

Chapter 1—Context and compositionality

1 The context principle

This is a philosophical inquiry into the relation between language and the world. Like all inquiry, it must start somewhere: I start—as a number of philosophers do, both classic and contemporary¹—with the existence of linguistic communication, the conveying and understanding of meaningful items of discourse. Meaningful sentences have meaningful subsentential components, and the meaningfulness of these components is governed, as I take it (an argument will be provided shortly), by what is generally known as the context principle. This principle has been formulated in a variety of ways, but in its most familiar, Fregean version it states that words mean (*bedeuten*) something only in the context of a sentence.² We might gloss this as saying that words are *made for* sentences; sentences are what give words their *function and point*. Like Frege, I shall focus here on the declarative sentence, making the assumption that other kinds of sentence can have their functions either reduced to or explained in terms of declaratives.³ Since Aristotle the declarative sentence has been defined to be a group of words that is true or false⁴—assuming that contextual parameters are fixed, we must add, and waiving a few complications (Aristotle himself, it would appear, wished to except future contingents from the rule).⁵ I follow that definition here, *modulo* the proviso about the role played by context (but without making an exception of future contingents). The focus on declarative sentences is not just a matter of convenience: more importantly, the connection with truth and falsity, mediated by declaratives, is both essential to and distinctive of language.

I do not insist that a group of suitable words must actually be *uttered* before it can take a truth-value.⁶ I thus work with a semantic, not a syntactic, definition of the sentence: this is standard practice in modern linguistic theory, and indeed it is hard—actually, I think, impossible—to see how a syntactic, or purely structural, definition could have the requisite generality to ensure complete capture of the relevant entities.⁷ Sentences are syntactic entities, but the criterion for whether a given syntactic construction can serve as a

¹ On Wittgenstein, see Lear 1984, p. 230; for a contemporary example, see McGee 2006, p. 179.

² Frege 1884, §62; cf. §60; see Dummett 1997, pp. 12–13; Hale 2010, pp. 408–9.

³ See Lewis 1983, pp. 171–2, 220–6; Davidson 1999a, pp. 113–14; Gaskin 2008, pp. 6–8.

⁴ Aristotle, *De Interpretatione* 17a2–3.

⁵ See Gaskin 1995, chs. 1–12.

⁶ Cf. Dummett 2000, pp. 9–10; MacFarlane 2014, p. 48. *Contra*, e.g., Collins 2017, p. 152.

⁷ See Keenan and Stabler 2003, p. 11 with p. 39 n. 2; Gaskin 2008, pp. 25–9; A. Moore 2019, pp. 64–6.

sentence must be semantic—a matter of whether the construction can perform the communicative function of a sentence. By ‘words’, the components of sentences that are governed by the context principle, I shall understand morphemes in general, so including word roots and semantically significant affixes (such as ‘un-’ and ‘-able’).⁸

Frege mentions the context principle several times in the *Foundations of Arithmetic*, and clearly regards it as significant, but he says very little about it either there or elsewhere. For more detail we must look back to Jeremy Bentham’s *Essay on Language* and forwards to Wittgenstein’s *Tractatus* and *Philosophical Investigations*. Bentham not only anticipated Frege’s version of the context principle; he also defended it by pointing to the absurdity of supposing that words can be anything other than an abstraction from sentences. In effect, for Bentham, sentences are the *given* in linguistic communication; words and their meanings are *derived*, that is, theoretically secondary. The mistake in the Aristotelian tradition, according to Bentham, was to treat terms (words) as conceptually prior to sentences. Once this fateful step has been taken, we are then naturally led to speculate that ‘finding these terms endowed, each of them, somehow or other, with a signification of its own, at a subsequent period some ingenious persons took them in hand, and formed them into propositions’⁹—that is, into (as we would say now) sentences. But the speculation is outrageous: words, after all, do not come neat; they come with an implicit syntax. They are, as used to be said, ‘parts of speech’: the English word ‘cat’ is a noun, the French word ‘naître’ a verb, the German word ‘wenn’ a conjunction, and so on. Words come essentially packaged with *functions*, and these are *sentential* functions. So Bentham’s ‘ingenious persons’ would have had nothing to do; that is his point. The words they found would already be governed by an implied syntax; they would already be geared up to figure in sentences. So Bentham against the atomism of the tradition.¹⁰

The sentence enjoys conceptual priority over words because it is the smallest unit of discourse with which one can, as Wittgenstein said, ‘make a move in the language game’; above all, it is the smallest unit of discourse that can bear a truth-value.¹¹ No kind of word is exempt from the context principle: in particular, proper names are not exempt. This was a point on which Wittgenstein insisted in the opening sections of the *Philosophical Investigations*:

⁸ Cf. Horwich 1998b, p. 15.

⁹ Bentham 1843, vi, §1.

¹⁰ Bronzo 2014, pp. 1–2.

¹¹ Wittgenstein 1977, I, §49; Dummett 1981, p. 194.

Only someone who already knows how to do something with it can significantly ask after a name. . . . Naming is so far not a move in the language-game—any more than putting a piece in its place on the board is a move in chess. We may say: *nothing* has so far been done, when a thing has been named. It has not even *got* a name except in the language-game. This was what Frege meant too, when he said that a word had meaning (*Bedeutung*) only as part of a sentence.¹²

Further, as Wittgenstein hints, the context principle applies not merely to words in respect of sentences, but also to sentences in respect of languages (a point I shall revisit in §3).¹³

Note the structure of the dialectic that I have presented so far: the phenomenon of linguistic communication is taken as a given; the context principle, by contrast, I do not assume but have argued for, following Bentham and Wittgenstein, on the basis of the fact that a word comes together with grammatical properties as a matter of its essence. Note, too, the connection that has been drawn between sentences and truth-values: this connection, I have said, is what makes the declarative sentence fundamental to discourse. Since the advent of the Fregean context principle, sentences have enjoyed a privileged status in theorizing about language and world. Partly that is because sentences, as Michael Morris puts it, ‘simply strike us as units’ (2017, p. 10). But there is a reason for their so striking us; there is a reason for the theoretical centrality of the (declarative) sentence in metaphysical investigations. That reason is given by the fundamental connection between sentences and truth-values.

I have emphasized, following Wittgenstein, that names are fully as subject to the context principle as other kinds of word. This point was denied by Peter Geach:

An act of naming is of course not an assertion; it may be correct or incorrect, but not, strictly speaking, true or false; it does, however, ‘express a complete thought’, as grammarians say about sentences—it has a sort of independent sense. Naturally the sense of a name used this way is not independent of the language, or of the situation that makes such use appropriate; but it is independent of any verbal context expressed or understood—it is not like the sense of a fragmentary expression that answers a spoken or unspoken question. (1950, p. 462)

But how can an act of naming be correct or incorrect, if it is independent of any verbal context, *expressed or understood*? What would make a bare utterance or inscription of ‘Socrates’ correct or incorrect? It seems that there is no possibility of answering this

¹² Wittgenstein 1977, I, §§31, 49, tr. Anscombe.

