3

Compositionality, Direct Compositionality, and the syntax/semantics interface

As noted in the preceding chapter, the program of semantics often takes as its point of departure the adage "To know the meaning of a (declarative) sentence is to know what it would take to make it true." There is a second foundational adage: "The meaning of the whole is computed in some predictable way from the meaning of its parts." This is also often known as *Frege's principle* or *the principle of compositionality*. Taken in its broadest sense, this is fairly weak: obviously speakers of a language know the basic units (words, or more accurately, morphemes) and know some system to combine these into larger meaningful expressions. Surely at least most of language has to be compositional, or there would be no way to systematically combine units into larger meaningful expressions which have never been uttered (or heard) before.[1] In other words, we can see the grammar of a natural language as a specification of the set of well-formed expressions of the language, combined with rules which map each well-formed expression

[1] But there is a question about whether everything is compositional. Idioms are a classic case of items that at least appear to be non-compositional.

into a meaning (that is, some object built out of our model-theoretic toolbox, which so far consists only of truth values, worlds, and times).

But while it is generally accepted that the meanings of larger expressions are computed in some predictable way from the meanings of smaller ones, what is not so clear is just what that predictable way is and how complex it is. A very simple and elegant view on this is as follows. The syntax consists of statements that predict the existence of well-formed expressions on the basis of the existence of other well-formed expressions. At the base of this is the lexicon; the list of basic units (for our purposes, words). The semantics works in tandem with the syntax: each syntactic rule which predicts the existence of some well-formed expression (as output) is paired with a semantic rule which gives the meaning of the output expression in terms of the meaning(s) of the input expressions. This is what we mean by **Direct Compositionality**, and is the point of view that will mainly be pursued in this book.

3.1. Building a fragment: First steps

Every expression of a language, including the basic expressions (the words), can be seen as a triple of <[sound], syntactic category, [[meaning]]>. Borrowing notation from phonology, we enclose the sound part in square brackets (although we simply use English orthography rather than phonetic transcription), and the meaning part is enclosed in double square brackets [[. . .]]. A rule is thus something which takes one or more triples as input and yields a triple as output. A reader who is used to thinking of the syntax as consisting not of a number of very specific "rules" (such as phrase structure rules) but a set of much more general rules (or "principles") such as the X-bar schemata need not worry: Chapter 6 will give the rules in far more general form.[2] Since we will often be adopting rules which are temporary and will later be revised, simplified, or generalized, any rule not part of the

[2] A certain amount of modern linguistics has eschewed the term "rule" on the grounds that the grammar hopefully does not consist of a list of a large number of "rules" rather than having the apparatus stated in much more general format. Nonetheless, most agree that the grammar contains statements that predict the set of well-formed expressions and assign them a meaning, and here the term "rule" is being used in the most general sense here to simply mean these statements. We will

final fragment will be notated as TR (followed by a number). (Rules which do remain in the ultimate fragment are simply labeled with R.)

Assume that the set of syntactic categories includes S which we use to mean a sentence.[3] We continue here to assume that all Ss have as their truth value 1 or 0 (i.e., none are undefined in some worlds). Moreover, for now let us go back to ignoring the world (and time) parameters, and give a purely extensional semantics. Then we begin with the following:

TR-1. If α is an expression of the form <[α], S, [[α]]> and β is an expression of the form <[β], S, [[β]]>, then there is an expression γ of the form <[α-and-β], S, 1 if and only if [[α]] = 1 and [[β]] = 1>. (Hereafter we use the standard *iff* to abbreviate "if and only if.")

There are a variety of other ways this rule could be written. For example, we could write the syntax as a *(context-free) phrase structure rule* and pair this with a semantic part. This notation was used in much of the literature within Generalized Phrase Structure Grammar (GPSG) (Gazdar, Klein, Pullum, and Sag 1985):

TR-1′. $S_1 \rightarrow S_2$ and S_3; $[[S_1]] = 1$ iff $[[S_2]] = 1$ and $[[S_3]] = 1$.

A phrase structure rule by itself, then, is a rule which gives the phonological and syntactic part of a rule such as TR-1 but not the semantics. When supplemented with the semantics—as in TR-1′—this is just TR-1 in a different format. Note also that TR-1′ uses subscripts in the syntactic part of the rule, but these are not actually meant as a part of the syntactic category.[4] These are needed just to state the semantics; this is avoided in the notational system in TR-1.

indeed be moving towards a view of grammar which contains just a few rather general statements to this effect, but there is no reason not to call these rules.

