# 2 Denotation, Truth, and Meaning

## 1 Introduction

We have outlined what we think the empirical coverage of a theory of meaning should be. This will help us in directly addressing the question, What is meaning? Answers should be evaluated on the basis of how well they account for the phenomena singled out in chapter 1.

The question of what meaning is, is important to any discipline concerned, directly or indirectly, with cognition, that is, with how humans process information. To indicate where we stand with respect to some of the traditional views of meaning, it is convenient to classify approaches to meaning in three groups.

The first family of theories can be labeled "referential" or "denotational." This kind of theory is outward looking; its main emphasis is on the informational significance of language, its aboutness. Meaningfulness lies, according to this view, in the relations of symbols and configurations thereof to objects of various kinds. The study of meaning is the study of such relations. This tradition is the basis of the semantic techniques that have been developed within mathematical and philosophical logic.

It seems reasonable to maintain that the study of the relation of symbols to what they stand for must indeed be part of an account of meaning. For otherwise, how could we understand the fundamental fact that configurations of symbols carry information about all the diverse aspects of our experience?

A second family of theories of meaning might be labeled "psychologistic" or "mentalistic." Theories of this sort are inward looking and focus on the cognitive significance of language. The meaning of a configuration of symbols, according to this view, lies in what we grasp when we manipulate them; that is, it lies in the internalized representation of their retrievable content. The study of meaning is the study of how contents are mentally represented, the study of semantic representations. This tradition is the basis of much semantic work in psychology and artificial intelligence.

It seems reasonable to maintain that a given configuration of symbols has meaning for us only if we are able to grasp its content,

which involves mentally representing it. If such representations are crucial in mediating between symbols and their content, we must not exclude them from semantics.

A third family of theories might be labeled "social" or "pragmatic." Its emphasis is on communication as a social activity. According to this view, meaningfulness lies essentially in the way agents use symbols in the course of their interactions with each other.

Again, it seems indubitable that we actually *do* things with words (saying "I promise to behave" constitutes, under the right circumstances, making a promise) and that key semantic notions like referring or making sense of some set of symbols involve activities. The way we actually use symbols, what we do with words, must play a central role in semantic considerations.

We believe that these three perspectives are by no means incompatible. On the contrary, meaning has all three aspects (namely, the denotational, representational, and pragmatic aspects). Any theory that ignores any of them will deprive itself of a source of insight and is ultimately likely to prove unsatisfactory.

Suppose that we adopted an approach of the second kind, an approach that studied meaning by relating symbols to mental representations or mental procedures of some sort, and stopped there. That would amount to limiting the domain of semantics to the relations between a language, which is a form of representation, and another representation. In other words, one would be relating two representations, translating one into the other (for example, translating our public language into an internal mental code, our "language of thought," say[1]). But how can mapping a representation onto another representation explain what a representation *means*, that is, what its information content is? Representations, routines, and procedures that manipulate symbols are precisely the kinds of things that have meaning. Mapping a representation $A$ onto a representation $B$ will not in itself tell us what representation $A$ means. It will simply transform the problem of what $A$ means into the problem of what $B$ means. Only if we know what $B$ means, what information $B$ carries, will mapping $A$ onto $B$ help. In other words, even if our interaction with the world is always mediated by representation systems, understanding such systems will eventually involve considering what the systems are about, what they are representations of.[2]

Thus, what is needed is some way of talking about what a representation represents, that is, a theory of the information content of a



system of symbols. We have to understand how information flows when we interact in certain ways. Only that, we think, can give to a theory of semantic representation its actual semantic bite.

The denotational perspective (the first one of those outlined above) seems to be promising in connection with the problem of explaining the link between symbols and their information content, in connection with the aboutness of language. In a nutshell, from a denotational point of view, symbols stand for objects. Consequently, configurations of symbols can be used to encode how objects are arranged and related to one another. We believe that this simple idea can be further articulated and developed into a full-fledged theory of what we are calling "information content." We will try to argue that such a theory leads to valuable insights about the structure and role of semantic representations and also meshes well with a view of language as a social activity. We hope, however, that even the reader who is not fully convinced by these arguments will find in what follows a battery of puzzles, techniques, and ideas crucially relevant to semantic analysis.

## 2 Denotation

It is often said that a name, *Pavarotti*, say, refers to or denotes its bearer (the popular singer). We shall use *denotation*, *denotatum*, *reference*, and *semantic value* for what a name (or some other expression) denotes.

The significance of a name does appear to consist largely of its being related to a given semantic value, a certain individual, say. Conceivably, the same paradigm might be extended to kinds of expressions other than proper names; perhaps it might be extended to expressions of any kind whatsoever. If that turned out to be the case, the denotation relation might constitute the most fundamental semantic relation.

### 2.1 Denotation and the foundations of semantics

Other noun phrases (NPs) besides proper names seem to derive their significance or semantic power from their reference. For example,

(1) *a.* It is a pencil.
*b.* This is yellow.
*c.* The tallest man in the world lives in Los Angeles.

