

1 The Fundamental Question

Semantics is the study of meaning. Those who make this field the center of their academic lives need to enjoy being seen as unimportant, because many people they meet during daily life believe that the study of meaning cannot be a serious, much less a scientific, pursuit. For the general public, an issue is “just semantics” if it has to do with mere matters of form, and accusing someone of focusing on semantics is a way of saying they want to avoid the heart of the matter. An interest in semantics sounds like a deep concern for details of etiquette: perhaps useful in certain contexts, like the diplomatic party or ivory tower, but something which tough-minded realists will brush right aside.

The reputation of semantics within its broader field of linguistics is the opposite of this popular one in many ways. Semanticists tend to use a lot of tools drawn from logic and even mathematics as they go about their jobs as university professors (and almost all of them are university professors; the rest mostly work for computer companies). Because of this, they tend to write down their ideas using all sorts of funny symbols (such as λ , \exists , and \forall). We call this approach to language *formal*, meaning that it is couched in these logical/mathematical languages. The formalism of semantics makes the field virtually impossible for the uninitiated to understand. It also makes it very difficult for those who don't have a native talent for abstraction and logic to become initiated. In other words, semantics is thought to be really technical and hard. (You can imagine what this means for the image of semanticists.)

Though semantics as practiced by the specialist is formal and abstract, at least it is this way for a reason: Formalization allows the construction of very precise theories, and precise theories are better because they don't allow the theorists to fudge the data quite so easily as less precise theories do. Semanticists have to be especially cautious in this respect, I believe,

because discussion of meaning can very quickly turn into pointless contests between vague but strongly held opinions – exactly the sort of situation which the popular view of “it’s just semantics” rightly mocks.

Let me give an example: suppose we wish to understand the difference in meaning between a simple past tense sentence like (1) and a present perfect sentence like (2):

- (1) Mary received the most votes in the election.
- (2) Mary has received the most votes in the election.

After thinking about the two examples for a while, an intuition about the difference may emerge. One common intuition is that (1) simply reports a past event, while (2) reports both a past event and a current result of that event, such as that Mary will be the next president. Armed with this intuition, one can then start looking at other present perfect sentences, and the idea that they report the present results of past events may seem better and better (*Shelby has finished his dinner* – so he’s no longer hungry, *I haven’t slept for days* – so I’m very tired, . . .). Then suppose we encounter the conversation in (3):

- (3) Speaker A: Will Mary be able to finish Dos Passos’ *USA* trilogy by the next book club meeting? It’s so long!
 Speaker B: Well, she has read *Remembrance of Things Past*, and it’s even longer.

There is a current fact which speaker B is pointing out: that Mary might indeed finish the *USA* trilogy. Our intuition seems to be confirmed. However, we think about it some more, and there’s a problem. The fact that Mary might finish the book isn’t really a result of her reading *Remembrance of Things Past*. The fact that she read *Remembrance of Things Past* is merely evidence for what she might be able to read in the future. Is there something else which speaker B’s sentence is trying to report, some other situation which really is a result of her reading *Remembrance of Things Past*? One suggestion: the result is that she has read a really long book before. But notice that this alleged result is itself reported with a present perfect sentence (*She has read a really long book before*), suggesting that our explanation is getting circular, and in any case this “result” is rather ephemeral. Any other possibilities? There are plenty of concrete results of her reading *Remembrance of Things Past*: she has confidence in her ability to read long books; she knows who Albertine is; and infinitely more.

But this raises another problem: How does speaker A know which of these speaker B has in mind? This debate isn't going to get anywhere unless somebody provides a precise explanation of what they mean by the term "result." A formal theory of events and results and time might well be of help here.

Returning to the main thread of discussion, formal theories have some other advantages as well. They are good for implementation in computational settings and for exploring the relations between semantics and the other sub-disciplines of linguistics which have a formal flavor, like syntax, as well the interdisciplinary field known as cognitive science, in particular computer science, psychology, philosophy, and neuroscience. While indirect, these are also good reasons to study semantics in a formal way.