¹³ See Davidson 1984, p. 22; Reck 1997, pp. 171–6; Pelletier 2012; Button 2020, p. 37 n. 11.

question unless we can supply a suitable sentential context, reconstructed from explicit or tacit verbal features of the occasion of utterance or inscription, into which the name ‘Socrates’, as uttered or inscribed on that occasion, slots. In fact Geach seems to concede this very point in the above passage when he allows that the sense of a name ‘is not independent . . . of the situation that makes [its] use appropriate’. Elsewhere, indeed, Geach agrees that some one-word utterances abbreviate sentences, but he claims that the following are independent uses of names: ‘nouns in the vocative case used as greetings; labels stuck on things, e.g. “poison” on a bottle or the name-labels sometimes worn at conferences; ejaculations like “Wolf!” and “Fire!” ’ (1962, p. 26). But it is surely obvious that these uses also abbreviate sentences—indeed abbreviate declarative sentences. Is there always a declarative sentence lurking behind such one-word utterances? Ian Rumfitt suggests not: ‘when Rodolfo exclaims “Mimi” just before the final curtain of *La Bohème*, he performs the complete speech act of invoking the late seamstress, even though he expresses nothing capable of assessment as true or as false’ (1995, p. 857). But, as I said above, non-declaratives can be reduced to, or explained in terms of, declaratives. In many cases this reduction goes via performatives, such as ‘I promise’, ‘I banish’, ‘I baptize’, or, as here, ‘I invoke’. Performative statements are special inasmuch as their truth is guaranteed by their being uttered (in the right context, by an authorized person, etc.); regular declaratives, by contrast, require co-operation from the extra-linguistic world for their truth. But that subtlety does not spoil the point, which is that Rodolfo *does* (perhaps indirectly) express something capable of assessment as true or false.

The conceptual priority of the sentence holds notwithstanding the likelihood that words came first in the historical development of language. For, even assuming the temporal priority of single words, those items of ‘protolanguage’ must have had an implicit syntax.¹⁴ This would be so, even when such words were uttered in isolation;¹⁵ *a fortiori* when they were uttered in combinations. Wolfram Hinzen has argued that complex noun phrases could have evolved independently and in the absence of sentences, and even that such independent noun phrases could have been employed in assertoric speech acts.¹⁶ I agree that that that could have happened—perhaps did happen. But the possibility that language evolved in that way does not show, as Hinzen supposes, that sentence structure and a correlative notion of truth-value are not essential to assertion, or that it is wrong to

¹⁴ Burling 2005, pp. 19, 43, 128–30.

¹⁵ Cf. Hurford 2012, pp. 100, 164–5, 348–62, 590, 596–7.

¹⁶ Hinzen 2007a, pp. 87–8; cf. 2007b.

accord declarative sentences a fundamental status in semantics: for, to put it compendiously and to anticipate somewhat, the contents of assertions, even if they are effected by the use of isolated noun phrases, are true or false *propositions*, propositions being constitutively the *referents* of declarative sentences. (I shall examine the purport of these terms of art in more detail in Chapters 3–5.) In fact, as I and others have argued, the meanings of complex noun phrases are *propositionally* structured.¹⁷ So the apparent syntactic independence of the complex noun phrase from the declarative sentence is superficial: semantically, the former depends on the latter.

James Hurford has suggested, in the manner of Geach, that interjections such as ‘Ouch!’ and ‘Damn!’ have no descriptive content, do not abbreviate sentences, and are mere relics of animal cries.¹⁸ But the plausible conjecture that such cries antedated the utterance of complete sentences in the evolution of language does not entail that we should exempt them from the context principle. On any occasion when such a cry is emitted, there is no difficulty in assigning to it descriptive content—content which is given by a complete sentence. Of course, it is often the case that those who emit such cries cannot *themselves* articulate that sentential content. So, for example, human infants and non-human animals cannot articulate that content; and we are here presuming that our ancestors could not do so either, at least if we go back to a sufficiently remote stage of human development, when (as we are hypothesizing) only isolated words of protolanguage were uttered. But these facts, assuming that they are facts, do not derogate from the conceptual priority of the sentence. The meanings of one-word utterances are and always have been conceptually dependent on the meanings of whole sentences: these sentences *we* can express, even if other beings, who—as we are forced to allow, at least in the most favourable cases—entertain those contents, cannot.¹⁹ Compare one-word commands: these stand alone, and in that sense are ‘syntactically unintegrated’, as Hurford puts it (2012, p. 221), just like one-word interjections; but one would not call them syntactically unintegrated in any deeper sense than that.²⁰ Such imperatives *presuppose* a sentential syntax. You might as well say that *any* stretch of stand-alone discourse was syntactically unintegrated; you could say that the whole of uttered human language, from beginning to end, was syntactically unintegrated. If you make the unit of discourse large enough, you will find that that unit is

¹⁷ Carstairs-McCarthy 1999; Gaskin 2008, pp. 394–405.

¹⁸ Hurford 2007, pp. 175, 236; 2012, pp. 205–6, 220–1.

¹⁹ See further here Gaskin 2006, ch. 4.

²⁰ See Gaskin 2018, pp. 302–3.

unintegrated. But it is no strike against the context principle to observe that, if you widen the domain far enough, you will sooner or later run out of context. The context principle is not committed to the indefinite extensibility of contextualization.

So words of protolanguage are, like all words, made for sentences: they are essentially syntactic entities. Admittedly it might be futile to attempt to attach determinate syntactic categories to these isolated words, if we could get hold of any.²¹ Still, such words must unpack into sentences; it must be possible—for hypothetical speakers, if not for the speakers themselves of protolanguage—to express the content of the individual words of protolanguage, as uttered on particular occasions, in complete sentences, that is, in complex linguistic structures that (if declarative) take a truth-value. Speakers of protolanguage do have access to sentential content—and thereby to linguistic meaning—but only by piggy-backing on the actual or possible use of complete sentences to express their meanings. This is a transcendentalist point: that is, the meaningfulness of one-word utterances, including not only utterances of recognized words such as ‘wolf’ and ‘fire’, but also the cries of infants, animals, and our own human ancestors before sentences were formed, depends on the *possibility* of the expression of their content in complete sentences.

Bentham tells us that words were formed from sentences by a process of analysis and abstraction: ‘Every man who speaks, speaks in propositions, . . . —terms taken by themselves are the work of abstraction, the produce of a refined analysis:—ages after ages must have elapsed before any such analysis was ever made’.²² In effect, as I have mentioned, Bentham is telling us that words are the product of a distinctively *theoretical* activity. Putting it in terms made familiar by Quine, we might say that words are a theoretical *posit*. I mean this quite literally: it is not that we have a theoretical conception of words *in addition to* an ordinary, non-theoretical conception; the conception of a word is through and through theoretical. We do, of course, talk about words in quotidian discourse, but what that shows is not that we have a non-theoretical conception of words, but rather that ‘ordinary’ chat about words is itself, if only inchoately, *theoretical*. This move—that is, the move of insisting on the theoretical status of semantically significant subsentential components—can be thought of as giving expression to an implication of the Fregean context principle as filtered through the earlier and the later Wittgenstein.²³ In the

²¹ Hurford 2012, p. 609.

²² Bentham 1843, vi, §1.

²³ See P. Sullivan 2000, pp. 78–9, 84–5; MacBride 2003, pp. 119–20; Bronzo 2017, p. 1350; A. Moore 2019, pp. 66–7.

Tractatus we are told that if everything in the symbolism behaves as if the signs had meaning (*Bedeutung*), then they *do* have meaning (§3.328). Note here in passing that Wittgenstein is hypothesizing a scenario in which everything in the symbolism *really does* behave as if the signs had meaning; and that is a fully objective matter. Whether signs behave as though they had meaning is not constituted by whether any particular speaker(s) *think* that they so behave. There is an appearance/reality distinction to be observed. But the crucial point is that the analysis of a sentence into smaller symbolic units is already a theoretical move: so for my purposes it is helpful to read Wittgenstein as saying that, in the metaphysically and conceptually initial position, where we are confronted with unanalysed sentences, if everything about a sentence behaves as though it had meaning, then it does have meaning. At which point the theoretical work of analysing sentences into smaller meaningful components—the positing of a syntactically pregnant lexicon—can begin.

