

Excerpts from Portner, Paul. 2005. What is meaning? Fundamentals of formal semantics. Blackwell.

### 1.3 We Should Think of the Meaning of Sentences in Terms of Truth-Conditions

So far we've thought in a bit of detail about the meanings of some nouns like *dog* and *Confucius*. What about other types of language? The traditional next move in building up a comprehensive semantic theory is to think about the meaning of complete sentences. Following this tradition, and assuming that meanings are part of language- and mind-external reality, we will now ask what sort of thing a sentence-meaning is.

As English speakers, we know the meaning of the sentence:

*The circle is inside the square.*

With this knowledge, we can display a certain ability. If I show you the picture on the left in diagram 2, you can tell me that the sentence is true, and if I show you the one on the right, you can tell me it's false.

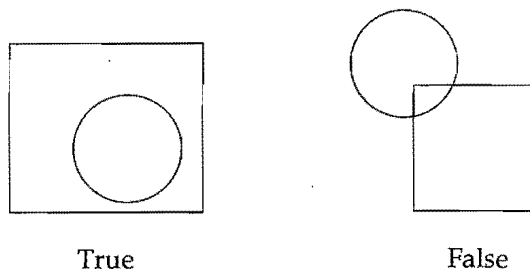


Diagram 2

More generally, provided with a range of scenarios, you can divide them into two classes. Calling these the "true set" and the "false set," you can draw a circle around the true set, as in diagram 3.

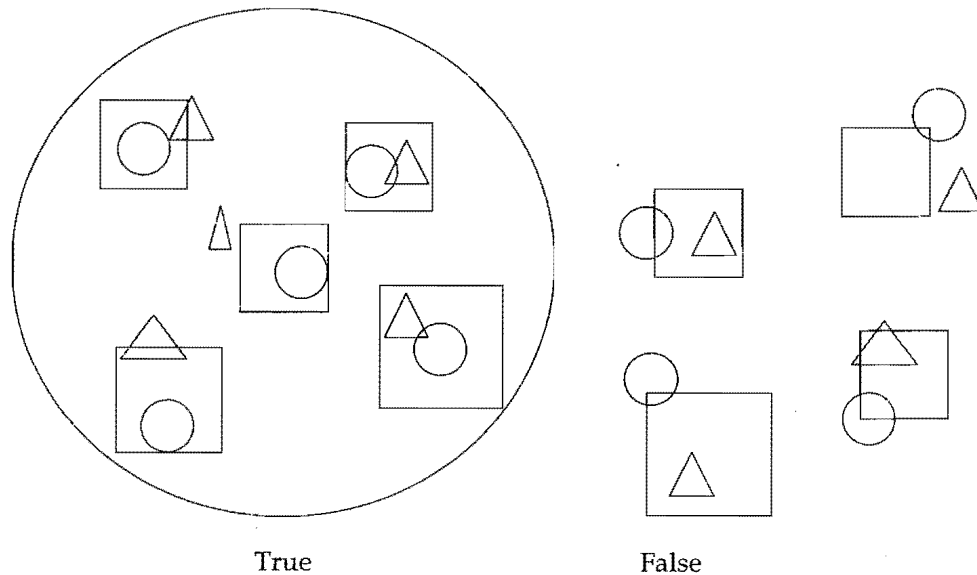


Diagram 3

One very important way of thinking about meaning is to take these kinds of abilities as crucial clues as to the nature of meaning. The knowledge of meaning involves (at least) the knowledge of the conditions under which a sentence is true, and those under which it's false. So let's begin our semantic investigation by focusing on this particular aspect of meaning as if it is all there is to the semantics of sentences. It's worth seeing where that gets us.<sup>11</sup> A theory which says that all there is to the meaning of a sentence is its truth-conditions is a *truth-conditional theory*. This might seem kind of odd, but I'll spend the rest of this chapter giving a number of reasons why this odd idea has a lot to recommend it. Perhaps it's even right – many formal semanticists think it is! But whether or not it's ultimately right, we'll see in this book that we can use it to understand many aspects of language better than we did before.

One common misunderstanding of truth-conditional semantics should be dealt with right away. Knowing the meaning of a sentence amounts to knowing its *truth-conditions*. It has nothing to do with knowing whether it is in fact true or false (what semanticists call its *truth-value*). You can know the meaning of a sentence without knowing whether it's true or false, or even having prospects for ever finding out. For example, you know

what it would take for the sentence *The third closest star to earth has six planets, one of which is inhabited by intelligent creatures* to be true, but you'll probably never find out if it's actually true or not. Yet, what you know – just the truth-conditions, not the truth-value – seems sufficient to say you know what it means.

Next, some terminology: there is a family of theories of semantics which we can count as just one theory for the purpose of this book. These theories go by names like *truth-conditional semantics*, *formal semantics*, *model-theoretic semantics*, *possible worlds semantics*, and *situation semantics*. These theories can be combined to some extent, so that one can practice model-theoretic possible worlds semantics, for example. From the perspective of the professional semanticist, there are important differences among these theories, but for the purposes of this book, I will treat them all as one, since they share the same central intuition about the nature of sentence-meaning. Since "formal semantics" is the most general term among these, this book is best described as an introduction to formal semantics. My discussion will freely borrow from the terminology of all of them, as it is convenient for making clear what I'm trying to make clear, but for the most part my discussion will be given from the perspective of the most popular flavor of formal semantics, the one known as possible worlds semantics. By and large, the ideas about language which I'll be presenting are compatible with any of the other flavors as well.

The little scenarios represented in diagrams 2 and 3 are called, in the technical terminology of formal semantics, *possible worlds* or *possible situations* (just "worlds" or "situations" for short). In diagram 3, I've represented nine different possible worlds or situations. There are infinitely many other worlds or situations which I didn't draw, but you can intuitively tell how they would be added into a more complete diagram. The terms "world" and "situation" are typically used in somewhat different ways, with "situation" suggesting a very incomplete scenario, a part of the universe bounded in space and/or time. For example, everything which is enclosed within the room as you read this sentence is a situation, as is everything enclosed by the boundaries of the District of Columbia on March 29, 2002, at 10:15 a.m. The pictures in diagram 3 are representations of very little situations. The term "world" is used when people have in mind a complete way in which the world could be. A possible world is a possible history of the universe – the kind of thing that often comes up in science fiction. If the pictures in the diagrams are thought of as representing worlds, then each one represents only very impoverished worlds only inhabited by a few shapes, and with no change over time (or perhaps no time at all). Or you can think of them as being merely partial depictions

of some more ordinary worlds, with lots of details left out, so that each actually corresponds to many different worlds (a different one on each way of filling in the details).

The notion of possible world or situation may seem metaphysical and so disconnected from reality as to be a pointless place to start a scientific investigation of anything. But really, it's quite simple and familiar. We think about possible worlds all the time. Suppose we are investigating a murder, and have two hypotheses about who did it. We first imagine the scenario in which suspect no. 1 is guilty, thinking through what would have happened in that case, and then looking for evidence of whether those things did in fact happen. Then we imagine the scenario in which suspect no. 2 is guilty, and go through the same process. Roughly speaking, each imagined scenario can be thought of as a possible world. This way of thinking about what might be or might have been is quite common, and so the story about the murder investigation makes clear why the notion of alternative possible worlds is not really all that unfamiliar or odd.

