Compositionality

Josh Dever

October 1, 2003

The following sentence of English has never before been produced:

- Having but recently exhumed the vicar’s absurd nephew, the garrulous author of several bombastic political pamphlets slumped against a nearby statue commemorating the discovery of oxygen by the Danes, and 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, or atomic, 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?

Looking at some examples, however, suggests that English may not in fact have a compositional semantics. Consider the phrase ‘large wheel of pungent blue cheese’ from the above 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 an easy task. The difficulty is that appropriate standards of largeness seem to vary depending on the type of object in question. What’s large for a cheese may be quite small for a person.\(^2\) 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.

\(^1\)This is only a rough sketch of compositionality; see section 3 for a more careful formulation of the principle.

\(^2\)The Octuple Gloucester of [88] notwithstanding.
On the other hand, while these examples of \textit{non-intersective adjectives} raise problems for compositionality\footnote{That is, problems for giving a compositional semantics for a particular language (in this case, English). I will use the phrase ‘problems for compositionality’ in this way throughout.}, 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:

- $⟦\text{cheese}⟧ = \{x : x \text{ is a cheese}\}$
- $⟦\text{large}⟧ = f, f(X) = \{x \in X : x \text{ is larger than the average size of } X\}$

Then the meaning of ‘large cheese’ can be derived from $⟦\text{large}⟧$ and $⟦\text{cheese}⟧$ through functional application, which will yield the set of all cheeses above average in size.

The same procedure will yield an appropriate meaning for ‘large person’ as well, and will allow 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 English and other 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 sections 1 and 2, by setting out two historical episodes in which compositionality plays a significant role in semantic theorizing. In section 3, a more precise definition of compositionality is developed in great detail, and in section 4 some technical consequences of that definition are explored. Section 5 examines several compositionally-problematic semantic phenomena, and proposed compositional treatments thereof, and section 6 closes by asking why one might \textit{want} a compositional meaning theory, and attempting to explain the philosophical significance of compositionality.

\section{First Historical Prologue}

In his thirteenth century treatise \textit{Summa Lamberti}, Lambert of Auxerre advances the following rule governing the supposition (roughly, reference/denotation) of terms:

\begin{quote}
\text{[Rule 1]} A substantial or accidental common term that is not restricted by any other means and that serves as the subject or the predicate of a present-tense verb that has no ampliative force of its own or from anything else is restricted to suppositing for present things if it has appellata; but if it does not have appellata, it reverts to nonexistent. (\cite{68}, 116)
\end{quote}

Abstracting from the terminological details of mediaeval semantics, Lambert’s rule requires that denotational terms in a present-tensed sentence denote the \textit{current} satisfiers of their associated descriptive conditions, rather than past or future (or merely possible) such satisfiers. Later rules of the \textit{Summa Lamberti} require denotational terms in past-tensed sentences to denote the \textit{past} satisfiers of their associated descriptive conditions,
and in future-tensed sentences the future satisfies. A quick glance at the data supports the correctness of Lambert’s rules. In:

- The United States president ordered the bombings of Hiroshima and Nagasaki.
- The United States president regards Tony Blair as an important ally.
- The United States president will be a Democrat when the Social Security system goes bankrupt in 2037.

the subject term picks out a former president (Truman), a current president (George W. Bush) and an (unknown) future president as the tense of the verb shifts.

Lambert then proposes an objection to his rule:

It seems that a present-tense verb cannot restrict a common term serving as its subject, because the restrictor and the restricted must be in the same place. This is evident when one says ‘A white man is running’; ‘man’ is restricted by ‘white’ because ‘man’ and ‘white’ are in the same place. On the other hand, if one says ‘A man is white,’ in this case ‘man’ is not restricted by ‘white,’ because they are not in the same place. But the term serving as subject and the verb are not in the same place, because one is in the subject and the other in the predicate. Therefore, the term cannot be restricted by the verb.

Again, if the term is restricted by the verb, that happens only because of the tense of the verb. But it cannot happen because of the verb’s tense. Proof: Tense is the consignification of a verb, which is accidental to the verb; but in a verb its signification is prior to its consignification. A verb does not restrict a common term by reason of its signification, however. For when one says ‘A man is running,’ ‘man’ is not restricted to those who are running. Therefore, neither will a verb restrict in virtue of its consignification (as is apparent by the Topic from the greater), and so it will not restrict by reason of tense. ([68], 121)

Suppose the atoms of ‘Every man runs’ are assigned semantic value as follows:

- $\text{\texttt{\{every\} = f : \wp(D \times T) \mapsto \wp(\wp(D \times T)), \ f(X) = \{Y : X \subseteq Y\}}$
- $\text{\texttt{\{man\} = \{<x, y> : x \text{ is a man at } y\}}$
- $\text{\texttt{\{PRES\} = f : \wp(\wp(D \times T)) \mapsto [T, F]) \mapsto (\wp(\wp(D \times T)) \mapsto [T, F]), \ f(g) = g' : g'(X) = T \iff (\forall Y \in X \forall <z, w \in Y w = NOW \land \exists X'(YY' \in X' \exists Y \in X \forall <z, w \in Y' (w = NOW \rightarrow <z, w \in Y)) \land g(X') = T)}$

4 A more careful look, of course, reveals that the data is not so univocal; see footnote 12 for more discussion.
5 Where $D$ is the domain, $T$ is the set of times in the model, and $NOW$ is the home time in the model, and where I have assumed the presence of a tense marker PRES in the sentence. Obviously this is (a) only one form of semantic analysis among many, and (b) a form the details of which cannot be attributed to Lambert. I have given a roughly Montagovian presentation to make the mode of semantic combination simple (functional application) given a certain picture of the path of semantic construction for the sentence.
• \([\text{runs}] = f : \wp(\wp(D \times T)) \rightarrow \{T, F\}, f(X) = T \iff \exists Y \in X Y = \{< z, w >: z \text{ runs at } w\}\]

In order to get the right truth conditions, PRES needs semantic impact on ‘man’ as well as ‘runs’; otherwise, the set of all men past, present, and future will be assigned to ‘man’, and the PRES-modified value of ‘runs’ will map this set to F.\(^6\) Lambert’s suggestion, though, is that achieving this semantic impact is problematic. Why?

Assume a simple syntactic analysis of ‘Every man runs’:\(^7\)

\[
\begin{array}{c}
\text{S} \\
\text{DP} \quad \text{VP} \\
D \quad \text{NP} \quad I \quad V \\
\text{A man} \quad \text{PRES} \quad \text{runs}
\end{array}
\]

Make Lambert’s notion of ‘the same place’ more precise by stipulating that X and Y are in the same place if one c-commands the other. Then the present tense marker PRES and the noun ‘man’ are not in the same place. Lambert’s worry is then that this syntactic separation between the two will prevent the semantic impact of PRES from reaching the subject term ‘every man’. Why would distance of separation matter? The apparent answer is that Lambert has an implicit compositional picture of meaning.

On this picture, supralexical semantic value develops in the following manner: the meanings of PRES and ‘runs’ will combine to assign the VP a function mapping a set of ordered pairs of objects and times to T if and only if every member of that set runs at the present time, and the meanings of ‘every’ and ‘man’ will combine to assign the DP the set of all sets containing the set of past, present, and future men, and then the meanings of the DP and the VP will combine to assign to the sentence as a whole ‘true’ if the function assigned to the VP maps the set assigned to the DP to T, and ‘false’ otherwise. But then ‘every man’ will denote all men past, present, and future, rather than (as required by Lambert’s [Rule 1]) only present men, and incorrect truth conditions will be assigned.\(^8\)

\(^6\) Type-shifting will be needed to apply PRES to ‘man’; this is an artifact of the Montagovian presentation and could be avoided by assigning both verbs and nouns sets of objects-at-times, having PRES in both cases filter such sets to present objects only, and having different rules of semantic combination for different grammatical combinations.

\(^7\) See below for a less simple analysis.

\(^8\) Nevertheless, proper truth conditions for ‘Every man runs’ can still be achieved by altering the mode of semantic composition. If the combination of I and \(\bar{V}\) proceeds not by functional application of [I] to [\(\bar{V}\)], but instead yields an ordered pair of that application and [I], and if the combination of D and NP produces simply an ordered pair of [D] and [NP]:

\[
\begin{align*}
\llbracket I \bar{V} \rrbracket & = ([I](\llbracket \bar{V} \rrbracket)), [I] > \\
\llbracket D \text{ NP} \rrbracket & = ([D], [\text{NP}]) >
\end{align*}
\]

then semantic composition at the S level can proceed by the rule:

\[
\llbracket S \text{ DP VP} \rrbracket = \text{arg}_1(\llbracket \text{VP} \rrbracket)(\text{arg}_1(\llbracket \text{DP} \rrbracket)(\text{arg}_1(\llbracket \text{VP} \rrbracket)(\text{arg}_1(\llbracket \text{DP} \rrbracket))^{-1}))
\]

where + and – are type-shifting operators defined by:

\[
\begin{align*}
X^+ & = f : \wp(\wp(D \times T)) \rightarrow \{T, F\}, f(Y) = T \iff X \in Y \\
X^- & = Y : \forall Z \ni YX(Z) = T
\end{align*}
\]
Lambert thus accepts in his semantic theorizing the constraint of compositionality. In addition requiring a theory to match some range of data, he also requires it to take a particular form – a form in which, roughly, the meanings of complexes are built out of the meanings of their parts. In order to achieve compositionality, Lambert needs to reshape his theory in some way. His diagnosis of the problem is:

In a verb there are these two: the action and the composition. The action considered in itself is indifferent with respect to any time whatever, but the action is restricted by time (which is the measure of motion) to being in that time which is its measure. Thus the action of a present-tense verb is placed in the present, the action of one in the past tense in the past, and the action of one in the future tense in the future. In this way, then, the action of a verb is restricted by the verb’s time [or tense]. And because the time that is the measure of the action is the measure of the composition of the parts, the composition is restricted when the action has been restricted. And because the action is something that cannot be expressed without the verb, [the action] is even said of the subject by means of the composition that unites [one] extreme with [the other] extreme in an expression. And so the subject of the action is placed in the same time in which the action and the composition are placed, and in that way the subject is restricted to the time to which the action and the composition are restricted. But the subject of the action is a common term serving as the subject of the verb. It is evident, therefore, that the verb restricts a common term serving as its subject, and so restriction is brought about not directly but indirectly by the time [or tense] of the verb. For first the action of the verb is restricted by the time [or tense], then the composition, and finally the common term serving as the verb’s subject. ([68],122)

There is no unique (and likely no wholly faithful) way to craft an explicit formal theory out of these remarks, but we might (anachronistically) attribute an event-based semantics to Lambert here. On a simple version of such a semantics, verbs will introduce event variables, and other components of the sentence will be woven into the sentence.
via various thematic relations to the event variable. Setting aside tense momentarily, the heart of:

\(1\) Every man runs.

analyzes to:

\((1e)\) \(\exists e \forall x (\text{man}(x) \rightarrow \exists e' (e' \sqsubseteq e \land \text{run}(e') \land \text{theme}(x, e'))))\)

The somewhat idiosyncratic use of a single master event, subsuming smaller events consisting of individual runnings, is necessary in order to make sense of Lambert’s thought that there is something which is the action of the verb.\(^{11}\) Now suppose, following on Lambert’s assertion that ‘the action is restricted by time’, that all event variables must be in the scope of the tense of the verb. This yields:

\((1'e)\) \(\text{PRES}\exists e \forall x (\text{man}(x) \rightarrow \exists e' (e' \sqsubseteq e \land \text{run}(e') \land \text{theme}(x, e'))))\)

This analysis places the subject noun ‘man’ inside the c-command of the present tense inflection, thus placing both in the same ‘part’ of the sentence, in Lambert’s terminology.\(^{12}\) A standard ‘inside-out’ semantic composition procedure for the first-order claim \((1e')\) will thus allow the present-tense inflection to affect, and appropriate limit, the

\(^9\)Where \(\sqsubseteq\) is the relation of event parthood.

\(^{10}\)Apologies for the sudden shift from a Montagovian to a semi-Davidsonian semantic framework. The event semantics can be given a Montagovian implementation, but only at the cost of a considerable loss of readability. Since compositionality is a constraint which is applicable across a wide range of semantic frameworks, some variety in system is perhaps didactically desirable.

\(^{11}\)See [92] for a similar analysis of the interaction between quantified noun phrases and event variables. Although the use of a single master event is particularly useful when treating collective quantification, \((1e)\) is intended to give a distributive reading. The analysis given in the main text will not work in general, as shown by examples with monotone decreasing quantifiers:

\((1')\) No man runs.

\((1'e)\) \(\exists e \neg \exists x (\text{man}(x) \land 3e' (e' \sqsubseteq e \land \text{run}(e') \land \text{theme}(x, e'))))\)

\((1'e)\) is trivially true, given the existence of events containing no running parts whatsoever. The master event variable is thus perhaps better read as a free choice variable, perhaps constrained by contextual factors. I do not assert, though, that this proposal can be made precise, or indeed that any unproblematic theory can be extracted or extrapolated from Lambert’s remarks.

\(^{12}\)There remains, however, the question of how to get from ‘every man runs’ to this analysis; in particular, how to get the inflection PRES to have scope over the whole sentence, given its initial position inside the VP. Two suggestions:

1. Perhaps PRES does not in fact have scope over the whole sentence, and the proper analysis of \((1)\) is:

\((\text{FN2})\) \(\exists e \forall x (\text{man}(x) \land x \sqsubseteq e) \rightarrow (\text{PRES}(e) \land \exists e' (e' \sqsubseteq e \land \text{run}(e') \land \text{theme}(x, e'))))\)

In this example, PRES acts as an event predicate, rather than as a sentential operator, and is assigned without taking scope over the subject. However, the subject men are also required to be parts of the master event (this requires extending the notion of event parthood to include participants as well as subevents; the resulting picture is similar to the situation semantics of [6]). If we then add the metaphysical assumption that present events can have as participants only things which presently exist, the men quantified over will, as desired, be restricted to the present men. In this analysis, PRES and the subject DP remain semantically isolated, but achieve a back-door interaction via the use of events as metaphysical carriers of information from one part of a sentence to another.

2. The syntactic analysis of \((1)\) given earlier was rather simple-minded, and we might replace it with a more sophisticated one. For example, we might analyze \((1)\) by:
This analysis could be motivated, along Minimalist lines, by assuming (a) that inflections take $V_{\text{MAX}}$ as complements, (b) that inflections need to feature-check against subjects (a natural assumption if Agr is a component of I), and (c) that feature checking occurs via Merge, requiring $D_{\text{MAX}}$ to be raised to the specifier position for I. In this tree, the immediate adjoining of ‘every man’ to the head PRES inflection will place the subject in adequate syntactic proximity to that temporal marker to allow a straightforward compositional story whereby temporal restrictions are placed on the relevant men (I assume here that category heads have semantic impact throughout their maximal projections). The more sophisticated syntactic analysis also allows us to note, and explain, the fact that the earlier-cited past and future examples ‘The United States president ordered the bombings of Hiroshima and Nagasaki’ and ‘The United States president will be a Democrat when the Social Security system goes bankrupt in 2037.’. Each of these has, as noted earlier, a reading on which the relevant president is selected by the tense of the verb, but also has a reading on which the relevant president is the one of the time of utterance. This ambiguity is predicted by assuming that past and future inflections include a complementizer element (‘will’ and ‘-ed’) and by assuming that $D_{\text{MAX}}$ phrases optionally raise to the specifier position of complementizers. Thus the simpler ‘Every man will run’ receives two analyses:

On the first reading, ‘every man’ is outside the maximal projection of the future inflection, and hence its interpretation is governed by the time of utterance, rather than by the future operator. On the second, ‘every man’ is governed by FUT. Thus the two desired readings are (compositionally)
contribution of ‘man’. Insisting on a compositional semantic theory, then, may change the basic shape of that theory – here, from primarily objectual, type-hierarchical, and functional, to primarily eventual and conjunctive\textsuperscript{13} – to reveal a truer semantic picture.\textsuperscript{14}

Lambert’s semantic remarks, of course, fall considerably short of the rigour of a contemporary theory. But even in this relatively primitive form, we can see the outlines of:

- The nature of the compositionality constraint (meanings of wholes must be built out of meanings of parts).
- The ability of this constraint to rule out certain semantic theories.
- Some of the moves available for achieving compositionality (alteration of the underlying syntax, the lexical semantic values, or the procedures for combining those values at higher levels).
- At least one reason for caring about compositionality (its ability to reveal hidden commitments of our semantic practice).

Our goal is to sharpen the details within these outlines.

2 Second Historical Prologue

Fast forward some six to seven centuries. Frege has published the \textit{Begriffsschrift} and thereby started us down the path toward rigorous formal semantics. Late in his career, generated by two different syntactic structures.

The syntactic justification for the requirement that sentences have subjects is surprisingly difficult to give in full detail – it is frequently taken that θ-role assignment controls the introduction of arguments into syntactic trees, and it is often difficult to see (a) how verbs can assign θ-role to the external subject position and (b) how subjects can be mandatory in sentences with verbs which assign no agent role. Suggestions include having θ-role assigned by V, rather than by \(\bar{V}\) (see [77], [14]) and endorsing an \textit{Extended Projection Principle} which explicitly mandates the presence of a subject at all levels of syntactic representation (see [12], [13]). Given the difficulty in properly locating the subject in syntactic structures, Lambert’s difficulty in compositionally accounting for the semantic behaviour of the subject is perhaps not surprising (and is indicative of the role of compositionality in enforcing a tight relation between syntax and semantics).

Lambert’s problem is further complicated by examples in which subject (and other term) interpretation is not governed by verbal inflection, or by time of utterance.\textsuperscript{4} considers examples involving verbs of creation and destruction, such as:

- A powerful thermonuclear device vaporised a VW. We saw its shape thermally etched on a wall upon our visit to the site of the explosion.

Here the interpretation of ‘its’ seems to require inherited temporal control from the time of the previous utterance, which at least looks like a violation of compositionality. Similar examples can be constructed without anaphora, such as:

- The hungry prisoners were wholly sated after the banquet celebrating their release.

Thanks to Nicholas Asher for bringing these examples to my attention.\textsuperscript{13}See [87] for detailed discussion of the differences between the ‘Functionalist’ and ‘Conjunctivist’ approaches to semantics, and of ramifications of these differences.

\textsuperscript{14}To call the resulting picture \textit{truer}, of course, is simply to \textit{presuppose} that a compositional semantic theory better reveals something deep about our semantic practice than a noncompositional theory. See section 6, especially section 6.3, below for discussion of the plausibility of this presupposition.
in ‘Compound Thoughts’, he says:

> 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. ([39], 1)

In this passage, Frege endorses a principle of compositionality for natural language. The meaning of an entire sentence is built up out of the meanings of the parts of that sentence.\(^\text{15}\) Frege also gives us a reason to want compositionality: to account for our ability to understand language. We need compositionality in language, Frege suggests, in order to account for our ability to comprehend an infinite number of sentences using a finite cognitive capacity. We are able to understand the thoughts expressed by novel sentences because those thoughts are built up out of the thoughts expressed by component parts of the sentence.\(^\text{16}\)

Similar statements can be found in Frege’s unpublished ‘Logic in Mathematics’\(^\text{17}\), where he says ‘the sense of the complex expression must be yielded by that of its parts’ ([38], 226) and:

> As a sentence is generally a complex sign, so the thought expressed by it is complex too: in fact it is put together in such a way that parts of the thought correspond to parts of the sentence. So as a general rule when a group of signs occur in a sentence it will have a sense which is part of the thought expressed. ([38], 207-208)

In addition to such explicit endorsements of the principle of compositionality late in his career, Frege earlier gives a compositionality-friendly semantics in his *Begriffsschrift* and famously appeals to compositionality constraints in arguing for the distinction between Sinn and Bedeutung. Frege is thus frequently taken to be originator of talk of compositionality in semantics, and the constraint that a semantics be compositional is often called Frege’s Principle.\(^\text{18}\)

Although the example of the previous section shows that Frege’s innovative role should not be overstated, it is certainly true that his work is deeply influential in establishing the contemporary significance of compositionality.

\(^\text{15}\)In fact, Frege seems to endorse the stronger principle that the method of building up is mereological composition – that the meanings of the parts of the sentence are in fact parts of the meaning of the whole sentence. Compositionality *simpliciter* does not require that the meanings of the parts persist in the meaning of the whole; it, for example, is compatible with the view that the meaning of the whole is the result of functional application of the meaning of one part on the meaning of another part, as in the Montagovian model discussed above. See the discussion of strong parallelism in section 3.1.5 below.

\(^\text{16}\)See discussion in section 6.2.1 of the learnability argument. The core thoughts here, that a central task of semantic theorizing is to bridge the gap between finite semantic knowledge and infinite semantic capacity and that a part-to-whole mode of semantic composition is crucial to achieving this task, have become truisms in the wake of Frege. See, for example, [20], [21] for extensive discussion of the best formulation, and consequences, of the view that ‘a satisfactory theory of meaning must give an account of how the meanings of sentences depend upon the meanings of words’ ([21], 17).

\(^\text{17}\)Written in 1914.

\(^\text{18}\)The extent of Frege’s commitment to compositionality is matter of ongoing debate, especially given his endorsement in The Foundations of Arithmetic of the ‘context principle’.
2.1 Compositionality and the *Begriffsschrift*

Frege’s *Begriffsschrift* does not give a fully explicit semantics for his newly-introduced logical language, but enough is said to work out a semantics which is plausibly faithful to his intentions. The *Begriffsschrift* contains four logical primitives: negation, the conditional, universal quantification, and identity. Of the two sentential connectives, Frege says:

If a short vertical stroke is attached below the content stroke, this will express the circumstance that *the content does not take place*. So, for example,

\[ \neg \text{A} \]

means “A does not take place”. ([40],17)

and:

If A and B stand for contents that can become judgments, there are the following four possibilities:

1. A is affirmed and B is affirmed;
2. A is affirmed and B is denied;
3. A is denied and B is affirmed;
4. A is denied and B is denied.

Now never to ask for the meaning of a word in isolation, but only in the context of a proposition. ([41], s)

The context principle is frequently taken to contradict, or at least be in tension with, the principle of compositionality, because it suggests that meanings of wholes are, in some important sense, 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 section 4.1.2 below for a more explicit formulation of the worry). See [26], 3-5 for one influential attempt to reconcile Frege’s apparent commitment to both context and compositionality, and [84] and [60] for more recent discussions, including overviews of various interpretative positions taken in the literature.

Frege hints at an explicit principle of compositionality at one point in the *Begriffsschrift* – ironically, at the very point that he abandons compositionality. When introducing identity to his system, he says:

Identity of content differs from conditionality and negation in that it applies to names and not to contents. Whereas in other contexts signs are merely representatives of their content, so that every combination into which they enter expresses only a relation between their respective contents, they suddenly display their own selves when they are combined by means of the sign for identity of content; for it expresses the circumstance that two names have the same content. ([40], 20, emphasis added)

Outside of identity contexts, this passage suggests, meanings of complexes are relations of meanings of their parts, satisfying compositionality. An identity of the form:

\[ \text{A} \equiv \text{B} \]

however, expresses the claim that the symbols ‘A’ and ‘B’ have the same content, and hence is not built out of the content(s) of ‘A’ and ‘B’. I will ignore the noncompositionality of identity in the discussion of this section.

Setting aside the judgement and content strokes.