For an utterance of (1*a*) to be felicitous, there must be some salient object in the context that is taken as the semantic value of the pronoun *it*. Similar considerations apply to (1*b*), where the reference of the demonstrative *this* might be individuated in our perceptual space by means of an act or demonstration. Sentence (1*c*) is an example that contains a definite description. The reference of the subject NP in (1*c*) is determined by whoever satisfies or fits the descriptive content expressed by the nominal *tallest man in the world*. Typical properties of definite descriptions are that they sound odd if nothing or more than one thing satisfies their descriptive content, as illustrated by the following examples:

(2) *a.* The present queen of France is smart.
*b.* The book that Agatha Christie wrote is about Hercule Poirot.

What is strange about utterances of these sentences is that there is no present queen of France and that Agatha Christie has written more than one book about Hercule Poirot. A theory of definite descriptions would have to account for these oddities.

Let us go back to referential NPs in general. To convince oneself that the notion of reference is central for the NPs in (1), it is sufficient to ask the following simple question: Could we say that we understand the meaning of the NPs in (1) if we didn't know what they referred to? Hardly, it would seem. The NPs in (1*a*) and (1*b*) clearly convey no information by themselves. The NP in (1*c*) does, yet its semantic role is to create an appropriate referential connection with some entity, and there is something distinctly odd (as we have seen in connection with (2)) if such a referential connection cannot be established. Thus, the notion of reference appears to be a fundamental component of what the NPs in question mean.

Of course, to grant this is not enough. Even if we believed that what makes the NPs in (1) meaningful is their relation to a denotation, one would still need an account of, for example, the much more direct role that the context plays in, say, fixing the reference of (1*a*, *b*) as compared to (1*c*). Even remaining within the limits of referential NPs, there is a wide variety of issues that a theory of reference faces.

NPs can refer not just to individuals but also to pluralities or collections of individuals:

(3) *a.* The students in my class are American.
*b.* The students in my class outnumber those in yours.



In (3*a*) the subject NP refers to a plurality of students in a distributive way. That is, the property of being American is attributed individually to each student in the relevant class. In contrast, in (3*b*) the subject NP refers to a plurality in a collective way. No individual student in the relevant class outnumbers anything; only the students as a group do. NPs can also refer to substances, actions, and abstract entities:

(4) *a.* Gold is expensive.
*b.* Running is healthy.
*c.* Justice should be prized.

They can refer to fictional characters:

(5) Bond is my hero.

What these examples suggest is that saying that the meaning of NPs such as those we have been considering consists of their relation to some denotatum is not saying much. Even at this preliminary level one can see that this view needs to be supplemented by theories of pluralities, abstract entities, fictional characters, etc.

In fact, there is a general point that is appropriate to make in this connection. To say that an NP like those we have been considering refers to an individual does not commit us to any preconceived view of what individuals are, nor does it presuppose that the notion of an individual is unproblematic. This comes up in an obvious way when we deal with, say, abstract nouns, as in (4*c*), but it is true of ordinary physical objects as well. Physical objects form causal patterns whose individuation across time or whose location in our perceptual space raise very interesting puzzles. To use a classical example, all the material parts of a table can be gradually changed in subsequent repairs and yet the table might be regarded as the same object before and after such repairs. Thus what we must mean by *the table* cannot be simply identified with the sum of portions of matter that make it up at a given time. Questions of this sort actually turn out to have direct semantic relevance.[3]

In spite of all the problems that there are, it is hard to see how in semantics one could dispense with the notion of an individual or with the notion of reference. Among other things, such notions seem to support a compositional theory of semantic relations (such as entailment or presupposition), even if all they do is link semantics to theories of how objects of various sorts are conceptualized. We hope that this idea will become clearer as we go along.

The following considerations add a further dimension to the problem of reference. Take the example in (6).

(6) A/some/every/no student in my class is blond.

In all these cases we have a nominal (*student in my class*) that is combined with what is generally called a determiner (*a, some, every, no*, etc.). What could the resulting NPs in (6) denote? It is far from clear. One is tempted to say, Nothing at all. In fact, these NPs are often called nonreferential.

There just is no obvious simple way to find a reference for NPs like *every student*. One might try to argue that *every student* denotes the class of students. Then the sentence "Every student outnumbers the professors" should have a sensible meaning that is roughly paraphrasable as "The class of students outnumbers the professors." But it doesn't. Furthermore, one would expect that a sentence like "Every Italian doesn't like Pavarotti" should be unambiguous and mean roughly "The class of Italians does not like Pavarotti." But such a sentence has (at least) two readings: "Not every Italian likes Pavarotti" and "Every Italian dislikes Pavarotti." Arguments in a similar vein can be constructed against other similar attempts to find a straightforward and intuitively simple denotation for the other NPs in (6).