As a philosophical aside, it's worth pointing out that it's not entirely accurate to say that the imagined scenarios are simply examples of possible worlds. When we imagine a scenario, we don't bother to be specific about each and every detail. But possible worlds are specific in every detail. The real world is a possible world after all, and it is quite specific in details that we never even consider, like how many leaves are on a particular tree in a particular forest on a particular day. For this reason, a scenario like our murder scenario is better thought of as a set of possible worlds. For example, we may imagine suspect no. 1 committing the crime at *about* 10 a.m., but what we imagine isn't specific as to whether it's *exactly* 9:59, 10:00, or 10:01. Suppose that these differences in the time of the murder are not significant to our investigation. What we imagine is compatible with a possible world where the murder happened at 9:59, and with one where it happened at 10:00, and with one where it happened at 10:01 (as well as all the ones with times in between). We don't care about the differences among these possibilities, and our imagination doesn't distinguish them. So, a scenario is more like a set of possible worlds such that the differences among the worlds is unimportant to whoever is imagining the scenario. End of aside.

On to the next piece of terminology. The meaning of a sentence is called a *proposition*. We say that a sentence *expresses* or *denotes* a proposition.<sup>12</sup> According to the truth-conditional view of meaning, the proposition expressed by a sentence amounts simply to its truth-conditions. The proposition denoted by *The circle is inside the square* is the one indicated

in diagram 3 (relative to my laziness in just drawing nine scenarios, when in reality there are of course many more true and false possibilities). Within the terminology of possible worlds, a proposition is a set of possible worlds, in diagram 3 the set indicated by the big circle on the left. Thinking of propositions as sets of possible worlds captures the idea that the meaning of a sentence is its truth-conditions, since knowledge of what it takes to make a sentence true is exactly what you need in order to decide if a given possible world is in the "true set." Informally, we can think of the meaning of a sentence as parallel to the meaning of a common noun, in the following way: The noun *dog* describes certain things (the dogs) and not others, and so we can explain the meaning of *dog* by saying it denotes the set of dogs. Likewise, a sentence describes certain possible worlds (those in which it's true) and not others, and so we can explain its meaning by saying it denotes the set of possible worlds in which it's true.

### 1.3.1 Three reasons why truth-conditions are a central part of meaning

As promised above, our next task is to bring out some reasons why this weird truth-conditional view of meaning is worth pursuing.

#### 1.3.1.1 Reason 1: The semantics of logical words

Thinking of meaning as truth-conditions lets us give a pretty good semantics for logical words like *and*, *or*, and *not*. These are called "logical words" because of the important role they play in determining what patterns of reasoning are valid, the traditional concern of logic. If you know that a sentence of the form *p and q* is true, you also know that *p* is true – this is a logical fact. For example, the following sentence is made up of two sentences joined by *and* (so it is of the form *p and q*). If we know that it is true, we can conclude that the circle is inside the square (i.e., that *p* is true).

(4) [*The circle is inside the square*] and [*the circle is dark*].

Suppose we indicate the truth-conditional meaning of each of the two component sentences in a possible-worlds diagram like the kind we've seen above. In diagram 4, the proposition expressed by each component sentence is indicated by a dotted circle. Then, it's easy to see what the truth-conditional meaning of the whole sentence is, in terms of the meanings

of the two sub-sentences: the overlap between the two dotted circles. (The overlap between two sets is called the *intersection* of those two sets.)

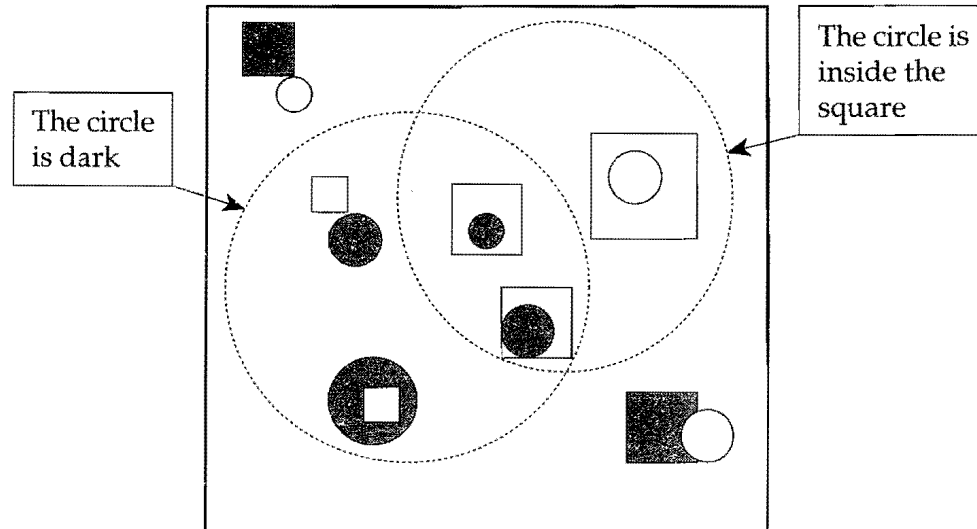


Diagram 4

From this diagram, you can easily see why (4) implies that the circle is inside the square, and that the circle is dark. In general, a sentence of the form *p and q*, where *p* and *q* are any sentences, describes the worlds in the intersection of the worlds described by *p* and the worlds described by *q*. In our example, *p* = *The circle is inside the square* and *q* = *The circle is dark*.

Give a semantics for sentences of the form *p or q* and *It is not the case that p*. What you come up with should look a lot like what we've just seen for sentences of the form *p and q*. As you think about this exercise, work with particular sentences for *p* and *q*, e.g. *The circle is inside the square and the circle is dark*. But make sure that your answer would work for any sentences you could fill in for *p* and *q*, not just these particular ones.

This exercise has an answer, no. 1, in the appendix.

At this point, most books on formal semantics would begin to get more formal, indicating the meanings of *and*, *or*, and *not* with symbols.

Right now we're focusing on the meaning of sentences. As an aside, though, notice that thinking of the meaning of *and* as intersection works for phrases other than sentences. Consider the following:

Mary is [a student] and [a baseball fan].

Let's assume that *a student* describes the set of students, and that *a baseball fan* describes the set of baseball fans. Then, what does *a student and a baseball fan* describe? Answer: the set of things which are both, that is, the intersection of the two sets. Saying that Mary is a student and a baseball fan therefore says that she is in the intersection of the set of students and the set of baseball fans, which implies that she is a student and that she is a baseball fan. Since this is exactly what the sentence should imply, it looks good for our idea that *and* means intersection.