In modern notion: \( \neg \text{A} \).
stands for the judgement that the third of these possibilities does not take place, but one of the other three does. ([40], 13-14)

The meanings of both negations and conditionals, then, are built out of the meanings of their component parts, by way of application of appropriate truth functions to the truth values of their parts.

Frege’s semantic treatment of the sentential fragment of the *Begriffsschrift* is an early version of what is now the standard truth-functional semantics for sentential logic. That semantics serves as a particularly clear paradigm of compositionality. Suppose we have a language which consists of (a) some small number of sentential connectives, each of which receives as semantic value a function from ordered tuples of truth values to a truth value, (b) atomic sentences, each of which receives as semantic value a truth value\(^23\), and (c) grouping markers. Complex sentences are then assigned truth values by application of the truth function of their main connective to the truth values of their largest sentential parts. For example, suppose a sentential language such that:

\[\langle \neg \rangle = f : \{T, F\} \to \{T, F\}, f(T) = F, f(F) = T\]
\[\langle \land \rangle = f : \{T, F\}^2 \to \{T, F\}, f(T, T) = T, f(T, F) = F \text{ otherwise}\]
\[\langle \lor \rangle = f : \{T, F\}^2 \to \{T, F\}, f(T, F) = T, f(F, x) = F \text{ otherwise}\]
\[\langle \to \rangle = f : \{T, F\}^2 \to \{T, F\}, f(T, F) = F, f(x, y) = T \text{ otherwise}\]

Consider the sentence \(\neg (p \to \neg (q \lor r))\), and let \(\langle p \rangle = \langle q \rangle = T\) and \(\langle r \rangle = F\). Then the meaning (truth value) of \(\neg (p \to \neg (q \lor r))\) is worked out piecewise as follows:

1. \(\langle q \lor r \rangle = \langle \lor \rangle (\langle q \rangle, \langle r \rangle) = \langle \lor \rangle (T, F) = T\)
2. \(\langle \neg (q \lor r) \rangle = \langle \neg \rangle (\langle q \lor r \rangle) = \langle \neg \rangle (T) = F\)
3. \(\langle (p \to \neg (q \lor r)) \rangle = \langle \to \rangle (\langle p \rangle, \langle \neg (q \lor r) \rangle) = \langle \to \rangle (T, F) = F\)
4. \(\langle \neg (p \to \neg (q \lor r)) \rangle = \langle \neg \rangle (\langle (p \to \neg (q \lor r)) \rangle) = \langle \neg \rangle (F) = T\)

The meaning of each complex expression, then, is derived exclusively from the meanings of its parts (and the way in which those parts form the whole, since \(p \to q\) needs to be semantically distinguished from \(q \to p\), despite having the same parts). Truth-functional sentential logics provide ideal examples of compositionality.

The quantified fragment of the *Begriffsschrift*, however, is not compositional. Frege says the following about the interpretation of his universal quantifier:

---

\(^{22}\)In modern notation: \(A \to B\).

\(^{23}\)in-a-model, if we prefer.

\(^{24}\)In Fregean notation: \(\text{\small 10} -\text{\small 11}\).
In the expression of a judgment we can always regard the combination of signs the right of as a function of one of the signs occurring in it. If we replace this argument by a German letter and if in the content stroke we introduce a concavity with this German letter in it, as in

\[ \Phi(a) \]

this stands for the judgment that, whatever we may take for its argument, the function is a fact. ([40], 24, emphasis in the original)

There are two compositionality difficulties with this rule. First, the meaning of the whole is built up out of more than the meanings of the parts. In addition to the meaning of the function \( \Phi() \), we also need the meanings of the various terms available as arguments for \( \Phi() \). If \( \Phi(a) \) is true just in case \( \Phi(A) \) is true, and \( \Phi(B) \) is true, and so on for all syntactically appropriate terms, then the meanings of \( A, B, \) and so on are factors in determining the meaning of \( \Phi(a) \).

Second, the Fregean interpretative rule does not allow piecewise semantic analysis of sentences with multiple quantifiers. Consider:

\[ \Psi(a) \]

This sentence is true if every sentence of the form:

\[ \Psi(A) \]

(for appropriate choice of \( A \)) is true. Such a sentence will be true unless \( \Phi(A, b) \) is true and \( \Psi(A) \) is false. \( \Phi(A, b) \) in turn will be true if every sentence \( \Phi(A, B) \) (for appropriate choice of \( B \)) is true. This is a perfectly sensible semantic analysis, but it is not a compositional one. The meaning of the target sentence is constructed not out of the meanings of its immediate parts \( \Phi(a) \) and \( \Phi(a, b) \), or even out of the meanings of these parts plus the meanings of appropriate arguments for the functions (in the wake of the first problem above), but out of the meaning of \( \Phi(a) \) combined with the meaning of a sentence \( \Phi(A, b) \) which is simply not a part of the target sentence. The Fregean approach to the quantifiers works essentially from the outside (the whole) to the inside (the parts), rather than the other way around. Were we to try to give the meaning of the target sentence compositionally, we would have to start by giving the meaning of \( \Phi(a, b) \). But the meaning of

\[ (\forall x \Phi(x)) \]

25 In modern notation: \( \forall x \Phi(x) \)

26 (59), 433 poses this as a general problem for the compositionality of substitutional quantifiers. Whether this counts as a violation of compositionality will depend on the precise wording of the compositionality constraint. The meaning of \( \Phi(a) \) is in a sense a function of the meaning of \( \Phi() \), where the function in question is that of conjoining the truth values of all saturations of \( \Phi() \). The discussion of section 3.1.5 below bears on this issue. [59] does not consider treating the meaning of the quantified claim as this sort of very broad function of the meaning of the quantified matrix, viewing instead the meaning of the quantified claim as a function of the meanings of the many substitutional instances of the quantified matrix, which instances clearly are not parts of the original sentence.

27 In modern notation: \( \forall x(\forall y \Phi(x, y) \rightarrow \Psi(x)) \).
this sentence will be derived from (negation plus) the truth value of $\Phi(a, B)$ for appropriate choices of $B$, and such sentences have no truth values, since they lack one of the two required arguments for the $\Phi$ function. Frege’s attempt to analyze the meanings of quantified sentences via the meanings of sentences formed by saturation of quantified argument positions leaves him unable to analyze compositionally sentences with nested quantifiers, since such sentences contain component quantified parts which do not yield (closed) sentences when appropriately saturated.

A properly compositional semantics for Frege’s quantified logic would have to await the work of Tarski ([99]), who introduces the notion of satisfaction to reverse the direction of the Fregean analysis. Let a sequence be a function from the set of variables in the language to the set of objects. Then a predicate $\Phi(x)$ is assigned as meaning the set of sequences satisfying it, where $f$ satisfies $\Phi(x)$ if $f(x)$ is $\Phi$. We can now say:

- $\langle \forall x \Phi(x) \rangle = \{ f : \forall g(\forall y \neq 'x'g(y) = f(y) \land g \in \langle \Phi(x) \rangle) \}$

Sentences with nested quantifiers can now be analyzed compositionally. In $\forall x \forall y R(x, y)$, $\langle R(x, y) \rangle$ is a collection of sequence each of which assign to $x$ and $y$ R-related objects. $\langle \forall y R(x, y) \rangle$ is then derived from $\langle R(x, y) \rangle$, by taking the collection of all sequences all of whose $y$-variants are in $\langle R(x, y) \rangle$. Finally, $\langle \forall x \forall y R(x, y) \rangle$ is derived from $\langle \forall y R(x, y) \rangle$ by taking the collection of all sequences all of whose $x$-variants are in $\langle \forall y R(x, y) \rangle$.

### 2.2 Compositionality, Sense, and Reference

In [44], Frege advances perhaps the most famous argument in the history of semantics:

Equality gives rise to challenging questions which are not altogether easy to answer. Is it a relation? A relation between objects, or between names or signs of objects? In my Begriffsschrift I assumed the latter. The reasons which seem to favour this are the following: $a = a$ and $a = b$ are obviously statements of differing cognitive value; $a = a$ holds a priori and, according to Kant, is to be labelled analytic, while statements of the form $a = b$ often contain very valuable extensions of our knowledge and cannot always be established a priori. . . . Now if we were to regard equality as a relation between that which the names ‘$a$’ and ‘$b$’ designate, it would seem that $a = b$ could not differ from $a = a$, i.e., provided $a = b$ is true. A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing. . . . If the sign ‘$a$’ is distinguished from the sign ‘$b$’ only as an object (here, by means of its shape), not as a sign (i.e. not by the manner in which it designates something, the cognitive value of $a = a$ becomes essentially equal to that of $a = b$, provided $a = b$ is true. A difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation of the thing designated. ([44], 25-26)

Frege here suggests that the assumption:
Terms have their semantic values exhausted by their referents. is incompatible with the existence of a semantic difference (difference in cognitive value) between \(a = a\) and \(a = b\), and hence must be given up. This argument is valid only if a principle of compositionality is presupposed. With such a principle, given \((M=R)\) \(a = a\) and \(a = b\) are assembled out of parts with the same semantic value put together in the same way, and hence must have the same meaning themselves. Since they do not, \((M=R)\) must be given up (in the face of the presupposed inviolability of compositionality). Abandoning \((M=R)\) leaves Frege room to assign terms sense as well as reference, and thus room to have \(a = a\) and \(a = b\) differ in meaning.\(^2\)

Frege, under pressure from compositionality, changes his picture of lexical meaning. On the new picture, terms carry not only a referent, but also a sense, understood as a way in which the referent is presented to the speaker. The addition of sense gives Frege’s semantics a more epistemological character: two terms \(\alpha\) and \(\beta\) will differ in sense just in case there are some sentences \(\Phi(\alpha)\) and \(\Phi(\beta)\) such that a rational and linguistically competent agent can be justified in accepting one and rejecting the other. Compositionality requires Frege’s decision to treat cognitive value as a semantic feature of agents’ linguistic practice to be reflected in his picture of lexical semantics. As with Lambert’s puzzle, compositionality serves as a tool for making semantically relevant features of linguistic practice explicit on the ground level of lexical meaning.

3 What is Compositionality?

Having examined the role of compositionality in two historical examples of semantic theory formation, it is now time to give a more precise formulation of what compositionality is. Compositionality can be thought of as 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.

\(^2\)Along with the \textit{Begriffsschrift} metalinguistic analysis of facing terms take themselves as semantic value in identity contexts, which Frege now finds does not account for our ability to talk about objects, rather than language, in making identity claims.

\(^2\)The same argument form emerges again later in [44] as Frege argues that the \textit{Bedeutung} of a sentence is its truth value:

If we now replace one word of the sentence by another having the same \textit{Bedeutung}, but a different sense, this can have no effect upon the \textit{Bedeutung} of the sentence. Yet we can see that in such a case the thought changes; since, e.g., the thought in the sentence ‘The Morning Star is a body illuminated by the Sun’ differs from that in the sentence ‘The Evening Star is a body illuminated by the Sun.’ Anybody who did not know that the Evening Star is the Morning Star might hold the one thought to be true, the other false. The thought, accordingly, cannot be the \textit{Bedeutung} of the sentence, but must rather be regarded as its sense. ([44], 32)
(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:

(2) Lois Lane believes that Superman can fly.

from being determined in part by the word ‘Superman’ (its phonetic, morphological, historical-causal properties, etc.), rather than the meaning of the word ‘Superman’ (whatever that meaning turns out to be).\textsuperscript{31} Semantic Locality, on the other hand, prevents the meaning of the occurrence of ‘Superman can fly’ in (2) 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 \textit{determination}. Two versions of determination dominate most discussion of compositionality: the \textit{functional} analysis and the \textit{substitutional} analysis.

3.1 Compositionality as Functionality

3.1.1 First Draft

Suppose we have a language $L$, considered as a set of expressions, and a parthood relation $\sqsubseteq$ which gives a partial ordering of $L$.\textsuperscript{32} Suppose we also have a function $\llbracket \cdot \rrbracket$ which assigns meanings to expressions of $L$, where $\llbracket \cdot \rrbracket$ maps from $L$ to some set $M$ of meanings.\textsuperscript{33} Then $\llbracket \cdot \rrbracket$ is a compositional meaning function for $L$ if and only if:

\textsuperscript{30}See the conception of words defended in [64].

\textsuperscript{31}Assuming, of course, that the word ‘Superman’, or its phonetic, morphological, historical-causal, etc. properties are not part of the semantics of ‘Superman’. Semantic Closure is only as strong a constraint as our notion of peculiarly \textit{semantic} property is rich.

\textsuperscript{32}In what follows, I will always assume that the parthood relation is given by the syntax of the language (that the parts of an expression $\epsilon$ are the expressions which combine, via syntactic rules, to yield $\epsilon$). However, this assumption is not essential for considering compositionality.

\textsuperscript{33}Notice that, in order to maximize generality, no constraint is placed on the set $M$. In particular, there is no need for $M$ to match perfectly the set of meanings used by any particular semantic theory. Thus:

- In a bilevel semantic theory like Frege’s, which makes a distinction between sense and reference, compositionality constraints can be imposed separately on each level, by taking $M$ first to be the set of referents and second to be the set of senses. Or we can require sense to depend compositionally on sense and reference to depend compositionally on both sense and reference, by taking $M$ first to be the set of senses and second to be the union of the sets of referent and senses (in this second case, the range of $\llbracket \cdot \rrbracket$ would be limited to that subset of $M$ containing only referents).

- If a semantic theory assigns meanings of multiple ‘types’, for some notion of type (such as: truth value, property, set of possible worlds, etc.), [91]’s first notion of compositionality:

  \begin{quote}
  The applicability of a semantic attribute $A$ to a complex expression $E$ is completely determined by the applicability of this particular attribute $A$ to the component expressions of $E$ plus the way they are combined into $A$. ([91], 49)
  \end{quote}

can be imposed by taking multiple sets $M$, each containing all meanings of a given type, and imposing compositionality constraints on the language for each resulting set $M$. The consequent notion of compositionality is very strong, and does not seem to have been endorsed at any point in the literature.
• There is some G such that for all ϵ ∈ L, if Γ = {α : α ⊑ ϵ}, then ⟦ϵ⟧ = G(⟨⟦α⟧ : α ∈ Γ⟩).

That is, if and only if there is some function which gives the meaning of any complex expression in terms of the meanings of its parts.

Every language is compositional given some parthood relation. If we take ⊑ to be the identity relation on L, then it will be trivial that the meaning of every expression of L is a function (namely, the identity function) of its parts. For this reason, we will typically want, minimally, ⊑ to be irreflexive on some subset of L (intuitively, on the complex expressions of L). Strictly speaking, a language should be called compositional (or not) only relative to a parthood relation.

This first draft, however, does not give us what we want. Consider the pair of sentences:

(3) Socrates kicked Plato.
(4) Plato kicked Socrates.

Assuming a plausible parthood relation, both sentences have the same parts, and thus by the first draft attempt at functional compositionality must have the same meaning. But, of course, they do not, and their failure of synonymy does not seem indicative of any violation of compositionality. The difficulty, of course, is that meanings of complex expressions are functions of both the meanings of the parts of those expressions and the way in which those parts are put together. A more sensitive notion of compositionality is required.

3.1.2 Second Draft

Suppose we have a language L, considered as a an ordered pair of a set L of expressions and a set Δ of constructive operations (paradigmatically, modes of syntactic composition, such as concatenation or the application of a particular phrase structure rule), where the members of Δ are functions δ_i : L^n_i ↦→ L. Suppose Δ induces a parthood relation ⊑ by the definition:

• α ⊑ β ≡ β = δ_i(..., α, ...) for some δ_i and some additional arguments for δ_i.34

Finally, suppose as before that we have a function ⟦·⟧ which assigns meanings to expressions of L. Then ⟦·⟧ is a compositional meaning function if and only if:

• There is some G such that for all ϵ ∈ L, if ϵ = δ_i(α_1, ..., α_n), then ⟦ϵ⟧ = G(< δ_i, ⟦α_1⟧, ..., ⟦α_n⟧ >).

The notion of compositionality can be generalized beyond semantics by appropriate choice of M. If, for example, we take M to be the set of syntactic properties of expressions of a language L (and take ⟦·⟧ to be a function assigning syntactic properties to expressions), a notion of a compositional syntax results. Semantic compositionality thus falls into a genre of mereological reducibility.

34We can guarantee the transitivity of ⊑, if desired, by taking Δ to be closed under functional composition.
That is, if and only if there is some function which gives the meaning of any complex expression in terms of the meanings of its parts plus their mode of combination. Note that this definition of compositionality requires that, if there is more than one way of constructing $\epsilon$ out of applications of members of $\Delta$ to $L$, then for each such way, the meaning of $\epsilon$ must be a function of the meanings of its parts under that way. Thus, for example, if:

$\text{(5) Some philosopher fears Socrates}$

receives the rather flat-footed syntactic analysis:

```
S
   ┌── NP ── VP ┘
     │     │    
   ┌─ D ── N ┐  ┌─ V ── NAME ┐
     │ Some ]  │ fears ]  │)
     └─────φ ───┘    └────NAME ┘
```

and expression parthood follows syntax in having constructive functions:

- $\delta_1$ for the rule $NP \to D N$
- $\delta_2$ for the rule $VP \to V NAME$
- $\delta_3$ for the rule $S \to NP VP$
- $\delta_4$ for the rule $S \to [NP D N ] [VP V NAME ]$

then $\llbracket (5) \rrbracket$ must equal $G(<\delta_3, \llbracket \text{some philosopher}\rrbracket, \llbracket \text{fears Socrates}\rrbracket >)$ as well as $G(<\delta_4, \llbracket \text{some}\rrbracket, \llbracket \text{philosopher}\rrbracket, \llbracket \text{fears}\rrbracket, \llbracket \text{Socrates}\rrbracket >)$

### 3.1.3 Third Draft

An expression might be constructible in various ways because it is syntactically ambiguous – the sentence ‘every man loves some woman’, for example, might be constructible in two ways corresponding to the two different scopings of the quantifiers. ‘Some philosopher fears Socrates’, as analyzed above, however, is not syntactically ambiguous – it receives different analyses simply because its unique tree structure can be characterized at different levels. The proper way to characterize compositionality in an ambiguous language is a difficult and contentious matter, and one which has received attention in the literature recently. We will thus restrict ourselves to unambiguous languages, in the following sense:

- $L$ is unambiguous if and only if there is some subset $\Phi$ (the basis set) of $\Delta$ such that:

---

35 See, for example, [83] for an argument that ambiguity blocks the possibility of compositional semantics, [27] and [28] for attempts to show how ambiguity and compositionality can be combined, and [103] for mathematical examination of notions of compositionality designed to take ambiguity into account.
− $\Delta \subseteq \Phi^n$ 36
− $\forall \epsilon \in L \exists ! \delta_i \in \Phi \exists ! \epsilon_1, \ldots, \epsilon_k \epsilon = \delta_i(\epsilon_1, \ldots, \epsilon_k)$

Suppose, then, that $L$ is unambiguous. Given an expression $\epsilon$ of $L$, we will say that one syntactic analysis $\delta_1(\epsilon_1, \ldots, \epsilon_k)$ of $\epsilon$ is a refinement of another syntactic analysis $\delta_2(\eta_1, \ldots, \eta_k)$ of $\epsilon$ just in case every $\eta_i$ is the result of applying some $\delta \in \Delta$ to some of the $\epsilon_j$'s. We can then prove the following result:

(Theorem 1) If $\Lambda_1$ is a collection of syntactic analyses $\delta_1^j(\eta_1^j, \ldots, \eta_m^j)$ for every $\epsilon \in L$, and $\Lambda_2$ is a collection of syntactic analyses $\delta_2^j(\theta_1^j, \ldots, \theta_m^j)$ for every $\epsilon \in L$, such that for $\epsilon \in L$ the analysis of $\epsilon$ in $\Lambda_1$ is a refinement of that in $\Lambda_2$, then if:

- There is some $G$ such that for all $\epsilon \in L$, $\|\theta\| = G(\delta_2^j, \|\theta_1^j\|, \ldots, \|\theta_m^j\|)$

then:

- There is some $G$ such that for all $\epsilon \in L$, $\|\theta\| = G(\delta_1^j, \|\theta_1^j\|, \ldots, \|\theta_m^j\|)$

If $\Lambda$ is a collection of syntactic analyses of expressions of $L$, then we say that $L$ is compositional at level $\Lambda$ just in case the meaning of each $\epsilon \in L$ is a function of the meanings of its $\Lambda$-given parts and their $\Lambda$-given mode of syntactic combination. The claim, then, is that if $L$ is compositional at a given level, it is also compositional at any refinement of that level. 37

Proof: Suppose $L$ is compositional at level $\Lambda_2$. Then $\delta_2^j$ together with the $\|\theta_i^j\|$'s are enough to determine $\|\epsilon\|$. Since $\epsilon$ is unambiguous, it receives a unique syntactic analysis using some subset $\Phi_1$ of the basis set of syntactic operations. Since $\Delta = \Phi^n$, $\delta_1^j$ must be a member of $\Phi_1$, so there must be some basis operations $\Phi_2$ which compose to make $\delta$. Now consider the syntactic analysis $\delta_1^j(\eta_1^j, \ldots, \eta_m^j)$ of $\epsilon$ from $\Lambda_1$. There must, as above, be some basis elements $\Phi_1$ which compose to make $\delta_1^j$. Moreover, it must be the case (a) that $\Phi_2 \subseteq \Phi_3$, and (b) that the functions $\delta$ by which the $\theta_i^j$ are constructed out of the $\eta_i^j$ (since $\Lambda_1$ is a refinement of $\Lambda_2$) are constructable by composition out of $\Phi_3$. Thus the value of $\Phi_3$, together with the $\|\theta_i^j\|$-values of the $\eta_i^j$, fix (a) $\delta_2^j$ and (b) the $\|\theta_i^j\|$'s, which then by $G$ give $\|\epsilon\|$. Thus the $\Lambda_1$ analysis of $\epsilon$ gives sufficient information functionally to determine the $\|\epsilon\|$-value of $\epsilon$ in terms of its $\Lambda_1$-analysis and the meanings of its $\Lambda_1$ parts.

Roughly speaking, if expressions in $L$ have tree-like syntactic analyses, then if the language is compositional at one level of the trees, it is also compositional at all lower levels. Two extreme cases are:

(Strong Compositionality) $L$ is strongly compositional if every expression has a coarsest non-trivial syntactic analysis $38$, 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 coarsest analysis. 39

36 Where $X^n$ is the closure of $X$ under functional composition.
37 Where the refinement of a level $\Lambda$ is a collection of refinements of each member of $\Lambda$.
38 If a non-trivial syntactic analysis at all; syntactic atoms, obviously, will not.
39 This definition is equivalent to [75]'s notion of strong 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.

If L is strongly compositional, it is compositional at any level, and if it is compositional at any level, it is weakly compositional. Typically, of course, many intermediary notions of compositional strength are available.

