:

- \[ \text{⟦most⟧} = f : \wp(D) \mapsto \wp(\wp(D)), \quad f(X) = \{Y : |X \cap Y| \geq \frac{|X|}{2} \} \]
- \[ \text{∗⟦most⟧} = f : \wp(D) \mapsto \wp(\wp(D)), \quad f(X) = \{Y : |X \cap Y| \geq \frac{|Y|}{2} \} \]

4 Why Might One Want Compositionality?

The examples of the previous section show that the question of whether English and other natural languages have compositional semantics is a difficult one to answer. What remains to be seen is why that question is one worth answering. In this final section, we turn to reasons for caring about compositionality. Three types of reasons will be considered:

1. Claims that natural languages are observably compositional, and hence that a semantic theory faithful to the data must take a compositional form.

2. Claims that compositionality is a consequence of or has as a consequence some other property which is of independent interest, and hence that compositionality is a derivatively desirable feature of a semantic theory, via its connection with that other property.

3. Claims that compositionality is a methodological principle for semantic theorizing, or a consequence of proper semantic methodology.

4.1 Observational Compositionality

Enough has been said already to dispel the attraction of the following naive line of reasoning:
A natural language like English must be compositional. For it to be compositional is for the meanings of its sentences to be functions of the meanings of their component words and their syntactic arrangement. But a sentence just is its component words and their syntactic arrangement, so there is nothing else that its meaning could be a function of.

This argument fails twice: once in its blindness to the possibility that the meaning of a sentence depends on non-semantic features of its constituent words (failure of semantic closure), and again in its blindness to the possibility that the meaning of a sentence depends on features (semantic or otherwise) of things which are not parts of it (failure of semantic locality). Commission of the second failure is encouraged by thinking of a semantic theory as assigning meanings to expression types, since doing so isolates the type from contexts of its instantiation, and thereby makes it difficult to identify factors relevant to its meaning outside its own parts. However, an adequate meaning theory is available on the type level only if the tokens are compositional in their semantic behaviour, so presupposing a type-level semantic theory begs the question.

More sophisticated variants on the naive argument, however, are available. Thus Fodor and Lepore ([20]) cite two aspects of our experience with language:

- **Apparent Compositionality (AC):** Practically all competent speakers of a natural language who understand all the parts of an expression $\epsilon$ also understand $\epsilon$.

- **Apparent Reverse Compositionality (ARC):** Practically all competent speakers of a natural language who understand an expression $\epsilon$ also understand all parts of $\epsilon$.

AC and ARC, Fodor and Lepore argue, are straightforwardly observable semantic facts. But they are best explained by the assumption that natural languages are compositional (in fact, strongly parallel), so by inference to the best explanation we should endorse that assumption.
AC and ARC are intended to be observable features of natural languages, but it is in fact not obvious that either is true. One can, for example, know the meaning of ‘squid’ and of ‘chair’, but still find the complex expression ‘squid chair’ obscure. However, its meaning might become clear (and clear in different ways) when imbedded in a larger context. Similarly with reverse compositionality – one can understand:

(32) Greta Garbo possessed a certain *je ne sais quoi*.

without knowing that ‘je’ means *I*, and one can understand ‘telephone’ without knowing that ‘tele’ means *far*.\(^{54}\)

Even granting AC and ARC, compositionality may not be the best explanation. Suppose a speaker of English, familiar with the meanings of ‘dogs’ and ‘bark’, understands the sentence ‘dogs bark’. Compositionality follows only if the meanings of ‘dogs’ and ‘bark’, together with syntax, are the only facts reliably available to the speaker. But, of course, they are not. Minimally, the speaker also has facts about the morphology and phonology of the words ‘dogs’ and ‘bark’, combined with facts about the context of utterance, and these facts may be pivotal in sentential interpretation. We know that *some* collection of information made available in communicative utterances suffices for understanding; the argument from AC could thus succeed only when bolstered by a further argument that only a compositional language could explain the very phenomenon of language learnability. ACR follows from the innocuous assumption that speakers of a language typically understand most of the words in that language, which in turn follows from the assumption that meanings of words are often partially determinative of meanings of sentences, which (finally) falls short of the compositional requirement that meanings of words are always fully determinative of meanings of sen-