---

**Exercise 1** Assume that *a woman* denotes an arbitrarily chosen woman. What problems does this assumption run into? (Hint: consider what sentences like "In my class, a woman is blond and a woman is red-haired and ..." and "Every man loves a woman" would be expected to mean under the assumption in question.) Assume that *no woman* denotes a class that contains no women. Argue for or against such an hypothesis.

---

So, if we wish to pursue the idea that meaning can be accounted for in terms of a relation between expressions and their denotata, then the problem of nonreferential NPs constitutes a formidable challenge.

Now as pointed out in the introduction to this chapter, part of the appeal of the idea that semantics is essentially denotational lies in the fact that it would enable one to explain the aboutness of language, how it is that expressions have content. As a very rough first



approximation one might say, for instance, that "Pavarotti is cute" conveys information about Pavarotti, because the name occurring in that sentence refers to that singer. Thus, what we say using the name *Pavarotti* will be understood as being about that particular entity. However, we have argued in chapter 1 that there are other empirical phenomena that a theory of meaning should account for, such as what we have called the productivity of meaning and judgments of semantic relatedness. How can we extend the denotational approach that we are considering so as to account for these other phenomena?

To account for productivity, it seems that we need two things: first, a way to determine what expressions of syntactic categories other than that of NPs denote, second, a procedure to determine how the reference of complex expressions depends on the reference of their components.

Let us see what the problems involved are by looking at a simple example.

(7) Pavarotti is an Italian singer.

Sentence (7) is generated by combining the NP *Pavarotti* with the verb phrase (VP) *is an Italian singer*. We might say that the VP *is an Italian singer* has a property as its semantic value. Properties can be predicated of individuals. The result of predicating a property of an individual is something like a state of affairs or situation. So sentence (7) might be regarded as having a situation (or a state of affairs) as its semantic value, intuitively, one in which *Pavarotti* has the property of being an Italian singer.

It might be possible to extend this strategy to more complex constructions. Forming an hypothesis concerning the denotation of other categories besides NPs, and in particular concerning the denotation of sentences, might help us in this task. To see this, let us take a first stab at the hard problem of such nonreferential NPs as *every woman* or *no woman*. We might try to analyze such NPs along the following lines. Let us say that these NPs indeed lack a denotation. This does not mean that they do not play any semantic role. The semantic role of, say, *no woman* would be that of combining with a property (such as, say, the one associated with the VP *smokes*) to yield a situation or state of affairs in which no woman smokes. The idea is to specify the semantic role of nonreferential NPs indirectly via the contribution that they make to the specification or description of the state of affairs associated with the sen-

tences in which they occur. The same strategy might be applied to other nonreferential expressions (such as *and*, *because*, etc.).

Of course, to pursue this line of analysis we would have to overcome many more problems (for example, nonreferential NPs combine with expressions of many different syntactic categories, prepositions, for example, as in "I walked with every student, " etc.). But the above considerations do lend some preliminary support to the idea that our overall strategy of providing a referential analysis for various kinds of expressions may be viable. If so, we could have an arguably elegant account for the productivity of meaning in terms of a primitive denotation relation.

Here is how we proceed. We can classify objects in various semantic categories (say, individuals, properties, situations, etc.), and we can individuate various ways of combining those objects (for example, predication combines individuals with properties to give states of affairs). Expressions of different syntactic categories would be associated with objects of different semantic categories (or types); syntactic modes of putting expressions together would correspond to ways of combining the objects that those expressions denote. In this way one could always compositionally figure out the object that any given expression denotes in terms of objects that its component expressions denote and the way in which they are put together. This also explains more precisely how configurations of symbols carry information about arbitrarily complex states of affairs.

This program is thus of potential interest, for there are grounds to believe that it might account for both the aboutness of language and the productivity of meaning, two important desiderata for a semantic theory. There are, however, some further problems that call for attention. We can bring them into focus by considering in more detail the kinds of entities that we need to assume as semantic values for expressions of categories other than NP. Let us consider in particular sentence denotations, what we have intuitively called "situations" or "states of affairs."

First notice that the notion of a situation or state of affairs that we need to support the notion of a sentence denotation is itself quite problematic. To see this consider the following examples:

(8) *a.* Pavarotti is French.
*b.* If Pavarotti sings "O che gelide manine," I want to be there.



What states of affairs can (8*a*, *b*) denote? There is no actual state of affairs or situation that corresponds to (8*a*). Perhaps we might say that (8*a*) denotes a "hypothetical" or "possible" situation. Similarly, what sort of a state of affairs or situation can (8*b*) denote? It must be some kind of "conditional" state of affairs.

But notations like "possible state of affairs" or "conditional situations" are quite abstract and not immediately clear; in particular, they do not appear to be any clearer than the notion of a sentence denotation, which is what we want to explain. And objections of this kind are not confined to sentence denotations. They also apply, for example, to the notion of "property," that is, what we have indicated as a possible candidate for the role of VP denotations.