[3] In some current theories of syntax (Minimalism and some of its predecessors) this category is instead called TP (*Tense Phrase*) and in slightly earlier versions of that theory it is called IP (*Inflectional Phrase*). As with the label NP, we are using the terminology standard in many other theories and in many other related disciplines; readers familiar with the terms TP or IP instead can make the appropriate translation.

[4] A *context-free phrase structure rule* specifies the well-formedness of an expression of some grammatical category X in terms of concatenations of strings of other expressions in the language. These other expressions are described either listed on a case-by-case basis (like *and* above) or are described by their grammatical category. The set of grammatical categories is finite, and it does not include things like S_1 etc.

To forestall one objection: one should not be misled by the fact that the word "and" is used in the semantic part to give a meaning for English sentences containing *and*. The word "and" in the first part of the triple in TR-1 is the English word [ænd] where English is the *object language*—the language whose syntax and semantics is being modeled. But we need a language in which to describe the grammar, and we have chosen English for that purpose as well. So here English is also being used as the *metalanguage*. The confusion that arises by using English as the metalanguage to model English as the object language would go away if we were using English to, e.g., model the grammar of Swahili—or if we wrote the semantics for English using some sort of purely symbolic notation (whose meaning had been agreed on). There are other ways to alleviate the potential confusion; we could just as well have written the semantic part of the rule by means of what is known as a truth table. This is illustrated below where we rewrite the semantic part of TR-1 by listing the possible combinations of values for each of the α and β followed by the value that this gives for the string [α-and-β]:

(1)

[[α]]	[[β]]	[[α-and-β]]
1	1	1
1	0	0
0	1	0
0	0	0

From now on we use the format in TR-1 as our official notation, although at times will use phrase structure rule format or other notational systems as convenient. Incidentally, TR-1 introduces *and* directly as part of the rule rather than treating it as an item in the lexicon with a grammatical category. This is known as treating it *syncategorematically*. This decision to treat *and* not as a lexical item but rather as only something introduced by a rule is not essential (and is almost certainly incorrect), and will be revised later.

The rule for *or* is analogous (here we will highlight the lack of need to use the English word "or" in stating the semantics).

(Of course we could posit such a category, as long as there are a finite number of them, but that is not the intent in the rule above.) In the rule above, the idea is that the syntactic rule is just S → S and S. The reason that the subscripts are there in the rule is because the semantics needs to refer to them. Note that this is just an artifact of this particular notational system; the rule in TR-1 doesn't use any subscripts in the syntax.

Comp. by: PG4144 Stage : Revises1 ChapterID: 0002108609 Date:16/5/14 Time:06:42:31
Filepath://ppdys1122/BgPr/OUP_CAP/IN/Process/0002108609.3d46

TR-2. If α is an expression of the form <[α], S, [[α]]> and β is an expression of the form <[β], S, [[β]]>, then there is an expression γ of the form <[α-or-β], S, 0 iff [[α]] = 0 and [[β]] = 0 >.

3.2. Implicatures vs truth conditions

Before continuing with the fragment, a digression is in order about the meaning of *or*. (More extensive discussion is found in Chapter 10.4.) According to the semantics in TR-2, the sentence in (2) is true if Carol skydives and doesn't bungee jump, if she bungee jumps and doesn't skydive, and also if she does both.

(2) Carol will go sky dive on her vacation, or she will go bungee jumping on her vacation.

This is what is known as *inclusive or*: both sentences can be true. But is that result correct? One might at first think not—there surely is a strong temptation to read (1) as saying she will do one or the other but not both. On the basis of this it is tempting to think that English *or* actually has as its meaning what is known as *exclusive or*—either conjunct could be true but not both. If that were correct then of course TR-2 needs revision.

But we can show that English *or* is actually inclusive. Although we often take sentences like (2) to convey that both events will not happen, this is not a part of its actual truth conditions. Note first that if I say (2) to you and it turns out that Carol goes completely wild and does both, you can hardly accuse me of lying. And it is easy to construct cases where there is no exclusive *or* suggestion and where things would not make sense if the meaning of *or* were "one or the other but not both." So, take a linguistics department which is very strict about enforcing prerequisites. There is a course called Phonetics and Phonology II (Phon/Phon II). The course brochure contains the language in (3):

(3) Students may not enroll in any course unless they have had the prerequisite(s). You will be allowed to enroll in Phon/Phon II if and only if you have taken Phonology I or you have taken Phonetics I.