There are problems, though. We can use *and* to combine words which don't intuitively describe sets, as in [Mary] and [John] *bought a dog*. If it is combining things which aren't sets, *and* can't mean intersection (since intersection is a way of combining sets and nothing else). Some semanticists think that *and* is ambiguous, sometimes meaning intersection and sometimes meaning something else more appropriate for *Mary and John*; others think that, contrary to naïve appearances, *Mary* and *John* do describe sets; still others think that *and* is, contrary to naïve appearances, not joining together *Mary* and *John*. We'll drop this issue for the time being, but you might want to think a bit about how each of these ideas would work before your thinking is corrupted by more education.

Maybe *Mary* and *John* do denote sets in the following way: *Mary* denotes the set containing only Mary, i.e., {Mary}, and *John* denotes the set containing only John, {John}. Explain why this does not solve the problem posed by our desire to understand the *and* in *Mary and John* as intersection.

### 1.3.1.2 Reason 2: Definitions of intuitive semantic relationships

Thinking of meaning as truth-conditions lets us define some basic semantic concepts: synonymy, contrariety, entailment, contradiction, tautology.

Two sentences are *synonymous* if they have the same meaning. Intuitively, (5) and (6) are synonymous:

(5) The square is bigger than the circle.

(6) The circle is smaller than the square.

Suppose we draw a box to indicate the set of all possible worlds (diagram 5). Call this box "W". (Semanticists often use "w" to stand for a

single possible world, and "W" to stand for the set of all possible worlds, as here.) Then we draw a circle within the box W to indicate those possible worlds in which (5) is true. Call this "5." Then draw another circle to indicate those in which (6) is true, called "6." Circles 5 and 6 are the same, showing that the truth-conditional view of meaning can capture the sense in which (5) and (6) are synonymous.

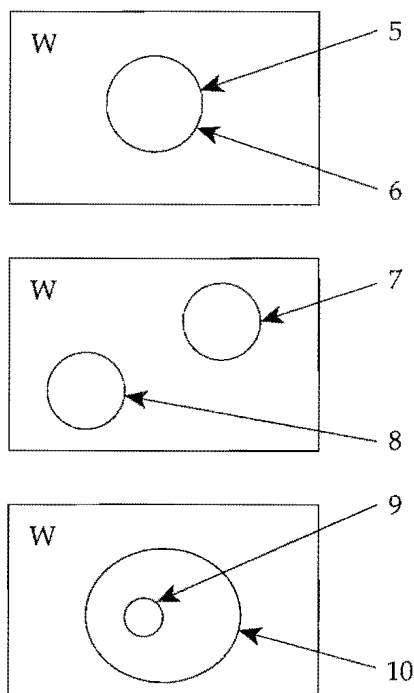


Diagram 5

Two sentences are *contrary* if both can't be true:

(7) The square is bigger than the circle.

(8) The circle is bigger than the square.

In terms of truth-conditions, these are contrary because the set of worlds in which (7) is true is completely disjoint from the set where (8) is true. A stronger notion than contrariety is *contradictoriness*. Two sentences are contradictory if they can't both be true and they can't both be false. (Since sentences (7)–(8) are clearly contrary, they are contradictory if it's impossible for both to be false as well. It may be that they can both be false in a situation in which there is no circle or no square, or it may be that they



are neither true nor false in that kind of situation. See chapter 5 for more details on the meaning of phrases introduced by *the*.)

A sentence *p* *entails* another sentence *q* if the truth of *p* guarantees the truth of *q*:

- (9) The circle is inside the square.
- (10) The square is bigger than the circle.

As seen in diagram 5, (9) guarantees the truth of (10) because the set of worlds in which the former is true is completely contained within (a subset of) the set where the latter is true. That is, if a situation is in the (9)-set, it's guaranteed to be in the (10)-set too.

(a) A sentence is a *contradiction* if, based on its meaning, it can never be true. A sentence is a *tautology* if, based on its meaning, it must be true. Two sentences are *compatible* if they are not contradictory. How would these terms be explicated in a possible worlds Venn (set) diagram?

(b) In terms of your semantic analysis for sentences of the form *p* and *q*, *p* or *q* and *It is not the case that p*, show that:

- *p* entails *p* or *q*
- *p* and *It is not the case that p* are contradictory
- *p* and *q* entails *p*
- If *p* entails *q*, and *q* and *r* are contradictory, then *p* and *r* are contradictory

One way to do this is to work with Venn diagrams, which let you represent the relations among *p*, *q*, and *r* pictorially.

This exercise has an answer, no. 2, in the appendix.

Possible worlds also let us define some fancier semantic properties in a way quite similar to synonymy, entailment, and the like. For instance, as we'll see in chapter 8, they come in very handy when we try to understand *modality*, the semantics of words like *must*, *may*, *can*, *necessary*, and *possible*.

### 1.3.1.3 Reason 3: Meaning and action

Thinking of meaning as truth-conditions fits into a plausible story about the usefulness of language in daily life.<sup>13</sup> Why do we talk to one another,

anyway? One simple, intuitive answer to this question is that language lets us pass on information about the world, so that we can benefit from the each other's experiences. This brings up another question: how do we benefit from information which comes from the experiences of others? An answer: by using that information to help determine which actions are most likely to lead to outcomes we desire.

Put a little bit more precisely, we can say that communication helps us refine our beliefs about what the world is like, and this lets us choose our actions in a rational way. An action is rational to the extent that it tends to maximize the satisfaction of our desires, given our beliefs. (Note that talking about desire here doesn't imply selfishness. One may have altruistic desires.) This may sound very philosophical, but the idea is quite simple. John tells Mary that it is raining outside, and so now she believes something about the world that she did not believe before. This belief helps her determine that it's a good idea to take an umbrella when she goes out, since this will maximize the chance that she'll stay dry (which she desires) given that she now believes it is raining.

We can describe this situation in terms of possible worlds, illustrated in diagram 6 (p. 22). Let's begin by thinking about her desires. Throughout, she wants to stay dry. This is indicated by the dashed line in the diagrams, which only contains worlds where she can avoid getting wet, either because it's sunny out or because she has an umbrella. (Notice that she doesn't desire all sunny worlds or all umbrella worlds, simply because some of these worlds may have other problems. For example, I believe that in the uppermost sunny world, she gets stung by a bee – we can't be bothered to put every detail of the world in our pictures, can we? – so she doesn't desire that world at all.)

Next consider Mary's beliefs. Before the weather report, she believes it will be sunny. This is indicated by the solid line, which on the left-hand diagram only contains sunny worlds. However, on the weather report John says "It will rain today," and this sentence denotes the set of possible worlds indicated by the dotted line. Since Mary believes John, she needs to shift her beliefs to a set which only contains worlds in the proposition expressed by *It will rain today*. This gets us to the right-hand diagram. As you can see in the right-hand diagram, the only worlds which match both Mary's beliefs and her desires are ones where she takes an umbrella. So she'll take an umbrella.