3.1.4 Fourth Draft

A meaning theory⟦·⟧for language L is thus compositional relative to a parthood relation ⊑ and a level of structural analysis {δ2(θ1, ..., θm) | ϵ ∈ L}, where the θi’s are ⊑-parts of ϵ, if there is some function G such that for all ϵ ∈ L, ⟦ϵ⟧ = G(< δ2, ⟦θ1⟧, ..., ⟦θm⟧>). Compositionality in this sense, however, is a very weak requirement, as shown by:

(Claim) If for any distinct ϵ1, ϵ2 ∈ L, ⟦ϵ1⟧ ≠ ⟦ϵ2⟧, then ⟦·⟧ is a compositional meaning function.40

Proof: Let G be the collection of ordered pairs < x, y > where y is the ⟦·⟧-meaning of an expression ϵ, and x is an ordered n-tuple of the syntactic rule used in constructing ϵ and the ⟦·⟧-meanings of the parts combined via that syntactic rule. G can only fail to be a function if it contains ordered pairs < x, y > and < x, z > with y ≠ z. If y = ⟦ϵ1⟧ and z = ⟦ϵ2⟧, then ϵ1 ≠ ϵ2 by the assumption on ⟦·⟧. If x = < δ, η1, ..., ηm >, then ϵ1 and ϵ are the result of applying rule δ to expressions having meanings η1, ..., ηm. But since no two distinct expressions can have the same ⟦·⟧-meaning, this means that ϵ1 and ϵ2 are built up from the same expressions by the same syntactic rule, and hence are identical. This contradicts the above, so G is a function.

Thus if a language contains no synonyms – no expressions agreeing in µ-meaning – it is automatically compositional in the above sense. But surely this is much too weak. Consider the following extreme example: in Lenglish, sentences have structured propositions as meanings, and a proper name contribute its referent to the structured proposition of a sentence in which it occurs, unless the sentence contains another name beginning with the same letter, in which case the name contributes a designated natural number (where each proper name is designated a different natural number). Thus the meaning of:

(6) Socrates kicked Plato.

might be <Socrates, <KICKED, Plato>>, but the meaning of:

(7) Socrates kicked Plato and Sophocles kicked Aristophanes.

could be < <17, <KICKED, Plato>>, AND, <29, <KICKED, Aristophanes>>>. Lenglish looks to fails as badly as is possible to be compositional: it violates semantic closure,

40See [105] for a brief note of this fact, and [98] for an argument that it shows that the functional approach does not capture the principle of compositionality.
since the meanings of expressions depend on morphological, in addition to semantic, properties of their parts (namely, initial-letter matching), and it violates semantic locality, since the meaning of a sentence which is part of a larger sentence may depend on features of distant parts of the larger sentence. Nevertheless, Lenglish will be compositional so long as it contains no two coreferential proper names.

A language with rules as wildly noncompositional as those of Lenglish should not be able to be made compositional just by deleting a few synonyms from the lexicon. The undesirable conclusion that it can be, fortunately, can be avoided through a richer conception of function. The above line of reasoning assumes that the meaning functions of a language are given in extension (thus allowing Lenglish to achieve compositionality by deleting choice ordered pairs from that extension), but the nature of our semantic competence suggests that meaning functions are in fact given in intension. Semantic mastery of a language includes not only the ability to determine the meanings of expressions in the language, but also the ability to determine what the meanings of complex expressions of the language would be, were they to include newly-introduced lexical items of specified meaning. Semantic mastery of English requires of me that I know the sentence:

(8) Sophia plays chess.

expresses the claim that that person plays chess.\footnote{Where, of course, the demonstrative ‘that person’ picks out Sophia. The form of my knowledge, though, needn’t be explicitly demonstrative; it need only fix on the relevant individual.} Semantic mastery of English also requires that I know that, were I to name my newborn, and as-yet-unnamed, daughter ‘Maeve’, then the sentence:

(9) Maeve plays chess.

expresses the claim that that daughter plays chess. Within certain bounds\footnote{Adding a new word with a meaning inappropriate for its syntactic category might, for example, exceed those bounds. My mastery of English is in no way impugned if I find myself baffled by the meaning of the sentence ‘He was an old man who fished alone in a skiff in the Gulf Stream quand he had gone eighty-four days now without taking a fish’ where ‘quand’ is a newly-introduced sentential connective which is given as meaning not (for example) some binary truth function, but rather an arbitrary concrete object such as the Eiffel Tower.}, someone who expresses bafflement about the meanings of complex expressions containing a newly-introduced word of known semantic value betrays a prior lack of understanding of the language – the rules of the language already determined what sentences in extensions of the language would mean.

This line of thought suggests the following modification to the functional definition of compositionality:

- A meaning theory $\llbracket \cdot \rrbracket$ 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 $L$, and (iii) a level of structural analysis $\{\delta^{\epsilon}_{L_i}(\theta^1_{L_i}, \ldots, \theta^m_{L_i}) \mid \epsilon \in L_i\}$ for each $L_i \in L^+$, where the $\theta^i_{L_i}$'s are $\sqsubseteq_{L_i}$-parts of $\epsilon$, and where again the levels of structural analysis agree on $L$, if for each $L_i \in L^+$ there is some function $G_i$ such that for all $\epsilon \in L_i$, $\llbracket \epsilon \rrbracket_{L_i} = G(<\delta^{\epsilon}_{L_i}, \llbracket \theta^1_{L_i} \rrbracket_{L_i}, \ldots, \llbracket \theta^m_{L_i} \rrbracket_{L_i}>)$, and the $G_i$'s agree on $L$. \footnote{Adding a new word with a meaning inappropriate for its syntactic category might, for example, exceed those bounds. My mastery of English is in no way impugned if I find myself baffled by the meaning of the sentence ‘He was an old man who fished alone in a skiff in the Gulf Stream quand he had gone eighty-four days now without taking a fish’ where ‘quand’ is a newly-introduced sentential connective which is given as meaning not (for example) some binary truth function, but rather an arbitrary concrete object such as the Eiffel Tower.}
Even if Lenglish contains no synonymous names, it will on this modified definition still not be compositional relative to a class of extensions some of which do contain synonyms. Of course, every language will be noncompositional relative to some class of extensions\textsuperscript{43}, so the crucial point will be what classes of extensions matter to compositionality considerations. The suggestion from above is that speakers’ semantic mastery will provide 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 will depend on why one is interested in compositionality.\textsuperscript{44,45}

3.1.5 Final Draft

Szabo, in [98] and [97], presents the following challenge to the adequacy of the functional conception of compositionality:

Let $L$ be a rich compositional fragment of English which contains the sentences ‘Elephants are gray’, and ‘Julius Caesar was murdered on the ides of March’. Let the function that assigns meanings to expressions of $L$ be $g$. $L'$ is a language with the same expressions as $L$, but a different interpretation. The meanings of expressions in $L'$ can be obtained from the corresponding meanings in $L$ via some permutation $p$. $p$ leaves every meaning as it is in English, except that it exchanges the meanings of sentences synonymous with ‘Elephants are gray’ and the meanings of sentences synonymous with ‘Julius Caesar was murdered on the ides of March’. $L'$ is interpreted in accordance with the principle of functionality: there is a function $g'$ which assigns meanings to the expressions of $L'$. For any expression $e$ of $L'$, $g'(e) = p(g(e))$. On the other hand, if the sentence ‘Elephants are gray’ is a genuine complex expression in $L'$, $L'$ violates the principle of compositionality. The assumption that $L'$ is compositional implies that the meanings of ‘elephant’ and ‘gray’ plus pluralization and predication determine the meaning of ‘Julius Caesar was murdered on the ides of March’ in $L$, and hence in English as well, which is absurd. ([98], 19-20; see also [97], 484-488)

$L'$ is, as Szabo notes, compositional by the functional standard, via composition of the meaning function for (the fragment of) English with the permutation $p$. If $L'$ ought

\textsuperscript{43}Consider the extension $L'$ of language $L$ which adds a new expression $\alpha$ synonymous with some existing expression $\beta$ of $L$, but which sets \texttt{⟦Σ(α)⟧'} ≠ \texttt{⟦Σ(β)⟧'} for some syntactically appropriate frame $\Sigma$. No meaning function will be able to derive the distinct meanings of $\Sigma(\alpha)$ and $\Sigma(\beta)$ from their identical syntactic structure and part meanings, so $L$ will be noncompositional relative to the extension of $L$.

\textsuperscript{44}Regarding which see section 6 below.

\textsuperscript{45}[98],[97] claim that the triviality of the substitutional conception of compositionality (on which see section 3.2 below) shows the inadequacy of that conception for capturing the intuitive notion of compositionality; that triviality worry is a direct analog of the triviality worry for the functional conception given the close logical relation between the functional and substitutional conceptions spelled out below. The relativization of compositionality to a class of extensions provides an answer to that worry, one which can be thought of as a generalization of Szabo’s suggestion that compositionality be treated as a modal supervenience relation with the modality ranging over the space of possible human languages.
not be compositional, this is a problem for the functional standard. But why should we think \( L' \) is not a compositional language? According to Szabo, because this assumption will entail that in \( L \) and English, the meanings of ‘elephant’ and ‘gray’, along with the syntactic rules of pluralization and predication, determine the meaning of ‘Julius Caesar was murdered on the ides of March’. What, though, does it mean to say that factors \( X \) determine the meaning of a sentence \( \sigma \) to be \( Y \) in a language \( L \)? If it means that the procedures that language actually uses to combine meanings of parts to produce meanings of wholes act on \( X \) to produce \( Y \) as the meaning of \( \sigma \), then Szabo’s argument requires the following implicit premise:

- \( L \) and \( L' \) have the same procedures for combining meanings of parts (plus syntax) into meanings of wholes.

In fact Szabo does accept this premise, endorsing the following strong conception of compositionality:

(Supervenient Compositionality) For all possible human languages \( L \), for any meaning property \( M \) and for any complex expression \( E \), if \( E \) has \( M \) in \( L \), then there is a constitution property \( C \) such that \( e \) has \( C \) in \( L \), and for any possible human language \( L' \) if any complex expression \( e' \) in \( L' \) has \( C \) in \( L' \) then \( e' \) has \( M \) in \( L' \). ([98], 25)

But this seems an overly strong requirement.\(^46\) Prima facie, for example, a human language could be such that the syntactic operation of predication induced the semantic effect of negation. That is, there could be a language English’ in which the sentence ‘Elephants are gray’ means what ‘Elephants are not gray’ means in English, and so on. The default way in English of using a predicate serving to pick out a property is to ascribe that property to some object(s), with other methods of use such as denying that property to some object(s) requiring the use of further syntactic devices such as overt negation, but a human language needn’t follow English in this default pattern, and could instead take the denial of a property to some object(s) as the default way of using a predicate.\(^47\),\(^48\)

Another example: in Romance languages, some adjectives can appear either to the left or to the right of the noun they modify, with differing semantic impact depending on their location. In such cases, the adjective modifies the noun in a literal way when postposed and in a figurative way when preposed, as in:

\[\text{‘Some elephant is gray’} \Rightarrow \text{‘Some gray elephant exists’}\]

\[\text{‘Some gray elephant exists’} \Rightarrow \text{‘Some elephant is gray’}\]

\(^46\)However, if it is wanted, it can be imposed in the current framework by taking as an extension of \( L \) a language \( L' \) which is the union of all possible human languages (with, perhaps, some lexical alterations (e.g., subscripting) performed to avoid gratuitous ambiguity). \( L' \)’s compositionality relative to the extension \( L' \) in the functional sense will entail its compositionality in Szabo’s supervenience sense.

\(^47\)Strictly speaking, Szabo does not deny that such languages are humanly possible – he merely endorses a standard of compositionality which entails that if such languages are humanly possible, then no language is compositional.

\(^48\)It could be denied that such a language is a possible human language on the grounds that in such a language there would be no reason to take the predicate ‘is gray’ (for example) to pick out the property gray rather than the property not gray. As discussed in section 4.1.2 below, sufficient tinkering with lexical semantics can save the compositionality of any meaning theory, potentially depriving this line of response of interest. In any case, the location of the negation in the syntactic construction of predication, rather than in the predicate itself, might be overt via different truth conditions for ‘Some elephant is gray’ and ‘Some gray elephant exists’.
Suppose, then, that Romance languages have a rule according to which the position of the adjective determines whether its meaning contributes literally or figuratively. By Szabo’s condition for compositionality, such languages will be compositional only if all possible human languages follow this rule. Were there a language in which preposed adjectives generated literal readings while postposed adjectives generated figurative readings, the Romance language semantic pattern would not be constant across all human languages. And since Szabo requires constancy across all possible human languages, the prima facie possibility of such a “swapped” language is evidence against the compositionality of Romance languages. This is a very strong condition on compositionality.

Even if we reject Szabo’s requirement that $L$ and $L'$ combine meanings of parts (plus syntax) into meanings of wholes in the same way, however, we might still feel that 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, no matter what method of construction we use. While you may be able to make an omelette from eggs and milk using one recipe (language) and a creme brulée 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.\footnote{Or, perhaps, the ways of combining them that do yield roast venison no longer count as cooking – ‘Lure a deer to a target site with a bowl of milk. Drop a dozen eggs on the deer from a height of 1000 meters, stunning it. . . . ’ – just as some ways of combining meanings no longer count as semantics.} 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.

Restrictions on what a genuinely semantic mode of meaning combination can achieve should be made cautiously – negation, after all, can be thought of as a semantic operation which takes us from one region of semantic space to a wholly disconnected one, so it is certainly possible for semantic operations to take meanings from one ‘place’ to another – but if we do indeed think that certain kinds of operations (such as Szabo’s permutation function $p$) are not semantically permissible, this conviction can be incorporated into the definition of compositionality by requiring that the function $G$ come from a class of permissible functions.\footnote{\cite{109}’s notion of systematicity, which requires that the meaning function belong to a “given class”, follows along these lines.}

This creates another element of relativization in the definition:

- A meaning theory $\mu$ for language $L$ is compositional relative to (i) a class $L^+$ of extensions of $L$, (ii) a parthood relation $\sqsubseteq_L$ for each $L_i \in L^+$, where the parthood relations for the extensions in $L^+$ agree on $L$, (iii) a level of structural analysis $\{\delta^*_L(e, \theta^1, \ldots, \theta^m) \mid e \in L_i\}$ for each $L_i \in L^+$, where the $\theta^i$’s are $\sqsubseteq_L$-parts of $e$, and where again 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 $e \in L_i$, $\mu_{L_i}(e) = G_i(\delta^*_L(e, \mu_{L_i}(\theta^1), \ldots, \mu_{L_i}(\theta^m)))$, and the $G_i$’s agree on $L$.
If, for example, complexes have something like structured propositions as meanings, and the class $G^+$ of permissible functions is constrained so that no function $G \in G^+$, when acting on structured meanings containing objects $O_1, \ldots, O_n$ as constituents, produces a structured meaning containing an object $O'$ distinct from the $O_i$'s, then a language which has a claim about Julius Caesar expressed in the words ‘Elephants are gray’, where those words mean what they do in English, cannot be compositional relative to $G^+$.

By making the class $G^+$ of permissible composition techniques small, the corresponding notion of compositionality can be made quite strong. Some examples:

- If $G^+$ contains only the technique of functional application, so that $G(<\delta, \mu(\epsilon_1), \mu(\epsilon_2)>) = \mu(\epsilon_1)(\mu(\epsilon_2))$\(^{51}\), where $G \in G^+$, then the corresponding compositionality constraint requires semantics to be done in a Montagovian style.

- Frege appears to adopt the very strong view, which Szabo in [98] calls strong parallelism, that the sense of a complex actually contains the senses of its parts as parts:\(^2^{52}\)

  A proposition consists of parts which must somehow contribute to the expression of the sense of the proposition, so they themselves must somehow have a sense. Take the proposition ‘Etna is higher than Vesuvius’. This contains the word ‘Etna’, which occurs also in other propositions, e.g. in the proposition ‘Etna is in Sicily’. The possibility of our understanding propositions which we have never heard before rests entirely on this, that we construct the sense of a proposition out of parts that correspond to the words. If we find the same word in two propositions, e.g. ‘Etna’, then we also recognize something common to the corresponding thoughts, something corresponding to this word. Without this, language in the proper sense would be impossible. ([42], 127)

If $G^+$ contains only functions which extend mereological fusion, then compositionality relative to $G^+$ will implement strong parallelism.

- Without any constraint on $G^+$, a compositional semantics can treat apparently quite similar constructions very differently. A language in which the meaning of a predicate-term construction $\Pi \tau$ is true iff $[\tau] \in [\Pi]$ except when $[\tau] = \text{Abraham Lincoln}$, in which case $[\Pi \tau]$ is true iff Kurt Gödel proved the incompleteness of arithmetic, can be compositional relative to an unconstrained $G^+$. If, however, we require elements of $G^+$ to treat all (meanings of) members of each syntactic category the same (with an adequately constrained sense of what counts as “the same”), such semantics can be ruled out. Some of the intuitive

\(^{51}\)Assuming, without imposing any real constraint on $G$, that the first argument is higher on the type hierarchy.

\(^{52}\)The strong parallelism constraint, however, is not meant to apply to the level of reference. While the referent of ‘the capital of England’ is determined by the references of the component words, England itself is no part of London, the referent of ‘the capital of England’. (See [43], 275 for this point.) It is thus clear that compositionality can be satisfied without endorsing strong parallelism.
dissatisfaction with the ‘Elephants are gray’/‘Julius Caesar was murdered on the ides of March’ permutation can thus be dealt with.

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. Having made these distinctions, we will proceed largely to ignore them in the subsequent discussion, assuming unless otherwise noted:

1. that the parthood relation is given by the syntactic theory of the language.
2. that the coarsest-grained (non-trivial) structural analysis is used (‘strong compositionality’).
3. that no extensions of the language are appealed to, but that the language in question contains a robust collection of synonyms.
4. that any meaning composition function is admissible.

3.1.6 Compositionality as Homomorphism

Assuming that the parthood relation $\sqsubseteq$ reflects a collection $\Delta$ of syntactic constructions, compositionality can be thought of 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 corresponding 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 [5] called the rule to rule hypothesis:

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

The rule to rule hypothesis is often described formally as the requirement that syntax and semantics be homomorphic: where both the syntax and the semantics of $L$ are thought of 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), we require that there be a homomorphism between the two algebras preserving the structure imposed by the operations of each.\(^{53}\)

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 $G$ will clearly be a composition function for all of $L$. If, on the other hand, $L$ has a composition 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) >).\(^{54}\)

\(^{53}\)See [79], [58], [48], [109], [59], and [105] for examples of this approach.

\(^{54}\)See Theorem 4 of [55] for a slightly more precise statement of this result.
3.2 Compositionality as Substitutability

Suppose that the meaning of a proper name is exhausted by its referent. Suppose also that Lex Luthor believes that Superman is harmed by Kryptonite, but that Lex Luthor does not believe that Clark Kent is harmed by Kryptonite. Then compositionality has a problem. The two sentences:

(12) Lex Luthor believes that Superman is harmed by Kryptonite.
(13) Lex Luthor believes that 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, 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.

More generally, suppose that 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 \( \mu \) 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 \( i \), either \( \eta_i = \theta_i \) or \( \mu(\eta_i) = \mu(\theta_i) \), we have \( \mu(\epsilon_1) = \mu(\epsilon_2) \).

A language is compositional, that is, if substitution of synonyms always preserves synonymy.

Some care is needed, however, in saying what counts as the substitution of synonyms. Szabo ([98], 16-18) gives the example of the prima facie synonymous sentences:

(14) Plato was bald.
(15) Baldness was an attribute of Plato.

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

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

---

55 This can, of course, be denied. Triadic theories of attitude verbs typically take sentences like (12) and (13) 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.
If the substitutional conception of compositionality is to be a reasonable one, it must not count (17) as a substitution instance of (16) with (14) swapped for (15).\(^{56}\) Substitution thus cannot be simple replacement of strings of words. Instead, substitution needs to respect syntactic structure. Suppose that (16) has the structure:

(16-T)

\[
\begin{array}{c}
S \\
\vdots \\
DP \\
\\vdots \\
D \\
\text{The} \\
\vdots \\
NP \\
\\vdots \\
N \\
\text{philosopher} \\
\vdots \\
S \\
\\vdots \\
DP \\
\\vdots \\
VP \\
\text{was bald} \\
\vdots \\
VP \\
\text{was Plato} \\
\vdots \\
\end{array}
\]

Then it is not the result of applying some syntactic operation to ‘Plato was bald’ and some other parts, assuming that the syntactic operations simply build up the tree structure. The definition of substitutional compositionality given above requires that when \(\alpha\) is substituted for \(\beta\), it is substituted in an expression which is derived from \(\beta\) (and other expressions) by some syntactic rule, and results in an expression which is derived from \(\alpha\) (and the same other expressions) by the same syntactic rule. Roughly: substitution must preserve full syntactic structure, not just overt word order.\(^{57}\)

\(^{56}\) Or it must deny the synonymy of (14) and (15). As Szabo points out, (a) there are plausible reasons for denying that synonymy, but (b) the unacceptability of this case as a counterexample to compositionality ought to be identifiable independent of settling the synonymy of (14) and (15). Similar examples can be generated with almost any complex phrase, so if any complex phrases are synonymous, this worry for substitutionalism must be dealt with.

\(^{57}\) Szabo takes the substitutional notion of compositionality to be too strong because a language can be compositional without counting (16) and (17) as synonymous. He agrees, however, that this undue strength can be avoided by ‘defining substitution in such a way that a non-structure-preserving replacement of one expression by another within a third would not count as substitution’; such a definition is incorporated in the substitutional principle given here. Szabo then worries that this approach ‘would have the unfortunate consequence that the innocent notion of substitution would become theoretically loaded: we no longer have a test to decide whether a particular sentence is the result of a substitution within another one.’ ([98], 18) More precisely, we no longer have a test which is independent of syntactic theorizing. Whether it is unfortunate or not, this consequence seems unavoidable – any plausible attempt to see if the meaning of a sentence derives from the meanings of its parts must include an attempt to see what the parts of that sentence are, and if contemporary syntax is at all on the right track (and if the relevant parthood relation is syntactic), then any
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 much 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 different expressions formed from parts of those meanings had different meanings. Let \( \Sigma(\alpha) \) and \( \Sigma(\beta) \) be two such expressions. Then \( \alpha \) and \( \beta \) have the same meaning, but \( \Sigma(\alpha) \) and \( \Sigma(\beta) \) do not, violating substitutional compositionality. Thus \( L \) is indeed compositional in the functional sense. The functional and substitutional definitions of compositionality, then, can be used interchangeably. The substitutional definition, however, has two advantages that occasionally make it desirable:

- **Application of the substitutional definition to a particular language is often easier than application of the functional definition.** To determine whether language \( L \) meets the functional definition, we must have in hand a meaning function \( \mu \) for \( L \), and then determine whether a composition function \( G \) can be fitted to \( \mu \). Since we rarely are so fortunate as to have a meaning function for a language in hand, this avenue of approach is rarely practical. Determining whether \( L \) meets the substitutional definition, on the other hand, does not require knowing anything about the structure of \( \mu \) for \( L \) – we need only be able to recognize synonyms in \( L \), and then search for substitution instances which do not preserve synonymy.

- **The functional approach presupposes a reified conception of meaning.** For a language to be compositional in the functional sense, there must be a function \( \mu \) mapping expressions to things, in some suitably broad sense of things, serving as the meanings of those expressions. Functional compositionality, then, rests on an external notion of meaning as a relation between expression and thing meant.

---

Plausible such attempt must look beyond the mere word order of the sentence.

Szabo also claims that the substitutional version of compositionality is too weak, because it is trivially satisfied in a language without synonyms. This is exactly the problem considered for the functional account in section 3.1.4, and the same answer is available here.

58 This step in the proof illicitly supposes that the two expressions differ only by intersubstitution of a 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 [55] to guarantee that large-scale substitutions could be achieved by a sequence of single-pair substitutions. See footnote 85 below for more details on the latter.