What theoretical purpose is served by the positing of words? The answer to this question is familiar. Words are required to explain the productivity and systematicity of language use—the fact that we have the capacity to understand and form a potentially infinite range of novel sentences, on the basis of a relatively short training in a relatively small (and certainly finite) number of sentences; and the evident fact that our understanding of language is systematic and predictive, not *ad hoc* and piecemeal. The obvious way to explain these facts is to credit the speaker with a compositional and recursive understanding of sentences and their construction. The form of this explanation is abductive: we are aiming to *model* the phenomena; a *causal* account is not in question.²⁴ That takes us to the principle of compositionality, which is also found in Frege’s writings, though he did not consistently adhere to it.²⁵ By thinking of sentences as being composed of already familiar words—words abstracted from previously encountered sentences—combined in permitted ways by iterable operations, speakers are enabled to form, and understand, new sentences in a systematic way.²⁶ Commentators sometimes try to deny that language has the kind of systematicity that is embodied in the compositionality principle, on the grounds that it is fertile and productive in unpredictable ways.²⁷ But while

²⁴ See here Szabó 2000b, p. 75. But Szabó’s reservations about the cogency of the explanatory inference (pp. 76–80) seem to me unpersuasive.

²⁵ See Janssen 2012, p. 36 and *passim*. On other historical antecedents of the principle, see Hodges 2012, pp. 245–7. In general on Frege’s attitude to both context and compositionality principles, see Pelletier 2001.

²⁶ See Evans 1985, ch. 11; Lepore and Pelletier 2012, p. 413; Pelletier 2012, p. 167; Recanati 2012, pp. 175–6; Szabó 2012a, pp. 74–7.

²⁷ See, e.g., Sampson 2017.

particular uses of language are often unpredicted, it would be implausible to hold that they were in principle unpredictable, and indeed normally language use (even innovative use) follows well-worn tracks. Now if, as theorists, we are to explain not merely the formation but also the understanding of sentences, we must do more than simply posit words as abstractions from sentences: we must attach meanings to these words. We can then explain the meanings of complete sentences on the basis of the meanings of their component words and the way those words are put together. Starting from the idea of the meaningful sentence, as the fundamental ‘move in the language game’—as the fundamental unit of communication—we derive words in one dimension (horizontally, say), and in another dimension (vertically) their meanings, as theoretical posits required to explain the productivity and systematicity of language.²⁸

2 The principle of compositionality

The principle of compositionality states that, barring a finite number of exceptions, the meanings of complex expressions are determined by the meanings of their constituents, taken individually, and by the manner in which those constituents are combined syntactically.²⁹ The idea is that, in all but a (relatively) small number of exceptional cases, the meanings of complex expressions are determined in a functional way from the meanings of their (semantically significant) component parts; the exceptions are idioms, which must be learnt on a case-by-case basis. Despite the intuitive plausibility of the principle of compositionality, it has met with not inconsiderable resistance. The objections to it generally centre on one or another form of context sensitivity. So, for example, in connection with adjective + noun (AdjN) compounds, Ran Lahav writes:

What it is for a bird to count as red is not the same as what it is for other kinds of objects to count as red. For a bird to be red (in the normal case), it should have most of the surface of its body red, though not [necessarily] its beak, legs, eyes, and of course its inner organs. Furthermore, the red colour should be the bird’s natural colour, since we normally regard a bird as ‘really’ red even if it is painted white all over. A kitchen table, on the other hand, is red even if it is only painted red, and even if its ‘natural’ colour underneath the paint is, say, white. Moreover, for a table to be red only its upper surface needs to be red, . . . not necessarily its legs and its bottom surface. Similarly a red apple . . . needs to be red only on the outside, but a

²⁸ Cf. Gaskin 2008, p. 43; Sandu 2012, p. 269; Bloor 2018, p. 342.

²⁹ Pelletier 1994, p. 11; Partee 2004, p. 153; Dever 2006, p. 638; Szabó 2012a, p. 71.

red hat needs to be red only [on] its external upper surface, a red crystal is red both inside and outside, and a red watermelon is red only inside. (1989, p. 264)

And so on. We can omit most of the rest of the paragraph, which goes through the differences in what it takes to be a red book, a red newspaper, a red house, a red car, a red star, a red glaze, a red mist, a red powder, and a red pen, and proceed straight to the conclusion:

In short, what counts for one type of thing to be red is not what counts for another. Of course, there is a feature that is common to all the things which count (non-metaphorically) as red, namely, that some part of them, or some item related to them, must appear wholly and literally reddish. But that is only a very general necessary condition, and is far from being sufficient for a given object to count as red. (ibid.)

Actually, what Lahav says about the painted bird case might be contested in such a way as to support his general thesis. Charles Travis in effect does this:

Pia's Japanese maple is full of russet leaves. Believing that green is the colour of leaves, she paints them. Returning, she reports, 'That's better. The leaves are green now.' She speaks truth. A botanist friend then phones, seeking green leaves for a study of green-leaf chemistry. 'The leaves (on my tree) are green,' Pia says. 'You can have those.' But now Pia speaks falsehood. (2017, p. 129)

Noun + noun (NN) compounds equally exhibit the semantic creativity of AdjN compounds. In a study testing speakers' intuitions on the intelligibility of novel NN compounds, Pamela Downing found a notable absence of constraints: 'Even compounds paraphrasable in terms of negative relationships, or temporary and fortuitous relationships, are acceptable in the appropriate context' (1977, p. 841).

What should we make of these apparent refutations of the principle of compositionality? The obvious strategy for one who wishes to defend the principle against these purported counterexamples, I think, is to start from the concession of Lahav's quoted above: what all red and only red things have in common is that they are, in some way or other, *red*—as opposed to being blue, or square, or divisible by three without remainder, or loved by Jane, or investigated by economists.³⁰ (Of course they might be some of these things as well, but if so that will not be in virtue of their being correctly called 'red'.) Lahav dismisses this feature of red objects as a mere necessary condition, but it is a

³⁰ See McDowell 2007, p. 364 n. 4.

necessary *and sufficient* condition. Moreover, to say that an object is red is to supply substantial empirical information, which in a real situation might be (for example) life-saving. It is just that, as Lahav says, ‘what counts for one type of thing to be red is not what counts for another’. Indeed we might add that what counts for a single type of thing to be red *in one context* might not be what counts for it to be red *in another context*. That is the moral that Travis is aiming at with his example of Pia and her painted leaves. The idea is that the characterized object need not change; it may be rather that what it takes to be an *F* varies from context to context.³¹ The examples mentioned by Lahav and Travis imply that there are both linguistic and pragmatic aspects to context sensitivity: one might then suggest that the difference between a pink hat and a pink watermelon is semantic, while that between naturally green leaves and ones that have been painted green is pragmatic.³²

However that may be, these points surely do not undermine the principle of compositionality. Rather, they require us to be more specific about how that principle works. The principle predicts that compounding an adjective, say ‘red’, with a noun, say ‘hat’, produces a composite, ‘red hat’, whose meaning is a function of the meanings of ‘red’ and ‘hat’. And that is just what we find: an object to which the compound description ‘red hat’ truly applies is correctly said to be *a hat*, whatever else it is, and it is correctly said to be *red*, whatever other properties it has. Moreover, in attributive-adjective constructions like ‘red hat’, understanders of that phrase will know that for something to be a red hat, it must (normally: see below) be red *for* a hat, and they must also know what it is for hats to count as red in standard situations. To understand ‘red *F*’ in its everyday range of application, one needs to know what it is to be red *for* a car, *for* a book, *for* an apple, and so on. In the case of ordinary applications of ordinary adjectives like ‘red’, the information that the understander needs to command can be thought of as front-loaded into the lexicon. This does not entail that the demands placed on the understander will be ‘encyclopaedic’, as John Collins suggests (2011, p. 24). That is because red objects will in practice fall into a relatively small number of similarity classes, and within each class the learner has only one task to perform: having learnt what it is for a car to be red, you do not need to start all over again in order to find out what it is for a bus to be red, for a train to be red, for an aeroplane to be red, and so on. So the task of acquiring competence in the use of an attributive adjective like ‘red’—given that, normally, the head noun will determine what it

³¹ See here García-Carpintero 2008, pp. 144–5.