Nevertheless, even if you grant that abstraction and formalism are excellent things for those working at the frontiers of research, you may still want to think about the nature of meaning from a scientific perspective while not taking on the task of learning a lot of formal logic. The goal of this book is to help you do that. It is also designed for those of you who may be undertaking to learn the technical side of semantics, and feel you need a bit more guidance as to what those formulas are really all about.

In order to study meaning, we have to begin with some basic understanding of what sort of thing a meaning is. Trying to answer the question, "What is a meaning?" in its most general sense is a scary task mostly undertaken by philosophers! But we need to begin this book with at least a few basic considerations.

1.1 What is a Meaning?

In ordinary life, we sometimes find ourselves talking about the meanings of words and sentences. For example, I am at home reading, but find myself confused. I don't know what something means. If I ask my wife the meaning of a word, she will give me an answer:

"What does 'kakapo' mean?"

"It's a kind of parrot."

Or,

"What does this character mean?" → 狗

"It means 'dog'."

When I ask what a word means, I typically get more words – perhaps in the language I used in my question, perhaps in another, but nonetheless more words. Can meanings be words? The answer is obviously “no,” if we want to approach meaning as scientists. Because, supposing that the meaning of “kakapo” is “a kind of parrot,” what about the meanings of “kind” and “parrot”? More words. Eventually identifying the meaning of a piece of language with more language is bound to become circular, as a word is defined in terms of some of the very words which it helps define.

Pick a simple word and look up its definition in the dictionary. Then look up the crucial words in that definition, and so forth. Do the definitions become circular?

A more sophisticated view similar to this one is known as *meaning holism*. Most famously supported by the philosopher Quine,¹ the theory of holism claims that the meaning of a word or phrase or sentence depends on its relationships with other words, phrases, and sentences. For example, it might be argued that part of what makes up the meaning of *tall* is that it's opposed to the meaning of *short* (something that seems quite plausible). More precisely, holist theories tend to be *functional* in the sense that it is some aspect of the use of a piece of language which makes for its meaning, so we should really say that part of the meaning of *tall* is that if you call something *tall*, you should not at the same time call it *short*, and if you call something *tall* you should be willing to also call it *not short*. The big issue for holism is to find a way to say which of the relations among words, phrases, and sentences are important to semantics. Radical holism takes the position that there is no line to draw between connections of this plausible sort (*tall* with *short*) and all of the connections among words, phrases, and sentences (*tall* with *I like beans and so I plan to make red bean soup for dessert*). In that case, the semantic system of a language will be a complex, interconnected network, and all meaning will be relative to the whole system. It's difficult to see how meaning can be studied in a scientific way from this perspective. (I should say that it's hard to assess whether holism is in general incompatible with scientific linguistic semantics because over the years a wide variety of theories have been labeled as “holistic.”²)

Linguists who study meaning don't believe that the study of meaning should be unscientific. They feel this way in the first instance (I believe) not because they are better philosophers than Quine and his followers, but because their experience with language shows them that the way languages

express thought is not as arbitrary as the holist's way of looking at matters would lead you to expect. Instead, they find overwhelming evidence for deep and consistent patterns in how languages express meaning, patterns which are in need of scientific explanation. We'll see some of these patterns in the chapters to come. Thus, though the initial intuition that we started with, that the meanings of words involve their relations with other words, is probably correct in some sense, it does not by itself provide a basis for the scientific study of linguistic meaning. Therefore it has not been incorporated much into the thinking of formal semanticists. It is more relevant to the practice of the field of semiotics, the study of symbolic systems generally (including language to the extent that it has something in common with such things as the "meaning" of foods and clothes).