59 See [55], Theorem 4, for more precise versions of each direction of this proof.
Substitutional compositionality, on the other hand, rests on an internal notion of meaning as a (synonymy) relation between expressions of the language. No commitment to reified meanings is needed to employ the substitutional test for compositionality. The substitutional conception of compositionality is thus useful for considering the compositionality of semantic theories (such as Davidsonian ones) which eschew talk of ‘meanings as entities’ ([21], 20).

3.3 Some Related Concepts

The logically equivalent functional and substitutional definitions of compositionality make precise the principle of compositionality (in fact, make precise many principles of compositionality, for many different choices of relativizing parameters). In this section we will consider briefly several other concepts which have, or have been taken to have, important relations to compositionality. Looking at these related concepts will help sketch the boundaries of the principle of compositionality, and some of the related concepts will be of importance in section 6 when we turn to consideration of reasons for wanting compositionality in a language.

3.3.1 Computability

A language $L$ is computable just in case its meaning function $\mu$ is a computable one, either in the formal sense of being calculable by a Turing machine or in the informal sense of being the output of some (ideally) humanly performable calculating procedure. A natural language, because its speakers are in principle able to determine meanings of any of its well-formed expressions, must have a computable meaning function. Compositionality might then be taken to explain the computability of a meaning function, on the assumption that the functional dependence of complex meanings on part meanings will yield a particular method of computing the meanings of complexes (picturesquely, by working one’s way up the syntactic tree from the semantic atoms). However, computability and compositionality turn out to float quite free of one another:

- A language can be compositional without being computable. Let $L$ be a language with a single syntactic operation of concatenation, and suppose expressions of $L$ take natural numbers as meanings, and let $M$ be an arbitrary noncomputable function on the natural numbers. If $\mu(\alpha \circ \beta) = M(\mu(\alpha), \mu(\beta))$, then $L$ is compositional, but $\mu$ will not typically be a computable function. Having meanings of complexes built out of meanings of parts does not make the meanings of the complexes computable unless the way they are built out of the meanings of the parts is itself computable.

- A language can be computable without being compositional. Suppose again that $L$ is a language with concatenation as its only syntactic operation, and with

---

60 Of course, the equivalence of the substitutional and functional senses of compositionality holds only for languages with reified meanings.

61 [59], Appendix A, gives another list of concepts related to the principle of compositionality, focusing primarily on alternative formulations of that principle and corollaries of it.

62 The meanings of the $\subseteq$-minimal parts will be trivially computable if they are finite in number, as they will be in any natural language.
expressions taking natural numbers as meanings. Suppose that the meaning of \( \alpha \sim \beta \), when it appears in the context \( \gamma \sim (\alpha \sim \beta) \), is \( (\mu(\alpha) + \mu(\beta)) - \mu(\gamma) \). The meaning of any expression is then straightforwardly computable, but the computation proceeds in a noncompositional way, since it violates semantic locality.

Compositionality is thus neither necessary nor sufficient for computability. Compositionality via a computable composition function gives computability, and computability by a compositional procedure gives compositionality, but nothing more substantive than this can be said.

3.3.2 The Context Principle

In the preface to The Foundations of Arithmetic, Frege sets out three guiding principles for his investigation, the second and most famous of which is the so-called ‘context principle’:

Never to ask for the meaning of a word in isolation, but only in the context of a proposition. ([41], x)

Interpretation of the context principle is a vexed matter, but for our purposes we will focus on a reading of it on which it endorses a principle of top-down semantics, on which the data to which a semantic theory is accountable constrains only the meanings of certain complexes (for example, sentences), leaving the semantic values of parts to be introduced as theoretic entities to account for the pretheoretic data. Thus if two words always make the same contributions to the meanings of complexes, they themselves must have the same meaning. Formally:

(\text{CP}) Language \( L \) satisfies the context principle if and only if \( \alpha \) and \( \beta \) are such that 
\[
\mu(\Sigma(\alpha)) = \mu(\Sigma(\beta)) \text{ for any grammatically permissible frame } \Sigma,
\]
than \( \mu(\alpha) = \mu(\beta) \).

Standing in opposition to top-down semantics is bottom-up semantics, on which a semantic theory is accountable to data directly constraining the meanings of individual lexical items. Given that compositionality presents a semantic picture of meanings of complexes being built out of meanings of parts, it is natural to take it to require the meanings of the parts to, as it were, really be there in the first place, and thus to take the context principle and the principle of compositionality to conflict and represent two incompatible modes of doing semantics. Thus, for example:

Compositionality requires that words in isolation have a meaning and that from these meanings the meaning of a compound can be built. The formulation of contextuality given above disallows speaking about the meaning of words in isolation and is therefore incompatible with compositionality. ([59], 420)

On the other hand, [55] has recently claimed that compositionality provably follows from the context principle. The discussion of section 4 below will show that, in a sense, both of these views are correct.
3.3.3 Systematicity

Starting with [37], Fodor and others have, in a series of papers, argued that various accounts of content (both linguistic and mental, and including connectionism, prototype theories of concepts, and inferential role semantics) fail because those accounts cannot explain the **systematicity** of various representational systems, and also that compositionality is crucial to the achievement of systematicity. Systematicity is:

Roughly, the fact that any natural language that can express the proposition $P$ will also be able to express many propositions that are semantically close to $P$. If, for example, a language can express the proposition that $aRb$, then it can express the proposition that $bRa$; if it can express the proposition that $P \rightarrow Q$, then it can express the proposition that $Q \rightarrow P$; and so forth.\(^{63}\) ([33])

Compositionality might then be taken to explain systematicity on the grounds that, if meanings of complexes are functions of meanings of parts plus structure, then making small alterations in the structure (reversing subject and object, for example), or making small alterations in the meanings of parts will yield small alterations in the meanings of wholes. This might at least suggest that, if a compositional language comes well-stocked with syntactic structures and with lexical items of a wide range of meanings, systematicity will result. However, compositionality and systematicity turn out to be independent issues:

- A language can be compositional without being systematic. Consider a toy example: a language $L$ with two lexical items $a$ and $b$, and a single rule of concatenation which applies only to distinct and atomic inputs (so that $a \sim b$ and $b \sim a$ are the only complex expressions of $L$). Suppose that:
  
  - $\mu(a) = \text{Jane Austen}$
  - $\mu(b) = \text{Charlotte Bronte}$
  - $\mu(a \concat b) = \text{the proposition that Socrates admires Plato.}$
  - $\mu(b \concat a) = \text{the proposition that Godard directed \textit{Alphaville}.}$

  $L$ is not systematic, since, although it can represent the proposition that Socrates admired Plato, it cannot represent the semantically close proposition that Plato admires Socrates.\(^{64}\) However, $L$ is compositional – trivially so, since it contains no synonyms (thus the function $G = \{<<\text{Jane Austen},\text{Charlotte Bronte}>\}$, that

---

\(^{63}\)Systematicity should not be confused with the claim that if (for example) ‘$aRb$’ is meaningful in a language $L$, ‘$bRa$’ is also meaningful in $L$. Supposing that this latter claim is fleshed out with enough provisos about syntactic categorization to make it substantive, it (a) is quite distinct from systematicity, since it carries no guarantee that the close syntactic variants will have similar meanings, (b) does not require for its explanation any particular assumption about the semantics of the language, being instead open to explanation in terms of the cognitive efficiency of grammatical rules which are phrased in terms of syntactic category rather than in terms of particular lexical items, and (c) may well be false, as [33] notes when observing that ‘John calculated the answer’ can express a proposition in English without ‘The answer calculated John’ also doing so.

\(^{64}\)This feature of $L$, moreover, is not an artifact of the toy size of the language. $L$ could be expanded to a fully productive language and still be unable to express that proposition.
Socrates admires Plato, Charlotte Bronte, Jane Austen, that Godard directed Alphaville is a composition function for L.

More generally, the point is that a mapping from part meanings to whole meanings needn’t preserve any feature like semantic distance – inputs which are semantically quite close to each other (such as different orderings of the same meanings) can be mapped to outputs arbitrarily far from one another. Systematicity requires not just compositionality, but compositionality with a systematic meaning function.

- A language can be systematic without being compositional. Let $M$ be a set of meanings closed under the relation of semantic closeness, and let $L$ be a language consisting of:
  - For each $m \in M$, a syntactic atom with $m$ as meaning.
  - Two syntactic atoms $\alpha$ and $\beta$ with the same meaning (in $M$), each of which can serve as the input to some syntactic rule yielding expressions $\Sigma(\alpha)$ and $\Sigma(\beta)$ which are distinct in meaning.

$L$ is thus immediately systematic, and given that it is built to violate the substitution constraint, it is not compositional. Adding additional constraints to $L$ (such as productivity or finite learnability) looks unlikely to help in securing compositionality; the general worry here is that once a language is systematic, its compositionality can be taken away just by introducing a new semantic procedure by which parts of the same meaning can yield complexes of different meaning. Compositionality requires not just systematicity, but also a compositional implementation of systematicity.65

65The above examples might seem unfair to Fodor and Lepore, who appear to favor an understanding of compositionality as strong parallelism:

If a sentence $S$ expresses the proposition that $P$, then syntactic constituents of $S$ express the constituents of $P$. ([33])

If $m$ is part of the meaning of ‘a’ and ‘a’ consists of ‘b’, then $m$ is part of the meaning of ‘b’. ([35])

It’s hard to see, for example, how primitive meanings could be birds or chairs since, whatever complex meanings may be it’s hard to see how birds or chairs could be parts of them. ([36])

Strong parallelism, of course, is a much stronger constraint than simple compositionality, and if systematicity indeed requires it, then the axe of the Fodorian argument can be wielded much more broadly than it currently is, since almost no current semantic project seeks to respect strong parallelism. However, moving to strong parallelism doesn’t obviously help the situation. Since strong parallelism is strictly stronger than compositionality, systematicity will no more entail it than it did compositionality. And a language can be strongly parallel without being systematic: if the parts $\mu(a)$, $\mu(b)$, and $\mu(R)$ are put together in very different ways, and perhaps with very different stuff, when calculating $\mu(aRb)$ than when calculating $\mu(bRa)$, then there is no reason to think that ‘aRb’ and ‘bRa’ will express semantically close propositions (and hence no reason to think that propositions semantically close to those expressed by either of these sentences will be expressible in the language).
3.3.4 The Autonomy of Syntax

A syntactic theory for a language $L$ is *autonomous* from the semantic theory for $L$ if the syntactic rules can be best phrased without reference to semantic properties of $L$. Consider, for example, the range of complements taken by various epistemic verbs:

(18) Albert knew where the conference was.
(19) Albert knew the location of the conference.
(20) Albert wondered where the conference was.
(21) *Albert wondered the location of the conference.
(22) Albert admitted where the conference was.
(23) Albert admitted the location of the conference.
(24) *Albert conceded where the conference was.
(25) *Albert conceded the location of the conference.
(26) *Albert inquired where the conference was.
(27) *Albert inquired the location of the conference.

An adequate syntactic theory needs to ensure that the right epistemic verbs allow the right kinds of complement constructions. One approach is to have epistemic verbs explicitly subcategorize for the appropriate construction types in their lexical entries. Since subcategorization information is part of syntactic theory, this approach respects the autonomy of syntax. Another approach is to seek semantic features of the epistemic verbs which control which complement constructions are permissible. Should controlling semantic features be available, and should a syntactic theory be best formulated in terms of those features, then the resulting theory would not be autonomous.\(^{66}\)

Whether syntax is indeed autonomous from semantics remains an open question. If, however, it is not, then the mere fact that a particular construction is grammatically permissible can yield information about the semantic environment in which a particular expression occurs. That information can then be used to work around difficulties for semantic locality. Suppose, for example, that in $L$ the expression $\epsilon$ appears to mean $m_1$ when in the scope of an expression meaning $M_1$, but to mean $m_2$ when in the scope of an expression meaning $M_2$. $\epsilon$ thus appears to violate semantic locality, since its meaning is a function of semantic information not dwelling in its parts. But suppose that the syntactic theory of $L$ is not autonomous. Then we can posit two homonymous terms $\epsilon_1$ and $\epsilon_2$, one with meaning $m_1$ and one with meaning $m_2$, and then have semantically-sensitive syntactic rules of the form:

- $\Sigma(\epsilon_1)$ is well-formed only if $\mu(\Sigma) = M_1$

\(^{66}\)See [45] and [25] for discussion of the above examples and their relevance to the autonomy of syntax. [45] argues that no semantic factors properly explain the observed distribution of permissible complements, while [25] claims that the properties of having positive epistemic commitment and having negative epistemic commitment make the requisite distinction.
• $\Sigma(\epsilon_2)$ is well-formed only if $\mu(\Sigma) = M_2$

The terms $\epsilon_1$ and $\epsilon_2$ will thus allow the language to behave compositionally, while mimicking the ability of $\epsilon$ to ‘look outward’ to determine its meaning.\(^67\)

3.3.5 Context Sensitivity and Insensitivity

Compositionality requires semantic closure, and hence requires meanings of sentences to depend only on semantic facts (plus structural facts). But meanings of sentences also depend on nonsemantic contextual facts. Thus:

• The meaning of ‘I am a philosopher’ depends on who the speaker is.
• The meaning of ‘Plato has not yet written the *Timaeus*’ depends on when the utterance is produced.
• The meaning of ‘He wore a red hat’ depends on whether it is immediately preceded by ‘A man walked in a park’, or ‘A mathematician walked into a bar’, or ‘Not every man didn’t walk in a park’, or ‘A woman walked in a park’, or by nothing at all.
• The meaning of ‘Every shelf was empty’ depends on whether it is immediately preceded by ‘I looked in the refrigerator for some food’ or ‘Nobody’s refrigerator provided anything to eat’ or ‘The bookstore had been ransacked’.
• The meaning of ‘I know that’s a zebra’ depends on whether the possibility of cleverly painted mules has been explicitly raised in the conversational context.
• The meaning of ‘Every author admires some playwright’ depends on whether it is followed by ‘He’s usually either Shakespearare or Marlowe’ or by ‘He recently died, saddening them all’.

These claims range from the truistic to the highly controversial\(^68\), but everyone should accept enough of them to make the point.

Does the context sensitivity of language doom compositionality from the start, or, equivalently, does compositionality require language to be context insensitive? Some authors answer affirmatively:

It’s generally agreed that the crucial condition that must be satisfied in order that compositionality should hold for a class of expressions is that

\(^{67}\)\([11]\) objects to compositionality on the grounds that it threatens the autonomy of syntax. He argues that compositionality requires a tight connection between syntax and semantics, which then allows semantic anomalies to influence the formation of syntactic theory. He cites in particular the semantic difficulties and diversities presented by bare plural constructions, such as the apparent difference in content between:

(28) Beavers build dams.
(29) Dams are built by beavers.

despite the apparent syntactic similarity between these two sentences, and more generally despite the apparent syntactic homogeneity of simple bare plural sentences.

\(^{68}\) And I don’t mean to be endorsing all, or any in particular, of them.
the satisfaction conditions of their constituents should be context independent. So, the fact that ‘brown dog’, ‘green dog’, ‘brown cat’, and ‘yellow cat’ are all compositional in English is part and parcel of the facts that: ‘brown’ means the same in the environment ‘…dog’ that it does in the environment ‘…cat’ ‘dog’ means the same in the environment ‘brown …’ that it does in the environment ‘green …’. Etc. Likewise, mutatis mutandis, for ‘brown poodle dog’ and the like. That compositionality requires context independence is, to repeat, the consensus view, and we simply take it for granted in what follows. ([35])

while others answer negatively:

It is an open question whether all alleged cases of context-dependency that arise in possible human languages can be treated in one of the two ways mentioned [earlier]. If they are all either metalinguistic – where context plays a role in determining what expression a given sequence of phonemes stands for, but not in determining what that expression means, or lexical – where context plays a role in determining what the simple constituents of a complex expression mean, but no additional role in fixing the meaning of the complex – then context-dependency is not a threat to compositionality. ([98], 91)

Formally, the latter authors have the right of it. Compositionality requires the meanings of complexes to depend on nothing but the meanings of their parts (plus structure), but it places absolutely no constraint on the meanings of the parts – these can depend on the meanings of other lexical items quite distant under the parthood structure, or on nonsemantic facts, or on nothing whatsoever. If, for example, ‘tall’ when prefixed to ‘man’ requires more than six feet of height, but when prefixed to ‘tree’ requires over one hundred feet of height, then we might assign to ‘tall’ a function from (the meaning of) the modified $N$ to a height requirement. Such a meaning might appear to violate semantic locality, but that constraint is limited in scope to complex expressions, and ‘tall’ is atomic. If, on the other hand, the combination of ‘very’ with ‘tall’ imposes different height requirements when prefixed to different $N$’s, despite no corresponding change in the meanings of ‘very’ or ‘tall’, then the complex expression genuinely violates semantic locality and hence compositionality. Context sensitivity threatens compositionality only when it cannot be localized in lexical items; the ‘very tall’ example shows how difficult it may be to find compelling examples of unlocalized context sensitivity, since the obvious temptation here is to make either ‘very’ or ‘tall’ appropriately context-sensitive.69

It is natural to cease requiring semantic locality and semantic closure at some level. While it is possible to satisfy these constraints indefinitely, either by having lexical items carry their semantic properties as brute and irreducible features or by having matching syntactic and semantic decomposition continue downward indefinitely, the more obvious picture is that eventually (and likely roughly lexically) semantic features

---

69 Since local context sensitivity poses no threat to compositionality, some recent work on context, such as [95], [96], and [98], ch. 5, has attempted to argue that all forms of context-sensitivity can be traced to the semantic influence of context-sensitive lexical items.
derive from other sources such as conventions, intentions, causal connections, and so on. Any such analysis will violate semantic locality, and any concomitant attempt to reduce the semantic to the non-semantic will violate semantic closure. Compositionality is intended as a constraint on the behaviour of the semantic realm once properly entered; not as a constraint on the entrance into that realm.

The danger of undermining the substance of a compositionality constraint through appeals to local context-sensitivity, however, should not be overlooked. As we will see in section 4.1.2 below, if a sentence contains even a single lexical item which can be context-sensitive in a wholly unconstrained way, then compositionality can always be satisfied for that sentence. More specifically, apparent violations of the substitution principle can be finessed by having two prima facie synonymous terms take on different meanings only when in the syntactic context witnessing the substitution failure. Frege’s theory of indirect reference is, on the semantic level of reference, an approach of this sort. Semantic innocence\(^70\), which is a ban on one sort of local context-sensitivity, can be imposed as a further requirement to avoid too trivially satisfying compositionality.

4 The Mathematics of Compositionality

With a precise definition of compositionality in hand, we turn to consideration of the implications of that definition. Before turning to the empirical, in section 5 below, we consider some \( a \ priori \) issues. In section 4.1, three arguments are considered, each reaching the conclusion that compositionality is a surprisingly weak constraint, one that any meaning theory does or can meet. In section 4.2, an argument in the other direction is considered, one claiming that compositionality is a surprisingly strong constraint, in that it (under minimal assumptions) fixes the meanings of all subsentential expressions in a language.

4.1 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 some two complex expressions differing only via intersubstitution of the now-synonymous lexical items. A theory of English, for example, which took the Millian line of assigning proper names only referents as semantic values, and which hence took ‘Clark Kent’ and ‘Superman’ to be synonymous, but which assigned:

\[(12) \text{Lex Luthor believes that Superman is harmed by Kryptonite.} \]
\[(13) \text{Lex Luthor believes that Clark Kent is harmed by Kryptonite.} \]

both (a) syntactic analyses which made one a substitution instance of the other and (b) different meanings, would be a noncompositional theory.\(^71\)

\(^70\)See \[19\].
\(^71\)It is in part for this reason, of course, that Millians are under pressure to accept that proper names intersubstitute \textit{salva significatio} in propositional attitude contexts.
Despite the ready availability of non-compositional semantic theories, some recent work has claimed that compositionality is a trivial requirement. The triviality claim obviously cannot be that all possible languages are compositional – instead, it is typically 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 simply giving a wholly new meaning theory (the latter of which can, of course, always be done compositionality), and such care is rarely given in the literature. 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.

4.1.1 Tinkering with the Top

Horwich, in [56] and [57], claims that ‘the compositionality of meaning places no constraint at all on how the meaning properties of words are constituted’ ([57], 154). While this is not intended as a complete triviality claim about compositionality – Horwich says that he is ‘not maintaining that compositionality is so trivial that no view about meaning could be in tension with it’ ([57], 159) – it is sufficiently trivializing ‘to be called “deflationary” . . . [and it] shows that the compositionality of meaning is much easier to explain than we have often been led to believe.’ ([57], 158). The heart of Horwich’s trivialization argument is:

The fact that “dogs bark” means what it does == or, as I will put it (maintaining the convention of capitalizing an English expression to obtain a name of its meaning), the fact that “dogs bark” means DOGS BARK – is constituted by whatever is the complex fact regarding its mode of construction and the meanings of its constituents. This turns out to be the fact that the sentence results from putting words meaning what “dog” and “barks” mean, into a schema meaning what “ns v” means: that is (employing my convention for referring to meanings), the fact that it results from putting words whose meanings are DOG and BARK into a schema whose meaning is NS V. Thus the meaning property

• x means DOGS BARK

consists in what I shall call the ‘construction property’

• x results from putting terms whose meanings are DOG and BARK, in that order, into a scheme whose meaning is NS V.

([57], 155-156)

Suppose you are given a language L, complete with atomic and complex expressions and a parthood relation between the two. Suppose you are given meanings for all of the 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 utterly 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 the meanings of atomic expressions, regardless of what the meanings of the atoms are.

Complete freedom to tinker at the top, then, gives a clear sense in which compositionality places no constraint on the meanings of words (compositionality here would be a leash tying a dog to nothing at all). The formal point is straightforward, but it also threatens to be uninteresting. One cannot, in general, construct a useful semantic theory via utter indifference to the meanings of complex expressions. Horwich does not allow sentence meanings to be wholly arbitrary, but only because he does not think there is, in the end, anything which is sentence meaning. The semantic properties of a language begin and end with the properties of the lexicon and syntactic constructions – "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" ([57], 155). As a technical device Horwich uses the fact that the sentence is constructed from parts with specified meanings as the meaning of the sentence, allowing a trivial compositional mapping; he might just as well have used the ordered n-tuple of lexical and syntactic-rule meanings, or an interpreted logical form labelled on the bottom level only, in which cases composition functions would also be trivially available.

The philosophical burden on a position which achieves compositionality on the cheap via a highly permissive view of sentential meaning is to show that the resulting semantic theory still has the resources to do the jobs that we want done. Horwich attempts to meet this burden either by arguing that the more constrained lexical meanings are sufficient to the task (as with sentential understanding above, or as with inferential relations) or by holding that various of these jobs are much easier done

---

72Imagine, for example, ‘discovering’ that all sentences are true by examining a meaning theory designed to achieve compositionality by matching pretheoretic intuitions about English word meanings, and then assigning all sentences true as meaning.

73See [74] or [75] for the idea of interpreted logical forms.

74Constrained, of course, not by compositionality, but by (in Horwich’s case) features of our acceptance practices.

75Thus:

Consider, for example, the sentence “All emeralds are green”. Suppose (enormously oversimplifying) that the meaning of “emerald” is given by the fact that

- There is a disposition to accept “Something is an emerald if and only if it has characteristics, $f_1, f_2$, and $f_3$”.