³² Cf. Pietroski 2005, pp. 260–4; Collins 2011, p. 25.

is for each type of red object to count as red—is quite tractable. Outlier cases can then be dealt with—that is, learned and produced—on an *ad hoc* basis.

It seems to me that the strategy of distinguishing between central uses of, say, ‘red’, thought of as grouped into a manageable number of similarity classes, is superior to an alternative response one might be tempted to deploy against the Lahav/Travis argument. Their argument aims to undermine compositionality by providing counterexamples—cases where the truth-value of a sentence, and therefore also its meaning, changes from one context to another even though its grammar and the meanings of its parts remain constant. My response is to distinguish between central uses of a word that impose a small, and certainly finite, number of tasks on the language learner, on the one hand, and *ad hoc* extensions of those central uses, that have to be—but also easily can be—learned on an individual basis as and when they arise. But an alternative response would defend the principle of compositionality by (ironically enough) taking contextualism further than the principle’s opponents: if the meanings not only of whole sentences but also of individual words are radically contextual, then the purported counterexamples can be disarmed, for the variation in meaning of whole sentences across contexts can now be put down to the variation in the meanings of their component words, and so no breach of compositionality is in the offing.³³ In effect, and to concretize the proposal, the response to Travis’s counterexample envisaged by this alternative strategy would be to distinguish between two homonymous words corresponding to ‘green’, one of which (*green_a*) applies to objects that appear a certain way, or reflect light of a particular wavelength, in ordinary viewing conditions, and the other of which (*green_b*) means something like ‘*green_a* in its natural condition’. Then Pia’s sentence ‘The leaves on my tree are green’ is ambiguous as between ‘The leaves on my tree are *green_a*’, which is what she meant when she uttered the sentence to herself, and is true, and ‘The leaves on my tree are *green_b*’, which is what the biologist took her to mean in their telephone conversation, and is false. Since ‘*green_a*’ and ‘*green_b*’ have distinct meanings, Travis’s example entails no breach of compositionality: the differently truth-valued and so non-synonymous sentences that Pia comes out with in her two utterances are analysed as being composed of, in part, distinct bits with distinct meanings. However, though this response might be suitable in particular cases, it cannot work as a general strategy, because the price it pays for salvaging compositionality is too

³³ So, e.g., Szabó 2010, pp. 265–71. Cf. Kennedy and McNally 2010.

high: it threatens to distend the lexicon beyond learnability. Since the very reason why we wanted compositionality in the first place was to model the capacity of finite creatures with finite resources to master a system of infinite scope, the proposed alternative solution is too radical.

Context sensitivity actually goes further than we have so far indicated, in two respects. My response to Lahav was framed in terms of attributive uses of ‘red’ (red *for* a car, red *for* a book, etc.). But there are contexts in which such adjectives function predicatively. For example, although ‘a big mouse’ usually means a mouse that is big *for* a mouse, there are contexts in which it may mean a mouse that is big absolutely, or at any rate big by some *other* standard of comparison than the usual size of mice. If the earth is invaded by vicious elephant-sized mice, and you run towards me shouting ‘Watch out! There’s a big mouse coming’, I do not expect to be confronted by a plump dormouse as opposed to a starveling field mouse.³⁴ In addition, context sensitivity can be an even more *ad hoc* matter than I have hinted so far. In ordinary contexts ‘red apple’ will mean an apple that is (naturally) red-skinned over most of its exterior surface, all the component words here being taken literally, but in non-standard contexts—imagine a children’s party game—the phrase could have any number of weird and wonderful applications, including to things that were not literally red, and/or not literally apples, at all. But I do not think that these phenomena raise a difficulty for compositionality. As far as the first point goes, we may suppose that it is front-loaded into the lexicon that a red *F* will standardly be something that is (literally) red *for* an *F*: this is, in effect, a kind of indexicality.³⁵ It is plausible that the lexicon does indeed carry information about such default values: ‘red hat’ does not *simply* mean ‘hat that is red in some way or other’;³⁶ a hat which has a red label on the inside but is otherwise black is not a red hat. Once we have a default value in place, a single general clause can specify that this value may, given appropriate contextual signals, be cancelled, so that the adjective reverts to predicative status. In this latter case the comparison class (if relevant) will be determined in some other, contextually salient way, as in the mouse scenario above: there a big mouse was a mouse that was big by comparison with animals in general. The second type of context sensitivity I have mentioned, in which, for example, ‘red apple’ is used in a game to mean, say, a white ping-

³⁴ See here Szabó 2000b, p. 105; Recanati 2012a, pp. 179–84.

³⁵ Cf. Reimer 2002, pp. 188–90, deploying Kaplan 1989. Collins’s criticism of Reimer at 2011, p. 24, strikes me as off-target.

³⁶ Pace Sainsbury 2012, p. 246.

pong ball marked with a black ‘R’, is handled by the inclusion of a further clause that allows lexical items to receive one-off novel definitions and applications. These novel applications have to be introduced, and learnt, on an *ad hoc* basis; but though the number of contexts in which such constructions can be introduced is potentially unlimited, an understander only needs to cope with them as and when they arise, a task which is eminently practicable on the basis of a merely finite training in a language with a finite lexicon and grammar.

Nothing stops particular functions like the one collected by the attributive adjective ‘red’ from being put together into one large function, specifying meanings for the compound ‘red’ + $\ulcorner F \urcorner$ context by context.³⁷ And, as far as the demands of the principle of compositionality go, such bumper functions may be infinitistic: that is, it would not derogate from the scope and validity of the principle if they were infinitistic. For first, as I have said, it is plausible that the contexts which such functions appeal to will fall into a small number of more or less natural similarity classes. But, secondly, even if this is not the case, we can be sure that *ordinary* uses of, say, ‘red apple’ *will* be restricted to a finite, and small, number of possible meanings, notwithstanding context sensitivity. The potentially indefinite range of distinct contexts across which the meaning of ‘red apple’ can vary is not a problem for a finitistic theory of meaning, because prior understanding need only track ordinary uses of the phrase; as we have said, non-standard uses are then generated, and understood, on an *ad hoc* basis. So, for example, to understand ‘red apple’ it is plausible that I need to know: (i) what it is for something literally to be apple in ordinary contexts of use; (ii) what it is for an apple to be literally red in ordinary contexts of use; (iii) that in non-standard contexts of use the adjective may function predicatively rather than attributively; and that (iv) temporary definitions of what it is for something to count as red (in general), or as an apple (in general), or as (specifically) a red apple may be extended on an *ad hoc* basis. Of these clauses, (i)–(iii) are clearly finitistic, and although the number of possible non-standard contexts alluded to in (iv) is infinite, these are added to my understanding of ‘red apple’ on a case-by-case basis (and then forgotten when they have served their temporary purpose), so that, as far as the demands placed on understanding are concerned, (iv) turns out to be finitistic as well.³⁸

An important fact about attributive AdjN and NN constructions is this. Unlike analytical compounds in which both elements have equal weight and can be presented in

³⁷ Partee 2004, p. 161.