Suppose you are a student in love with all of phonetics and phonology, so of course you have taken both Phonetics I and Phonology I. You now go to enroll in Phon/Phon II. Surely you would be shocked if the Registrar told you could not enroll in this course. Anyone reading this course brochure would certainly come to the conclusion that someone who had both courses is eligible for the more advanced one. But if *or* meant "one or the other but not both" then that is exactly what (3) would be saying.[5]

Of course there is still one other possibility: that *or* is ambiguous and that (2) uses "exclusive *or*" and (3) uses "inclusive *or*." But if it were ambiguous, then each of these sentences should be too—it should be possible to read (3) in such a way that the Registrar would be entitled to block you from enrolling. But (3) has no such understanding; the Registrar simply could not legitimately come back to you and say "Oh, sorry you misunderstood. You've taken both Phonetics I and Phonology I; so you do not fit the eligibility requirement for enrollment in Phon/Phon II."[6]

[5] This point is amusingly illustrated by the following true story. An Ivy League institution which will remain nameless has a teaching center as a resource for graduate students to hone their teaching skills, and gives a certificate to any student who completes a program that includes a series of workshops and other activities (including a "micro-teaching" session). In February of 2011, students in the program received the following e-mail:

Attendance in Workshop #5 is MANDATORY if you belong in any of the following categories:

a. You haven't completed workshop #5, or
b. You haven't completed the Micro-Teaching requirement, or
c. You haven't completed workshop #5 AND you haven't completed the Micro-Teaching requirement.

Of course anyone reading this would be completely amused and wonder why on earth (c) was included here. Surely, no one who had both not completed the workshop and not completed the Micro-Teaching requirement would really think that they were exempt. The fact that we react to this e-mail as being a silly instance of bureaucratic language confirms the point.

[6] The argument is not quite complete, because one might still think that *or* is ambiguous, but that the reason the exclusive interpretation does not emerge is simply because it would be strange in this situation. Our knowledge of how prerequisites work makes this a highly unlikely interpretation in this scenario. But we can tweak the scenario to make the "exclusive" interpretation one that *might* make sense, but is still absent. Suppose that the linguistics department is also very unhappy with students taking courses that overlap too much with what they already know. And suppose that Phon/Phon II overlaps in some ways with Phonetics I and in some ways with Phonology II. In other words, the department is concerned not only about underqualified students, but also overqualified ones. Yet even in this situation, you would be genuinely outraged if the Registrar blocked your enrolment in the course.

Thus while an example like (2) often carries with it the suggestion that only one of the two conjuncts is true, this is just a suggestion and is what is known as an *implicature* rather than being part of the truth conditions. Sometimes the exclusive implicature is quite natural, as in (2), and sometimes it disappears, as in (3). In fact, when embedded under *if* (as in (3)) the exclusive suggestion associated with *or* systematically disappears. Ultimately one would hope for a single, stable meaning, where other principles can predict when the exclusive *or* implicature arises and when it doesn't. Indeed, by assigning it the inclusive semantics in TR-2 combined with a theory of language use (pragmatics), this can be done. (Many readers will be familiar with this under the rubric of *scalar implicature*.) This is the subject of Chapters 10.4 and 16.6.

3.3. Folding in worlds (and times)

TR-1 and TR-2 are obviously inadequate in that they are given purely *extensionally*: they both ignore the fact that the value of a sentence is not 1 or 0 but a function of type <w,t> (or really <s,t>). For convenience, this book generally gives the rules in purely extensional fashion, but this is always just for expository simplification. We can easily write these rules in a more adequate fashion. Thus let us rewrite TR-1 as what we will call INT-TR-1 (INT here for intensional) as follows (actually this includes just the world argument and not the time; the full intensional version would replace "w" with "s" (a world–time pair):

INT-TR-1. If α is an expression of the form $<[\alpha], S, [[\alpha]]>$ and β is an expression of the form $<[\beta], S, [[\beta]]>$, then there is an expression γ of the form $<[\alpha\text{-and-}\beta], S,$ for any world w, $[[\gamma]](w) = 1$ iff $[[\alpha]](w) = 1$ and $[[\beta]](w) = 1>$.

Very often the world argument is given as a superscript, so one can rewrite the semantic part of this rule as given below:

$[[\gamma]]^w = 1$ iff $[[\alpha]]^w = 1$ and $[[\beta]]^w = 1$

Indeed, if the intent were to keep out both the underqualified and the overqualified student, the brochure would have to be revised to say "if you have taken Phonology I or you have taken Phonetics I and you have not taken both." So even in a scenario designed to bring out the "exclusive *or*" interpretation, it is not there.