Suppose that the meaning of “green” is given by the fact that

- There is a disposition to accept “That is green” in the presence of something clearly green, and otherwise to deny it.

And suppose that the meaning of “All As are Bs” is given by a tendency to conform to certain classical rules of inference, including, for example,

- From “All As are Bs” infer “All CA$s$ are Bs”.

These facts constrain the use of “All emeralds are green”. They tell us something about what will be inferred from it, and about the circumstances in which it will be accepted. They tell us, for example, that we will infer “All big emeralds are green” from it; and that, given the
than we might have thought (as with accounting for truth conditions, via Horwich’s minimalist views on truth). Whether Horwich’s burden-shouldering is successful lies outside the scope of this discussion; the role of compositionality is merely to present the semantic check, not to arrange for a payment schedule. 76

4.1.2 Tinkering with the Bottom

Suppose $L$ is an arbitrary language, consisting of a set $\mathcal{L}$ of expressions closed under the syntactic operation of concatenation. 77 Suppose $m$ is a meaning function, assigning members of $\mathcal{L}$ to various meanings. This meaning function can be wholly arbitrary, and in particular can be as noncompositional as desired. Nevertheless, Zadrozny, in [109], has claimed 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. ([109], 329)

How is this possible? Consider a small example using a minimally noncompositional language. Let $L$ contain atomic elements $a$, $b$, and $c$, and assuming the following facts about $m$:

- $m(a) = 1$
- $m(b) = 1$

But the plausibility of the claim that inferential propensities associated with particular words will suffice to characterize inferential propensities of sentences is perilously closely tied to the plausibility of the claim that the language has a compositional semantics by means of which the semantic properties of whole sentences are fixed by semantic properties of words. Our confidence (what we have) that inference rules associated with individual quantifiers can handle sentences with quantifier-alternation patterns, or [50]’s:

- Some relative of every townsman and some relative of every villager hate each other.

But with its branched quantifiers, is based on a background confidence that the inferential properties of sentences do indeed derive from the inferential properties of their constituent words. Horwich’s trivial compositionality, or at least the satisfactoriness thereof, may therefore rest on an underlying substantive compositionality.

The claim that understanding all the words in a sentence suffices for understanding the sentence would also rely on a background compositionality assumption. Horwich’s stronger claim that collective lexical understanding just is sentential understanding seems mysterious, appearing to entail:

- that were I to define a new transitive verb ‘frozzles’ using a meaning appropriate to a noun, such as Horwich’s clause for ‘emerald’:
  
  - The meaning of ‘frozzle’ is given by the fact that there is a disposition to accept ‘Something is an emerald if and only if it has characteristics $f_1$, $f_2$, and $f_3$.’

  I would thereby be in a position to understand the sentence ‘Socrates frozzled a small crab’.

- that were adjectival modification of determiners grammatically permissible in English, I would understand the ‘sentence’ ‘Socrates saw blue several tall trees’.

76 For more detailed criticisms of Horwich’s views on trivial compositionality, see [36] and [47].

77 I follow [109] in considering only languages with only a single syntactic operation. Nothing in the subsequent discussion will depend crucially on this restriction.
\[ m(c) = 2 \]
\[ m(a \sim c) = 3 \]
\[ m(b \sim c) = 4 \]

\( L \) 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 \( m \)” ([109], 330). Zadrozny’s \( \mu \) function will satisfy the following two constraints:

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

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

- \( \mu(a) = \{ <a, 1 >, <\mu(c), \mu(a \sim c) > \} \)
- \( \mu(b) = \{ <b, 1 >, <\mu(c), \mu(b \sim c) > \} \)
- \( \mu(c) = \{ <c, 2 > \} \)
- \( \mu(a \sim c) = \{ <a \sim c, 3 > \} \)
- \( \mu(b \sim c) = \{ <b \sim c, 4 > \} \)

The two unsolved terms easily resolve to:

- \( \mu(a) = \{ <a, 1 >, <\{c, 2 \}, \{a \sim c, 3 \} > \} \)
- \( \mu(b) = \{ <b, 1 >, <\{c, 2 \}, \{b \sim c, 4 \} > \} \)

Inspection shows that the resulting \( \mu \) meets both the agreement and the compositional-ity constraints. More generally, \( \mu \) is constructed from \( m \) by solving the set of simultaneous equations given by \( \mu(\epsilon) = \{ \epsilon, m(\epsilon) > \} \bigcup \{ \mu(\eta), \mu(\epsilon \sim \eta) > : \epsilon \sim \eta \in L \}.^{78} \)

The compositional \( \mu \) function matches the noncompositional \( m \) function in that \( \mu \), when applied to an expression \( \epsilon \) produces not the \( m \)-meaning of \( \epsilon \), but another function which, when it is applied to \( \epsilon \), produces \( m(\epsilon) \). Thus, in the above example, \( m(\epsilon) \neq m(\epsilon) \) for every expression in the language. The resulting sense of “agreement” is very thin. In particular, it is sufficiently thin that it fails to preserve synonymies – \( m(a) = m(b) \), but \( \mu(a) \neq \mu(b) \). In fact, the construction of \( \mu \) guarantees that no distinct terms of a language \( L \) will 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 \( \mu \) such that (a) it is compositional and (b) \( m \)-meanings can somehow be extracted from it. For example:

---

78 A solution will always be available in the set theory AFA, with an anti-foundation axiom (see, e.g., [3] for formal details on AFA). The use of AFA is essential for Zadrozny’s result, since the system of equations will lack a solution in ZFC if \( L \) has two elements which can be concatenated in either order, but the generalizations of Zadrozny’s result discussed below do not require AFA.
• Take $\mu(\epsilon) = \{< \eta, m(\eta) > : \eta \in \mathbb{L}\}$ for all $\epsilon \in \mathbb{L}$. Then the $\mu$-meaning of each term in the language encodes the $m$-meanings of every term – compositionality is then trivial (via the identity function), and $\mu(\epsilon)$’s encoding of $m(\epsilon)$ is just a special case of the general encoding.

• Define $\mu(\epsilon) = < m(\epsilon), \epsilon >$. Encoding of $m$-meanings is obvious, and compositionality is guaranteed because no expressions will be $\mu$-synonymous.\(^{79}\)

The lesson is the same as that of the previous section: 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. 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.\(^{80}\) If, however, the meaning function must meet significant constraints on the meanings of the atoms, then this route to cheap compositionality is blocked.\(^{81,82}\)

### 4.1.3 Tinkering In Between

Suppose one is committed, in constructing a meaning theory for English, to assigning the same meanings to ‘Clark Kent’ and ‘Superman’, but different meanings to:

12 Lex Luthor believes that Superman is harmed by Kryptonite.

13 Lex Luthor believes that Clark Kent is harmed by Kryptonite.

Thus the methods of tinkering at the top and bottom, discussed in the previous two sections, are unavailable. Nevertheless, the possibility of compositionality has not yet been eliminated. Suppose a syntactic theory of English is endorsed according to which one of the parts of ‘Lex Luthor believes that Superman is harmed by Kryptonite’ is the expression ‘caped superhero’, while one of the parts of ‘Lex Luthor believes that Clark Kent is harmed by Kryptonite’ is ‘mild-mannered reporter’. (Perhaps these expressions fill the subject position of the complement S at deep structure, but are deleted in the transformations leading to surface structure.) On the plausible assumption that ‘caped superhero’ and ‘mild-mannered reporter’ differ in meaning, (12) and (13) are

\(^{79}\)See [24] for more details on alternative ways of producing Zadrozny’s result.

\(^{80}\)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, just by having that item carry case-by-case instructions on what each sentence is to mean. Value loading, of course, need not be so crude. The proposal in the introduction for treating ‘large’ as a function from $\mathbb{N}$ extensions to extensions is a more refined version of value-loading, with ‘large’ carrying upward instructions on how to interact with different modified nouns.

\(^{81}\)One obvious example: if the meaning function is required to respect pretheoretic judgements of synonymy, then compositionality will be impossible if the distribution of such pretheoretic judgements violates the substitution principle.

\(^{82}\)See [24], [59] [67], and [105] for additional discussion of Zadrozny.
then built out of parts of different meanings, so their lack of synonymy poses no threat to compositionality.

This approach generalizes, in a result due to Janssen.\(^{83}\) Let \(\mathbf{L}\) 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 \(\mathbf{L}\) compositional. Most trivially, this can be done by setting \(\sqsubseteq\) to be the identity relation on \(\mathbb{L}^2\). 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 \(\epsilon \in \mathbb{L}\), \(|\{\eta \subseteq \epsilon \subseteq \eta \} \land |\{\eta \subseteq \epsilon \land |\{\eta \subseteq \epsilon \cup \epsilon\}|\), then the pretheoretically-given constraints on parthood can always be extended to a parthood relation \(\sqsubseteq\) making the language compositional. If \(\mathbf{L}\) 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 \(\mathbf{L}\) 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 largely or entirely 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.\(^{84}\) Thus consider a hypothetical view according to which linguistic practice imposes constraints only on sentential meanings, leaving word meanings to be assigned as internal theoretical entities, and according to which linguistic practice is insufficiently rich to fix a single assignment of meanings to sentences. Any attempt to use compositionality to select among the range of available theories, each compatible with the impoverished data constraining sentence meanings, must fail, since the freedom granted by treating word meanings as unconstrained will make compositionality trivially available for every theory in the range.

### 4.2 The Strength of Compositionality

Having seen the weakness of the compositionality constraint in the ways in which it can be trivially satisfied, we will now examine the strength of that constraint by setting out a surprising consequence of it due to recent work by Hodges in [55].\(^{85}\) Suppose

83. See [59] and [58]. 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.

84. If, on the other hand, compositionality is just a claim about the language which may or may not be true, then the ease with which it is satisfied will be a matter of indifference.

85. 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, where compositionality allows intersubstitution of arbitrary numbers of synonymous pairs, while 1-compositionality allows intersubstitution only of a single synonymous pair. The two are equivalent under the assumption that the meaning functions are Husserlian.
we have a set \( L \) of expressions, and two meaning functions \( \mu_1 \) and \( \mu_2 \) whose domains are (possibly 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 three conditions are met:

1. If \( \mu_2(\varepsilon) = \mu_2(\tau) \), then if \( \Sigma(\varepsilon) \) is in the domain of \( \mu_1 \), \( \Sigma(\tau) \) is as well. (\( \mu_2 \)-synonyms are \( \mu_1 \)-meaningfully intersubstitutable, although not necessarily with preservation of \( \mu_1 \)-meaning).

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

3. If \( \mu_2(\varepsilon) \neq \mu_2(\tau) \), then either:
   - There is some \( \Sigma \) such that \( \Sigma(\varepsilon) \) is \( \mu_1 \)-meaningful and \( \Sigma(\tau) \) is not, or vice versa.

or:
   - There is some \( \Sigma \) such that \( \Sigma(\varepsilon) \) and \( \Sigma(\tau) \) are both \( \mu_1 \)-meaningful but \( \mu_1(\Sigma(\varepsilon)) \neq \mu_1(\Sigma(\tau)) \)

(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.)

Roughly speaking, 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 \( L \). A Fregean cover of \( \mu \) whose domain included subentential expressions of \( L \) would treat subentential meanings as driven by the need to account for the sentential data, assigning different subentential meanings only when there was a corresponding difference in sentential behaviour.

We can now state the following result, which Hodges reads as stating that “Frege’s context principle implies compositionality” ([55], 19):

(Lemma 1) Suppose \( \mu_1 \) and \( \mu_2 \) are meaning functions with domains \( D_1 \) and \( D_2 \) of \( L \) and that \( \mu_2 \) is a Fregean cover of \( \mu_1 \). Then \( \mu_2 \) is compositional.

**Proof:** We will show that \( \mu_2 \) obeys the substitution principle. Suppose that \( \mu_2(\varepsilon) = \mu_2(\tau) \), and consider \( \Sigma \) such that \( \Sigma(\varepsilon) \) and \( \Sigma(\tau) \) are both in \( D_2 \). We need to show that \( \mu_2(\Sigma(\varepsilon)) = \mu_2(\Sigma(\tau)) \). Suppose not. Since \( \mu_2 \) is a Fregean cover, it assigns terms different meanings only when necessary to explain \( \mu_1 \)-differences in substitution instances, so there must be some larger context \( \Omega \) such that \( \mu_1(\Omega(\Sigma(\varepsilon))) \neq \mu_1(\Omega(\Sigma(\tau))) \) (or one of the

\* \( \mu \) is Husserlian if, for all \( \varepsilon, \tau \) such that \( \mu(\varepsilon) = \mu(\tau) \), and for all frames \( \Sigma \) if \( \Sigma(\varepsilon) \) is in the domain of \( \mu \), so is \( \Sigma(\tau) \).

I will thus simply presuppose (rather unjustly, given Hodges’ care in highlighting the importance of the condition) that all meaning functions are involved are Husserlian.
two is not in the domain of \( \mu_1 \). But \( \Omega(\Sigma(\epsilon)) \) and \( \Omega(\Sigma(\tau)) \) differ only by intersubstitution of \( \epsilon \) and \( \tau \), and terms synonymous under a Fregean cover must be intersubstitutable according to the covered meaning function, so this is a contradiction. Thus \( \mu_2 \) must assign the same meanings to \( \Sigma(\epsilon) \) and \( \Sigma(\tau) \), and hence is compositional. 

The basic idea of the lemma is straightforward: if the Fregean cover makes terms synonymous exactly when complex terms built out of them behave the same, then intersubstitution of synonyms will always preserve synonymy. The next step is to show that Fregean covers are always available:

(Lemma 2) Suppose \( \mu_1 \) is a meaning function on some subset \( D_1 \) of \( \mathbb{L} \). Then there is a meaning function \( \mu_2 \) on \( D_2 = \{ \epsilon \in \mathbb{L} : \exists \eta \in D_1 \epsilon \subseteq \eta \} \) such that \( \mu_2 \) is a Fregean cover of \( \mu_1 \).

Proof: Call two expressions \( \epsilon \) and \( \tau \) co-categorical for \( \mu_1 \) if, whenever \( \Sigma(\epsilon) \) is \( \mu_1 \)-meaningful, \( \Sigma(\tau) \) is as well, for \( \Sigma(\epsilon) \) and \( \Sigma(\tau) \) in \( \mathbb{L} \). 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 \( D_1 \) \( \mu_2 \), then, assigns to each term the set of expressions which are co-categorical and \( \mu_1 \)-intersubstitutable (in \( D_1 \)) with it. Then \( \mu_2 \) is a Fregean cover for \( \mu_1 \):
  - Suppose \( \mu_2(\epsilon) = \mu_2(\tau), \) and \( \Sigma(\epsilon) \) is in \( D_1 \). Clearly \( \alpha \in \mu_2(\alpha) \) for any \( \alpha \) in the domain of \( \mu_2 \), so \( \epsilon \in \mu_2(\tau) \), and thus \( \epsilon \) and \( \tau \) are co-categorical for \( \mu_1 \). Thus \( \Sigma(\tau) \) is in \( D_1 \).
  - Suppose \( \mu_2(\epsilon) = \mu_2(\tau), \) and \( \Sigma(\epsilon), \Sigma(\tau) \) are both in the domain of \( \mu_1 \). Then, as before, \( \epsilon \in \mu_2(\tau) \), so \( \mu_1(\Sigma(\epsilon)) = \mu_1(\Sigma(\tau)) \).
  - 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 \( \mu_1(\Sigma(\tau)) \neq \mu_1(\Sigma(\eta)) \). Then:
    - If the former, then \( \epsilon \) and \( \tau \) are not co-categorical for \( \mu_1 \) (since \( \epsilon \) and \( \eta \) are, and since co-categoricity is transitive). But if \( \epsilon \) and \( \tau \) are not co-categorical for \( \mu_1 \), then there is some \( \Sigma \) such that \( \Sigma(\epsilon) \) and \( \Sigma(\tau) \) are not both \( \mu_1 \)-meaningful, satisfying one prong of the third clause of the Fregean cover definition.
    - If the latter, then since \( \mu_1(\Sigma(\tau)) = \mu_1(\Sigma(\eta)) \), we immediately have \( \mu_1(\Sigma(\epsilon)) \neq \mu_1(\Sigma(\tau)) \), satisfying the second prong of the third clause of the Fregean cover definition.

Furthermore, Fregean covers are almost unique. Call two meaning functions equivalent if they agree in their synonymies:

- \( \mu_1, \mu_2 \) are equivalent if \( \mu_1(\epsilon) = \mu_1(\tau) \) if and only if \( \mu_2(\epsilon) = \mu_2(\tau) \).

Then:

(Lemma 3) If \( \mu_1 \) is a Fregean cover of \( \mu \) defined on a set \( \mathbb{L} \) of expressions, then \( \mu_2 \) defined on \( \mathbb{L} \) is a Fregean cover of \( \mu \) if and only if \( \mu_1 \) and \( \mu_2 \) are equivalent.
Proof: Left to right: suppose \( \mu_1(\epsilon) = \mu_1(\tau) \), but \( \mu_2(\epsilon) \neq \mu_2(\tau) \). Then either (a) there is some \( \Sigma \) such that only of of \( \Sigma(\epsilon), \Sigma(\tau) \) is \( \mu \)-meaningful, or (b) there is some \( \Sigma \) such that \( \mu(\Sigma(\epsilon)) \neq \mu(\Sigma(\tau)) \). But (a) is impossible: since \( \mu_1 \) is a Fregean cover and \( \epsilon \) and \( \tau \) are \( \mu_1 \)-synonymous, \( \epsilon \) and \( \tau \) must be \( \mu \)-meaningfully intersubstitutable. And (b) is also impossible: again since \( \mu_1 \) is a Fregean cover and \( \epsilon \) and \( \tau \) are \( \mu_1 \)-synonymous, \( \epsilon \) and \( \tau \) must be intersubstitutable in all \( \mu \)-meaningful contexts without change of \( \mu \)-meaning. Thus \( \mu_1 \) and \( \mu_2 \) cannot disagree in their synonymies, and are equivalent.

Right to left: Since \( \mu_2 \)-synonymy implies \( \mu_1 \)-synonymy, the first two clauses in the Fregean cover definition are immediate. Since lack of \( \mu_2 \)-synonymy implies lack of \( \mu_1 \)-synonymy, the third is as well. ..

The several almost unique Fregean covers 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, such that \( \mu \) is fully characterized by:

- \( \mu(a \sim b) = 1 \)
- \( \mu(c \sim d) = 1 \)
- \( \mu(c \sim b) = 2 \)
- \( \mu((a \sim b) \sim b) = 3 \)
- \( \mu((c \sim b) \sim d) = 4 \)
- \( \mu((c \sim d) \sim b) = 5 \)

Applying the construction of Lemma 2 to obtain a Fregean cover for \( \mu \) defined on the closure of \( \mu \)'s domain under the parthood relation yields:

- \( \mu_1(a) = \{a\} \)
- \( \mu_1(b) = \{b\} \)
- \( \mu_1(c) = \{c\} \)
- \( \mu_1(d) = \{d\} \)

Given Lemma 3, the crucial point is the pattern of synonymies, so any meaning function (on the requisite domain) which makes \( a \sim d \) and \( c \sim b \) synonymous, \( (a \sim b) \sim b \), \( (a \sim b) \sim d \), and \( (c \sim d) \sim b \) synonymous, and \( (c \sim d) \sim d \), and \( (c \sim d) \sim d \) synonymous 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:

(Thorem 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 \( D_2 = \{ \epsilon \in \mathbb{L} : \exists \eta \in D_1 \iff \eta \} \) such that for all \( \epsilon \in D_1 \), \( \mu_2(\epsilon) = \mu_1(\epsilon) \).

Proof: Left to right: The existence of a Fregean cover for \( \mu_1 \) is given by Lemma 2, so we need only show that there is a Fregean cover which extends \( \mu_1 \). Let \( \mu \) be an
arbitrary Fregean cover of \( \mu_1 \). Suppose \( \epsilon \) and \( \tau \) are \( \mu_1 \)-synonymous. Then, since \( \mu_1 \) is compositional, \( \Sigma(\epsilon) \) and \( \Sigma(\tau) \) are \( \mu_1 \)-synonymous for all \( \Sigma \). Since \( \mu \) assigns different meanings to terms only when they contribute differently to meanings of complexes, it will also treat \( \epsilon \) and \( \tau \) as synonymous. Suppose, on the other hand, that \( \epsilon \) 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.\(^{86}\) Thus \( \mu \) must have exactly the same synonymies as \( \mu_1 \) on \( D_1 \). By Lemma 3, any meaning function equivalent to \( \mu \) is a Fregean cover for \( \mu_1 \). Thus 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 by Lemma 1. 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 is attempting to construct a meaning theory for language \( L \). His task does not take place in a vacuum – there are certain given meaning facts about \( L \), and his task is to work these facts into a comprehensive theory. Suppose that, in particular, the semanticist is given meanings for all of the sentences of \( L \), and asked to find a way to assign 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 \( L \) is compositional on the sentential level\(^{87}\) and the semanticist wants the 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.

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. It’s thus inevitable that meaning theories constructed in accord with the Context Principle are unique up to equivalence. If the starting fragment is compositional, then it will fit into the semantic agenda dictated by the Context Principle, so the final pattern of synonymies will contain the starting pattern as a fragment. Of course, semantic theories may well place constraints on lexical meaning other than systematic derivation of proper sentence meaning, in which case a suitable meaning function may be finer-grained than the Context Principle requires, and hence not uniquely determined by that principle. Even given the determinative role of the Context Principle, settling the synonymies of the language falls far short of fully answering the concerns of the semanticist. Suppose, for example, that, in attempting to extract ontological commitments

\(^{86}\) A Fregean cover can only break, not introduce synonymies.

\(^{87}\) This is a non-trivial requirement if sentences can contain sentences as proper parts.
from our semantic practice, we come to wonder whether the lexical item ‘gavagai’ should be understood as meaning rabbit or undetached rabbit part. Given meanings for all of the sentences, adherence to the Context Principle will dictate what ‘gavagai’ is synonymous with, but it will not tell us what it means. If, as hypothesized, we do not know the meanings of any of the subsentential expressions, this synonymy information will do nothing to settle the ontology acquired via the use of the term ‘gavagai’. Or suppose that a language with verbs of propositional attitude, such that any sentence of the language can appear in complement position to a propositional attitude verb, is under study. Suppose also that (perhaps in light of [78]) the semantic theory is committed to an assignment of sentence meanings such that, for any pair of words α and β, there is some pair of sentences Σ(α) and Σ(β) which differ in meaning. Then adherence to the Context Principle will tell us only that no two words of the language have the same meaning. Clearly, many conceptions of semantic theory will regard such information only as the barest of starts toward construction of an adequate theory.

5 Some Problem Cases for Compositionality

While the results of section 4.1 show that when semantic theories are sufficiently unconstrained, compositionality can be cheaply obtained, the question remains open of whether a satisfactory semantic theory for English, or another natural language – a theory properly responsive to a variety of prior constraints on semantic and syntactic facts – can be given in a compositional form. In this section we will look at five problem cases for the construction of compositional semantics. In each case we will examine data which resist a compositional treatment, and then consider ways of overcoming that resistance. The goal is not to settle definitively 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 in fact influence semantic theorizing.