³⁸ Contrast Recanati’s account of these contexts at 2012a, pp. 190–1.

either order, such as ‘love–hate relationship’, ‘good cop–bad cop routine’, ‘Bolzano–Weierstrass Theorem’,³⁹ an attributive construction consists of a head noun together with a modifier (adjectival or nominal): the head noun determines what type of thing is in question, and the modifier then characterizes that thing in more detail. This is a case of what is called endocentricity.⁴⁰ Further, the structure is rule-governed: for example, in modern English, the modifier generally comes first and the head noun second. So a shoe-tree is a kind of tree, not a kind of shoe, a motorway is a kind of way, not a kind of motor, a weather forecast is a kind of forecast, not a kind of weather, a screwdriver is a kind of driver, not a kind of screw, and so on; in Downing’s study ‘bullet hole’ and ‘pea princess’ generated a large range of alternative interpretations (1977, p. 820), but a bullet hole was always a kind of hole, not a kind of bullet, and a pea-princess a kind of princess, not a kind of pea.

Now endocentricity is an important feature of syntactic compositionality, but it has one very significant exception, which we have in effect already registered, namely the sentence. The sentence is an exception because although it inherits *some* of its semantically significant properties, such as its tense, from its parts, and in particular from the main component VP, it does not inherit from its components *all* of its semantically significant properties: in particular, the sentence’s capacity to take a truth-value is a crucial feature of it that it does not inherit *at all*. As far as that goes, the sentence exhibits *exocentricity*: its ability to be true or false, and so to make a distinctively sentential move in the language game, supervenes on the fully formed structure, and is not projected up the syntactic tree from any subordinate node. This fact has caused syntacticians some discomfort,⁴¹ and they have tried to get rid of the sentence’s exceptional status, at least symbolically, by dropping the old ‘S’ label from their inventory of abbreviations—for that label, attached to the root node of the syntactic tree, appeared to pop up out of nowhere—and replacing it with an explicitly endocentric alternative, such as TP (‘tensed phrase’), IP (‘inflectional phrase’), or CP (‘complement phrase’), dominating T, I, or C nodes. The idea in each case was that the rebranded sentence should inherit, and be seen to inherit, its tensed (inflected, complementary) status as a TP (IP, CP) from a component which possessed that property *de iure*, so to speak, and which then projected it up the tree to the root. (Remember that syntactic trees are upside down.) But this whole strategy is misconceived. The old

³⁹ Downing 1977, p. 824 n. 9.

⁴⁰ On which see, e.g., Haegeman 1995, pp. 1–51.

⁴¹ See here Collins 2007, pp. 811–12.

approach, in which the label ‘S’ attached to the root node emerged suddenly and out of non-sentential in puts—there was no SP dominating an S node—was right, because the sentence *does* emerge suddenly from bits that lack its most vital semantic property, taking a truth-value. The root of the tree *is* thus distinct from its subordinate nodes; its key semantic property is *unprecedented*. The old, unprojected ‘S’ label made that point clear; the projected replacements obscure it.

Opponents of the principle of compositionality typically underestimate how much information can be, and is, built into the lexical entries of words whose combination determines the meanings of complex linguistic expressions. (Front-loading of syntactic information into the lexicon—‘radical lexicalism’—is a particular feature of categorial grammar.)⁴² So, for example, Ray Jackendoff (2012, pp. 68–9) suggests that compositionality is refuted by considerations such as the following. In the sentence ‘Joe jumped until the bell rang’ we understand Joe to have jumped *repeatedly* until the bell rang; in the sentence ‘Joe jumped when the bell rang’ we understand Joe to have jumped just once, at the moment when the bell rang; in the sentence ‘Joe slept until the bell rang’ we think of Joe as sleeping not once or repeatedly but continuously. Where, Jackendoff asks, does the idea of repetition, present in the first sentence but not in the second or third come from? We cannot understand it to come from the meaning of the verb ‘jump’, he implies, because then it would have to be present in the second sentence too: the repetition seems to supervene on the first sentence as a whole, rather than being derived from any of its parts. But then it seems not to be an aspect of the *way the sentence is composed*, either, since in the relevant sense the second and third sentences are composed in just the same way as the first. Hence we have an apparent counterexample to compositionality. To this it can be replied that the meaning of a verb like ‘jump’ (though not of a verb like ‘sleep’), when combined with the meaning of a conjunction like ‘until’ (though not ‘when’) is *such that* our first sentence means that Joe jumped repeatedly. (Jumping just once can count as a degenerate case of jumping repeatedly: if, for instance, Joe is on the surface of the moon, he might only manage one jump before the bell rings.) There is nothing to prevent the meaning of a word from being *disjunctive*. Jackendoff, considering something like this option, tells us that it ‘violates Fregean compositionality, because the extra meaning doesn’t come from the words’ (ibid., p. 69); but quite evidently it can, and does, come from the words. So long as lexical entries are understood to be sufficiently rich, cases like

⁴² See Dowty 2007, pp. 41–2, 52.

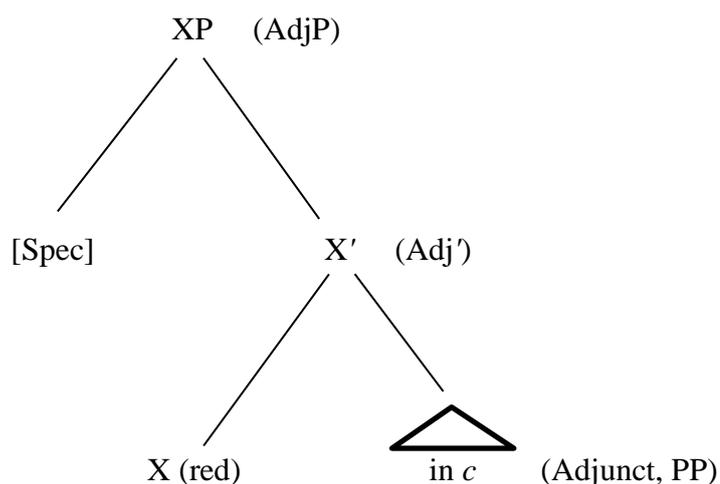
the one considered present no difficulty to compositionality. As we have said, much of the information that needs to be conveyed will be of a general nature. An efficient theory generalizes so far as possible, and in the present case that means that it will generalize over similarity classes (the meanings of ‘red car’, ‘red bus’, etc.). So the technique of front-loading as much context-sensitivity into the lexicon as possible, though it might initially appear to trivialize compositionality, does not do so: for the meaning of a word can be thought of as an abstraction from the similarity classes it falls into, and since these will be small, and certainly finite, in number, with exceptions and extensions handled on an *ad hoc* basis, the possibility of language mastery by finite beings is not thereby undermined. In effect, what we are doing here is adducing the context principle to support compositionality: I shall say more about this point in the next section.⁴³

From a technical point of view, the move I have been canvassing involves applying the method of ‘reverse Schönfinkelization’ (or ‘uncurrying’). By parametrization, any implicit parameters governing the correct application of a function can be brought within the pale of a new function as explicit arguments; so if, in a context c , we have a monadic function f such that $f(x) = y$, then there exists another, dyadic function, g , which is such that, absolutely, $g(x, c) = y$. In syntactic terms, the context variable figures as an *adjunct* to the head word. It is convenient to express this point in an X-bar-theoretic framework, which employs a specifier-head-complement configuration in its analysis of syntactic structures, and in which structurally significant non-atomic constituents find representation as intermediate projections. (This framework was actually implicit in my discussion above of sentence exocentricity.)⁴⁴ What we find when we do this is that an adjective like ‘red’ has the following implicit syntactic structure:⁴⁵

⁴³ On the general question of triviality, see further the discussions in Zdrozny 1994; Kazmi and Pelletier 1998; Westerståhl 1998; Pelletier 2012, pp. 160–2; Sandu 2012, pp. 266–8.

⁴⁴ For elucidation and motivation of X-bar theory see any good textbook on syntax, e.g., Carnie 2013, Part 2.