\(^{54}\)Are such expressions idioms? (See [17].) Such a defense of ACR requires minimally a standard of idiomaticity other than the trivializing one that idioms are expressions one can understand without understanding the meanings of their parts.
Compositionality might also be argued for inductively, on the grounds that successful semantic theories have tended to be compositional. This line of argument is at most as strong as its inductive evidence base, and the examples of §3, as well as many others, cast substantial doubt on that base. Also, if semanticists tend to prefer compositional theories for reasons other than theoretical adequacy, then the sample space may be illegitimately biased toward such theories, with many potentially successful non-compositional theories never given sufficient consideration. The form of the inductive argument, however, is unobjectionable, and if the worries about the evidence quality can be addressed, it can ground a rational confidence in the compositionality of natural languages.

4.2 Consequentialist Compositionality

Suppose the compositionality of a language \( L \) is not a directly observable or inducible feature, but that there is some other feature \( X \) which \( L \) observably or inducibly possesses, or which we would like \( L \) to possess. \( X \) might be first-order, or hyperintensional, or of subject-predicate form, or systematic, or bivalent, or admitting of adherence to the Gricean maxims, or any number of other features. If compositionality can be shown to be inferentially related to feature \( X \), a reason for wanting compositionality then emerges. Two versions of this argument style are available:

1. If compositionality is a necessary condition for feature \( X \), then the presence of \( X \)

---

55Note that even if ACR is strengthened to state that speakers understand meanings of words by virtue of understanding meanings of sentences containing those words, compositionality still does not follow. The strengthened ACR remains compatible with the assumption that word meanings plus other factors determine sentence meanings and with the assumption that speakers are nascently aware of the semantic processes by which they calculate sentence meanings, even if the sentence meanings themselves weed out some of the richness of those processes.
2. If compositionality is a sufficient condition for feature \( X \), then the presence of \( X \) in \( L \) offers no guarantee that \( L \) is compositionality, but compositionality may still serve as a plausible explanation of the \( X \)-ness of \( L \), allowing for an inference to the best explanation of the compositionality of \( L \).

In the ideal case, compositionality is both necessary and sufficient for \( X \), but features so closely inferentially related to compositionality, but nevertheless independently verifiable in \( L \), are hard to come by. Both styles of argument are legitimate, but care should be taken to distinguish the two. Both styles of argument can be thought of as consequentialist – compositionality is endorsed either because of its consequences, or because of what it is a consequence of, rather than on its own merits. The crucial question in evaluating consequentialist arguments thus becomes the tightness of the inferential relation between compositionality and the chosen \( X \).

We will focus on one instance of a consequentialist argument for compositionality – the learnability argument. This overview of compositionality began with an example of a novel sentence which was immediately comprehensible to any competent speaker of English. All natural languages have an infinite number of grammatical and meaningful sentences, and \( a \textit{fortiori} \) an infinite number of such sentences never encountered by a given speaker. Nevertheless, linguistic competence gives one the capacity to understand all of these sentences. This feature of linguistic competence stands in need of an explanation. A newborn lacks the capacity to understand any sentences; some five years later, after taking in a finite body of information, he has gained the capacity to understand an infinite number of sentences. How can this infinite capacity be finitely learnable?

Frege famously answers this question as follows:

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a
terrestrial being for the very first time can be put into a form of words which will be understood by somebody to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought. ([21], 1)

English has a finite vocabulary, and a finite collection of syntactic rules for forming complex expressions. These features of English are thus finitely learnable. If English is compositional, Frege suggests, this finite information accounts, via recursive reapplication, for our infinite linguistic capacity. Thus we have reason to think that English is compositional. Frege puts the argument in terms of the necessity of compositionality for learnability; an analogous argument could also be framed in terms of sufficiency.