If meanings aren't words, our next guess might be that meanings are something in the mind: concepts, thoughts, or ideas. When you understand the meaning of the word *dog*, your mind (and brain) change in certain ways. At some point you form a concept of dogs (let's indicate the concept with capital letters: DOG). Then, you associate this concept with the English word *dog*, and from then on you have an ability to use the word *dog* whenever the concept DOG is active in your thoughts. From here it's a short step to saying that DOG is the meaning of *dog*. This theory would explain the patterns in how languages express meaning in terms of the nature of concepts, and perhaps ultimately in terms of the way the brain is structured. Let's call this perspective the idea theory of meaning.³

One obvious challenge for the idea theory is to come up with a sound psychological theory of what concepts and ideas are. This psychological theory needs to provide a concept or idea for every meaningful piece of language. Thus, there will need to be ideas and concepts associated with each of the following (at least in any situation in which they are meaningful):

Dogs and cats
The picture of my wife
Three
Whatever
The president lives in Washington, DC.
Had been sleeping
Why
Who said that we had to be at the airport so early?
 -ed (the past tense marker)

The idea theory needs to say what idea is associated with *whatever*, *why*, or *three*, and this doesn't seem as easy a project as explaining what idea

is associated with *dog*. At least, the idea theory provides no quick and easy path to a complete theory of meaning. But even if it's not going to be easy, the idea theory may work. Certainly, something is going on in our minds when we use words and phrases, so in some sense there are ideas associated with all meaningful language. Don't we just have to discover what they are (hire more psychologists!) and use them to explain meaning?

The question which the scientist of meanings needs to ask is not simply whether our concepts and ideas play a role in how we use language in a meaningful way – of course they do. The real issue is whether those concepts and ideas have the right properties to explain everything we need to explain about meaning. In other words, we can consider what we know about meaning already, and then check out whether the idea theory is consistent with that knowledge. Well, what do we know about meaning? Here are some basic points:

- i Sometimes pieces of language have the same meaning – they are SYNONYMOUS.

Dog

Canis familiaris

Mary kissed John.

John was kissed by Mary.

- ii Sometimes pieces of language conflict with each other in terms of their meanings – they are CONTRADICTORY.

The pig is on top of the turtle.

The turtle is on top of the pig.

- iii Sometimes the fact that one piece of language is an accurate description of a thing or state of affairs automatically guarantees that another is an accurate description of it too – the first ENTAILS the second.

Robin

Bird

The circle is inside the square.

The circle is smaller than the square.

The idea theory can say a bit about what it is for *Mary kissed John* and *John was kissed by Mary* to be synonymous. Suppose that in my mind I have ideas, or concepts, of Mary, of John, of kissing, and of “pastness.” These ideas are combined into some kind of aggregate idea, the one associated with the sentence *Mary kissed John*. The idea theory would then want to say that *John was kissed by Mary* is associated with same aggregate idea,

and this is why the two sentences are synonymous. The second sentence has all the same pieces as the first, put together in a different order, plus an additional one, the passive voice (the fact that the sentence takes the form . . . *was kissed by* . . . instead of . . . *kissed* . . .). In some way, the idea associated with the passive voice exactly undoes the effect of putting *Mary* as the subject of the sentence in one case, and *John* as the subject of the sentence in the other. In other words, the sentence without passive voice, *John kissed Mary*, is not synonymous with *Mary kissed John* – clearly – and it’s the passive voice which gets into the aggregate meaning of *John was kissed by Mary* and sets things right.

I think it’s clear that the meaning associated with the passive voice will not be the kind of thing that we typically call an “idea.” This meaning has a grammatical nature, having to do with the order in which John and Mary are mentioned in these sentences. This suggests that, if the idea theory is to work at all, ideas will have to have a language-like nature. That is to say, because the meanings of certain pieces of language are deeply tied into the grammar of that language, ideas themselves will need a grammar. If our ideas have a grammar, they are a language, and we are thinking of them as a *language of thought*.⁴

An important objection to the idea theory arises from the famous “Twin Earth thought experiments.”⁵ The basic idea of the Twin Earth thought experiments is that we can learn a lot about the nature of the mind and language by imagining a world which is exactly like our earth except for some specific differences, and then examining whether those differences seem relevant to how our minds or language work. The following kind of Twin Earth thought experiment is relevant to whether we should accept or reject the idea theory of meaning. One thing we’re absolutely sure about is that when we use a common term like *water*, we are referring to a very definite kind of thing in the natural world. In fact, when we use *water*, we’re referring to H₂O. Now, imagine some people inhabiting a planet very much like ours. This planet is, in fact, so very much like ours that if you went there, you couldn’t tell you weren’t on earth. Everything on Twin Earth is just like on earth. There’s even a copy of you there (and while you’re visiting Twin Earth, he or she is visiting our earth). This Twin Earth only differs from earth in one way, and that is that everywhere we expect to find water there, we find another substance, XYZ, which looks, feels, and acts just like water, but which is actually not water. So, obviously, when the Twin Earth people who speak a language very similar to English say *water*, they are referring to XYZ, and not H₂O. This implies that Twin Earth English and our earth English are not quite the same language. The word *water* differs in meaning between them.