What about intuitions of semantic relatedness? Here lies the heart of the problem for the kind of semantic approach we have been considering. To see this, take the following example:

(9) Someone is an Italian singer.

Clearly, (7) is related to (9). The information that (9) conveys is somehow implicit in (7), and this knowledge is part of what we know about the meanings of (7) and (9). So, for example, it is impossible to assert (7) and deny (9). The relation between (7) and (9) is that of entailment, discussed in chapter 1.

We know that any sentence can enter a potential infinity of such relationships. That is, any sentence entails and is entailed by a potential infinity of other sentences, and when confronted with a pair of sentences, we are in general able to judge what entails what. Appealing to properties, predication, situations, and the like will not suffice, unless these notions are able to support a theory of semantic relatedness, among other things, a theory of entailment. In particular, to enable us to characterize entailment, the structure of properties or situations must be rich enough to support a logic. Appealing to properties or situations without specifying their logic is, in Donald Davidson's words, labeling a problem rather than solving it.

Again, it should be intuitively clear that the above argument applies not just to properties or situations but also to sorts of things that we might want to assign to expressions as semantic values. Appealing to any kind of thing whatsoever will be of little help if the logical structure of such a kind is not specified, that is, if no theory of entailment comes with it.

Now it is appropriate to ask the following question: what would it mean for, say, a theory of situations (or states of affairs) to be able to support a characterization of entailment? Let's go back to examples (7) and (9). A plausible first guess would be to say that the situation or state of affairs that (9) refers to is somehow contained in the situation that (7) is associated with. Equivalently, we might say that whenever the situation described by (7) occurs, the one described by (9) must occur. This, in turn, is equivalent to saying that whenever (7) is true, (9) must also be: saying that the situation denoted by a sentence occurs is tantamount to saying that the sentence in question is true. These preliminary considerations suggest that the logic of notions of potential semantic interest is linked in some crucial way to the notion of truth. In section 3 we will begin to explore this line of thought.

But before addressing directly the relation of denotation to truth, we would like to point out another interesting puzzle that specifically concerns the notion of sentence reference. The solution to this puzzle, advocated by the mathematician and philosopher Gottlob Frege, appeals to the notion of sense. First, we turn to the puzzle.

## 2.2 Reference and sense

We have assumed so far that a sentence denotes something like a state of affairs, or a situation. We shall now argue that this assumption, along with two rather plausible principles, leads to highly counterintuitive results.[4] The principles in question are the following:

(10) *a.* Two expressions that entail each other (that are content-synonymous) have the same reference.
*b.* If we have an expression *A* containing an expression *B* and we replace *B* in *A* with an expression *C* that has the same reference as *B*, the reference of *A* does not change.

We claim that these principles constitute valid generalizations about referential NPs. Let us convince ourselves that (10*a*, *b*) are true of referential NPs. Consider

(11) *a.* the sister of John
*b.* the daughter of John's parents

Content synonymy has been defined so far only for sentences. However, it is possible to generalize it to expressions of other categories. In particular, we can say that a referential NP *A entails* another NP *B* whenever the sentence "*x* is *A*" entails "*x* is *B*." Clearly, "*x* is John's sister" and "*x* is the daughter of John's parents" entail each other. Thus, by the definition just given, the NPs (11*a*) and (11*b*) entail each other. And our semantic competence tells us clearly that they also *must refer* to the same individual (whoever that may be). Thus principle (10*a*) appears to be true as applied to referential NPs.



Consider now principle (10*b*). It is easy to see that this too is true of referential NPs. Relevant examples are of the following kind:

(12) *a.* the sister of John
*b.* the sister of Mary's husband

Suppose that John is Mary's husband; that is, suppose that *John* and *Mary's husband* have the same reference. Expression (12*b*) is the result of substituting "Mary's husband" for "John" in (12*a*). Using the schema given in (10*b*), (12*a*) is our *A*, *John* is our *B* and *Mary's husband* our *C*. Again, our intuitions are pretty sharp on this score: if John is Mary's husband, the reference of (12*a*) and (12*b*) *must* be the same, just as (10*b*) would predict.

So, the principles in (10) appear to characterize correctly two properties of the denotation (or reference) relation, as exhibited by referential NPs. We are trying to build our semantics on (some version of) such a relation by generalizing it from NPs to other categories of expressions. Thus, we ought to expect such principles to hold also with respect to the denotation of expressions different from NPs. There is no reason to expect these principles to be limited just to the denotation of NPs if denotation is a unitary semantic relation. In particular, these principles should apply to the reference of sentences. Let us see what happens.

Take two arbitrary sentences (say, "It snows" and "Pavarotti is cute") and suppose that the only thing that they have in common is that they happen to be both true or both false. Let us introduce some handy terminology. If a sentence is true, we say that its truth value is true (abbreviated as T). If a sentence is false, we say that its truth value is false (abbreviated as F). Consider now the following:

(13) *a.* Pavarotti is cute.
*b.* The truth value of "Pavarotti is cute" = T.
*c.* The truth value of "It snows" = T.
*d.* It snows.