The learnability argument is that most frequently cited in discussions of compositionality. Textbooks on formal semantics, for example, typically introduce compositionality via the learnability argument. Consider two examples:

We presumably understand a sentence like:

(1) I saw a pink whale in the parking lot.

because we know what the single words in it mean (what pink and whale mean, for example) and we have an algorithm of some kind for combining them. Thus part of the task of semantics must be to say something about what word meaning might be and something about the algorithms for combining those word meanings to arrive at phrasal and sentential meanings. ([5], 6)

and:

If there were no direct relation between lexical and sentential meaning, of course, the meaning of each sentence would have to be listed. Since the number of sentences that make up a language is infinite, this would
mean that no human being would be able to determine the meanings of all the sentences of any language due to the finite resources of the brain. This is absurd, of course, and just as sentences are defined recursively by syntactic rules, taking words (or morphemes) as their basis, so their meanings should also be defined recursively from the meanings ascribed to the lexemes they contain. ([3], 3)

Nevertheless, learnability provides no good reason for taking natural languages to be compositional. For a language to be learnable, it is necessary and sufficient that it have a computable meaning function. If the meaning function for $L$ is computable, then it gives a procedure, graspable by beings like us, by which meanings of complex expressions can be determined. $L$ is thus learnable if computable. If, on the other hand, $L$ has no computable meaning function, then any procedure, graspable by beings like us, for determining meanings of complex expressions fails to determine the meanings of some expressions of $L$. $L$ is thus unlearnable if uncomputable.

However, computability and compositionality turn out to float quite free of one another:

- A language can be compositional without being computable. Let $L$ have concatenation as its only syntactic operation, and have expressions taking natural numbers as meanings. Let $M$ be an arbitrary noncomputable function on $\mathbb{N}$. If $\llbracket \alpha \bowtie \beta \rrbracket = M(\llbracket \alpha \rrbracket, \llbracket \beta \rrbracket)$, then $L$ is compositional, but $\llbracket \cdot \rrbracket$ is not computable. Compositionality does not make for computability unless the mode of composition is itself computable.

- A language can be computable without being compositional. Let $L$ have concatenation as its only syntactic operation, and have expressions taking natural numbers as meanings. Suppose that the meaning of $\alpha \bowtie \beta$, when it appears in the context $\gamma \bowtie (\alpha \bowtie \beta)$, is $\llbracket \alpha \rrbracket + \llbracket \beta \rrbracket + \llbracket \gamma \rrbracket$. $\llbracket \cdot \rrbracket$ is then computable, but the computation proceeds in a noncompositional way, violating semantic locality.
Computability does not make for compositionality unless the mode of computation appeals only to compositionally-available features. Compositionality is thus neither necessary nor sufficient for computability. The two features are wholly orthogonal, and no evidence for the presence of the one can be derived from the presence of the other. Learnability is a feature which goes to the computational complexity of meaning functions, but compositionality is a feature of the topology of meaning functions. Without further assumptions, the one tells us nothing about the other.\(^{56}\)

### 4.3 Methodological Compositionality

When compositionality is not motivated using the learnability argument, it is most often introduced as a methodological principle governing semantic theorizing. Call Φ a methodological principle for an activity A if Φ either is, or is a logical consequence of, a claim whose truth is a constitutive feature of performance of A. Suppose, for example, that for the construction of a theory to count as the construction of a semantic theory, the theory constructed must obey the Context Principle. Perhaps this is the case, as Frege suggests in [22], because it is constitutive of a semantic theory to characterize objective features, independent of merely psychological facts about individual speakers, in virtue of which linguistic expressions convey information, and only with the Context Principle is it possible to avoid psychologism in characterizing the meanings of expressions whose meanings are not concrete particulars.\(^{57}\) Perhaps it is the

---

\(^{56}\) Szabo, in chapter 3 of [64], gives an extended and insightful critique of the learnability argument, drawing attention to additional assumptions about the relation between linguistic understanding and meanings as provided by semantic theories, without which assumptions the learnability argument cannot get started.

\(^{57}\) Thus:

> Only by adhering to [the Context Principle] can we, as I believe, avoid a physical view of number without slipping into a psychological view of it. ([22], §106)
case, as [57] suggests, because a semantic theory requires an ontology of meanings, an ontology of meanings requires a principle of individuation of meanings, and the Context Principle provides the only available such principle. In either case, the Context Principle is a methodological principle of semantic theorizing, and given that, as seen in Lemma 1 of §2.2, the Context Principle entails compositionality, compositionality is also a methodological principle.