5.1 Supertruth

The Problem: The word ‘bald’ is vague. There are many ways of saying what this vagueness amounts to, but one is this: there are objects of which it is inappropriate to say that they are bald, but also inappropriate to say that they are not bald. Again, there are many ways of saying what appropriateness amounts to, but one is to equate it with truth. On this approach (a ‘truth value gap’ analysis of vagueness), there are objects α

...In the wake, of course, of Quine’s discussion in [89], §12. 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 [51], [52], [53], [54], [55]; of compositionality and prototype theory in [31] (ch. 5), [34], [33], [63], [81], [94]; of compositionality and idioms in [66], [65], [80], [106]; of compositionality and ‘unless’ in [49]; [2], [102]; of compositionality and propositional attitude cases in innumerable places, but especially [90] (ch. 4), [93] (ch. 8), [71], [18]; of compositionality and ‘any’ and other negative polarity items in [73], [72], [9], [76], [69]; of compositionality and anaphora in [62], [46], and [1]. [59] also contains an overview of several compositionally-problematic semantic phenomena.
such that the claim ‘α is bald’ is neither true nor false, but the recipient of some third truth value I.

Suppose, then, that a predicate is assigned not a single set (of things satisfying the predicate) as extension, but rather a pair of sets: a positive extension and a negative extension. In the case of ‘bald’, we might have:

- \([\text{bald}]^+ = \{\text{Michael Jordan, Patrick Stewart, Jesse Ventura, Sinead O’Connor…}\}\)
- \([\text{bald}]^- = \{\text{Carrot Top, Artis Gilmore, Sarah Michelle Gellar}\}\)

Some people, however, will be in neither \([\text{bald}]^+\) nor \([\text{bald}]^-\), such as Sean Connery or George Carlin. These people are the borderline cases of baldness, to whom baldness is neither appropriately ascribed nor denied. Truth conditions for atomic sentences involving ‘bald’ can then be given as follows:

- \([\text{α is bald}] = T\) if \([\text{α}] \in [\text{bald}]^+\)
- \([\text{α is bald}] = F\) if \([\text{α}] \in [\text{bald}]^-\)
- \([\text{α is bald}] = I\) if \([\text{α}] \notin [\text{bald}]^+\) and \([\text{α}] \notin [\text{bald}]^-\)

More generally, an atomic sentence will be true if the n-tuple of its terms’ referents is in the positive extension of the predicate, false if that n-tuple is in the negative extension of the predicate, and indeterminate if it is in neither.

There will thus be people of whom it is neither true nor false to say that they are bald. The basic semantic approach can then be extended to molecular sentences using three-valued truth tables such as:

<table>
<thead>
<tr>
<th>φ</th>
<th>¬φ</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>F</td>
<td>I</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>φ</th>
<th>ψ</th>
<th>φ ∧ ψ</th>
<th>φ ∨ ψ</th>
<th>φ → ψ</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td>T</td>
<td>I</td>
<td>T</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>I</td>
<td>T</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>I</td>
<td>F</td>
<td>F</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
<td>T</td>
<td>I</td>
</tr>
<tr>
<td>F</td>
<td>I</td>
<td>F</td>
<td>I</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

But now a difficulty emerges. Suppose α and β are both borderline cases of baldness, but that α has more hair on his head than β. Then the given semantic theory is such that:

- \([\text{Either α is bald or α is not bald}] = I\)
- \([\text{α is both bald and not bald}] = I\)
If $\alpha$ is bald, then $\alpha$ is bald $= I$

If $\alpha$ is bald, then $\alpha$ is not bald $= I$

If $\alpha$ is bald then $\beta$ is bald $= I$

But one might want the first, third, and fifth of these to be true, and the second and fourth to be false.

Enter supervaluation semantics. Supervaluation, first developed by van Fraasen in [100] for treating nonreferential terms and applied by Fine in [29] to vagueness, allows a three-valued logic to recapture classical tautologies as well as what Fine called penumbral truths about vague concepts, such as: if one man has more hair than another, then if the first is bald, the second is also bald. Suppose we have a three-valued interpretation $M$ for $L$ as above, in which each predicate is assigned a positive extension and a negative extension which are not necessarily collectively exhaustive. An interpretation $M'$ extends $M$ ($M' \supseteq M$) just in case for every predicate $\Pi$ of $L$:

- $\llbracket \Pi \rrbracket_M^+ \subseteq \llbracket \Pi \rrbracket_{M'}^+
- \llbracket \Pi \rrbracket_M^- \subseteq \llbracket \Pi \rrbracket_{M'}^-

Extensions of interpretations, then, decide some of the undecided cases of the extended interpretation, without altering any of the decided cases. A model for $L$ will then be a collection $\mathcal{M}$ of interpretations such that:

- There is some $M_0 \in \mathcal{M}$ such that for all $M \in \mathcal{M}$, $M \supseteq M_0$. ($M_0$ is the base point of the model).
- For all $M \in \mathcal{M}$, there is some complete $M'$ such that $M' \supseteq M$, where an interpretation is complete if the positive and negative extensions of each of its predicates are collectively exhaustive.

Intuitively, a model is a minimal classification of objects (that given by the base point), together with all permissible ways of making that minimal classification more precise. Truth in a model is now supertruth – $\phi$ is true in $\mathcal{M}$ if it is true at all of the complete extensions of the base point of $\mathcal{M}$. Similarly, falsity in a model is superfalsity – $\phi$ is false in $\mathcal{M}$ if it is false at all of the complete extensions of the base point. If $\alpha$ is borderline for baldness, then $\llbracket \alpha \rrbracket_{M_0} \notin \llbracket \text{bald} \rrbracket_{M_0}$ and $\llbracket \alpha \rrbracket_{M_0} \notin \llbracket \text{not bald} \rrbracket_{M_0}$. Extensions of $M_0$ are thus free to put $\alpha$ into either the positive or the negative extension of ‘bald’. If both choices are realized in various extensions of $M_0$, then some complete extensions of $M_0$ will make ‘$\alpha$ is bald’ true and others will make it false, so the ‘$\alpha$ is bald’ will be neither supertrue nor superfalsely, and hence will be assigned $I$ by the model. But consider:

(30) $\alpha$ is bald or $\alpha$ is not bald.

Every complete extension is a classical bivalent interpretation, and hence makes true all classical tautologies, including (30). Thus (30) will be supertrue, and hence true, in $\mathcal{M}$. Similarly, ‘if $\alpha$ is bald, then $\alpha$ is bald’ will be true at every complete extension, and hence supertrue, and ‘$\alpha$ is both bald and not bald’ will be false at every complete
extension, and hence superfalse. And if $\mathcal{M}$ is chosen so that one interpretation in it extends another only if whenever the first adds some individual to the positive extension of ‘bald’, it also adds any other previously undecided individuals with less hair to the positive extension of ‘bald’, then if $\alpha$ has more hair than $\beta$, every complete extension will make true ‘if $\alpha$ is bald, then $\beta$ is also bald’, making that sentence supertrue.

The move to supervaluation and supertruth allows ascriptions of vague properties to borderline cases to take on truth value gaps, while also allowing certain conceptual generalizations about borderline cases to come out true. The price to be paid, however, is a loss of compositionality. Sentences in a supervaluationist semantics receive a semantic value of true (if supertrue), false (if superfalse), or indeterminate (if neither supertrue nor superfalse). But now suppose $\phi$ and $\psi$ are both indeterminate, and consider $\phi \lor \psi$. $\phi \lor \psi$ might be indeterminate, if for example $\phi$ and $\psi$ are two unrelated ascriptions of vague predicates to borderline cases, but it might also be true, if for example $\psi$ is $\neg \phi$. Similarly, $\phi \land \psi$ can be either indeterminate or true. The meanings of the parts, then, are not enough to determine the meaning of the whole.

A Solution: The technical difficulty here is clear. When $\phi$ is indeterminate, it is true at some but not all complete extensions, but the mere information that it is indeterminate tells us nothing about which complete extensions. But when two indeterminate sentences are combined by a Boolean connective, the resulting sentence will be true at a collection of complete extensions determined by the complete extensions at which the two combined sentences are true. Since that collection can at times be universal (yielding supertruth) or empty (yielding superfalsity), it is semantically essential that we be able to see what it is, and hence that we know its determining factors. Indeterminacy, then, is too coarse-grained a semantic value – it hides crucial details about the mode of indeterminacy.

To regain compositionality, then, we need finer grain. Suppose that $\mathcal{M}$ assigns a set of interpretation points, rather than supertruth values:

- $\llbracket \phi \rrbracket^+ = \{ I \in \mathcal{M} : I \vdash \phi \}$
- $\llbracket \phi \rrbracket^- = \{ I \in \mathcal{M} : I \dashv \phi \}$

Then the Boolean connectives can be defined:

- $\llbracket \neg \phi \rrbracket^+ = \{ I \in \mathcal{M} : I \in \llbracket \phi \rrbracket^- \}$
- $\llbracket \neg \phi \rrbracket^- = \{ I \in \mathcal{M} : I \in \llbracket \phi \rrbracket^+ \}$
- $\llbracket \phi \land \psi \rrbracket^+ = \llbracket \phi \rrbracket^+ \cap \llbracket \psi \rrbracket^+
- $\llbracket \phi \land \psi \rrbracket^- = \llbracket \phi \rrbracket^- \cup \llbracket \psi \rrbracket^-$
- $\llbracket \phi \lor \psi \rrbracket^+ = \llbracket \phi \rrbracket^+ \cup \llbracket \psi \rrbracket^+
- $\llbracket \phi \lor \psi \rrbracket^- = \llbracket \phi \rrbracket^- \cap \llbracket \psi \rrbracket^-$
- $\llbracket \phi \rightarrow \psi \rrbracket^+ = \llbracket \phi \rrbracket^- \cup \llbracket \psi \rrbracket^+$
- $\llbracket \phi \rightarrow \psi \rrbracket^- = \llbracket \phi \rrbracket^+ \cap \llbracket \psi \rrbracket^-$
Suppose \( \phi \) and \( \psi \) are sentences each of which has some, but not all, complete interpretations from \( M \) in its meaning. Then \( \phi \lor \lnot \phi \), given these rules, will always contain all complete interpretations in its meaning, while \( \phi \land \lnot \phi \) will never contain any complete interpretations in its meaning. \( \phi \land \psi \), however, may well again have some but not all complete interpretations in its meaning. The previously obscured difference in behaviour between indeterminate sentences is now laid bare.

Compositionality is achieved by a combination of tinkering at the top and at the bottom; in altering the meanings of sentences we alter both the meanings of the semantic termini and of their occasional sentential parts. The price paid is the abandonment of the distinct categories of true, false, and indeterminate – the reworked semantics makes clear that supervaluationism is at heart a modal system which replaces local evaluation for truth at a home point (the actual world) with global evaluation for modal status (two versions of necessity in supertruth and superfalsity, and one version of contingency in indeterminacy). The question thus becomes whether, for example, truth can be replaced with inclusion of all complete interpretations without undermining the point of the semantics – whether inclusion of all complete semantics can serve as a norm guiding assertoric practice in the same way that truth does, and so on. Supposing this question to be answered affirmatively, the supervaluationist has successfully discharged the burden imposed by compositionality.

5.2 What the Hell

The Problem: Compare the following two sentences:

(31) Who bought that book?

(32) Who the hell bought that book?90

These two sentences are roughly synonymous. While the addition of ‘the hell’ does alter the rhetorical impact of (32) (it may, for example, encourage the conversational implicature that it is surprising that the book was bought or that the speaker is displeased with the buyer), 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 needs to be compatible with the close semantic proximity of (31) and (32).

However, ‘who’ and ‘who the hell’, or, more generally, ‘wh-’ and ‘wh- the hell’ expressions, diverge in meaning in other contexts. Thus,91

- The minimal variant of adding a modal auxilliary to the previous examples causes a difference to emerge:

(33) Who would buy that book?

(34) Who the hell would buy that book?

---

90 These two examples are drawn from [23]. The behaviour of ‘the hell’ phrases was first noted in [86].

91 The first three of these examples are drawn from [23]; the last draws from [23] and [86].
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.\footnote{\cite{23} claims that only the indirect assertion reading of (34) 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.}

- 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 turn out to be grammatical only in negative contexts, whether overt:
  
  - 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 \textit{adversative attitude verbs}:\footnote{See \cite{72}.}
  
  - 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 contextu-
ally provided range of salient objects, do not allow ‘the hell’ modification:

– *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 (32), but much greater in the other cases set out above.\textsuperscript{94}

\textbf{A Solution}.\textsuperscript{95} A simple ‘who’ question can have its interpretation influenced by
linking the range of admissible answers to a contextually-provided domain. Thus con-
side r the following dialogue:

\par

(35) 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 that the semantic function of ‘the hell’ is to require that the range of ad-
missible 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 (31), adding ‘the hell’ will have no e
\textit{ff}
fect, because it will merely demand that what has already happened indeed happen – that
the question admit of novel answers. But when the wh-question is a linked one, adding
‘the hell’ will have a semantic impact. Thus consider:

\par

(36) A: 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 currently counts as A’s
friends), this new implicature is to be expected. The various effects of ‘the hell’ noted
above now fall out:

\par

\begin{itemize}
  \item ‘Wh- the hell’ phrases will refuse anaphoric linkage because that linkage dic-
tates the range over which the wh- phrase ranges, which contradicts the novelty
requirement imposed by ‘the hell’.\textsuperscript{96} ‘Which’ phrases, which \textit{always} require
anaphoric/contextual linkage, can thus never combine with ‘the hell’.
\end{itemize}

\textsuperscript{94}Whether a compositional semantics needs to 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’ will depend on the extent
to which we want grammatical phenomena explained by semantic facts. It is very tempting to think that
grammatical failures due to failures of anaphoric linkage, at least, have a semantic explanation.

\textsuperscript{95}The following solution is a simplified and modified version of the proposal of [23]. Any shortcomings
of it are due to the present alterations.

\textsuperscript{96}Although note the acceptability of:

\begin{itemize}
  \item Someone walked in the park, but I don’t know who the hell it was.
\end{itemize}

53
A question of the form ‘Who would buy that book?’ takes as answer pairs of people and possible situations. Given the broad total range of possible situations, such a question will typically be linked to a contextually-provided range of admissible situations. 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, and an alternative communicative explanation will be 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’ will be impossible to fulfill when the ‘wh- the hell’ phrase is 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. On the other hand, to say that I don’t know who the hell bought the book creates no conflict with the novelty requirement. 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’, will for similar reasons allow ‘the hell’ modification. The novelty requirement thus explains the distributional facts noted above.

The novelty requirement make ‘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.

Consider again:

– What the hell did everyone buy for Max?

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 will intervene between ‘what the hell’ and its licenser, violating the Immediate Scope Constraint. The unavailability of a reading wide-scoped for ‘everyone’ is thus explained.

Thus: ‘Albert, if it has a chapter on direct reference’, ‘Louisa, if autographed copies are available’, etc.
Thus ruling out answers such as ‘Brian, if we threaten to kidnap his dog if he doesn’t’.

See [76] for formulation and defense of the Immediate Scope Constraint.
5.3 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 half a percent of the world’s population, the following claim looks acceptable:

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

On reflection, however, the acceptability of (37). Fourteen, after all, is not by most natural standards many. Thus consider the oddity of:

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

(39) Many prime numbers less than a billion begin with a ’7’. Fourteen, in fact.

The acceptability of (37) seems to derive from having fourtee qualify not as many Scandinavians, but rather as many winners of the Nobel prize in literature, and thus from reading (37) as:

(40) 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:

(41) Scandinavians often win the Nobel prize in literature. (The Nobel prize in literature is often won by Scandinavians.)

and in sentences with generics:

(42) 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’, also exhibit similar behaviour. But all of these cases look like violations of the semantic locality component of compositionality: the role of ‘many’ in imposing a proper cardinality constraint involves semantic interaction between that determiner and the syntactically distant verb phrase.

A Solution: We’ll concentrate here on addressing the problem in its ‘many’ form, leaving the task of integrating these suggestions with adequate theories of adverbs of quantification.

---

100 Bjørnstjerne 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 (37) are both due to [104] in 1985. The intervening years have not been kind to the recognition of Scandinavian literature, and the number holds at 14.

101 See [22] and [17]

102 See [16] and [17]

103 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?
quantification and generics as a task for the reader. 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 $\kappa$.\textsuperscript{104} This sort of ‘many’ can be given a straightforward compositional semantics:

- $\llbracket \text{many} \rrbracket = f : \wp(D) \mapsto \wp(\wp(D))$, $f(X) = \{ Y : |X \cap Y| \geq \kappa \}$
- $\llbracket [DP \text{ many } X] [VP Y] \rrbracket = \text{true} \iff [Y] \in \llbracket \text{many} \rrbracket([X])$

‘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 $\kappa$ will account for the truth of (37).

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

(43) Many philosophers of mathematics have read Russell and Whitehead’s \textit{Principia Mathematica}.

(44) Many Brazilians have read Russell and Whitehead’s \textit{Principia Mathematica}.

Suppose the facts are that among the 2000 philosophers of mathematics, 800 have read the \textit{Principia}, and that among the 180 million Brazilians, again 800 have read the \textit{Principia}. Then (43) looks true and (44) false, which is impossible if ‘many’ simply imposes a cardinality constraint. We want (43) to be true despite the falsity of (44) because the 800 are a larger percentage of the philosophers of mathematics than they are of the Brazilians. This suggests the following alternative reading of ‘many’:

- $\llbracket \text{many} \rrbracket = f : \wp(D) \mapsto \wp(\wp(D))$, $f(X) = \{ Y : |X \cap Y| \geq \rho|X| \}$
- $\llbracket [DP \text{ many } X] [VP Y] \rrbracket = \text{true} \iff [Y] \in \llbracket \text{many} \rrbracket([X])$

where $\rho$ sets the threshold percentage for \textit{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 (37).

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 suggest that we don’t always want this result:

(45) Many penguins live in Antarctica.

(46) Many penguins live in my bedroom.

\textsuperscript{104}The required number of Y-ing X’s will surely be vague, but the idealization to a specific $\kappa$ will do no harm here.
Twelve penguins in a bedroom seems to suffice for *many*, but will hardly do for a whole continent. Here ‘many’ appears to mean 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:

- $\mathbb{[}\text{many}\mathbb{]} = f : \wp(D) \mapsto \wp(\wp(D))$, $f(X) = \{Y : |X \cap Y| \geq \frac{|X||Y|}{|D|}\}$
- $\mathbb{[}S_{\mathbf{DP}} \text{ many } X][\mathbf{VP} Y]\mathbb{]} = \text{true}$ iff $Y \in \mathbb{[}\text{many}\mathbb{]}(\mathbb{[}X\mathbb{]})$

This is a crude measure of expectation, of course, because it’s not just the smaller number of things in my bedroom that makes me expect fewer penguins there, but it will do for a start. The resulting semantics is symmetric, like the cardinality semantics, and will give the desired result for (37), 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 (37) 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 (37) 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.\textsuperscript{105} Thus St. Lucia, with its population of some 200,000 and a single Nobel laureate in literature\textsuperscript{106}, dramatically exceeds the expectation threshold. But the claim:

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

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:

(48) 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:

- The ‘proportionate’ semantics could be combined with a syntactic story according to which prior to semantic analysis the argument positions of ‘many’ sentences are swapped. Thus (37), 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 the desired result. Such an approach would use the techniques of section 4.1.3 to achieve compositionality. Any attempt to construct a plausible story along these lines, however, would have to explain the failure of the verb phrase of (37) to display certain characteristic semantic features of sentence subjects. For example, quantification over empty classes is typically pragmatically disfavoured; hence the peculiarity of:

\textsuperscript{105}I idealize here by assuming all Nobel laureates in literature are currently alive.
\textsuperscript{106}Derek Walcott
(49) Many Freedonians have won the Nobel prize in literature. However, as [17] observes, a sentence like:

(50) 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.

- A ‘reverse proportionate’ semantics could be given, by altering the class a minimal percentage of which needs to behave as required:

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

(37), under this approach, will require a certain minimal percentage of the Nobel laureates in literature to be Scandinavian.\(^{107}\) However, we will now need an explanation of why ‘many’ allows both proportionate and reverse proportionate semantics, while ‘most’ allows only the proportionate semantics:

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

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

5.4 Deep Inside This Paper

The Problem: Consider the sentence:

(51) Ardalion saw an airplane under the cloud.

Suppose this sentence has the following syntactic form:

(51-T)

\[\text{S} \rightarrow \text{DP} \rightarrow \text{VP} \rightarrow \text{V} \rightarrow \text{DP} \rightarrow \text{NP} \rightarrow \text{D} \rightarrow \text{N} \rightarrow \text{PP} \rightarrow \text{DP} \rightarrow \text{under} \rightarrow \text{the cloud}\]

\(^{107}\)Note that this approach allows (47) to be false.
If (51) is to have a composition semantics, its meaning needs to derive from the meanings of ‘Ardalion’, ‘saw’, and ‘an airplane under the cloud’. The meaning of ‘an airplane under the cloud’, in turn, needs to derive from the meanings of ‘an’, ‘airplane’, and ‘under the cloud’. ‘An’ and ‘airplane’ can be given at least preliminary treatments using standard methods, but what of the prepositional phrase ‘under the cloud’?

‘Under the cloud’ is a member of the class of locative prepositional phrases, which also includes ‘near the barn’, ‘before the party’, and ‘between El Paso and Albuquerque’. A tempting first analysis is that locative prepositional phrases pick out locations, and when those phrases modify nouns, they limit the extension of those nouns to objects in the relevant location. Thus, for example, [under] would be a function from an object (the denotation of its complement DP) to the set of locations (conceived of either as points or as regions) under that object. [airplane under the cloud] would then be the set of airplanes in the location [under][the cloud]). ‘Location’ will need to be broadly conceived here, in order to allow for temporal location with prepositions such as ‘before’ and ‘after’, and perhaps various sorts of metaphorical location in prepositional phrases such as ‘in naive set theory’ and ‘under current First Amendment jurisprudence’.

Locative prepositional phrases can be modified in various ways:

(52) two feet under the cloud, directly above the door, just outside the forest, deep in the argument, diagonally across the street

Again, if compositionality is to be satisfied, the meanings of a modified locative phrase must derive from the combination of the meaning of the locative phrase itself with the meaning of the modifier. But sets of locations as meanings of locative phrases look inadequate to this compositional task. Consider ‘two feet under the cloud’. If [under the cloud] is a set of locations, then [two feet] must supply (in the syntactic context) a method of narrowing that set of locations to a particular subset (those, of course, two feet under the cloud). The difficulty here is that the mere distance of twenty-four inches is not enough information to identify the proper locative subset. We cannot, for example, simply travel two feet from the boundary of the region [under the cloud], because the region has boundaries other than the cloud itself, and points two feet from one of these other boundaries needn’t be two feet under the cloud. Similarly, we cannot identify the region ‘diagonally behind the house’ without knowing more than the region, since the region alone will not give us the spatial orientation needed to determine what counts as diagonal. To create an explicit counterexample to compositionality, suppose that [before Christmas] and [after Christmas] both pick out the 364 days of the year other than December 25. Then compositionality requires the synonymy of

108Not all prepositional phrases are locative, as is seen in examples such as ‘without enmity’, ‘about movies’, and ‘in a charcoal-gray suit’. Even among prepositional phrases which deal with location in a broad sense, locative phrases should be distinguished (as they are in [112]) from directional prepositional phrases such as ‘to the park’ and [walked] ‘under the bridge’. The current taxonomy clearly leaves much to be desired, especially given the appearance of single prepositions in multiple taxonomic locations.