Since Twin Earth is just like earth, each person on earth has a twin there. And this twin is exactly like the true earthling in every respect (except that instead of H_2O , there is XYZ in his or her body). Importantly, the mental life of the earthlings and their twins on Twin Earth are identical. Despite the fact that their mental lives are identical, they don't mean the same thing by the word *water*. This implies that the meaning of *water* is not determined solely by the what goes on inside the head of people who use the language, and this is just to say that meanings are not ideas. Rather, what a word refers to is partially dependent on the environment in which people who use a language live. Given that you live surrounded by H_2O , *water* will refer to water and not XYZ. (If you want to make Twin Earth a bit more consistent, you can say that people who know a lot about chemistry there know that XYZ is common there, but H_2O is not, and conversely on real earth. All that matters for the thought experiment is that some people – not knowing much about chemistry – have no clue about what XYZ or H_2O are. These people and their twins will have the same internal mental concepts associated with the word *water*, but will mean different things by it.)

Another important conclusion about the idea theory can be drawn by considering the meanings of individual words like *dog*. I have formed a concept of dogs through my experience with them. In particular, when I think of dogs I often think of the fluffy gray-and-white, mid-sized ones called Keeshonds, because this is what my dog Shelby is. For me, DOG has many components drawn from Shelby's appearance and behavior. Most people, though, have never heard of or seen a Keeshond, as I can attest from Shelby's celebrity whenever we go for a walk through town. Let us call these people "the unfortunate ones." The concept of DOG held by the unfortunate ones lacks many of the most prominent features of my concept DOG. Since we have different concepts of dogs, and according to the idea theory the concept equals the meaning, this would seem to imply that the word *dog* means something different for me and for the unfortunate ones. So, when they ask me "What kind of dog is that?" their question actually doesn't mean the same thing for me and for them. This conclusion seems somewhat implausible, for even if two people with different experiences with dogs ask that question, they are nevertheless probably looking for the same kind of answer from me.

You might say that my concept of DOG and the unfortunates' concept are similar enough to count as the same because of the fact that they are concepts of the same things out there in the world, namely the members of the species *Canis familiaris*. In other words, all the unfortunates and I aim to use our concept of DOG to classify and think about the members of that species, and on this ground our concepts are different versions of

the same thing. But this concession weakens the idea theory tremendously. It turns the idea theory into a side-trip on the road of understanding meaning. As illustrated in diagram 1, if we're going to explain the meaning of *dog* in terms of the concept DOG, and the concept in terms of the animals which it describes, we might as well explain the meaning directly in terms of the things, avoiding the detour through ideas. This is the view which is adopted within the theory of meaning presented in this book.

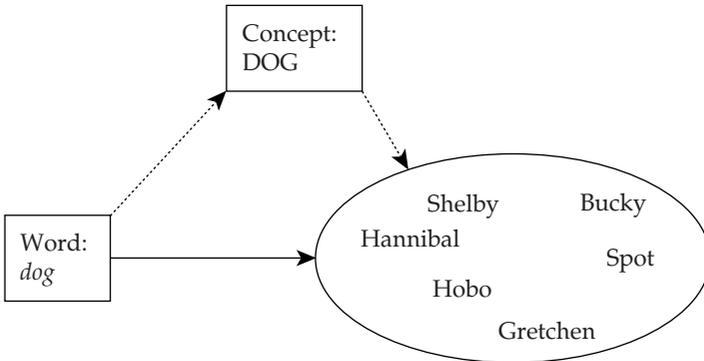


Diagram 1

The reason that the word *dog* means the same thing for you and for me is not that we have the same mental constructs relating to the word. Rather, it's because of our intention to apply the word *dog* to the same things out there in our environment, namely the dogs. (Of course we may not agree about precisely which things these are. You may honestly think that Shelby is a kind of small bear, and refuse to call him a dog. But still our intentions are the same: to call everything by the word *dog* which really is a dog. It's just that you are failing to do so, due to ignorance.)