This overview closes with a brief consideration of another attempted methodological justification of compositionality. Why do individual words of a language have meaning? If the goal of a semantic theory is to account for meanings at some terminal level (say, the level of whole sentences), then word meanings seem superfluous. Perhaps the productivity of language requires that sentence meanings be derived systematically from some prior information base, but there is no reason why this information base need be either semantic or lexical. If word meanings (plus syntax) determine sentence meaning, and words determine word meanings, and orthographic structure determines word, then orthographic structure (plus syntax) determines sentence meaning, despite the fact that orthographic structure is not semantically invested.

Consider the following example.58 Sentences of ‘donkey anaphora’, such as the classic:

(33) If a farmer owns a donkey, he beats it.

present two challenge to compositionality. First, an account of the anaphoric pronouns must be given without violating semantic locality. Second, an explanation of the semantic shift in the indefinites from existential in the null context to universal in the embedded context must be provided. Kamp’s Discourse Representation Theory (DRT)

58Other familiar examples of the same phenomenon: Frege’s desire for a compositional treatment of propositional attitudes leads to an ontology of senses; Davidson’s compositional solution to adverbial modification leads to an ontology of events.
DRT treats indefinite descriptions, such as ‘a farmer’ as introducing *discourse referents* into a discussion, and uses as a tool of semantic analysis discourse representation structures (DRS’s), which consist of a combination of discourse referents and conditions imposed on those discourse referents. For example, (33) receives the following DRS:

\[
\begin{array}{c}
X,Y \\
{\text{farmer}}(X) \\
{\text{donkey}}(Y) \\
{\text{owns}}(X,Y) \\
\Rightarrow \\
X,Y \\
{\text{beats}}(X,Y)
\end{array}
\]

The two indefinites ‘a farmer’ and ‘a donkey’ introduce the two discourse referents, and the conditions ‘*farmer*(X)’, ‘*donkey*(Y)’, and ‘*owns*(X, Y)’ are then imposed on them. The subsequent anaphoric pronouns pick up on these discourse referents, and impose the further condition *beats*(X, Y) on them. A DRS is true relative to a model if there is an assignment to the discourse referents satisfying the conditions. The accumulation in a single DRS of discourse referents and constraints from multiple clauses then guarantees that anaphorically-linked phrases receive the same interpretation. When two DRS’s are linked by a conditional, semantic interpretation requires that every variable assignment satisfying the conditions of the first DRS can be extended to an assignment satisfying the conditions of the second DRS. Indefinites in the antecedents of conditions thus have universal force.

While the matter is not perfectly straightforward, DRT is typically taken to be
a non-compositional semantic theory. Construction of a DRS for a sentence with anaphoric pronouns depends crucially on the way that DRS’s have been constructed for sentences elsewhere in the conversational context, and the mode of quantification of DRS variables cannot be determined locally, shifting from existential in null contexts to universal in conditional contexts. Non-compositionality might come as no surprise here, of course, given that the target phenomenon of cross-sentential anaphora seems in its very nature to involve violations of semantic locality.

Dynamic predicate logic (DPL) (see [24]) restores compositionality while capturing the key insights of DRT. A compositional treatment of anaphora is achieved by making a fundamental alteration in the kind of semantic values assigned. Whereas static quantified logics assign sentences satisfaction conditions, DPL assigns input-output pairs. A sentence of DPL does not simply receive an assignment and evaluate relative to it; it can, instead, change an assignment. Thinking of variable assignments as specifications of context indicating which objects are conversationally salient, the dynamic insight of DPL is that sentences can affect, as well as be affected by, context.

Syntactically, DPL is a standard first-order language. Semantically, DPL assigns to sentences sets of ordered pairs of variable assignments, thought of as input-output pairs. Atomic sentences have no dynamic effects; they merely pass assignments satisfying a descriptive condition:

\[
\llbracket \Pi \tau_1 \ldots \tau_n \rrbracket = \{ < g, h > : g = h \land g \models \Pi \tau_1 \ldots \tau_n \}\]

59 If DRS’s are part of the semantic analysis, then DRT is straightforwardly non-compositional. If, on the other hand, DRS’s are a syntactic prologue to semantic analysis, along the lines of logical form, then DRT is plausibly compositional, albeit with an idiosyncratic syntax.

60 Non-compositionality shows up in the inability to replace a sentence in a discourse with a logically equivalent one, as in:

(35) * Not every man doesn’t own a donkey. He beats it.