⁴⁵ The treatment I here present borrows (aspects of) a proposal made by Stanley and others concerning the general syntactic representation of context dependence at LF (see n. 47). See the essays collected in Stanley 2007 (including ones co-authored with Szabó and King); Pagin 2012, pp. 312–19. Stanley’s original suggestion (*ibid.*, chs. 2 and 3) that covert variables should be taken to ‘cohabit’ nodes with other expressions does not work (Collins 2007, pp. 828–31), and he subsequently moved to the proposal that they be treated as adjuncts (Stanley 2007, p. 222 n. 15; cf. pp. 248–9), which seems to me the right strategy. I am not persuaded by the objections to this strategy that I have seen, for example those raised by Recanati 2002 (supported by Collins 2017, pp. 161–7), Lepore 2004, pp. 50–9, and Collins 2007, pp. 831–4, but it would take us too far afield (and is anyway unnecessary for my purposes) to pursue that matter here.



The prepositional phrase (PP), including the contextual variable *c*, is unarticulated at the level of surface syntax. The triangle indicates that the adjunct entry is complex, but for present purposes that does not matter. The important point here is that we can do justice to the context sensitivity of a word like ‘red’ by recognizing the tacit presence of more structure than is articulated on the surface. For that we need a suitable level of syntactic representation, say LF.

As it is used in the literature (especially in the approaches to syntactic theory known as GB and Minimalism), the label ‘LF’, originally an abbreviation of ‘logical form’, is a technical term for the level of analysis in the generation of sentences from lexicon and grammar that is posterior to any movement or copying, and at which all and only semantically relevant syntactic and lexical material is displayed. At least, that is the core idea: sometimes more is built into the definition of LF depending on an individual syntactician’s approach—one might, for example, incorporate a Davidson-style analysis of action verbs at LF that discerns quantification over events⁴⁶—but this general statement will suffice for my purposes. Put more simply: LF is that level of lexical and syntactic construction at which sentential entities are fed directly into semantic interpretation.⁴⁷ One might surmise from the simpler statement that ‘semantic form’ would be a more perspicuous label for the relevant status than ‘LF’.⁴⁸ (Of course, syntacticians insist that we should not worry about the etymology of ‘LF’, but treat it as a mere label for the syntactic status I have attached to it.) That would be because, as one might suppose, there is no *a*

⁴⁶ Cf. Ostertag 2019, pp. 1492–3.

⁴⁷ See further May 1985, ch. 1; Hornstein 1995, chs. 1 and 4; Keenan and Stabler 2003, ch. 1; Szabó 2000b, pp. 96–7; 2012a, pp. 65–6; Jacobson 2012.

⁴⁸ Szabó 2012b, p. 105.

priori guarantee that semantic form, as discerned by some adequate theory of meaning for a language, will align with what philosophers might wish to regard as logical form.⁴⁹ After all, philosophers, from Russell onwards, have typically wanted the idea of logical form to model *inference patterns*, whereas linguists have deployed the notion of LF in *syntax theory*: why should satisfying these desiderata agree in result? Actually I am sceptical about whether these tasks—modelling inference, analysing syntax—are, in the end, genuinely distinct, for reasons that I shall discuss later on (§22); but we do not need to broach this question now. Returning to the depicted structure, conceived as a structure at LF, as I have defined that, we see that the PP ‘in *c*’ is adjoined to the relevant word (‘red’). With that structure in place, it will then be possible to generalize over contexts, in the process collecting similarity classes, as well as include a clause admitting of *ad hoc* exceptions and extensions to normal contexts. The inclusion of the contextual parameter in the lexical representation of ‘red’ allows us to treat the sensitivity of that word to context—whether that context be linguistic or not, and whether the realization of the adjunct variable be standard and predictable or exceptional and *ad hoc*—as flowing from an application of the principle of compositionality.

3 Forwards and reverse compositionality

Do the context principle and the principle of compositionality clash? The context principle prioritizes the sentence—it says that words are abstractions from sentences—while the principle of compositionality prioritizes the word (morpheme)—it says that sentences are constructed out of words. Which comes first, sentence or word? The two principles appear to deliver opposite answers to this question.⁵⁰ In his first book on Frege, Michael Dummett resolved the difficulty by saying that in ‘the order of explanation’ sentences are prior to words, whereas in ‘the order of recognition’ words are prior to sentences: that is, the *theoretical account* of what a word is proceeds in terms of its role in the sentence, whereas a *speaker’s understanding* of sentences is based on an understanding of the words they are constructed out of (and their syntax).⁵¹ This seems to me the right way to reconcile the two principles, and preferable to an alternative strategy that Dummett later preferred.⁵² (According to the revised story, context and compositionality principles apply to different

⁴⁹ See, e.g., Collins 2017, pp. 155–6.

⁵⁰ See here Nimtz 2002, p. 222; Hurford 2012, pp. 622–4.

⁵¹ Dummett 1981, pp. 4, 192–6. See also Davidson 1990, p. 300; Horwich 1998b, p. 165; Gaskin 2008, p. 257; Linnebo 2018, pp. 121–2.

⁵² Dummett 1991b, pp. 202–3.

grades of sentential complexity, the former principle to simple sentences, the latter to complex; but it is clear that both principles apply to both kinds of sentence.)⁵³ Varying the terminology, we might say that on this approach sentences are prior to words in a *conceptual* sense, whereas words are prior to sentences in a *practical* sense. When the two principles are relativized to these different points of view, the appearance of conflict vanishes. An equivalent way of resolving the apparent tension between the context and compositionality principles is to distinguish, following Manuel García-Carpintero (2010a, p. 284), between two levels at which they can be said to apply, namely a general and a particular level. The point will then be that the context principle applies at the general level, whereas the principle of compositionality does so at the particular level. Here is how the contrast operates. Words and their meanings are *abstracted* from sentences *in general*: meaning is use, and the meaning of a word is a matter of how it is used in *all* (or at any rate a large number) of the sentences in which it occurs; so a given word is abstracted from its occurrence not just in one sentence but in many. The *construction* of sentences, by contrast, proceeds case by case: a given *particular* sentence is constructed from its constituent words and their mode of combination. So we might say, of any particular word, that it is prior to each particular sentence in which it meaningfully occurs, but posterior to the set of *all* (or a significant number) of sentences in which it occurs. This is another way of saying that the word is *practically* prior, but *conceptually* posterior, to the sentence.⁵⁴

So far so good, but at this point a difficulty arises. It is standard in contemporary discussions of compositionality to distinguish between a principle of *forwards* compositionality and a principle of *reverse* compositionality. The principle of forwards compositionality (FC) states that the meaning of a complex linguistic expression is a function of the meanings of its component words and their (the words') mode of combination; that is what I have hitherto meant by 'the principle of compositionality'.⁵⁵ But we can also formulate a principle of reverse compositionality (RC), which states that the meanings of a given component word in a complex linguistic expression is a function of the meaning of that complex expression and its syntax.⁵⁶ If both principles are true, meaning is determined in both directions, forwards from the simple to the complex,

⁵³ Parallel remarks apply, *mutatis mutandis*, to Linnebo's suggestion (2018, p. 112) that the reconciliation might proceed in terms of the type-token distinction.

⁵⁴ Cf. Bronzo 2014, pp. 23–4, who presents a reconciliation along these lines as an alternative to Dummett's earlier resolution; but it seems to me substantially the same.

⁵⁵ Cf. Szabó 2000b, p. 11.