This little story helps support the idea that the meaning of a sentence should be thought of in terms of its truth-conditions because of the role played by John's sentence *It will rain today*. What John said was useful to Mary because it helped her decide to take an umbrella. The

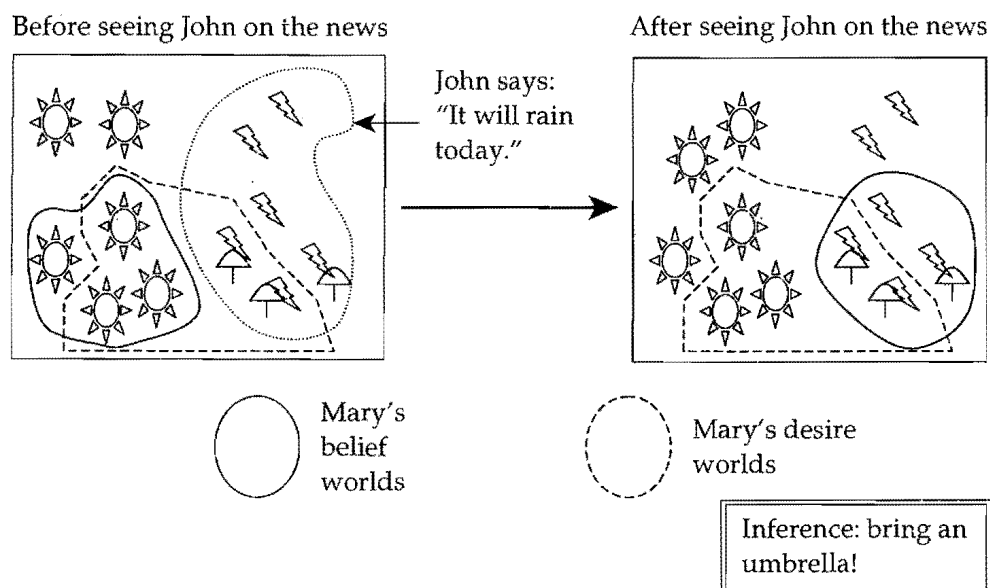


Diagram 6

truth-conditional aspect of its meaning was precisely what was needed to explain how it helped her in this way. If we think that the fundamental function of language is to help us share information and so make better decisions about what actions to take, it seems that truth-conditional meaning is the kind of meaning which underlies language's fundamental function. And if this is so, it makes sense that it is considered to be the first kind of meaning to study!

To think about: many linguists would argue that this function of language is by no means the only one, and even that it's not fundamental. For example, a lot of small talk seems useless from the perspective of facilitating rational action, and might better be described as designed to build or maintain social relationships. The truth-conditional view of meaning sees that function of language as secondary, arising not from the meaning of what is said, but rather from the higher-level awareness of the conversation which is taking place. What do you think?

It can be an interesting exercise to select random things you say during the day (say, by having an alarm go off at hourly intervals), and then try to categorize each as to whether its basic function is social, informational, or a combination of the two.

### 1.3.2 Non-declarative sentence types: interrogatives and imperatives

If the meaning of sentences is to be understood in terms of their truth-conditions, what of sentences which can't intuitively be described as true or false? It doesn't make sense to say that an interrogative sentence like (11), or an imperative sentence like (12), is true (or that it's false).

(11) Who did Sylvia visit?

(12) Draw a circle inside a square.

While they require us to expand our horizons a little bit, these non-declarative sentences nevertheless can fit naturally into the truth-conditional view of meaning.

First, interrogatives: a question is a request for information, and what the form of the question does is tell the hearer what sort of information is being looked for. In the case of (11), the speaker desires the kind of information that can be expressed with a sentence of the form "Sylvia visited person *x*." We might say that the fundamental role of a question is to tell the hearer what kind of answer is being sought. This has led to the major approach to understanding the meaning of questions in truth-conditional terms: the meaning of a question is defined in terms of its possible answers.<sup>14</sup> This can be made more precise in various ways, and the debates concerning them aren't important for us here. Hamblin, for example, says that the meaning of a question is the set of propositions which are possible answers to it. For example, suppose the only people relevant to the conversation are Lucia and Linna. (How it is determined that they are the only ones relevant is an important question, and clearly involves thinking about the situation in which (11) is used.) If only Lucia and Linna are relevant, and we assume Sylvia visited only one, the meaning of (11) would be the set of the following two propositions: the proposition that she visited Lucia and the proposition that she visited Linna (diagram 7, p. 24). In terms of this meaning, the function of a question is to provoke the hearer into picking the true answer(s) from this set.

Imperatives are a bit easier. Imperative sentences, like declarative sentences, categorize worlds into two kinds. Just as *The circle is inside the square* categorizes worlds in the way illustrated in diagram 3, sentence (12) categorizes worlds in a similar way. It's just that in the case of (12), we don't naturally call these the "true" situations and the "false" worlds. Rather, we might call them the "satisfactory" worlds and the "unsatisfactory" ones.

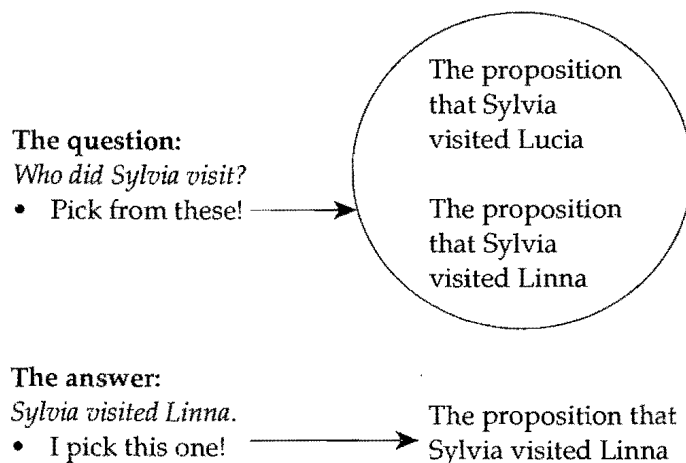


Diagram 7

Worlds in which the addressee draws a circle inside the square are satisfactory, and other kinds are not. Therefore, imperative sentences teach us something more terminological than substantive: we would be better off if we do not describe sentence meaning in terms of "true" vs. "false" worlds, but rather we should use a more general term. We might call them the "yes" worlds vs. the "no" worlds. In the case of a declarative, "yes" is understood to mean "true," while with imperatives, it's understood to mean "satisfactory."

### 1.3.3 Semantic meaning vs. speaker's meaning

The following conversation takes place at a party:<sup>15</sup>

- A: Most of the people here seem pretty glum.  
 B: Not everybody. The man drinking champagne is happy.  
 A: Where?  
 B: That guy! (pointing)  
 A: He's not drinking champagne. He's drinking sparkling water. The only person drinking champagne is crying on the couch. See?  
 B: Well, what I meant was that the first guy is happy.

The last thing B says is an explanation of what he meant when he said "The man drinking champagne is happy." He meant that the guy drinking sparkling water is happy. This points out the need to distinguish what a person means from what the words uttered by that person mean.

In saying "The man drinking champagne is happy," B meant that the guy drinking sparkling water is happy. And he meant something true. Nevertheless, what he said, "The man drinking champagne is happy," was false.