61 Where \( \models \) is classical truth-in-a-model.
Existential quantifiers and quantifiers, however, are dynamic. An existentially quantified sentence expands the set of assignments to include all individuals meeting the existentially quantified condition:

- $\llbracket \exists x \phi \rrbracket = \{ < g, h > : \exists k : k[x]g \land < k, h > \in \llbracket \phi \rrbracket \}$

$k[x]g$ asserts that $k$ differs from $g$ at most in the assignment to $x$. Conditionals are internally dynamic – the output assignment $h$ must equal the input assignment $g$, but an input passes through to become an output only if every possible output of it, passed through the antecedent, produces some output when passed through the consequent:

- $\llbracket \phi \rightarrow \psi \rrbracket = \{ < g, h > : g = h \land \forall k(< h, k > \in \llbracket \phi \rrbracket \rightarrow \exists j < k, j > \in \llbracket \psi \rrbracket) \}$

Thus consider:

(36) $\exists x F x \rightarrow G x$\textsuperscript{62}

Suppose we start with an assignment $g$ assigning Napoleon to $x$ (thus $x$ represents a discourse referent in the DRT sense, and Napoleon is a live option for the real identity of that discourse referent). When that assignment hits the conditional, the existential quantifier $\exists x$ temporarily erases this information, allowing $g$ to be replaced by any assignment which agrees with it in all non-$x$ positions and which assigns an $F$ to the $x$ position. These assignments then input to $G x$, with $x$ now representing possible identities of the discourse referent introduced by the existential quantifier (indefinite) of $\exists x F x$. $G x$ acts as a test on these assignments, passing them through unchanged if their $x$ value is $G$. If some internally-altered assignment passes through, then then $g$ is output, unchanged, from the conditional. Thus $g$ emerges unchanged if and only if every $F$ is $G$. DPL hence produces the desired universal interpretation of (33).

\textsuperscript{62}Where, importantly, the scope of the existential quantifier is limited to the antecedent.
The DPL semantics are compositional: every constituent of a formula receives local semantically evaluated. The requirement, felt in DRT, for information about the prior communicative context, is eliminated by having each sentence in its semantic interpretation specify how it interacts with context; thus (a) the way in which early sentences in a discourse alter its context is built into their semantic interpretation and (b) the way in which a later sentence in the discourse reacts to any given changed context is built into \textit{its} semantic interpretation. Simple context-sensitivity of language can seem a threat to compositionality; one way of eluding that threat is to shift from thinking of expression meanings as wholly context-free entities to treating them as functions from contexts to traditional meanings. Dynamic context sensitivity represents a further shift, and it can in turn be responded to with another shift in our conception of basic semantic value – now as a tool of updating context through language\textsuperscript{63}. This pattern suggests another methodological argument for compositionality, one based on protecting the philosophical role of subsential semantics.

Suppose that lexical meanings play the role of revealing the ontological commitments inchoately present in our linguistic practice. Fulfillment of this role requires that the structure of the linguistic practice be faithfully mirrored in lexical meanings. Here we find a role for compositionality. Compositionality demands a certain integration of the meanings of sentences and words (given background assumptions about the syntactic relation between the two). Suppose, for example, that intensional contexts are distinguished from extensional ones by the inability to substitute \textit{salva veritate} therein terms identical at one, extensional, level of meaning. Given compositionality, this failure of intersubstitutability signals the presence of a further, intensional, dimension of meaning possessed by those terms, and a corresponding sensitivity to that dimension of meaning on the part of the intensional operators. Without compositionality, however,

\textsuperscript{63}For more on the dynamic perspective, see (e.g.) [65].
failures of intersubstitutability are without special significance, and do not reveal anything about the semantics of the terms involved. Thus it is only with compositionality that a meaningful distinction between the extensional and the intensional can be made, and thus that modal commitments can be localized in particular parts of our linguistic practice (counterfactuals, deontic expressions, epistemic contexts, and so on). Without compositionality, semantic features of the linguistic practice can ‘float free’, not appearing anywhere in the lexicon but emerging non-compositionally as lexical items are combined. Compositionally thus enforces a variety of honesty in semantic theory construction. If the role of words is to allow ontological taxonomizing of the world, this honesty, and hence compositionality, is essential.

References


51