We are going to show that by principles (10*a*, *b*), (13*a*) and (13*d*) must refer to the same thing. First, notice that (13*a*) and (13*b*) entail each other. For suppose that (13*a*) is true. Then the truth value of "Pavarotti is cute" is T, and (13*b*) would be saying that the truth value of "Pavarotti is cute" (namely T) equals T, which, of course, is indeed the case. Suppose, on the other hand, that (13*a*) is false. Then (13*b*) will be false too, since the truth value of "Pavarotti is cute" would be F, and (13*b*) would be saying that F equals T, which is clearly false. Since (13*a*) and (13*b*) entail each other, they must have the same reference, by principle (10*a*).

Now (13*b*) and (13*c*) must also have the same reference, this time in virtue of (10*b*), since by hypothesis (13*a*) and (13*d*) have the same truth value and (13*c*) is obtained from (13*b*) by replacing in it the definite description *the truth value of "Pavarotti is cute"* with the coreferential definite description *the truth value of "It snows."* Finally, (13*c*) and (13*d*) must have the same reference, because they too entail each other (the reasoning here is fully parallel to that used to show the content synonymy of (13*a*) and (13*b*)).

Thus, if (10*a*, *b*) are true generalizations about reference, as they appear to be in the case of NP reference, then two *arbitrary* sentences with the same truth value must have the same reference, or denotatum.

But now look: in the cases where the truth value of sentences can be determined, there are going to be only two truth values (true and false). We have chosen in (13) two sentences that, aside from truth values, have nothing in common semantically. At the same time we have shown that they must have the same denotatum. But then what can this denotatum be? Clearly, the denotatum must be something that those sentences have in common, namely, their truth value. That is, if the principles in (10) are valid, then the denotation of a sentence must be its truth value. To put this in different terms, if we want to maintain the principles in (10*a*, *b*), and also the idea that sentences refer to states of affairs, we are forced to conclude that there can be at most two such things: the true state of affairs and the false one. But this seems counterintuitive at best.

What way out do we have? Perhaps there is something wrong with the principles in (10) as applied to semantic values of sentences. But it seems hard to tell what and why if the logical structure of the notion of a sentence denotation isn't spelled out more clearly. Or perhaps we can say that the denotation of sentences is



indeed their truth value. In other words, our strategy of developing a semantic theory in which referential NPs are the model for reference generally has as one of its consequences that sentences must be taken to refer to truth values. We do not in fact have an obvious pretheoretical understanding of a notion of sentence reference; our everyday talk is of sentences' *describing* situations, not referring to them. Still, many find it odder to think of sentences as referring to or denoting truth values than to think of them as referring to or denoting situations. However, such a result need not necessarily be regarded as a negative one if our theory delivers what it should (a theory of entailment, presupposition, etc.).

Nonetheless, if sentences denote their truth values, then there must be something more to sentence meaning than denotation, for we don't want to say that any two sentences with the same truth value have the same meaning. This is what led Frege to posit the notion of *sense*. Let us explore it briefly.

Frege proposes that sentences (and indeed, expressions of any category) have not only a reference (a standard translation of the German word *Bedeutung*) but also a sense (Frege's term was *Sinn*). The reference of an expression is what it stands for on a given occasion of its use. Its sense, Frege says, is the way in which the reference is presented. To illustrate the distinction, Frege uses an example along the following lines. Suppose we are looking at the moon by means of a telescope. The moon corresponds to the reference. The sense corresponds to the moon's image as projected on the telescope's lens. The image on the retina corresponds not to the sense but to its mental representation. The sense (like the image projected on the telescope's lens) is "objective." The retinal image is subjective and may vary from perceiver to perceiver.

More specifically, table 2.1 shows how Frege classified the sense and reference of expressions of the categories we have been considering. The reference of an expression depends on its sense and on what the circumstances are. For example, we can determine the reference of *the morning star* by finding out what fits that description, given what we understand of it and what the facts are. According to this view, meaning is to be analyzed along two complementary dimensions. The meaning of an expression *A* lies in the relation that *A* has with its sense and its reference.

Similar ideas have several historical antecedents and have also been elaborated by other researchers independently of Frege. For example, Ferdinand de Saussure (1916) has a distinction between

**Table 2.1** Frege's classification of sense and reference

| | EXPRESSION | REFERENCE | SENSE |
|---|---|---|---|
| Category | Referential NPs | Individuals | Individual concepts |
| Example | *the morning star* | Venus | The concept of the star that disappears last in the morning |
| Category | VPs | Classes of individuals | Concepts |
| Example | *is Italian* | The Italians | The concept of being Italian |
| Category | Ss | True or false | Thoughts |
| Example | "Pavarotti is Italian." | True | The thought that Pavarotti is Italian |

*signification* and *signifié* that appears to be conceptually similar to Frege's distinction between reference and sense.