The *semantic meaning* of a sentence is its literal meaning, based on what the words individually mean and the grammar of the language. The *speaker's meaning* of a sentence is what the speaker intends to communicate by uttering it. Often these two coincide, but in the party scenario they did not, due to a mistake of speaker B. Sometimes they might fail to coincide for other reasons. Irony, for one (the semantic meaning is the opposite of my speaker's meaning). Convenience, for another (I know that the woman in a queen-costume at a costume party is not a queen, but say "The queen is quite beautiful." Even though my sentence is literally false, I'm confident you'll understand what I mean).

What other situations can you think of where speaker's and semantic meaning would differ?

When we do semantics, we try to understand – no surprise here – semantic meaning. Speaker's meaning will become more of a direct concern in chapters 10 and 11, when we discuss the sister-field of semantics known as pragmatics. For now, we will try to put speaker's meaning aside, and concentrate on semantic meaning. But we'll have to work hard at this, because it is not always so easy to do. In some cases, it's not easy to figure out whether the meaning we see for a sentence is its literal (semantic) meaning, or a speaker's meaning. For example, a newscaster says "The people remember Tiananmen Square." Of course what she means is that the people remember certain events which took place in Tiananmen Square (and maybe other events which took place around the same time). Does the sentence also literally mean this, or is its literal meaning just that the people remember the physical square itself? It's hard to say. Until we know more basics of semantics, it's best to avoid examples of this kind, even though this means we have to avoid some aspects of how people actually talk. (Is it a problem that we avoid some aspects of how people actually talk? I think it's OK. We're doing science after all. Science typically steers clear of the complexities of our daily world in the hopes that from simplicity will come deeper truths. But science can often return to the daily world and be applied to real situations, and we must try to make sure that semantics eventually can do the same.)

## NOTES

- 1 Quine (1953; 1960).
- 2 See Fodor and Lepore (1992). Some holists certainly think that holism is compatible with semantics as it is practiced by linguists: see Block (forthcoming). Donald Davidson's approach to meaning is holistic, and he and his followers not only believe a scientific approach is possible, but have developed a formal semantic theory (see chapter 12 for more discussion). A recent paper interesting for linguists is Dresner (2002).

This little book doesn't pretend to offer a substantial discussion of the debates among foundational theories of meaning, and indeed you don't have to know much to understand what linguists who practice formal semantics do. I just hope to explain in a pretty intuitive way the formal semanticist's perspective on matters. The references in the notes to this chapter should provide the reader who has both linguistic and philosophical interests with some place to start. Other introductions to semantics would be useful too: Heim and Kratzer (1997); Chierchia and McConnell-Ginet (2000); Larson and Siegel (1995); Saeed (2003).

I thank Steve Kuhn for discussion of the material in this section.

- 3 I like this terminology from Martin (1987), which discusses many of these philosophical issues in a very clear way. Some starter references on versions of the idea theory are: Jackendoff (1992; 1990). On the theory known as cognitive linguistics, see: Lakoff and Johnson (1980); Lakoff (1987); Fauconnier (1985).
- 4 See Fodor (1975) on the idea of a language of thought. Steven Pinker (1994) makes the claim that language understanding is the translation of regular language into the language of thought.
- 5 On Twin Earth, see for example: Putnam (1975); Burge (1979; 1982).
- 6 There are ideas about semantics which have something in common with the idea theory, but which don't identify meanings with ideas. We might say, for example, that meanings are ideas plus something else which makes up for the deficiency of the idea theory. See for example Field (1977); Block (1986); Harman (1987). Some of these theories are functional, holistic views about the nature of meaning in the sense we discussed earlier, so whatever is added to the hypothesis that meanings are ideas would need to make up for the problems we identified with holism as well.
- 7 This idea originates with Wittgenstein (1953). For a recent development, see Brandom (1983; 1994; 2000). Thanks to Mark Lance for discussion of these paragraphs.
- 8 This point does not seem clear to many linguists, who have the feeling that a social practice theory of meaning necessarily undermines formal semantics.
- 9 This way of thinking is consistent with the possibility that what counts as a dog is socially constructed, rather than a natural category. The category "office

worker" is socially constructed. Still, I would say that the term *office worker* describes all the members of this category. Perhaps all categories are like "office worker." This would have nothing to do with the point that meanings are in the world.

- 10 One shouldn't be thrown off the track by thinking about words like *idea* or *concept* themselves. One might think: if meanings are parts of mind-external reality, wouldn't the word *idea* show that ideas themselves are part of mind-external reality? But then wouldn't the idea theory simply be one version of the theory that meanings are out in the world? The difference is that the idea theory takes all meanings to be ideas, while this argument simply shows that some meanings are ideas. The perspective that meanings are out in the world puts ideas on a par with dogs and ancient Chinese philosophers as far as their ability to serve as meanings goes, and, in contrast to the idea theory, doesn't give them a paramount role to play.
- 11 On this strategy, see for example: Davidson (1967a); Lewis (1970); Field (1977); Lycan (1984).
- 12 I will use these terms interchangeably, but sometimes they are used differently. Some scholars would say that a word, phrase, or sentence expresses its sense, and denotes its reference. See chapter 5 for an explanation of the difference between sense and reference.
- 13 See Lewis (1986) and Stalnaker (1984), for two good examples.
- 14 Hamblin (1973); Karttunen (1977); Groenendijk and Stokhof (1982; 1984).
- 15 This example is based on Kripke (1977).



## 2 Putting a Meaning Together from Pieces

We now have some idea of what sort of thing the meaning of a sentence is. The meaning of a sentence is its truth-conditions. *The square is inside the circle* denotes the set of situations where it is true. So, knowing the meaning of a sentence amounts to knowing what it would take for it to be true.

Let's take an even simpler example: knowing the meaning of *Shelby barks* amounts to knowing that it's true if (and only if) Shelby barks. You might say: "This is silly! I bother to study semantics, and all you can tell me is that *Shelby barks* is true if and only if Shelby barks. This is trivial!" This is a natural reaction, one that almost every student of semantics has. But it's not trivial, for two reasons.

First, think about what happens when you don't know a foreign language. Suppose that I tell you in Chinese *Xaio bi jiao*, and you have the impression I'm telling you something important. But since you can't speak Chinese, you begin to get worried. "What is Paul telling me?" you wonder. It would be useful to you to know that what I said is true if and only if Shelby barks.

The statement "*Shelby barks* is true if and only if Shelby barks" seems trivial because I'm using English in two ways. I am using English to talk about English. The italicized *Shelby barks* is the piece of language I'm talking about; we say that it is an expression in the *object language*. It is a piece of the English language, but it could just as well have been Chinese. In contrast, the rest of what I say "... is true if and only if Shelby barks" is being used to say something about the object language; it is an expression in the *metalanguage*. The metalanguage is also English, because that's the language I know best and can write a book in. If I use English to talk about Chinese, or Chinese to talk about English, a lot of the appearance of triviality disperses. But it's most convenient for me to use English to talk about English.