109I assume here that ‘under the cloud’ picks out something like a trapezoidal prism starting at the cloud and extending to the ground, although the details are not essential. Attempts to pick out the privileged regional boundary (that adjoining the cloud) via geometric features of the region will, in general, fail due to the possibility of picking out regions under oddly shaped objects (such as objects with a trapezoidal prism carved out of their undersides).
two days before Christmas and two days after Christmas. In fact, of course, these two phrases pick out December 23 and December 27 respectively.

**A Solution:** The above examples show that the meanings of locative prepositional phrases must supply information about orientations, as well as boundaries, of space. [111] and [112] thus propose assigning locative phrases sets of vectors. Since a vector is a directed arrow between two locations, interpreting locative phrases as sets of vectors will provide the desired orientation information. Thus, for example, before Christmas will be the set of (temporal) vectors starting at December 25 and directed backward, while after Christmas will be the set of vectors starting at December 25 and directed forward. Modifiers can then also pick out sets of vectors, and semantic composition of modifiers with prepositional phrases will proceed via set intersection. Thus two days will pick out all temporal vectors of length two days, so:

- \( \text{\{two days before Christmas\}} = \{v : |− → v| = \text{two days}\} \bigcup \{v : − → v \text{ starts at Christmas and is oriented backward in time}\} = \{v : − → v \text{ starts at December 25 and ends at December 23}\} \)
- \( \text{\{two days after Christmas\}} = \{v : |− → v| = \text{two days}\} \bigcup \{v : − → v \text{ starts at Christmas and is oriented forward in time}\} = \{v : − → v \text{ starts at December 25 and ends at December 27}\} \)

However, there are some complications. Consider ‘two feet under the cloud’. If two feet is the set of all vectors of length two feet, and under the cloud is the set of all vectors starting at the cloud and oriented downward, we get the wrong results. Let \( x \) be a spatial point one foot under the cloud, \( y \) be the point on the boundary of the cloud one foot directly above it, and \( z \) be an arbitrary point on the cloud \( \sqrt{3} \) feet from \( y \). Then the vector \( − → xz \) is in under the cloud, since it is a downward-oriented vector starting at the cloud, and is also in two feet, since \( |− → xz| = 2 \) feet. Thus the point \( x \) will count, undesirably, as two feet under the cloud, when it is in fact only one foot under the cloud.

Following [111] and [112], we therefore want the notion of a closest vector:

- Vector \( − → xy \) is a closest vector to object \( A \) if \( x \) is on the boundary of \( A \) and there is no point \( z \) on the boundary of \( A \) such that \( |− → zy| < |− → xy| \).

We then require that the meaning of a locative phrase be limited to the set of closest vectors to the relevant object. The vector \( − → xz \) is now eliminated from under the cloud, since it is not a closest vector to the cloud, being longer than \( − → xz \). Similarly, a point almost on the edge of a forest cannot count as deep inside the forest by virtue of lying at the end of a vector starting at the far edge of the forest, because such a vector will not be closest to the forest.

---

[107] extends the vector space approach to the analysis of scalar adjective phrases, such as ‘two centimeters taller than Mary’.

[111] Perhaps only vectors orthogonal to the referent of the complement of the locative prepositions should be in the set picked out by the locative phrase? But then (a) similar undesirable results will follow for objects whose bottom surfaces are not perpendicular to the preferred downward direction, (b) modifiers like ‘diagonally’ will always return the empty set, and (c) we lose the intuitive idea that ‘under the cloud’ picks out not only points directly below the cloud, but also points slightly to the side (and below) the cloud boundaries.
A fully adequate theory will require further complications. Suppose there is a building with a flat rectangular roof, and in one corner of the roof a rectangular tower, with a small footprint compared to the roof as a whole, which extends several hundred feet further into the air. Consider a bird a hundred feet directly above the roof of the main building and a few feet to the side of the tower. Surprisingly, this bird will be excluded by \( \text{above the building} \), because the closest to the building of the vectors ending at the location of the bird are oriented horizontally, starting from the wall of the tower, rather than vertically, starting from the roof of the building. If \( \text{above the building} \) picks not from the closest vectors to the building, but rather from the closest of the upward-oriented vectors to the building, then the bird is covered by \( \text{above the building} \), but not by \( \text{directly above the building} \), since the closest of the upward-oriented vectors will be a nearly-horizontal one from the tower, rather than a strictly vertical one from the roof of the building. Patches for this problem are imaginable, but the point here is that a compositional treatment of locative prepositional phrases will require elaborate details of our spatial thought to be reflected in the semantics of prepositions.

5.5 Apples and Bananas

The Problem: The following sentence is grammatically permissible in German:

\[
\text{(53) } \text{"Apfel isst der Hans drei und zwei Bananen.}\]
\[
\text{Apples eats the Hans three and two bananas.}\]

\[
\text{\‘Hans eats three apples and two bananas.\textsuperscript{112}}
\]

The conjunction ‘und’ joins ‘drei’ with ‘zwei Bananen’, but in doing so must give rise to a conjunction meaning three apples and two bananas. The difficulty for compositionality lies in getting the apples into the conjunction, despite the fact that ‘\text{"Apfel}’ lies at the beginning of the sentence, outside the syntactic reach of ‘und’.

A Solution: The difficulty here lies in the violation of semantic locality, and one tempting method for dealing with violations of semantic locality is to tinker with the syntax to relocate the nonlocal item to a more desirable neighborhood. Such tinkering can be done, for example, by appealing to a transformational syntax, according to which the string of words we see on the paper is a result of applying various movement rules to some sort of underlying structure. Using simple brute force, we can assign (53) the underlying structure:

\textsuperscript{112}The example is from [61], as is the syntactic analysis given below, although I have (over)simplified the syntactic discussion vastly.
and posit a transformational rule which allows arbitrary nodes of the tree to be raised to adjoin S, allowing the derivation of:

(53-S)

In the underlying form, ‘und’ joins ‘drei Äpfel’ and ‘zwei Bananen’, and if we impose the compositionality constraint on a parthood relation which reflects the constituent structure of the underlying form, then the conjunction will pose no special difficulty.\footnote{We will, of course, still require a reading of “und” other than the familiar truth-functional one, since in (53-U) it joins terms rather than sentences. But techniques of the sort discussed in [82] provide such readings unproblematically.}

The brute force method, however, is, as such methods tend to be, too crude. A transformational rule which allows any item to be adjoined to the top S node yields wildly unacceptable results, since it allows word order to be shuffled arbitrarily. The challenge
to the ‘underlying structure’ approach to achieving compositionality is to tell a story about the movement from underlying to surface structure in (53) which integrates into larger syntactic theories. A full narrative would divert us too far into syntax for current purposes, but a brief overview of one possible tale will give an indication of the difficulties faced.\footnote{The syntactic approach I sketch here is due to [61].} The overview relies on four assertions about an adequate syntactic theory, which will be listed but not justified:

1. Transformational rules are limited by the \textit{Coordinate Structure Constraint}, which prohibits movement of only one head expression or non-argument expression out of a coordinate structure (a structure formed by linking two expressions of the same grammatical category by a conjunction).\footnote{This formulation of the Coordinate Structure Constraint is idiosyncratic to [61]; standard formulations prohibit any one-sided movement out of a coordinate structure. Johnson argues that the prohibition is well-attested only for $A$ and head elements.} Thus the transformation from:

   (54) Sophocles wrote tragedies and Aristophanes wrote comedies.

to:

   (55) *Comedies, Sophocles wrote tragedies and Aristophanes wrote.

is barred, although transformation to:

   (56) Tragedies and comedies, Sophocles and Aristophanes wrote.

is permitted. The Coordinate Structure Constraint blocks simple movement of ‘Äpfel’ out of ‘drei Äpfel und zwei Bananen’, and thus presents the major obstacle to a convincing ‘underlying structure’ story.

2. The top levels of a syntactic tree for a sentence are given by the maximal projection of the verbal inflection (IP), which then serves as the complement to an optional complementizer, leading to a culminating maximal projection of the complementizer (CP). When such movement is permitted, tree elements can be \textit{topicalized} by movement to the specifier of C; and when no complementizer is present, and when such movement is permitted, verbs can be moved to C.

3. German requires verb-final word order in subordinate clauses, as in ‘daß der Hans drei Äpfel isst’. In some closely related Germanic languages, however, verb-final word order is not always required. [61] gives the example, from West Flemish, of the permissible ‘da Jan wilt dienen boek kuopen’, rather than the verb-final ‘da Jan dienen boek kuopen wilt’. To account for this variation, suppose that verb phrases serve as arguments to a functional phrase – call it FP. Suppose that F (the functional head of FP) requires movement either of (a) a verb to F or of (b) a VP to the specified of FP. The choice of movement will determine the presence or absence of verb-final position, with (a) blocking that position and (b) enforcing it.

The overview relies on four assertions about an adequate syntactic theory, which will be listed but not justified:
4. If the same constituent element occurs in both parts of a coordinate structure, then its occurrence in the second part can be deleted (‘gapped’). Thus ‘Francine read Ubu Roi and Maxwell read Ubu Cuckolded’ can be transformed into ‘Francine read Ubu Roi and Maxwell read Ubu Cuckolded’.

With these pieces in place, suppose (53) begins with the underlying structure:

(53-T₁)

With ‘Und’, at this level, joins ‘drei Äpfel’ and ‘zwei Bananen’, allowing for a smoothly compositional semantics. The second occurrence of ‘isst’ is then gapped, yielding:

(53-T₂)
F then requires movement of a VP to the specifier position, yielding:

(53-T₃)
The Coordinate Structure Constraint allows movement of VP, so this is permissible. Since (53-T₃) contains no complementizer, ‘isst’ can move to that position (now that it is out of the coordinate structure):

(53-T₄)
Finally, ‘Äpfel’ can be topicalized to yield:

(53-T\textsubscript{5})
which, using a simple left-to-right read-out procedure, gives (53). It remains to be seen, of course, whether this speculative syntactic story extends successfully to the rest of the language. Compositionality for a single sentence comes cheap, but the cost increases rapidly as the purchase extends to a wider portion of a language.\footnote{Johnson in \cite{61} cites as evidence in favor of his syntactic analysis its ability to explain the unacceptability of (53)'s minimal variant: \begin{equation} \text{Apfel wird der Hans drei und zwei Bananen essen.} \end{equation} The terminal occurrence of ‘essen’ requires that there be only a single verb phrase (since an earlier occurrence of ‘essen’ could not be gapped), giving the starting point: \begin{equation} (53V) \end{equation}}

6 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.

### 6.1 Observational Compositionality

Enough has been said by now to dispel the attraction of the following naive line of reasoning:

The movement of VP to the specifier of FP then yields:

(53V-T2)
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 naive 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 (although not impossible) to identify factors relevant to its meaning outside its own parts. A meaning theory which derives the meaning of the token occurrence of 'Hesperus is a planet' imbedded in 'Albert believes Hesperus is a planet' using, in part, the meaning of 'Albert' violates semantic locality and thereby compositionality, but if we are asked simply to give the meaning of the sentence type 'Hesperus is a planet', it becomes tempting to think that we have no tools available other than 'Hesperus',

But here 'Apfel' remains in a coordinate structure, and hence cannot be moved out by topicalization. With only a single VP, the F-forced movement cannot separate the coordinate structure as occurs in (53).
Compositional theory says, roughly, that its syntax and its lexical constituents determine the meaning of a complex expression; it’s thus part of the explanation of why practically everybody who understands ‘dogs’ and ‘bark’ understands ‘dogs bark’. But it also needs explaining that you practically never find people who understand ‘dogs bark’ but don’t understand ‘dogs’ or ‘bark’. What we’ll call ‘reverse’ compositional explains this by assuming that the meanings of constituent expressions supervene on the meanings of their complex hosts. If that’s right, then if you understand ‘dogs bark’, it follows that you know everything to determine the meanings of ‘dogs’ and ‘bark’. Fine so far, but now there’s a further puzzle: as far as anybody knows, compositionality and reverse compositionality always go together. … The explanation is obvious; the meaning of ‘dogs bark’ supervenes on the meanings of ‘dogs’ and ‘bark’ because the meanings of ‘dogs’ and ‘bark’ are parts of the meanings of ‘dogs bark’, and meaning of ‘dogs’ and ‘bark’ supervene on the meaning of ‘dogs bark’ for exactly the same reason. ([36])

Fodor and Lepore cite two aspects of our experience with language. To avoid loading the terminology unduly, call these features apparent compositionality (AC) and apparent reverse compositionality (ARC):

- **AC**: Practically all competent speakers of a natural language who understand all the parts of an expression $\epsilon$ will also understand $\epsilon$.
- **ARC**: Practically all competent speakers of a natural language who understand an expression $\epsilon$ will also understand all parts of $\epsilon$.

AC and ARC, goes the argument, 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:

(57) Be more careful when sorting the specimens – you put that octopus on the squid chair.

---

117 Of course, a compositional type-level theory compatible with various forms of context-sensitivity can always be constructed, by making meanings of expressions functions from contexts to (pre-theoretic) meanings. This is just to reiterate the lessons of sections 4.1.1 and 4.1.2.
With a hammer and some nails, I can assemble those frozen squid into a squid chair.

If you want to stop discussing the special effects of *Jurassic Park* and start discussing the special effects of *20,000 Leagues Under the Sea*, you should move from the dinosaur chair to the squid chair.

Similarly with reverse compositionality – one can understand:

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*.

It is also not obvious that understanding the meaning of ‘Hesperus is a planet’ as an isolated sentence contributes usefully toward understanding the meaning of ‘Hesperus is a planet’ as embedded in ‘Albert believes Hesperus is a planet’, or vice versa.

Even supposing the truth of AC and ARC, compositionality may not be the best explanation. Suppose a speaker of English, familiar with the meanings of ‘dogs’ and ‘bark’, is presented with the sentence ‘dogs bark’, and understands it. Compositionality follows only if the meanings of ‘dogs’ and ‘bark’, together with syntax, are the only facts 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 (both Kaplanian context facts, such as speaker, audience, and time and world of utterance, and broader contextual facts such as the background of common knowledge, even if that background is null), and these facts may be pivotal in sentential interpretation. We know, after all, 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 sentences.

Can ACR be rescued by claiming that ‘*je ne sais quoi*’, and perhaps trivially ‘telephone’, are idioms? Thus:

To say that you could master “pet fish” without mastering “pet” (or, mutatis mutandis, “red triangle” without mastering “red”) is tantamount to saying that “pet” isn’t *really* a constituent of “pet fish” after all; which is, in turn, tantamount to saying that “pet fish” is an idiom. Which, however, “pet fish” patently is not. ([32], 52)

If idioms just are expressions whose meanings are not determined by their parts, as this argument would suggest, and ACR is qualified to be limited to non-idiomatic expressions, then an argument for compositionality from ACR, within the domain of ACR, is clearly circular. Thus some independent standard of idiomaticity is required, and not provided, here. Furthermore, that standard must not be: idioms are expressions one can understand without understanding the meanings of their parts, or ACR becomes trivialized.

Regarding which, see section 6.2 below.

Note 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
Reasoning along similar lines, one might see compositionality in the possibility of on-line language interpretation:

In a translation program one would like to be able to interpret a text in an on-line manner, i.e., incrementally, processing and interpreting each basic unit as it comes along, in the context created by the interpretation of the text so far. ([46], 93)

Whether natural language interpretation does in fact occur on-line is a matter of ongoing debate in empirical linguistics, undermining the prospects for a powerful argument for compositionality. Compositionality, too, is at best a sufficient condition for on-line processing, not a necessary one, since on-line processing is compatible with the utilization of extensive non-semantic inputs in calculating meanings:

Although certainly not the only way to meet this requirement, compositionality is a most intuitive way to do so. As such, on-line interpretation does not preclude that in the interpretation of a unit of text, other things than the interpretation of the text sofar play a role. But it does require that at any point in the processing of a text we are able to say what the interpretation thus far is. ([46], 93-94)

An argument for compositionality here would thus need to take the form of an inference to the best explanation, and would require the dubious auxiliary premise that complex expressions are syntactically structured so that their component parts always occur consecutively at surface structure (which would rule out the possibility of any transformational syntactic component relevant to semantic structure).

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 section 5, 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. It will, however, like all of the observational arguments discussed here, at best tell us only that natural languages are compositional, without telling us why that compositionality is significant.

6.2 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 sentence meanings. Even if ACR is further strengthened to block this possibility, strong parallelism will not follow (although compositionality will), since the doubly-strengthened ACR will follow from the assumption that speakers are 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.
ses, or which we would like \text{L} to possess. \text{X} might be \textit{first-order}, or \textit{hyperintensional}, or \textit{of subject-predicate form}, or \textit{systematic}, or \textit{bivalent}, or \textit{admitting of adherence to the Gricean maxims}, or any number of other features. If compositionality can be shown to be inferentially related to feature \text{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 \text{X} in \text{L}, guarantees the compositionality of \text{L}.

2. If compositionality is a sufficient condition for feature \text{X}, then the presence of \text{X} in \text{L} offers no guarantee that \text{L} is compositionality, but compositionality may still serve as a plausible explanation of the \text{X}-ness of \text{L}, allowing for an inference to the best explanation of the compositionality of \text{L}.

In the ideal case, compositionality will be both necessary and sufficient for \text{X}, but features so closely inferentially related to compositionality, but nevertheless independently verifiable in \text{L}, are hard to come by. Both styles of argument (from necessity and from sufficiency) 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 \text{X}.

\subsection*{6.2.1 The Learnability Argument}

This overview of compositionality began with an example of a novel sentence which was nevertheless immediately comprehensible to any competent speaker of English. English, like all natural languages, has an infinite number of grammatical and meaningful sentences, and \textit{a fortiori} has an infinite number of such sentences which a given speaker has never encountered. Nevertheless, competence in English gives one the capacity to understand all of these sentences.\footnote{The capacity requires, of course, a certain amount of idealization. Given finite lifespan, memory, and cognitive capacity, one cannot in fact understand a sentence that is \(10^{100}\) words in length, but the structure of one’s linguistic competence is such that it in principle allows for grasping the meaning of such a sentence. The claim that linguistic competence is infinite in scope has received some resistance, such as \cite{110}, due to these idealization issues, but linguistic competence capacities do not differ from any other capacities in their need for idealization (see linking literature).} 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?

We have already seen one answer to this question in Frege’s famous passage from ‘Compound Thoughts’ quoted earlier:

\begin{quote}
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
\end{quote}
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. ([39], 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 will, via recursive reapplication, account for our infinite linguistic capacity. Thus we have reason to think that English is compositional. Call this the *learnability argument*. Frege puts the argument in terms of the *necessity* of compositionality for learnability, but 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. ([10], 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. ([8], 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.\footnote{122 Given Church's thesis.} 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 is computable. If, on the other hand, \( L \) has no computable meaning function, then any procedure, graspable by beings like us, by which meanings of complex expressions can be determined will fail to determine the meanings of some expressions of \( L \). \( L \) is thus unlearnable if uncomputable.
We saw in section 3.3.1, however, that there is no interesting inferential relation between computability and compositionality. A computable language can fail to be compositional, and a compositional language can fail to be computable. Thus that \( L \) is computable certainly does not guarantee that it is compositional, and it also provides no plausible grounds for an inference to the best explanation to the compositionality of \( L \), since compositionality is not sufficient for and hence would not in fact explain computability.

Compositionality is not sufficient for computability because the compositional function might not be computable. Hence the only auxiliary premise which can be added to that of compositionality to derive computability is computability itself. Similarly, computability is not sufficient for compositionality because the computable function might not be compositional in form. Hence the only auxiliary premise which can be added to that of computability to derive compositionality is compositionality itself. 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. The one tells us nothing about the other.

### 6.2.2 The Argument From Safety

An expressivist about moral language might want a semantic theory according to which a sentence such as:

\[
(61) \text{Boiling babies in oil is wrong.}
\]

has as meaning a con-attitude toward the boiling of babies in oil, which we might express as:

- \([\text{Boiling babies in oil is wrong}] = \text{Boo! boiling babies in oil!}\)

In order to capture the non-cognitivism of expressivism, \( \text{Boo! boiling babies in oil!} \) must not have, be, or determine, a truth-value or a truth-evaluable proposition. But now the expressivist encounters sentences such as:

\[
(62) \text{If boiling babies in oil is wrong, then sauteeing them in butter is also wrong.}
\]

If ‘if’ expresses a truth function, then the expressivist has a problem, since the truth function here appears to apply to \( \text{Boo! boiling babies in oil!} \) and \( \text{Boo! Sauteeing babies in butter!} \), neither of which, according to the heart of the expressivist position, is truth-related. This is the Frege-Geach problem: how to imbed expressivism, whose

---

123 Szabo, in chapter 3 of [98], gives an extended and insightful critique of the learnability argument. He draws 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. I have suppressed these additional assumptions because, in the light of the gap between computability and compositionality, they make no difference even if granted. Szabo draws a distinction between a Strong and Modest Principle of Understanding which corresponds to treating compositionality as necessary or sufficient, respectively, for learnability; his specific examples (algebraic chess notation and decimal representations of integers) of productive systems of complex expressions learnable without appeal to compositionality provide additional examples of the general observation above that compositionality is not necessary for computability.
basic semantic currency is the (pro- or con-) attitude, in the larger language, whose
basic semantic currency is, say, the truth value or the truth-evaluable proposition.

An expressivist has many paths of response to the Frege-Geach problem available.
Suppose that, unwisely, he chooses the following. The meaning of ‘wrong’ is a con-
attitude. When ‘wrong’ is used in simple subject-predicate sentences of the form ‘X
is wrong’, the result is the expression of that con-attitude toward the referent of ‘X’.
When, however, that same sentence ‘X is wrong’ is imbedded in a conditional of the
form ‘If X is wrong, then φ’, then it no longer provides an expression of a con-attitude
toward an object, but rather provides (say) a set of typical consequences, under some
notion of practical reasoning, of the holding of a con-attitude toward the referent of ‘X’.
The conditional can then be evaluated by comparing that set to the standard proposition
expressed by the consequent (if the consequent is free of moral language), in a manner
which is potentially compatible with non-moral conditionals.

The expressivist has thus opted for a non-compositional semantics. Simple sen-
tences with moral vocabulary have one sort of meaning (attitude expression) when
occurring in isolation, but another sort of meaning (determination of typical pragmatic
attitude consequences) when occurring in a conditional. These two options, further-
more, are unlikely to exhaust the non-compositionality of moral language. Different
contributions may be necessary when simple moral sentences are imbedded in nega-
tions, or modal operators, or propositional attitude contexts, or quantifiers. A difficulty
is thus created for the semantic theorist. Suppose that a semantic theory, motivated
by expressivist principles, has been given for some small segment L− of L. L− contains,
perhaps, simple subject-predicate moral sentences, as well as combinations of
such sentences under (what are normally considered) Boolean connectives, and is non-
 compositional along the lines sketched above. A semanticist now wants to construct
a meaning theory for some other portion of L, such as a simple fragment L₁ of the
portion of L using counterfactual conditionals. If L₁ contains moral vocabulary (for
example, counterfactuals with moral antecedents or consequents), then, given the prior
commitment to the non-compositionality of moral language, then the semanticist must,
in addition to working out the semantics of counterfactuals, also work out how to im-
plement the moral semantics in the counterfactual domain.