It is worth reiterating that for Frege senses are not to be thought of as mental or psychological entities. In particular, the sense of a sentence, say, "Pavarotti is Italian," is not what we grasp in hearing it, for the latter is intrinsically a subjective matter, and varies to a degree from individual to individual. Senses are what enable us to communicate with each other, and as such they must be intersubjective (or objective). So the notion of a thought for Frege should be construed as something like the information content that we grasp in understanding a sentence. Henceforth we will follow the common practice of using the term *proposition* for this purpose. A proposition is the sense of a sentence.

Of course, it is conceivable to adopt Frege's distinction without being radical Fregean objectivists about what sense are. For example, one could hold the view that senses are a characterization of the common structure that our semantic representations must share (given that communication is successful). But the question of the nature of senses has no easy answer. Luckily, as we will see, it is possible to do semantics even in the absence of a complete understanding of this.

In later formal work stemming from the tradition originated by Frege (see especially Carnap 1947), the sense/reference contrast is understood in terms of *intension* versus *extension*. Carnap's notion of the intension of an expression is intended as a more precise



rendering of what Frege called its sense; the extension is what Frege called its reference (or denotation). Sometimes we use Carnap's terminology, sometimes Frege's.

Frege put forth other arguments that point to the need for an appeal to sense (or intension) in semantic considerations. Here are two. The first is concerned with identity statements. Consider

(14) *a.* The morning star is the evening star.
*b.* The morning star is the morning star.

Both definite descriptions *the morning star* and *the evening star* happen to pick out the same entity, namely, Venus. So (14*b*) is derived from (14*a*) by replacing coreferential expressions. If reference is all there is to meaning, then (14*a*) and (14*b*) should have the same information content. But they clearly do not. Sentence (14*b*) is utterly uninformative: we know it a priori. Sentence (14*a*) is informative: in fact, it was an astronomical discovery. Using the notion of sense, we can account for this contrast. Sentence (14*b*) is uninformative because the two expressions being equated are identical, and thus both have the same sense and the same reference. The two expressions in (14*a*), on the other hand, have different senses, and it is an empirical fact that they happen to pick out the same reference, whence the informativeness of (14*a*).

The second argument has to do with what Frege called "indirect" and Quine "opaque" contexts:

(15) Sophia Loren believes that Pavarotti is French.

Sentence (15) attributes a certain belief to Loren. What Loren believes is somehow described by the sentence "Pavarotti is French." Consequently, it must be systematically recoverable from the meaning of the latter. But clearly the actual truth value of "Pavarotti is French" does not determine what Loren may believe. Thus, the notion of sense is needed to account for contexts such as these. Sentence (15) can be interpreted as saying that Loren bears the *believe* relation to the thought (or proposition) expressed by "Pavarotti is French." Examples such as these could be multiplied and elaborated upon in many ways.

At this point we should ask the following questions. Does appealing to the notion of sense really help us? Does it provide a base from which we can study meaning in natural language? Recall that what we want is a compositional theory of meaning that accounts for the properties discussed in chapter 1. In particular, such a theory

should account for our intuitions concerning semantic relations. If meaning is to be studied along two dimensions (the intensional and the extensional), we need a way to determine compositionally both the intension and the extension of an expression in terms of the intension and extension of its parts. We also need to know precisely how intensions and extensions are related. Moreover, they should provide an account of the various semantic relations, such as entailment and presupposition. In the absence of all this, appealing to intensions will not help us much. To say that "Pavarotti is French" has the thought (or proposition) that Pavarotti is French as its sense links the notion to be explained (namely, that of meaning) to the notion of a thought (or proposition), and this latter notion is equally in need of an account. Furthermore, this move in itself buys us nothing in terms of explaining the various semantic relations. This is precisely the criticism that we have leveled against accounting for sentence meaning in terms of the notion of a situation or state of affairs. It seems, therefore, that we have reached an impasse.

We started out by exploring the notion of reference or denotation and giving some general reasons why such a notion could play a central role in semantics. However, we have met some difficulties in extending such a notion beyond referential NPs; in particular, we have had trouble with the notion of sentence reference. We first saw some of the difficulties that arise from adopting the view that sentences denote situations or states of affairs. In essence, we argued that this claim is nearly vacuous if its connection to a theory of semantic relations (such as entailment and presupposition) is not made clear. We have also argued that generalizations that seem to be true of the notion of NP denotation lead to counterintuitive results if we try to maintain that the semantic content of sentences be analyzed in terms of situations or states of affairs. We then considered Frege's way out of these difficulties, which appeals to the notion of a sense (or intension). Such a notion, however, also appears to be nearly vacuous if its connection to a theory of semantic relations is not made clear.