The second reason that a statement like “*Shelby barks* is true if and only if *Shelby barks*” is not trivial is that it opens up another question: how does the sentence get precisely those truth-conditions? What contribution does each word make? Each makes a crucial contribution, since exchanging *Hobo* for *Shelby* or *sleeps* for *barks* would make a big difference. And what role does the grammar of English play? Consider a three-word sentence like *Shelby bit Hobo*. This does not mean the same thing as *Hobo bit Shelby*, so grammar is crucial too. We’re not so much interested in the fact that *Shelby barks* is true if and only if *Shelby barks*, since any English speaker can tell us that; we’re more interested in figuring out how it gets to mean what it does.

More fundamentally, the issue is that, because we intuitively know the meaning of *Shelby barks* or *Hobo bit Shelby*, it seems trivial to report on those meanings. But it only seems trivial. Just because something is intuitive and easy, this doesn’t mean that it’s simple. There are savants who know the answers to complex mathematical equations just by looking at them. But the fact that they find it easy doesn’t mean that the math is simple; it only means that somehow their brains let them do complex problem solving effortlessly. You could say that we’re all semantics-savants (or language-savants more generally), effortlessly solving complex meaning-problems. As semanticists come to understand the complexity of meaning, we can all appreciate just how amazing our effortless skill in creating meaningful language is.

In this chapter, we’ll focus obsessively on the most simple English sentences possible, two-word sentences like *Shelby barks*, and try to develop an idea about how they get to mean what they do. This may sound like it’s going to be pretty dull, but in fact there’s a lot to learn from two-word sentences. Moreover, it will lay the foundation for studying more complex sentences in subsequent chapters.

## 2.1 Names Refer

*Shelby barks* has two words in it, so let’s focus first on the first word. What contribution does *Shelby* make to the sentence’s meaning? We have already discussed the meaning of names in the last chapter, and concluded that their meaning is their reference.<sup>1</sup> *Confucius* refers to a particular ancient Chinese philosopher; *Shelby* refers to my dog. This is at least a major portion of their meaning, and perhaps all there is to their meaning. Their reference is what they contribute to the meaning of a sentence in which they occur.

If I am walking with Shelby and wish to tell you that he barks, I could just point at him and say “Barks.” By pointing at him, I would make clear that he’s the thing which I am describing with my verb. If language had predicates but no names, we would need to keep around us everything we might want to talk about – we’d point to a thing and use a predicate to describe it.<sup>2</sup> Fortunately language gives us the flexibility to avoid this. We have names for things, and the name brings the thing into a statement without the need to have the thing itself at hand.

## 2.2 Incomplete Propositions

Diagram 8 is a new way of representing a proposition. Within the big oval, we indicate that if the dog (that’s Shelby) makes noises like “Woof!” we have a true sentence, and if he doesn’t we have a false one. (This is equivalent to the way we’ve indicated the meaning of a sentence before, drawing a bunch of situations, then indicating some as the TRUE set and the rest as the FALSE set.) Obviously, I intend this diagram to represent the meaning of the sentence *Shelby barks*.

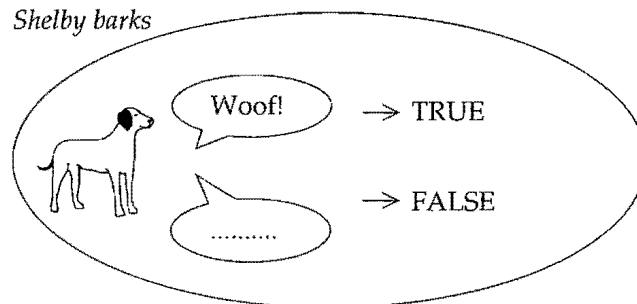


Diagram 8

Suppose we want the meaning of *barks* alone. One way to get at it is by subtraction: take the meaning of *Shelby barks*, and then remove the meaning of *Shelby*. What’s left should be the meaning of *barks*. Doing that to diagram 8, we get diagram 9. This is the same picture as in diagram 8, except that there is a hole where we’ve cut out Shelby. (This can be a fun game if you don’t mind cutting up the book, or making some copies. As you go through this book, you can create your own set of toy semantics building-blocks.)

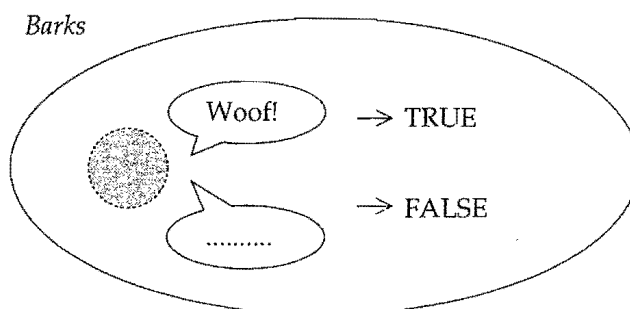


Diagram 9

The meaning of *barks* in diagram 9 is an incomplete proposition, a proposition with a piece missing. This is the basic idea behind the semantics of predicates: they are sentence meanings minus the contribution of the sentence's subject. Following the philosopher Gottlob Frege, semanticists frequently use the term *unsaturated*. A predicate is an unsaturated proposition. Another common term is *property*. A property is an unsaturated proposition, the kind of thing denoted by a predicate. This view of properties takes them to be the semantic core of the sentence, as can be seen by the fact that diagram 9 contains almost everything in diagram 8, minus only one little bit.

### 2.3 Predication is Saturation

Now that we've figured out what *Shelby* means and what *barks* means, we can work backwards and see what happens when someone takes a name and puts it together with a predicate to make a sentence. Somebody says *Shelby barks* to you. You think, "I know that first word. They are referring to Shelby." (This is indicated by the little circle in diagram 10.) Then you think, "I know that second word, they are denoting that property I am familiar with, the one from diagram 9." Finally you think, "The person talking to me put these two words together, so I suppose I'm supposed to put Shelby together with the property; I will saturate the unsaturated, letting Shelby fill in the missing piece of the property." The result is the proposition we saw earlier in diagram 8.

If we had another name instead of *Shelby*, the process would have been parallel, but the result is somewhat different. Suppose I had said *Bach barks*. Then instead of referring to Shelby, I would have referred to a particular composer. This might be represented by having the little circle picture