Semantic theory construction often proceeds on the assumption that a certain kind
of semantic modularity is possible. A semantic theory is constructed for a particular
phenomenon (quantifiers, modal operators, propositional attitude contexts, progressive
aspect, etc.) on the assumption that the semantically distinctive features of that phe-
nomenon can be fully characterized on a restricted linguistic domain. A larger-scale
semantic theory is then created by plugging together several modules. The modu-
lar combination may, of course, reveal aspects of the semantic behaviour of the phe-
nomenon targeted by a particular module, but this should be the exception, revealing
a failure in the original modular construction, rather than the norm. If creating a se-
mantic theory to cover multiple domains inevitably requires reworking the semantics
for the individual domains, in addition to designing the interfaces of the domains, then
semantic theory construction becomes prohibitively difficult.

A non-compositional theory, at least if its non-compositionality derives from con-
textual sensitivity to environment of embedding, makes modular theorizing
impossible, as the expressivist example above shows. In this example, there is in an

77
important sense no such thing as the theory of moral language; the semantics of moral language evolves and changes every time new vocabulary is added to the target domain. A compositional theory, on the other hand, requires that all sentences built up using moral vocabulary have meanings determinable from the meanings of that moral vocabulary, so once a compositional theory of a moral fragment of $L$ is given, a moral module is created, and the semantics of other constructions can be studied in their interactions with moral language without redesigning that module. Compositionality, then, contributes toward a valuable safety of semantic theory design, and a desire for that safety can, via a consequentialist argument, translate into a desire for compositionality.

6.3 Methodological Compositionality

When compositionality is not motivated using the learnability argument, it is most often introduced as a methodological principle governing semantic theorizing. What it means to say that something is a methodological principle is far from clear, and the term runs the risk of becoming little more than a label for a claim we would like to think is true, but lack a justification for, or a claim that we will stipulatively take as part of the definition of a field. However, a suitably precise characterization of a methodological principle can be given as follows:

$$(\text{MP}) \quad \Phi \text{ is a methodological principle for an activity } A \text{ if } \Phi \text{ either is, or is a logical consequence of, a claim whose truth is a constitutive feature of performance of } A.$$ 

Suppose, for example, that in order for the construction of a theory to count as the construction of a semantic theory (and hence for a theory to count as a semantic theory), the theory constructed must obey the Context Principle. Perhaps this is the case, as Frege suggests in [41], 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.\textsuperscript{124} Perhaps it is the case, as [85] 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 will be a methodological principle of semantic theorizing, and given that, as seen in Lemma 1 of section 4.2, the Context Principle entails compositionality, compositionality will also be a methodological principle. In seeing this, we will see why semantic theories must be compositional, and how compositionality contributes toward the very nature of semantics.

\textsuperscript{124}Thus:

We next laid down the fundamental principle that we must never try to define the meaning of a word in isolation, but only in the context of a proposition: only by adhering to this can we, as I believe, avoid a physical view of number without slipping into a psychological view of it. ([41], §106)
6.3.1 The Methodological Principle of Semantic Innocence

Suppose that, persuaded by [70], we take a proper names to carry a referent as its sole semantic value. In constructing a semantic theory, we thus start listing facts such as:

- \( \llbracket \text{Julius Caesar} \rrbracket = \text{Julius Caesar} \)
- \( \llbracket \text{Christopher Marlowe} \rrbracket = \text{Christopher Marlowe} \)
- \( \llbracket \text{Peter Greenaway} \rrbracket = \text{Peter Greenaway} \)

These axioms carry with them ontological commitments; we cannot endorse them without also endorsing the existence of Julius Caesar, Christopher Marlowe, and Peter Greenaway. The existence of these particular individuals, though, we might feel is a matter for history, broadly speaking to settle, not semantics. This feeling might be heightened when we attempted to complete the following semantic claims:

- \( \llbracket \text{Santa Claus} \rrbracket = ? \)
- \( \llbracket \text{Homer} \rrbracket = ? \)
- \( \llbracket \text{King Arthur} \rrbracket = ? \)

Construction of an adequate semantic theory, we might plausibly feel, ought not require settling questions about the authorship of epic Greek poetry or the government of the British Isles in the post-Roman era. A similar worry arises when we give meaning axioms for common nouns such as:

- \( \llbracket \text{electron} \rrbracket = \{ x : x \text{ is an electron} \} \)
- \( \llbracket \text{theta role} \rrbracket = \{ x : x \text{ is a theta role} \} \)

Endorsement of such axioms involves, if not ontological, at least ideological commitment to electrons and theta roles, commitments one might feel are the proper domain of physics and syntax, respectively, not of semantics.\(^{125}\)

The worry thus emerges that in order to give a semantic theory, we will in effect have to give a theory of absolutely everything – that all questions of history, physics, syntax, theology, and so on must be answered as a precondition for doing semantics. Szabo, in chapter 2 of [98], calls this Bloomfield’s problem after the following comment from Bloomfield:

The situations which prompt people to utter speech include every object and happening in the universe. In order to give a scientifically accurate definition of meaning for every form of a language, we would have to have a scientifically accurate knowledge of everything in the speaker’s world. The actual extent of human knowledge is very small compared to this. ([7], 139; via [30], 248 and [98], 30)

\(^{125}\)Again questions of the boundaries of proper commitment arise here. Should, for example, an adequate semantic theory endorse \( \llbracket \text{unicorn} \rrbracket = \{ x : x \text{ is a unicorn} \} \) or \( \llbracket \text{karma} \rrbracket = \{ x : x \text{ is karma} \} \)?
Szabo argues that semantics – or more properly, a subdiscipline of semantics he calls linguistic semantics, as distinguished from philosophical semantics – must, if it is to be a feasible field of study, proceed in an innocent fashion, so that formulation of a semantic theory does not require totalizing ontological and ideological commitments. He also argues that compositionality makes such innocence possible. If this is correct, it provides a methodological argument in favour of compositionality.

Szabo suggests that innocence can be achieved using what he calls hypothetical descriptive semantics. A hypothetical descriptive semantics has as axioms exclusively claims of the form:

\[ \forall f \forall \sigma \exists \epsilon_1, \ldots, \epsilon_n (([\epsilon_1] = \sigma(\epsilon_1) \land \ldots \land [\epsilon_n] = \sigma(\epsilon_n)) \rightarrow [f] = O(\sigma(\epsilon_1), \ldots, \sigma(\epsilon_n)) \]

where \( f \) is restricted to some particular syntactic or semantic category, \( \sigma \) provides a way of assigning semantic value to expressions, and \( O \) is some semantic operation. If, for example, \( L \) is a simple language consisting of subject-predicate expressions, with either unary or binary predicates, and some Boolean connectives, then a hypothetical descriptive semantics might be given as follows:

\[ \forall \Psi \forall \tau \forall \sigma ([\Psi] = \sigma(\Psi) \land [\tau] = \sigma(\tau)) \rightarrow ([\Psi \land \tau] = \sigma(\Psi) \land \sigma(\tau)) \]

Axioms of this form Szabo calls structural conditionals.\(^{126}\) A semantic theory built only from structural conditionals is, roughly speaking, the result of taking a standard semantic theory, removing the clauses for all of the semantic atoms, and then generalizing the clauses for complex expressions to range over possible values for expressions contributing to their meaning. A hypothetical descriptive semantics of \( L \), although in some senses very thin, not even determining what sentences of the language mean (‘It is not even committed to the claim that the sentence Julius Caesar was murdered on the ides of March means that Julius Caesar was murdered on the ides of March’ ([98], 55)), may for example allow determination of inferential relations among those sentences. In the absence of the clauses for the semantic atoms, however, the hypothetical descriptive semantics will be ontologically and ideologically neutral. It simply tells us what to do with meanings, whatever those meanings turn out to be, without saying anything about what they in fact are.

Szabo claims that the prospects for a hypothetical descriptive semantics, and hence for properly innocent linguistic semantics, ‘depend largely’ ([98], 56) on whether the principle of compositionality holds.\(^{127}\)

\(^{126}\)We will see below that this terminology threatens to beg important questions.

\(^{127}\)Szabo does not claim that compositionality is either necessary or sufficient for a hypothetical descriptive semantics. He says that if compositionality hold, ‘it might be possible to come up with a theory that entails all structural conditionals’ ([98]), and also says ‘I argued that if compositionality is true, there might be such
If the principle of compositionality . . . is true, we can understand how the structural conditionals can be true. Take the complex expression \( f \), whose constituents are \( \epsilon_1, \epsilon_2, \ldots, \epsilon_n \). Let \( \sigma \) assign meanings to these expressions. It is an instance of the principle of compositionality that \( \sigma(f) \) is determined by \( \sigma(\epsilon_1), \sigma(\epsilon_2), \ldots, \sigma(\epsilon_n) \) and the structure of \( f \), i.e. that there is a law-like connection in all possible human languages\(^{128}\) between \( \sigma(f) \) and \( \sigma(\epsilon_1), \sigma(\epsilon_2), \ldots, \sigma(\epsilon_n) \) constituted by the structure of \( f \). These law-like connections are the fundamental features of human languages that are responsible for the truth of the instances of [structural conditionals]. If the principle of compositionality is true, linguistic semantics can be an empirical investigation of the law-like connections that make the structural conditionals true. So, in a sense, it is the crucial principle in the attempt to meet Bloomfield’s challenge. ([98], 54)

The possibility of a hypothetical descriptive semantics turns out, however, to have little to do with compositionality. Suppose that semantic locality is violated, and that the meaning of some expression \( f \) depends (entirely or in part) on the meaning of some \( \epsilon \) that is not a part of \( f \). Since the definition of a structural conditional says nothing about a parthood relation, and imposes no requirement that the \( \epsilon_i \)'s be parts of the \( f \) in question, this violation of semantic locality in no way threatens the availability of appropriate structural conditionals.\(^{129}\) The goal of hypothetical descriptive semantics is to assign meanings to expressions only hypothetically, under the conditional assumption that certain other expressions have certain meanings, and hence to avoid categorical commitments which violate innocence; that goal does not require that those second expressions be parts of the first. Nor does the quest for innocence even require that the antecedent of the conditional be loaded with semantic facts. If the meanings of complex expressions can be hypothetically derived from the assumption that, for example, the atmospheric pressure is such-and-such, or the winner of the World Series was so-and-so, then innocence can also be achieved. Of course, if the hypothetical inferential basis becomes too rich, then innocence will fail, but semantic closure can be violated without going so far.

Suppose \( L \) has structural conditionals such as:

- If the context of utterance is \( c \), then \( \llbracket f \rrbracket = O(c) \).
- If \( \llbracket \Pi \rrbracket = \sigma(\Pi), \llbracket \tau \rrbracket = \sigma(\tau), \) and \( \llbracket \Sigma \rrbracket = \sigma(\Sigma) \), then \( \llbracket \Pi \tau \rrbracket \), when \( \langle \Pi \tau \rangle \) occurs in the context \( \langle \Sigma(\Pi \tau) \rangle \), is \( O(\sigma(\Sigma), \sigma(\Pi), \sigma(\tau)) \).
- If \( \llbracket \phi \rrbracket = \sigma(\phi) \) and the current distance from the Earth’s center of gravity to Mars’ center of gravity is \( x \) kilometers, then \( \llbracket \neg \phi \rrbracket = O(\sigma(\phi), x) \).

\(^{128}\)This requirement is a consequence of Szabo’s particular conception of compositionality, as discussed in section 3.1.5 above.

\(^{129}\)It is for this reason that the terminology ‘structural conditional’ is tendentious: it implies that the meanings of complex expressions are to be hypothetically derived in terms of their structures, when this particular choice of derivational basis is in fact irrelevant to achieving the innocence at which structural conditionals are aimed.
Then L has been given an innocent semantic theory – rather than reaching out to encompass the ontological and ideological commitments of all human theories, it requires instead only a few modest astronomic facts. That semantic theory, however, is non-compositional: the first and third structural conditionals violate semantic closure, and the second violates semantic locality. Not just any possible language can be given a hypothetical descriptive semantics – if every complex expression in a language is sensitive to every feature of the world, then the attempt will fail – but when it is possible, it is possible because of the relative paucity of the information needed to determine complex expression meanings, not because of the type (semantic closure) or location (semantic locality) of that information. The quest for innocence thus provides no methodological impetus toward compositionality.

6.3.2 The Methodological Principle of Ontological Revelation

Sentences of ‘donkey anaphora’, such as the classic:

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

have always presented challenges to compositionality. In a first-order quantified language, obtaining the right truth conditions for (63) calls for an analysis such as:

(64) \( \forall x \forall y ((\text{man}(x) \land \text{donkey}(y) \land \text{owns}(x,y)) \rightarrow \text{beats}(x,y)) \)

However, the constituent sentence ‘A farmer owns a donkey’ requires:

(65) \( \exists x \exists y (\text{farmer}(x) \land \text{donkey}(y) \land \text{owns}(x,y)) \)

and this first-order formula does not neither appears as a constituent of nor seems to contribute to the systematic determination of (63-1), given the shift from existential to universal quantification, and the extension of the quantifier scopes to cover ‘beats(\(x, y\))’.

Kamp’s Discourse Representation Theory (DRT) (introduced in [62]) aims at a more natural treatment of natural language anaphora. 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, the sentence:

(66) A farmer owns a donkey.

gives rise to the DRS:

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

130There is, I take it, no reason to think that a semantic theory can be given in complete isolation of all other fields of investigation. Certainly compositionality will provide no such reason.
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. When a pronoun is used anaphorically on an earlier indefinite, the same discourse referent is used. Thus:

(68) A farmer owns a donkey. He beats it.

becomes:

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

Notice that the DRS (67) is not a proper part of the DRS (69), just as the first-order formula (65) is not a proper part of the first-order formula (64). The larger DRS needs to connect ‘he’ and ‘it’ to the appropriate discourse referents, just as the first-order formula (64) needs to bring the variables ‘x’ and ‘y’ within the scope of the relevant quantifiers, and in both cases this can be done only by referring to the nature of ‘A farmer owns a donkey’ while analyzing ‘He beats it’, in violation of semantic locality.

Conditional sentences, such as (63), are given DRS’s by embedding two DRS’s, joined by a conditional arrow, within a larger DRS:

\[
\begin{array}{c|c}
X, Y & \\
\hline
\text{farmer}(X) & \\
\text{donkey}(Y) & \\
\text{owns}(X, Y) & \\
\text{⇒} & \\
\end{array} \quad \begin{array}{c|c}
X, Y & \\
\hline
\text{beats}(X, Y) & \\
\end{array}
\]

A DRS is true, given a model specifying a domain of interpretation and extensions for the conditions, if there is an assignment to the discourse referents satisfying the conditions. Since multiple sentences, with pronouns anaphoric on the same indefinite, collapse into a single DRS, this method of semantic interpretation guarantees that anaphorically-linked phrases will receive the same interpretation (since they are all, in the DRS, occurrences of the same discourse referent). 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 (adding assignments to new discourse referents where necessary) to an assignment satisfying the conditions of the second DRS. Indefinites in the antecedents of conditions thus have universal force, since every extendable interpretation must now be checked.
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), as introduced in [46], offers a way of restoring compositionality while capturing the key insights of DRT. DPL provides a compositional treatment of an apparently non-compositional phenomenon like anaphora by making a fundamental alteration in the kind of semantic values assigned by the theory. Whereas traditional, static, quantified logics assign sentences satisfaction conditions – functions from variable assignments to truth values – DPL assigns sentences input-output pairs. A sentence of DPL does not simply receive an assignment and evaluate relative to it; it can, instead, change an assignment. If variable assignments are thought of, in the style of DRT, as specifications of context indicating which objects are conversationally salient, then the dynamic insight of DPL is that sentences can affect, as well as be affected by, context.

Syntactically, DPL is a standard first-order quantified 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 simply check that assignments satisfy a descriptive condition and ‘let pass the assignments which satisfy, and block those that don’t.’ ([46], 46). Thus:

- \[
\{\langle g, h \rangle : g = h \land g \models \Pi \tau_1 \ldots \tau_n\}\]

Existing quantifiers, however, are dynamic. An existentially quantified sentence does not return the same assignments it receives, but rather expands the set of assignments so that all individuals meeting the existentially quantified condition become live options for the assignment to the quantified variable. Formally, we have:

- \[
\{\langle g, h \rangle : \exists k : k[x]g \land \langle k, h \rangle \in \phi\}\]

where \(k[x]g\) asserts that \(k\) differs from \(g\) at most in the assignment to \(x\). Thus consider:

(73) \(\exists x F.x, Gx\)

131If 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 (which then occurs in the discourse referent assignments), along the lines of logical form, then DRT is plausibly compositional, albeit with an idiosyncratic syntax.

132Non-compositionality shows up in the inability to replace a sentence in a discourse with a logically equivalent one; thus the infelicities of:

(71) * Not every man doesn’t own a donkey. He beats it.
(72) * Either a farmer doesn’t own a donkey or he beats it.

133Where \(g \models \phi\) means that the assignment \(g\), given the background model, makes true \(\phi\) in the classical sense.
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). The existential quantifier $\exists x$ 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. The input to $Gx$ will then be these assignments, with $x$ now representing possible identities of the discourse referent introduced by the existential quantifier (indefinite) of $\exists x Fx$, and $Gx$ will act as a test on these assignments, passing them through unchanged if their $x$ value is $G$. The net effect, then, is that an assignment is output if its $x$ value is both $F$ and $G$, which captures the idea that existential quantification in effect extends beyond sentential bounds to all anaphorically-linked pronouns. The dynamic reach of the existential quantifier will thus be potentially unlimited on the right, not ending until another existential quantifier of the same variable resets pronouns. The dynamic reach of the existential quantifier will thus be potentially unlimited on the right, not ending until another existential quantifier of the same variable resets the assignments.

Conditionals also have a dynamic effect, and are interpreted as follows:

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

Conditionals are only internally dynamic. The output assignment $h$ must be the same as the input assignment $g$, but in checking whether an assignment successfully passes through, alterations of that assignment induced by the antecedent are checked against the consequent. An assignment passes through if and only if every possible output of it, passed through the antecedent, produces some output when passed through the consequent. Thus consider:

(74) $\exists x Fx \rightarrow Gx^{134}$

Let $g$ be an arbitrary incoming assignment. When that assignment hits the conditional, we consider outputs of it through the antecedent. The antecedent, with its existential quantifier, dynamically alters the $x$ value of assignments to an $F$ object. Each such alteration is then checked against the consequent. Since the consequent is an atomic sentence, it acts only as a check on assignments, passing them through if their $x$ value is $G$. If each dynamically-altered output of the antecedent passes this test, then $g$ is output, unchanged, from the conditional. The effect is that $g$ will emerge unchanged if and only if every $F$ is $G$. This captures the shift to existential force which characterized (63), and indeed DPL will produce the desired interpretation of that sentence.

The DPL semantics are easily seen to be perfectly compositional: every constituent of a formula can be locally 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 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

---

134 Where, importantly, the scope of the existential quantifier is limited to the antecedent.
shift in our conception of basic semantic value – now as a tool of updating context through language.\textsuperscript{135}

In sections 1 and 2, we saw how the pursuit of compositionality led Frege and (an interpretation of) Lambert of Auxerre to change their basic conception of semantic value: in Frege’s case, supplementing the world-oriented notion of \textit{Bedeutung} with the epistemically-oriented notion of \textit{Sinn}, and in Lambert’s case, adding verbally-induced quantification over events to talk of objects. DPL’s attempt to make DRT compositional leads to a similar shift, in the move to dynamic semantics. This pattern suggests another methodological argument for compositionality, one based on protecting the philosophical role of subsentential semantics.

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. Trivially, if word meanings (plus syntax, and plus other inputs if desired) determine sentence meaning, and words determine word meanings, and orthographic structure determines word, then by functional composition, orthographic structure (plus syntax, and perhaps plus other inputs) determines sentence meaning, despite the fact that orthographic structure is not semantically invested.\textsuperscript{136}

In \textit{The Foundations of Arithmetic}, Frege argues that because the numerical terms ‘zero’, ‘one’, and so on act as names in natural language, they must, in order to account for the full range of their expressive use, be taken as picking out objects, thus giving rise to an ontological commitment to numbers. He briefly essays an \textit{adjectival} approach, on which, for example:

\begin{align*}
\text{(75) } & \text{The number 0 belongs to a concept } F. \\
\text{(76) } & \text{The number 1 belongs to a concept } F. \\
\end{align*}

mean respectively:

\begin{align*}
\text{(77) } & \text{No object falls under concept } F. \\
\text{(78) } & \text{Not everything fails to fall under } F, \text{ and if } a \text{ and } b \text{ fall under concept } F, \text{ they are identical.}\textsuperscript{137}
\end{align*}

These definitions are rejected, however, because, as Frege puts it in §56, they fail to determine whether Julius Caesar is a number. One way of taking this objection is

\textsuperscript{135}For more on the dynamic perspective, see (e.g.) [101].

\textsuperscript{136}This argument is analogous to familiar causal overdetermination arguments against nonreductive conceptions of mental and other higher-level properties. The principal line of response to causal overdetermination arguments is to argue that important causal generalizations can be captured at the higher level which cannot, or cannot finitely, be captured at the lower level. In the current context, the analogous response is to appeal to lack of completeness, in which case there are inferential features of the language which elude (finite) syntactic capture. First-order, and modal first-order, fragments of natural languages are complete, and hence leave no room for uncapturable ‘higher-order’ lexicosemantic generalizations. Outside these fragments, completeness may fail and there may be a potential explanatory role for lexicosemantic generalizations, but it is unclear that any such generalizations feature in our semantic mastery or theorizing.

\textsuperscript{137}See §55 of [41] for these proposed definitions.
that the adjectival definitions, while they give sense to some constructions involving numerical terms (those involving attribution of numerical terms to concepts), they fail to give sense to other constructions, such as identities of the form ‘Julius Caesar is the number 0’. Since numerical terms occupy the syntactic category of names, they can appear in all the sorts of constructions that names appear in, and any semantic analysis of those terms must be faithful to that fact. Frege’s suggestion is that the only way to do so is to treat numerical terms in the same way that other names are treated: as picking out objects.

In §62 of the Foundations, Frege proceeds to show how the ontological debt incurred through the use of numerical terms can be paid off. If we understand a sufficiently wide class of sentences using numerical terms – Frege singles out the class of identity claims featuring numerical terms – then we can extract a semantics for those terms, and thereby an arithmetic ontology from that understanding, by using the comprehended class of sentences to construct a contextual definition of the numerical terms, or by explicitly defining the numerical terms as equivalence classes under an equivalence relation provided by the comprehended sentences.138

Suppose we generalize Frege’s project in the following way. Our linguistic practice, in which we affirm or deny various propositions, expressed by various sentences, in various situations, represents, in Wittgensteinian terminology, a ‘form of life’. Our view of the world we live in shows up in our linguistic practice. That worldview, however, at the level of sentential or propositional representation, lacks internal articulation. The role of lexical meanings is to decompose our linguistically represented form of life into various ontological categories. It is by having a syntactic category of names, and by treating names as systematically contributing to the meanings of sentences, that we come to represent the world as (partially) composed of objects. It is by having a syntactic category of verbs, and by treating verbs as systematically contributing to the meanings of sentences, that we come to represent the world as (partially) composed of events. Our lexical decomposition of our linguistic practice tracks our ontological taxonomy of the world into things, properties, actions, modalities, and so on.

If lexical meanings are to reveal the ontological commitments inchoately present in our linguistic practice, though, the structure of the linguistic practice must be faithfully mirrored in those 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 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, 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

138See §§62-69 of [41].
139‘To imagine a language means to imagine a form of life.’ ([108], §19)
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

[1] On an argument against semantic compositionality.