Notice, by the way, that none of this implies that we don't have concepts of dogs and other things – of course we do! – or that these concepts aren't crucial in how we use meaningful language in the correct way. It doesn't imply that we lack a language of thought (or that we have one). It doesn't mean that semanticists can't contribute to studying the psychology of language. In fact, I am convinced that semantics has an important part to play in the study of cognition. It simply says that meanings aren't ideas. Whatever the relation between meanings and ideas may turn out to be, they aren't the same thing.⁶

If meanings are not in the relations among pieces of language, and are not ideas, what else could they be? Another type of answer which is influential within philosophy says that they are social practices.⁷ The idea is that when

somebody says something, it should be thought of as a kind of move in a giant language game which we all play. This game has rules, and these rules imply things like "If somebody makes the move of saying 'What time is it?' to you, an appropriate move for you to make is to say 'It's 6 o'clock' (if it is 6 o'clock)." Of course this is, in fact, an appropriate response, so any type of semanticist is going to want an understanding of meaning which explains why it is appropriate. But an advocate of the social practice theory of meaning goes beyond this to argue that this is all there is to meaning. So an important task is to explain precisely how each meaningful piece of language comes to have the roles in the language game which it has. Take for example the response "It is 6 o'clock." It is an appropriate response only if you think it is in fact 6 o'clock, and this is somehow based on the fact that it is comprised of the pieces *it*, *be*, present tense, *6* and *o'clock* put together in a particular way. (If these same pieces were put together as "Is it 6 o'clock?" it would not be an appropriate response.) Perhaps we can say that there is a social practice (a rule of the language game) which says something like this:

If the previous move was an utterance of "What time is it",
 then an utterance of the form "It is X o'clock" is a candidate appropriate move, and
 if the clocks in the neighborhood look roughly like this: ⓐ,
 then it is an appropriate move to make the utterance with
 X filled in as "6."

This social practice theory has not had much impact on linguistically oriented semanticists for three reasons. First, there seems to be a fundamental conflict with one of the basic insights of modern linguistics (not just semantics) that our languages are not arbitrary social conventions, but rather reveal deeper universal patterns which spring from the way our minds are built. These patterns call out for scientific explanation, and it seems hard to explain them on the basis of the notion of game-like social practices. In this way this perspective on meaning suffers from the same problem as the first approach we looked at, the one which said that the meaning of an expression comes from its relations with other expressions of language. However, it is not clear to me that there is a fundamental conflict here. It could be that a follower of the social practice theory could say that language is a very special kind of social practice which is instinctual and which has rules that are to a large extent determined by a part of our brains which is responsible for making us follow them.

The second reason that the social practice theory has not been seen as relevant to linguistic semantics is that it does not seem to provide an

important role to the intuition that “It is 6 o’clock” is an appropriate response because it *is* 6 o’clock and you want me to believe this. It is this aspect of language which seems to distinguish it from true games. In baseball, a player will try to catch the ball before it hits the ground because this will help achieve the goal of winning, and that’s it. But you answer my question not just because the rules of the language game tell you that you have to (or else you “lose”), but because you recognize that I want to know a certain piece of information, which I can get if you say “It’s 6 o’clock.” Any other way of giving the same information, or other relevant information, would also be an appropriate move (think about *It’s dinnertime* and *I don’t know*). Gathering these moves together under a collection of rules which tell you what you can do in response to my making the move of saying “What time is it?” would just miss the reason why they are appropriate. So advocates of the social practice theory must provide a notion of “the information provided by a sentence” which is compatible with the fundamental idea that language is a social practice. This is a difficult task, and there is controversy about whether they can succeed. If they can, then it is possible that the social practice theory is compatible with the ideas about meaning presented in this book.⁸

The third reason that linguistically oriented semanticists tend not to pursue the social practice theory is that they feel that we already understand a great deal about how meaning works in language. The main point of this book is to introduce you to the most important of these insights. But these insights have not been achieved in terms of the social practice theory, or the other theories discussed above for that matter. Rather they have been achieved in terms of some ideas about meaning which I’m about to get around to introducing. Until some other way of thinking about meaning shows that it helps us understand something about how language works that we didn’t understand before, semanticists – like other scientists – will see little reason to change.