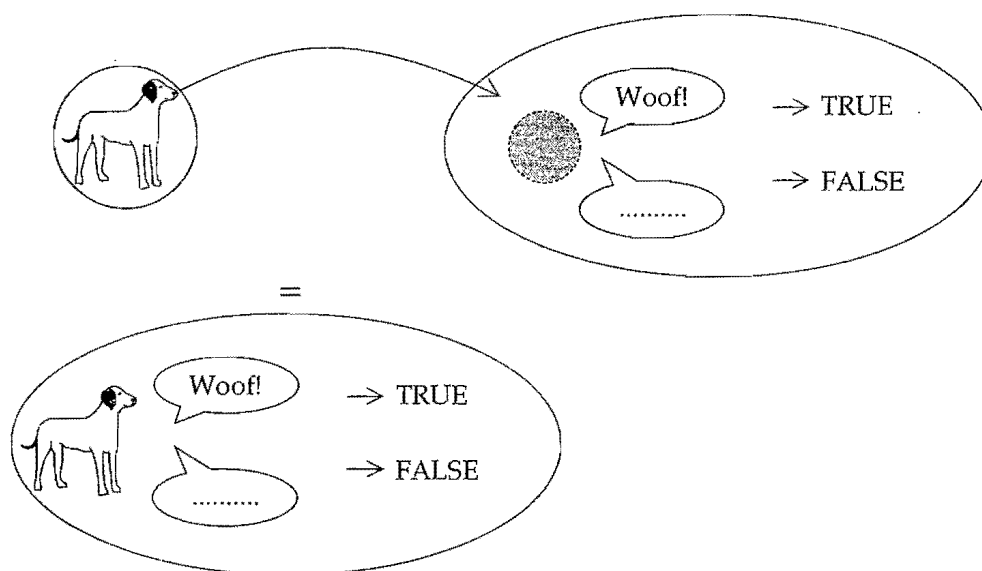


Diagram 10

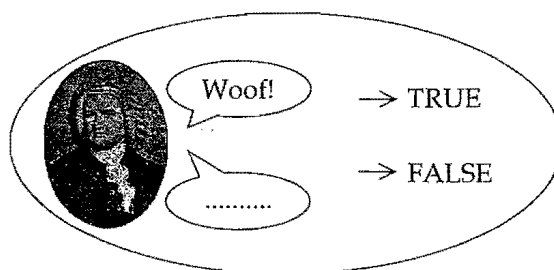


Diagram 11

Bach, rather than a dog. This guy would saturate the property, and we'd end up with the proposition indicated in diagram 11.

The relationship between a predicate and its subject is called *predication*. We say that *barks* "is predicated of" *Shelby*. As we've seen, the semantics of a predicate is a property, or unsaturated proposition; the semantics of a subject (at least when that subject is a name) is a thing referred to. The semantics of predication is saturation, as the property receives its missing piece from the thing referred to. The phrase which saturates a predicate is known as an *argument* of the predicate. *Shelby* is an argument of *barks*.

There's a lot of terminology flying around now. Let me summarize how I'm using some words. Terms which pertain to the form of a sentence (in

linguistics, what we call morphology and syntax) are in **bold**. Terms which describe the meanings (or semantics) of words or phrases are underlined.

- **Predicate** A kind of grammatical unit. *Barks* is an example. If you need more details, check out any grammar, or better yet syntax, book.
- **Name** Another kind of grammatical unit. *Shelby* is an example.
- **Noun phrase** A kind of grammatical unit centered on a noun. A name is a kind of noun phrase.
- **Subject** A grammatical function which a name or other noun phrase can have in a sentence. In English, it's typically the noun phrase which comes before the predicate.
- Property A semantic object. The sort of thing which a predicate denotes. An unsaturated proposition. The property of being something which barks is an example.
- Referent A thing, in a very broad sense, which serves as the semantic meaning of a name (and certain other kinds of noun phrases). My dog *Shelby*, for example, is the referent of *Shelby* when it is used in this book.
- **Predication** The grammatical relationship between a predicate and a subject.
- Saturation Making an incomplete semantic object (like a property) more complete by filling in a missing part. Predication brings about the saturation of a property.
- **Argument** A word or phrase whose referent saturates a predicate. A subject is an example.

## 2.4 Compositionality

A language is an infinite collection of phrases, sentences, and discourses. This can be easily seen from the fact that any sentence can be made longer by combining it with another sentence – in fact, it's possible to do this in very many different ways. (1) becomes (2) by adding another sentence using *and*. (2) becomes (3) by adding a relative clause *which is fat*. It's not hard to see that with tricks like this, there is no longest sentence.

- (1) The cat ate the rat.
- (2) The cat ate the rat and the bat ate the cat.
- (3) The cat which is fat ate the rat and the bat ate the cat.

Since humans are finite creatures, one of the main tasks of linguistic theory is to understand how their minds can give rise to the infinity of language. This is one of the main reasons for the linguistics industry known as syntax (the field which studies how words are combined into phrases and sentences), but the infinity of language is equally an issue for semantics. Each of those infinite number of pieces of language has meaning, and the theory of semantics must somehow be able to link each one to the right meaning.

Apart from the issue of infinity, semanticists (like other linguists) also have to consider the fact that language is creative. We are constantly hearing new phrases and sentences, ones we've never heard before, and yet we somehow manage to figure out what they mean. Semantic theory, considered from a psychological perspective, has to provide an account of the knowledge of meaning which language users have, and which allows for their ability to quickly and easily understand novel pieces of language.

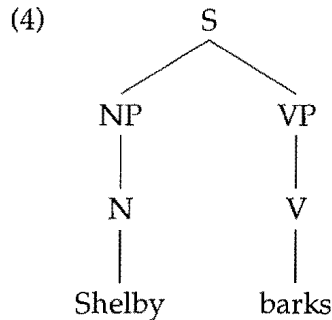
Semanticists seek to bring the infinity and creativity of language within the capacity of human minds by appealing to the *principle of compositionality*. In its simplest form, the principle says that the meaning of a piece of language is based solely on the meanings of its (linguistically relevant) parts, and the way they are put together. With the principle of compositionality, the infinity and creativity of language are comprehensible, since it means that we only have to know a finite number of basic things (the meanings of the smallest pieces of language and the methods of combining them), and this gives us enough knowledge to associate the right meaning with every bigger piece of language.

We've seen the principle of compositionality applied to simple subject-predicate sentences like *Shelby barks*. The meaning of this sentence is based on the meanings of the parts *Shelby* and *barks*, and the way in which they are combined, by predication: in the sentence *Shelby barks*, *Shelby* is the subject and *barks* is the predicate. *Shelby* refers to a thing, *Shelby*; *barks* gives us the property of barking; and predication tells us to saturate that property with that thing. The rest of our discussion of semantics is essentially about fleshing out this picture. We'll need to understand what each basic language-piece means (adjectives, determiners, various sorts of verbs, conjunctions, . . .), and what each way of combining pieces of languages says about how to combine those meanings.

## 2.5 Syntax and Semantics

The principle of compositionality tells us that semantics is always going to have a close relationship with the field of syntax. As mentioned above,

syntax is about how sentences are constructed. The internal structure of a sentence arises from how it is constructed, and this structure can be represented with a *phrase structure tree*, a kind of diagram like (4) which indicates the internal structure of the sentence:



In (4), S stands for "sentence," NP for "noun phrase," VP for "verb phrase," N for "noun," and V for "verb." The lines show how the phrases higher in the tree are composed of the pieces lower in the tree. The individual points in the tree connected by lines are known as "nodes," so that S, NP, N, etc. are all nodes. When two nodes are connected by a line, the lower one is known as a "daughter" of the higher one, its "mother."

Looking at the tree in (4), the principle of compositionality says that the meaning of S is derived from the meaning of the NP, the meaning of VP, and the method of combining NP and VP to make S. It says that the meaning of the NP is based on the meaning of the N and the method of construction ("make an NP from an N, if that N is a name"), and that the meaning of this N is based on the meaning of *Shelby* ("make an N from a word, if that word is a noun"), and so forth.