One will commonly find this notation in the literature (and this will be used from time to time in this book). Notice that $[[\alpha]]^w$ means exactly the same thing as $[[\alpha]](w)$ –the superscript notation has an implicit "for all worlds w" at the beginning of the whole rule. One should keep in mind that nothing (other than taste and expository ease) hinges on the choice of notation. The grammar is a system that proves expressions well-formed and assigns each expression a model-theoretic interpretation; the notation that one chooses to express the model-theoretic object being assigned is of no theoretical consequence.

3.1. Rewrite the intensional version of TR-2 (i.e., INT-TR-2) using both of the notations for dealing with the world argument shown above.

3.2. For any function f that characterizes a set, let us use the notation f_S to mean the set characterized by f. Now recall that we can think of the value of a sentence as a function of type <w,t> (a function from worlds to truth values) or—equivalently—as a set of worlds; the set mapped to 1 by the function. Thinking in set terms, what does *and* do? (That is, what operation on sets does it correspond to?) What about *or?* Using the notation suggested just above, write out the semantics for INT-TR-1 in set terms. Do the same for INT-TR-2.

3.4. Negation: A first pass

Readers familiar with first-order logic (or simple propositional logic) will recognize that the syntax and semantics of English sketched so far looks a bit like the syntax of simple logical languages in which two propositions p and q may be connected with ∧ ("and") and v ("or"). A third operator that one might be used to seeing from elementary logic is a negation operator which is prefixed in front of a proposition. (Thus, in the development of an elementary propositional logic, there is a rule saying that if p is a well-formed formula then so is ~p. The truth value of ~p is the reverse of that of p.)

If we try to model English directly using the tools from first-order logic in this way, we immediately reach a challenge for the Direct Compositional program. For a sentence like (4b) is the negation of (4a): if the first is true the second is false and vice versa:

(4) a. Olympia Snowe voted for Bill 1838 (the only time it came up for a vote).
b. Olympia Snowe didn't vote for Bill 1838 (the only time it came up for a vote).

Yet the syntax of English is completely different from the syntax of propositional logic. *Not* is not prefixed at the beginning of a sentence. It is not even prefixed at the beginning of a verb phrase. It generally shows up as a clitic on an auxiliary (as in *didn't* above, where it is a suffix attached to *did*.) Semantically, though, it seems reasonable to think of it as negating the entire sentence in which it sits.

There are two choices for resolving this apparent mismatch between the syntax and the semantics. One is to reject the Direct Compositional hypothesis. A common solution here is to imagine that actual English syntax is mapped into another level (called Logical Form, or LF) and that the grammar proceeds containing a rule or series of rules that map (4) into something like (5) (or perhaps something even more abstract), and that the compositional semantics interprets (5):

(5) not [Olympia Snowe voted for Bill 1838 (the only time it came up for a vote)]

Then *not* here can be given the interpretation exactly like the symbol ~ is in logic. The second possibility is to conclude that it is a mistake to try to copy logic in giving a meaning for English *n't*—its meaning might have something to do with "~" but perhaps it is a more complex packaging of negation. This is the strategy that we will ultimately aim for; it allows for a much simpler conception of the grammar.

Temporarily, though, we take a third tack: which is to postpone incorporating into our fragment the natural way to do negation in English, and instead pretend that English has a single word *it-is-not-the-case-that* which can be prefixed to a sentence to negate it. Hence we enrich our fragment with TR-3 and its intensional counterpart INT-TR-3:

TR-3. If α is an expression of the form <[α], S, [[α]]> then there is an expression β of the form <[it-is-not-the-case-that-α], S, 1 iff [[α]] = 0>.

INT-TR-3. If α is an expression of the form $<[\alpha], \text{S}, [[\alpha]]>$ then there is an expression β of the form $<[\text{it-is-not-the-case-that-}\alpha], \text{S}, [[\beta]]$ is a function from worlds to truth values such that for any w, $[[\beta]](\text{w}) = 1$ iff $[[\alpha]](\text{w}) = 0>$.

(or, using the shorthand, the semantic part would read $[[\beta]]^{\text{w}} = 1$ iff $[[\alpha]]^{\text{w}} = 0$).

For truth in advertising, while the final fragment will be closer to actual English, it will still not be a complete account of English *n't*, simply because space precludes an account of the English auxiliary system. But TR-3 will be improved upon.

3.3. As in Exercise 3.2, rewrite this in set terms. What is the operation on sets that corresponds to negation?