We don't think, however, that these difficulties constitute an insurmountable obstacle to constructing a semantics in which the notion of denotation plays a central role. Nor do we believe that they conclusively show that Frege's notion of sense has no use in semantics. In fact, in chapter 5 we argue that it does. But we do seem to need a different starting point.



## 3 Truth

### 3.1 Nur im Zusammenhange eines Satzes bedeuten die Wörter etwas

One of the recurrent problems that we observed in section 2 has to do with how to characterize sentence meaning. In particular, we have tried to focus on the reasonably clear name-bearer relation and adopt it as paradigmatic for all key semantic notions. But what sentences denote and how one gets to such a denotation remain outstanding problems. Perhaps before trying exactly to identify the denotation of words (or morphemes), we should try to make some progress toward a viable characterization of the semantic content of sentences.

In fact, it is not even clear that the notion of denotation can be understood independently of sentence meaning. This is arguably true even of the best understood referring expressions, like proper names. Consider how we could go about explaining the "meaning" of the name *Pavarotti* to someone who doesn't know it. Two obvious possibilities would be pointing at the famous singer and giving a description of him. We can of course combine these two possibilities in various ways as the circumstances require. But even if we use a simple ostension, or deixis (pointing), what our act expresses is something like a complete utterance with roughly the same meaning as "This person is Pavarotti." So it would seem that we are dealing with a propositional kind of knowledge. Moreover, for the pointing to make sense, we must already be able to distinguish and classify people from other objects. In other words, as Quine (1960) argued, the perceptual stimuli from which deixis can be drawn are insufficient to characterize the objects that constitute the frame of reference for our language. We can refer to something and individuate it within a given background only by using a conceptual system. It follows that in grasping the meaning of a word, any word, for the first time, we cannot get at it directly (whatever that may involve). We never deal with labels and objects in isolation. We are typically confronted with complex states of affairs in which objects stand in relations to other objects. Indeed, one can say that we arrive at objects via a process of abstraction that enables us to identify them as, say, causal structures, regularities across states of affairs.

What this suggests is that to get started, we should pursue units more complex than names (or words). Language, as an information code, provides an association between two systems: what signifies

and what is signified. Sentences, as opposed to whole texts, appear to be the smallest autonomous information units in a language (with some qualification having to do with context dependency—see below). Sentences comprise a category of well-formed structures capable of expressing thoughts that can stand on their own, of describing whole situations. Thus, perhaps getting at sentence meaning might be easier than getting at the meaning of other units. What we might try to do is to define "S means *p*" precisely, where S is a sentence. We might then be able to identify further crucial semantic notions in terms of sentence meaning. This, in fact, is one way of capitalizing on the famous (and controversial) dictum by Frege that we use as title for this section: "Only in the context of a sentence do words have meaning."

Before seeing how such a program might be pursued, we should clear up some obvious problems. First, what are sentences? How do we define them? For the time being, we will consider only ordinary declarative sentences. We hope to be able to convince the reader that the approach developed in connection with this kind of sentence does extend to the other kinds. Second, it is a commonplace observation that the content of (declarative) sentences can depend on the situation, or context, in which they are uttered. Consider, for example,

(16) I am tired.

What (16) can convey is going to depend partly on who the speaker is and when the sentence is uttered. And there are, of course, many other more complex ways in which what a sentence means depends on the context (ways having to do with intersentential anaphora, focus, presuppositions, etc.). Trying to address fully the issue of context dependency at this stage would complicate our task considerably. Therefore, we adopt a simplification known as the "fixed-context assumption." We assume that the context of use (who the speaker is, what the time of the utterance is, etc.) is a known quantity. Consequently, so-called indexicals such as *I* in (16) come to have a definite reference and behave just like other referential expressions (such as proper names). This assumption will then be abandoned when we specifically address the issue of indexicality.

Within these restrictions a conspicuous property of declarative sentences is that they can be true or false in a given situation or circumstance. Consider, for example, (17), and assume that its



context of utterance is known, say, September 10, 1986, in a classroom on the Cornell campus in Ithaca (as per the fixed-context assumption).

(17) The pope talked to Reagan between 3:00 and 4:00 P.M. on September 9.

When is (17) true? Of course, even if the context is fixed, truth depends on more: what the facts are. Sentence (17) is going to be true if, in fact, the two relevant people were talking at the specified time. We may never know whether such an event actually took place. Perhaps neither of the protagonists has been seen at the designated time. Perhaps they have been struck by amnesia concerning the event described in (17). However, even though we may lack actual knowledge of facts, we know, for example, that a transatlantic phone call could suffice for (17) to be true, but (17) would not be true if John Paul spent all the relevant time talking to Pavarotti. The important thing to notice here is that though we might not know what the facts are, we do know what they ought to be in order to make the sentence true. This knowledge, we claim, is semantic (and hence grammatical) in nature: it is constitutive of our knowledge of what (17) means.