The principle of compositionality leaves a lot of leeway in terms of precisely how we see the relationship between syntax and semantics, and because of this leeway, many different technical theories have been developed. One major division among the ways of thinking about the issue is between what are called *rule-by-rule theories* and *interpretive theories*.

A rule-by-rule theory sees the meaning of a sentence as being built up in parallel with its structure. It suggests a method like the following for building up the meaning of our simplest sentence, *Shelby barks*:

- (5)
- Take the noun *Shelby* from the lexicon (= the storage space for words), and also take its meaning (let's call that meaning *s*).
  - Take the verb *barks* from the lexicon, and also take its meaning *b* (a property).
  - Assign *Shelby* to the category N, and assign this N the meaning *s*.



- Assign *barks* to the category V, and assign this V the meaning **b**.
- Assign *Shelby* to the category NP, and assign this NP the meaning **s**.
- Assign *barks* to the category VP, and assign this VP the meaning **b**.
- Combine the NP and the VP into an S, and assign this S the meaning you get by saturating **b** with **s**.

Notice how the rule-by-rule theory makes semantics operate in tandem with syntax, in that the meaning of the sentence is built up as the sentence itself is built up. The theory is called “rule-by-rule” because as each syntactic rule is brought to bear (e.g. “combine the NP and the VP into an S”), so is a corresponding semantic rule (“saturate **b** with **s**”).

An interpretive theory has all of syntax happen first, and then has semantics work on a tree like (4) which is the output of syntax. The semanticist needn’t worry about all the details of syntactic theory. All she needs to know is that the result is (4). Then, a series of semantic rules applies to this tree, determining the meaning of the S in terms of the meanings of its parts in a compositional way. For example, the procedure might look like the following, starting at the bottom of the tree in (4):

- (6)
- Look up the meaning of *Shelby* in the lexicon (again, let’s call this meaning **s**).
  - Look up the meaning of *barks* in the lexicon (= **b**).
  - Because *Shelby* is the sole daughter of N, assign **s** as the meaning of N.
  - Because *barks* is the sole daughter of V, assign **b** as the meaning of V.
  - Because N is the sole daughter of NP, assign **s** as the meaning of NP.
  - Because V is the sole daughter of VP, assign **b** as the meaning of VP.
  - Because NP and VP are the two components of S, and one of their meanings is a property **b**, while the other’s meaning is a thing **s**, let **s** saturate **b** and assign the result as the meaning of S.

To a considerable extent, an idea which can be presented in terms of a rule-by-rule theory can also be presented within an interpretive theory, and vice versa. The fundamental difference between them has to do with whether every syntactic process corresponds to a something semantic.

The interpretive theory requires that the final tree derived by syntactic processes be given a meaning, but it doesn't require that every syntactic step along the way has semantic relevance. For instance, if during the process of building up a sentence, some phrase can be created and then destroyed, the interpretive theory doesn't require that the destroyed phrase have a meaning. The rule-by-rule, instead, requires that every syntactic step correspond to a semantic step.

Here's an analogy: suppose that business X wants a new office, and orders that a skyscraper be built. Let's consider two ways X could be charged for this building:

- The price could be based on the cost of each activity which takes place during construction: \$40 for each hour of worker time, the actual cost of all materials used by the construction company plus 10 percent, \$100,000 for design work, and so forth. Suppose that as it builds the office tower, the construction company puts up and later removes some scaffolding. According to this pricing scheme, business X will be billed for the scaffolding, and the time it takes to put it up and take it down.
- The price could be based on the specifications of the final building: \$100 for each square foot of floor space, for example. According to this pricing scheme, business X will not be directly charged for scaffolding. Its cost must be factored into the price of the final building.

The first scheme is analogous to a rule-by-rule theory, and the second to an interpretive theory.

A real linguistic example might be a sentence like (7):

(7) What did John sit on?

Many syntactic theories would propose that *what* is originally combined with *on* to form a prepositional phrase *on what* (parallel to *on the chair*), and that *what* is then moved away to the beginning of the sentence. If this is right, the rule-by-rule theory would say that *on what* should have a meaning, and that this meaning contributes to the meaning of the whole sentence, while the interpretive theory would not need to claim such a thing. The interpretive theory would just base its explanation of the meaning of (7) on the final structure, with *what* at the beginning of the sentence. We would require more detailed technical knowledge than we're going for in this book to pursue the question of which approach is

better. I'll frame the discussion in this book mostly in terms of the interpretive theory.

Let's put aside the distinction between the rule-by-rule theory and interpretive theory. Another goal which motivates how semanticists talk about the relationship between syntax and semantics is to simplify the format of the rules as much as possible. For example, looking at the procedures in (5) and (6), we can see the following pattern:

- (8) If a node has a single daughter, the meaning of that node equals the meaning of the daughter.

It's better to have a single general-purpose rule like (8) rather than a collection of more specific rules like "if an NP has a single daughter N, the meaning of the NP equals the meaning of N," "if a VP has a single daughter V, the meaning of the VP equals the meaning of the V," and so forth.

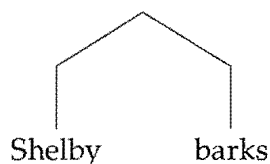
One important issue in the relationship between syntax and semantic is whether semantic rules ever need to pay attention to syntactic relationships other than what is the daughter of what. Consider the process of combining the NP and the VP to make an S, and the corresponding semantic operation of predication. Should this rule be stated as in (9) or (10)?

- (9) If an S has two daughters, one an NP and the other a VP, the meaning of the S equals the result of allowing the meaning of the NP to saturate the meaning of the VP.
- (10) If a node has two daughters, and the meaning of one of the daughters is a thing and the meaning of the other is a property, the meaning of the mother equals the result of allowing the thing to saturate the property.

In (9), we need to know that we have an S, an NP, and a VP in order to trigger the semantic process of saturation. In (10), we need to know that we have one daughter whose meaning is a thing and another whose meaning is a property. That is, in (9) we focus heavily on syntactic facts, while in (10) we focus primarily on semantic facts. Semanticists these days typically strive for rules like (10), and the resulting theories are known as *type-driven theories*.<sup>3</sup> (They are called "type-driven" basically because it is the types of meanings at hand which determine how semantics works, not the syntactic categories like NP and VP.) Type-driven theories may

be either of the rule-by-rule sort or the interpretive sort. If we're working with a type-driven interpretive theory, one nice feature is that it makes the node labels S, NP, VP, etc. in the syntactic tree irrelevant to how the meaning is arrived at. So, we can drop all those labels and use a simplified tree like (11) instead of (4):

(11)



#### NOTES

- 1 There are other views on the semantics of names. See chapters 5 and 6 below for discussion.
- 2 If we had enough predicates, we could be creative and get away without having names. Instead of saying *Shelby barks*, we could say *There is something which shelbies, and it barks* (by definition, Shelby is the only thing that “shelbies”). So the point made in the text is meant as a statement about how language actually works, not how it must work on basic logical principles.
- 3 Klein and Sag (1985).