1.2 Meanings are Out in the World

This discussion so far points to the conclusion that meanings are not internal to language, are not in the mind, and are not merely social practices. Rather, they are based in language- and mind-external reality. The meaning of the word *dog* implies that it describes all of those things that actually are dogs, regardless of our ability to define it with words or to formulate an appropriate mental concept.⁹ The point is even simpler to see in the case of names. The name *Confucius* refers to the ancient Chinese philosopher, and this is the basis of its meaning (indeed this may

be all there is to its meaning). The fact that people may have very different ideas about Confucius, including some very vague ones like “he’s just some ancient Chinese philosopher” which would not distinguish him from lots of other individuals, doesn’t prevent them from all meaning the same person when they say “Confucius.”¹⁰ Of course many names can be used to refer to more than one person. But in a given situation, a speaker intends to refer to just one of them, and if everything goes well, she will. The ability to do this does not depend on the speaker and hearer sharing some idea about the person which would serve to pick him or her out of a lineup. For example, someone could ask “Who was Alexander?” and refer to the ancient general, not any other Alexander, even though that person didn’t know any more about him than that he’s some long-dead guy.

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.

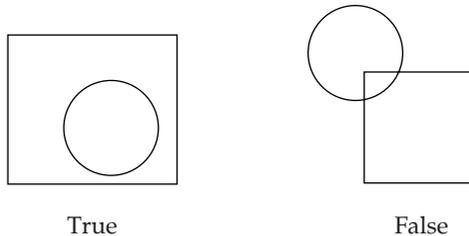


Diagram 2

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.

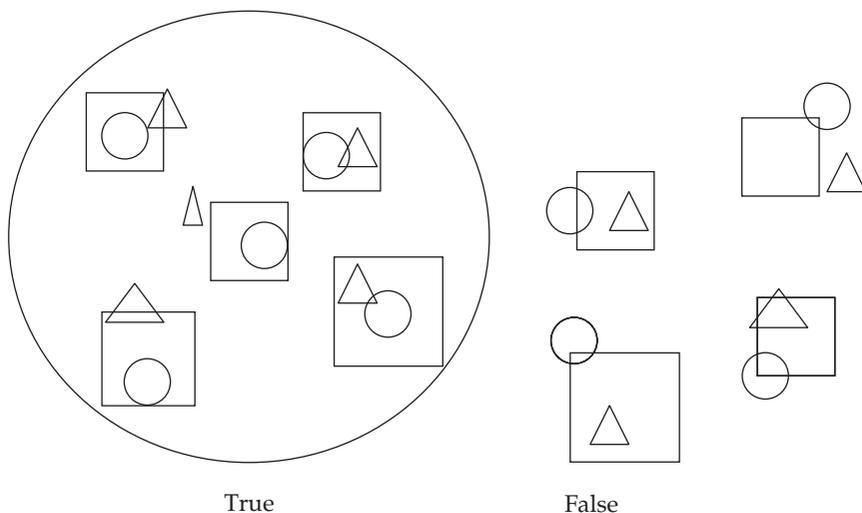


Diagram 3

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.¹¹ 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.¹² 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.)