⁵⁶ In general on both principles, see Szabó 2000a; Fodor and Lepore 2002, p. 59; Pagin and Pelletier 2007, pp. 41, 54–7; Pagin and Westerståhl 2010.

backwards from the complex to the simple. Should we accept RC? The difficulty I had in mind at the beginning of this paragraph is the following: it might seem as though, if we did accept RC, that would spoil the resolution I offered above of the apparent clash between the principle of compositionality—specifically, as we can now say, between FC—and the context principle. The context principle has it that the meaning of a word is determined by a (sufficient) *plurality* of sentences in which it occurs; but according to RC word meaning is determined contextually by each *individual* sentence in which it occurs. However, if the meaning of a word is not merely prior to the meaning of a given individual sentence (FC), and posterior to the meanings of sentences in general (context principle), but also posterior to the meaning of a given individual sentence (RC), we appear to be facing again a version of the conflict we thought we had resolved. Of course, one way out of the difficulty would be just to reject RC; and a number of commentators have urged this solution on us by way of advancing what they take to be counterexamples to RC.⁵⁷ Unfortunately, however, these purported counterexamples are unconvincing.

It has been suggested, for example, that someone can understand compound expressions like ‘nuclear submarine’ without knowing the meaning of, say, ‘nuclear’, and that someone might understand the use of ‘build’ in a sentential context without knowing that ‘build’ is a telic verb: that is to say, that one might understand ‘Mary was building a house’ without knowing that building (a house) is a process that has a completion, and that one does not count as *having built* a house *during* the building process but *only when* that end point has been reached; so the idea is that one might understand the sentence ‘Mary was building a house’ without knowing the meaning of ‘build’.⁵⁸ I think it must be conceded to the opponents of RC that, in some circumstances, a speaker who did not know the meaning of ‘nuclear’ might nevertheless acquire at least a *primitive* understanding of ‘nuclear submarine’, and that a speaker who did not know that ‘build’ was a telic verb (it is not necessary, of course, to be familiar with that particular, or indeed any, theoretical description of the phenomenon) nevertheless had a *basic* understanding of the sentence ‘Mary was building a house’. But it is hardly plausible that you could have an *adequate* understanding of ‘nuclear submarine’ or ‘Mary was building a house’ without knowledge of the meaning of ‘nuclear’ or of the meaning and telic status of ‘build’.

The reason for this has to do with an extended, even more holistic, version of the context principle than we have adduced so far, though it was briefly mentioned in §1 in

⁵⁷ See, e.g., Robbins 2005; Patterson 2005a; Johnson 2006.

⁵⁸ Patterson 2005a, p. 336; Johnson 2006, p. 42.

connection with the passage of Wittgenstein quoted there. The simple version of the principle states that a word has meaning only in the context of a sentence; the extended version adds that a sentence has meaning only in the context of a language. This entails that understanding a sentence is not an isolated confrontation between a speaker and that individual sentence, all else in abeyance, but rather something that is only possible if the speaker is able to fit that sentence into a network of sentences with which it is connected in various ways—inferentially, evidentially, and otherwise. So a speaker who could not use and understand the word ‘nuclear’ in other contexts than just ‘nuclear submarine’, or who could not use and understand ‘build’ in its telic sense in a range of other sentences than just ‘Mary was building a house’, would not count as understanding those particular complex expressions. The opponents of RC are simply underestimating how much context and background is presupposed to the understanding of a given complex linguistic expression; pleasingly enough, they make the reverse mistake to that made by opponents of FC, who, as we saw, underestimated the complexity of meaning that could be front-loaded into the lexicon. Setting aside RC is therefore not an option. That brings us back to the problem we had before: is a clash between context and compositionality principles now inevitable?

No. At the level of the *individual word* and the *individual sentence*, we may simply accept that a relation of meaning determination runs in both directions—from word to sentence and from sentence to word—without admitting that this symmetry involves any sort of clash of principle.⁵⁹ It is indeed the case that, at the level of the individual linguistic complex, the meaning of the whole is a function of the meanings of its semantically significant parts and their mode of combination, and the meaning of each semantically significant part is a function of the meaning of the whole and its syntax. The meanings of the parts and the meaning of the whole are *simultaneous*: there is no question of priority on either side. But this simultaneity leaves the status of the context principle undiminished, for that principle says that, however things may be with the *individual* sentence—even if there is no priority between part and whole here—nevertheless *in general* sentences are prior to words. So we continue to maintain that the context principle provides a *conceptual* sense in which the sentence is prior to the word, while the principles of compositionality (now taking FC and RC together) provide a *practical* sense in which the word is prior to the sentence. The context principle tells us that words are made *from and for* sentences: that is, they are abstractions from sentences, and they derive their function and point from

⁵⁹ Cf. Wiggins 1980a, p. 159 n. 13; Gaskin 2010, pp. 306–9; Szabó 2010, pp. 256–7; 2012a, p. 78; Bliss 2014, p. 248.

their role in sentences. The two principles of compositionality tell us that a given individual sentence not only *can* be understood on the basis of a knowledge of the meanings of its component words and the way they are put together (FC), but can *only* be understood on that basis (RC).

Douglas Patterson (2005a, p. 329) points out that assuming RC on its own, without the assistance of FC, would undermine the learnability of language because it would problematize the infinitistic generation of well-formed sentences from a finite base: given the creativity of language, there is no limit to the number of complex expressions any reasonably resourceful language can generate, so that if all we had were RC, in supposing that speakers operated on linguistic complexes we would be supposing that they had an infinitistically grounded understanding of their language. Or, in other words: adopting a bare RC, unaccompanied by FC, would force us to construe linguistic complexes as semantical primitives, and that would be inconsistent with the learnability of language. It follows that RC is *only* acceptable in the presence of FC. Moreover, it seems clear (*contra* Patterson, *ibid.*) that, just as RC requires FC—in the sense that RC would not plausibly model linguistic understanding without the assistance of FC—so also FC requires RC. The reasons for these opposite requirements are basically the same. RC needs FC because, as we have agreed, language would not otherwise be learnable: speakers would not be able to compute the meanings of an infinity of complex expressions compositionally, and would face the impossible demand of having to learn these meanings individually. FC needs RC because, if a language *is* compositional in the forwards sense, then it is a constraint on understanding complex expressions that one actually *discern* their complexity: speakers who understand a complex expression *E* must know how it has been assembled. If they lacked this knowledge, they would be unable, in accordance with the principle of *forwards* compositionality, to understand a sufficiency of *related* complex expressions to count as having understood *E* in the first place. It may help to spell out this latter point.

Suppose that *E* is composed of semantically significant parts *A*, *B*, and *C*. If speakers could understand *E* without (i) breaking it down into its parts *A*, *B*, and *C*, and (ii) knowing the meanings of those parts, then they could not understand *other* complex expressions that made use of some or all of these parts, and they could not understand how these other expressions were related—inferentially, evidentially, perhaps in other ways also—to *E*. But then it would not be plausible to say that they understood *E* in the first place. Consider again the ‘nuclear submarine’ example: anyone who is to understand this phrase must not only know what ‘nuclear’ means in *this* context, but also what it means in

other combinations. A speaker who failed to grasp these connections would have a merely idiomatic understanding of ‘nuclear submarine’, like someone’s understanding of ‘hoist with his own petard’ if all that that person knows is the metaphorical meaning of the phrase (‘confounded by one’s own devices’, *vel sim.*), and does not know what the phrase literally means. An idiomatic phrase simply forms an unstructured block of meaning: it is, to that extent, a semantic atom. However, ‘nuclear submarine’ is *not* semantically atomic, so that anyone who treats it as such does not really understand it. Patterson argues (2005a, p. 340) that since expressions like ‘microwave’ are not idioms, they cannot be misunderstood to be such; but they indeed can, and that is what is going on in the case of someone who uses this word more or less correctly but does not hear ‘microwave’ as being, at some level, composed of semantically significant bits. This person has a merely idiomatic understanding of ‘microwave’, and that means that he does not properly understand it, something that would emerge around the edges of common uses of the phrase (hence his usage is only ‘more or less’ correct, as I said above). For example, if we presented him with a machine that, to all appearances, looked and behaved and was operated exactly like a standard microwave oven, but in fact employed a quite different technology, and if we then remarked that this gadget was ‘not actually a microwave’, he would be mystified.