Conversely, someone who did not know what (17) means (for example, a monolingual speaker of Russian) could not make use of a specification of the facts to evaluate it. To judge whether (17) is true, one needs not only knowledge of the facts; one also needs to know what (17) means, to know something about the grammar of the language. If we didn't know what (17) means, we would have no clue as to what circumstances would make (17) true.

Notice that we are not trying to provide effective criteria for checking the truth of sentences. We don't think that semantics could or should aim so high. What we want to do is simpler. Are there criteria to determine when it is appropriate to say that a sentence is true? We think that there are. The examples illustrate them. A declarative sentence like (17) describes a corner of reality, claims that a certain condition (John Paul's talking to Reagan) obtains. Saying "S is true" amounts just to saying that the conditions that S claims to obtain do obtain. Thus we have at least a criterion of adequacy for the predicate *is true*. It may seem a trivial one, but consider that we don't have even that much for "S means *p*." The notion of truth, whatever problems it may have, is a little bit clearer than the notion of meaning.

Tarski (1935, 1944) has shown that we can draw further consequences from having a clear criterion of application for the truth predicate. To give a characterization of this predicate for a whole language, we need to have a theory that, for any $S$ in $L$ and any $v$, gives us the following:

(18) $S$ is true in $v$ if and only if (iff) $p$.

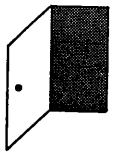
Here S is a structural description of a sentence of a language $L$, $v$ is a situation or a specification of the relevant facts, and $p$ describes the conditions that have to obtain for S to be true in $v$ (that is, the truth conditions for S). The reader may be worried by the fact that we are still relying in (18) on the notion of a situation (or circumstance), which gives rise to problems. We will show later, however, that the way we use this notion in giving a truth definition is quite unproblematic.

Sentences of the form (18) are called T-sentences. Now, if the language contains only a finite number of sentences, then one could simply list all the relevant T-sentences, and we could directly pair up all members of the syntactic category of sentences with their truth conditions. But if the language contains an infinite number of sentences, then a theory of truth must incorporate a mechanism for generating all of the correspondingly infinite number of T-sentences. Presumably, such a mechanism will have to be based on the generative device that characterizes the syntax of the language. In other words, a characterization of the truth predicate for an infinite language must be compositional. As we shall see, to obtain a compositional definition of truth for a sufficiently rich language is not exactly trivial.

We should perhaps point out that we are not claiming that meaning is completely exhausted by truth conditions. What we are claiming is that if we ignore the conditions under which S is true, we cannot claim to know the meaning of S. Thus, knowing the truth conditions for S is at least necessary for knowing the meaning of S. We cannot have the latter without the former. Suppose we did not know whether sentence (19*a*) is true or false in the situation represented in (19*b*).

(19) *a.* The door is closed.

*b.*





Could we be said to know what (19*a*) means? We think not. But then truth conditions must surely be a necessary component of sentence meaning: there is no meaning without truth conditions.

In fact, various philosophers have gone beyond this and argued that knowing the meaning of S is just knowing its truth conditions.[5] If that is so, one could propose the following definition:

(20) S *means* $p =_{df}$ S is true in $v$ if and only if (iff) $p$.

What we have on the left hand side is quite obscure: an intensional relation involving an entity whose nature is unknown ($p$, viewed as the meaning of S). What we have on the right hand side is a lot clearer: a biconditional between two sentences of our semantic metalanguage, "S is true in $v$" and $p$ (viewed as a sentence describing when this holds). From the perspective of definition (20), a theory of sentence meaning (for a language $L$) is just a formal device that compositionally generates all the T-sentences for $L$.

Perhaps before discussing this claim any further, we should see what such a formal device would actually look like. We do so by providing a phrase-structure grammar for an elementary fragment of English and developing a Tarski-style truth definition for it.

## 3.2 The fragment $F_1$

The syntax of $F_1$ is specified in terms of a very simple set of phrase-structure rules and hardly requires any comment. The semantics of $F_1$ corresponds essentially to the semantics of the propositional calculus. Its design, however, differs from what can be found in most introductory logic textbooks, as the emphasis here is on the actual linguistic applications of propositional logic. The simplest sentences in $F_1$ are made up of noun-verb (N-V) or N-V-N sequences. We shall call such sentences *atomic*. Complex sentences are obtained by conjoining, disjoining, and negating other sentences. Even though the grammar of $F_1$ is so simple, it generates an infinite number of sentences.

**3.2.1 Syntax of $F_1$** In specifying the syntax of $F_1$, we use more or less traditional grammatical categories (S for sentences, VP for verb phrases, $V_t$ for transitive verbs, $V_i$ for intransitive verbs, etc.). These categories are adopted purely for pedagogical purposes. Discussing syntactic categories and phrase structures goes beyond the limits of the present work. As far as we can tell, any of the major current theories of syntactic categories (such as X′ theory, or extended