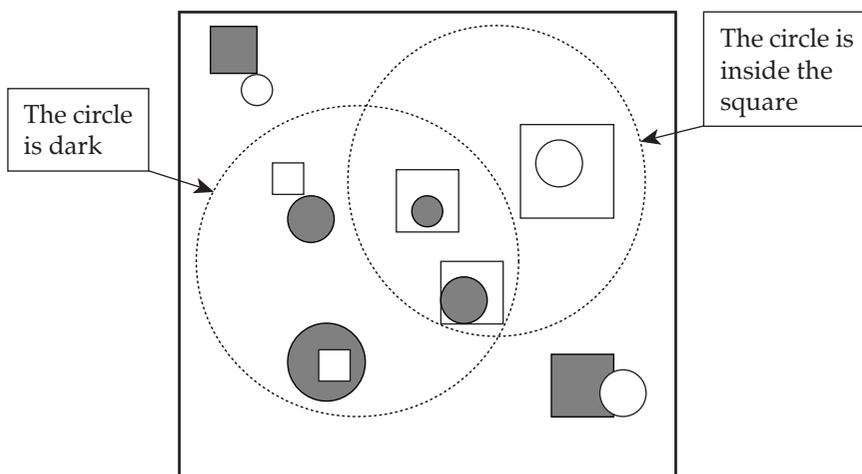


Diagram 4

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.

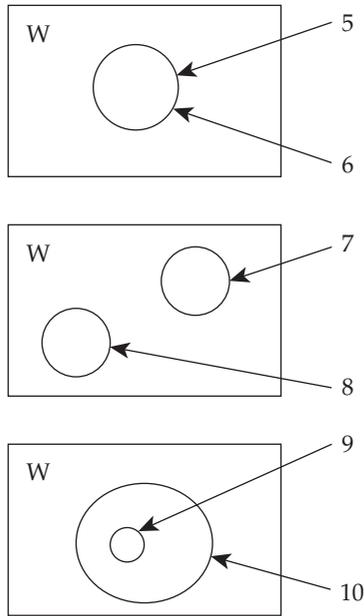


Diagram 5

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.¹³ 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

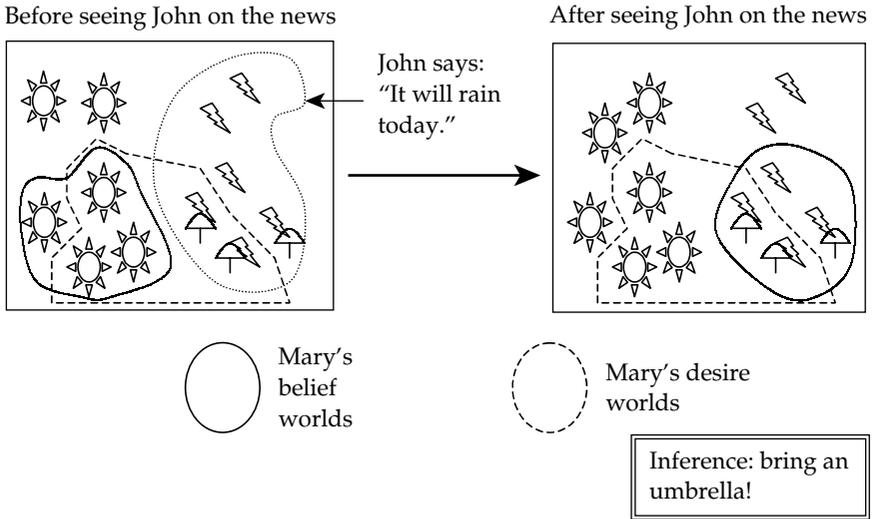


Diagram 6

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 it's 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.¹⁴ 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.

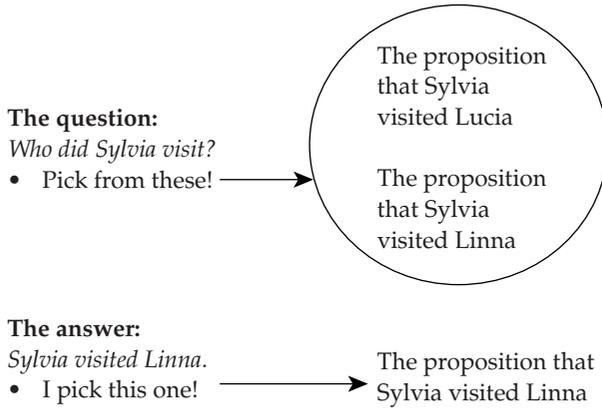


Diagram 7

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:¹⁵

- 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).