So FC needs RC just as much as RC needs FC. We can impress this fact more strongly on ourselves by imagining a scenario involving a metalinguistic permutation on the meaning specifications of object-language sentences.⁶⁰ (This type of scenario will occupy us more extensively in Chapter 3.) Suppose we start with a theory of meaning for a language that obeys FC. If we tack a suitable permutation onto the deliverances of this theory, the resulting theory will still obey FC: the relevant forwards function is now a composition of the original function with the adopted permutation, taken in that order. But, in the absence of RC, the language for which the resulting theory is correct will not be learnable, because, without RC, there will be no way in which speakers can compute the meanings of *new* complex expressions they encounter: depending on the permutation, these new expressions could mean just about anything. Of course this was *already* true of the original theory, before the permutation was applied: the permutation makes no difference to that; appeal to it only renders the point a bit more vivid. In either the original or the developed scenario, if we do not have RC the meanings of complex expressions must be learned on an individual basis, which sets merely finite beings an impossible task, given

⁶⁰ Cf. Szabó 2000a, p. 487; 2000b, pp. 19–25; 2012a, pp. 67–8.

that the language generates indefinitely many such complexes. But if RC is added in to the mix as a further constraint the learning problem is solved. The permutation in itself is, as I have indicated, neutral on the issue of learnability: it neither helps nor hinders that, but simply complicates the story. So, in our envisaged case, suppose we have an overall forwards compositional function, F , for a language L , and suppose that F is composed of the original compositional function f followed by a permutation g . Then, adding RC as a constraint will enforce the existence of an overall reverse compositional function, F^{-1} , which will be a composition of the inverse permutation, g^{-1} , and of the inverse of original function, f^{-1} , in that order. Given the existence of F then, so long as F^{-1} also exists (some purely technical manoeuvres may be necessary to ensure this), it follows that we have both FC and RC, and that the language will be in principle learnable. The fact that F and F^{-1} are themselves composite functions makes no difference to that; it is not something that the object-language speaker needs to know. In general—a point I shall explore further in Chapter 3—permutations (or other theoretical operations) in the metalanguage on the deliverances of a theory of meaning for an object-language make no difference to object-language speakers, who (truistically enough) do not register any such metalinguistic manoeuvres (and so these manoeuvres have no sceptical implications for object-language speakers' grasp of the meanings of their expressions).

Patterson also thinks (*ibid.*, p. 339) that, as far as understanding new complex expressions goes, it would suffice for speakers to dispose of a merely *syntactic* version of RC. The idea is that it would suffice if speakers were simply able to *parse* the novel expressions—work out what their semantically significant syntactic bits were—without having at this stage to *know the meanings* of those bits (as demanded by RC proper). Then, at a (temporally or logically) subsequent stage, those speakers could apply FC to compute the meanings of the complex expressions that they had previously (syntactically) decomposed. But this suggestion fails, for the reason that it is just not possible, with any plausibility, to distinguish a temporal or logical state at which a 'merely syntactic' version of RC is brought to bear on a sentence. We are liable here to be tricked by our knowledge of the language into supposing that there is less to parsing than there really is. In fact, breaking a complex unit down into its semantically significant syntactic parts requires that we already know the semantic significance of those parts—at least to a considerable extent. We can 'understand' Lewis Carroll's 'Jabberwocky', for example, but only because most of the words are familiar to us, and because the understood frame allows us to calculate the intended syntactic shape of the nonsense words—'gimble' must be an intransitive verb,

‘slithy’ an adjective, and so on—onto which we can then project suitable toy meanings. So there is no temporally or logically prior process of parsing sentences ‘merely syntactically’, before we have assigned meanings to their semantically significant parts. RC is necessarily *both* a syntactic *and* a semantic principle, and the syntactic and semantic operations are simultaneous. The same is true of FC.

In saying that anyone who understands a composite expression must be able to break it down into its *semantically significant* parts and must know the meanings of these parts, I stress the point that the relevant parts of a complex expression must *have* semantic significance. We thereby rule out as possible counterexamples to RC such proper names as ‘Lou Gehrig’s Disease’. Patterson asks whether a person who does not know who Lou Gehrig was also does not know what ‘Lou Gehrig’s Disease’ means, ‘even if he knows his uncle has it, what its symptoms are, and so on’ (*ibid.*, p. 341). To which the answer is: of course you can understand what ‘Lou Gehrig’s Disease’ means without knowing who Lou Gehrig was, just as you can understand ‘Dartmouth’ without knowing what the Dart is. These cognitive achievements are possible because ‘Lou Gehrig’ and ‘Dart’ are not—or are not any longer—*semantically significant* parts of the names ‘Lou Gehrig’s Disease’ and ‘Dartmouth’. These latter names may have started life as definite descriptions meaning ‘the disease first diagnosed in/by Lou Gehrig’ and ‘the town at the mouth of the river Dart’, or similar, but over time they have become fossilized into (genuine) proper names, that is, into names that do not abbreviate or otherwise go proxy for definite descriptions. And the crucial difference between genuine proper names and definite descriptions is that the former have no semantically significant parts, whereas the latter do. (I shall explore this point, and the terminology I am here using, in Chapter 4.) An indication of this is that, if we discovered that Lou Gehrig had nothing to do with the disease named after him, we would not judge that previous uses of the name ‘Lou Gehrig’s disease’ had been meaningless, or even faulty. We might of course decide to rename the disease for reasons of historical accuracy; but nothing in *semantics* would force us to do that. By the time that a genuine proper name is up and running, its descriptive aetiology, if it has one, is irrelevant to its successful use as a name.⁶¹

We must, therefore, embrace both FC and RC. Indeed, before this distinction was formulated, references to the ‘principle of compositionality’ often combined both ideas.⁶² Strictly, as we have noted, FC only says that the meaning of the complex is a function of,

⁶¹ See Mill, *System of Logic* I, 2, §5 (2006, vol. 1, p. 33).

⁶² Cf. Szabó 2012a, p. 77; Keller and Keller 2013, p. 314.

or supervenes on, the meanings of the parts: it does not just so far entail that the understander must grasp that complex meaning *as being actually composed* of those bits. For anything that FC says to the contrary, the meaning of a complex expression might be simple, and graspable as such. But usually what people have had in mind when they have appealed to the idea of compositionality is more than a sheer relation of supervenience of complex on simple: they have intended that the understander be able to decompose complex expressions into their semantically significant parts, cognize the parts' meanings, and recognize the meaning of the complex as being actually *made up* of those meanings—not merely as *functionally dependent* on them. For, though FC suffices to model a speaker's ability to *compute* an indefinite number of new, complex meanings on the basis of a finite stock of old, relatively simple meanings, it cannot model the ability to *comprehend* endless new, complex meanings. But for a language to be genuinely learnable, and for creative language use to have its intended range of application, the understander needs to be credited with both capacities—to compose simples into compounds and to decompose compounds into simples. And to make that requirement is, in effect, to insist that the theorist deploy both FC and RC. Or, alternatively put, to do full justice to traditional intuitions about the role of compositionality in linguistic understanding, we need to combine a functional model of sentence meaning with a part–whole model: the meaning of the sentence is not only a *function* of the meanings of its semantically significant parts and the way they are put together, but also a *whole* composed of those meanings as *its* parts.⁶³

⁶³ On combining these models see further §23 below.