Word Meaning

Throughout much of this book, I have waffled on the question of whether I am talking about concepts or words. The two seem to be closely related. I can talk about children learning the concept of sheep, say, but I can also talk about their learning the word *sheep*. I would take as evidence for a child having this concept the fact that he or she uses the word correctly. On a purely intuitive basis, then, there appears to be considerable similarity between word meanings and concepts. And in fact, much of the literature uses these two terms interchangeably. In this chapter, I will address in greater detail the relation between concepts and word meanings, and so I will have to be somewhat more careful in distinguishing the two than I have been. (Let me remind readers of the typographical convention I have been following of italicizing words but not concepts.)

To discuss the relation between these two, one ought to begin by providing definitions for these terms, so that there is no confusion at the start. By *concept*, I mean a nonlinguistic psychological representation of a class of entities in the world. This is your knowledge of what kinds of things there are in the world, and what properties they have. By *word meaning*, I mean quite generally the aspect of words that gives them significance and relates them to the world. I will argue that words gain their significance by being connected to concepts. It is important to note, however, that this is not true by definition but needs to be argued for on the basis of evidence and psychological theory. And that is the main goal of this chapter. More generally, the goal is to say how it is that word meanings are psychologically represented.

As the reader will see, there is less new empirical material here than in most of the chapters of the book. Instead, the chapter is more of an essay in which I attempt to argue for a specific relation between word meanings and concepts. I begin by discussing notions of word meaning from psychology and linguistics and then propose a particular kind of relation between concepts and meanings. I attempt to support this proposal with theoretical and empirical evidence. Finally, I discuss some important topics from lexical semantics that are usually not addressed in psychological
theories of word meaning, including polysemy and lexical modulation. These phenomena are critical to a complete explanation of how word meaning is related to comprehension, even though they have not received much attention in the psychological literature.

Different Conceptions of Word Meaning

The present chapter will not provide an introduction to the study of linguistic meaning. That is an enormous topic that has a huge literature in philosophy and linguistics, in addition to psychology, and a single chapter cannot even mention many of the topics in it, much less develop them in any detail. The focus of this chapter will be on how word meanings are psychologically represented. It will not address the meaning of larger linguistic structures (sentence, discourse, or story) except tangentially, and it will not consider controversies involving meaning in the philosophy of language. That said, I need to quickly review approaches to the study of meaning outside of psychology, in part to explain why it is that cognitive scientists from different perspectives have said things about meaning that seem at odds with what I propose. The reason, in brief, is that they have goals and assumptions that are not ours. In some cases, their enterprises are simply separate questions that do not conflict with the psychological ones. In other cases, however, it is likely that the two approaches will clash, and one will eventually have to give way (hopefully theirs). But resolving such disputes will not be the focus of this chapter.

One idea that people often have about word meaning is that it is a kind of “official” statement of the sort that is found in dictionaries. That is, when you have an argument with a friend over whether berm means a bank of earth or a small stream, you would look it up in the dictionary and find out that it means the former. Or you might have an argument over whether data is singular or plural, and check a style manual to decide the answer. From this perspective, the meaning is whatever is in the dictionary or other authoritative source. However, this kind of prescriptive meaning is not what we will be talking about, though it does have some importance. We will be talking about the meanings of words as people ordinarily use and conceive of them. So, if people typically misuse a word according to the dictionary or style manual, that is a problem for experts in usage. Psychologists are interested in the meaning of the word represented in the person’s head—not what people are supposed to do.¹

¹ In linguistics and philosophy, semantics (a synonym for meaning, usually) is taken to mean the relation between language and the world. Typically, we use language to
talk about real objects, people, events, and states. A theory of semantics, from this perspective, is one that explains how this connection between language and the world works. This approach is often called referential semantics, because it argues that words get their meanings by referring to real objects and events. One reason for framing the question of semantics this way is to explain the truth of sentences (see Dowty, Wall, and Peters 1981). That is, a statement is true just in case it corresponds to a situation in the world. Theories of semantics based on logic spell out that correspondence. For example, if I see you opening a gate and tell you “There’s a mad dog in that yard,” it would not be surprising if you were to step back and quickly close the gate. The reason for this is that there is a reliable (though not perfect) relationship between this sentence and a situation in the world in which a mad dog is to be found in a yard. The goal, then, is to explain the relationship between mad and various mad things, between dog and kinds of animals, between in and different spatial relations, and so on, such that the meaning of “There’s a mad dog in that yard” corresponds correctly to a real-world situation of a certain kind of animal located within a yard. (Or if it doesn’t correspond correctly, we could explain why you complained that I misled you.) For a good explanation of this aspect of meaning, see Chierchia and McConnell-Ginet (1990).

In order to explain the reliable connection between language and the world, many philosophical theories of semantics argue that word meaning is simply a relation between a word and the world. So, the meaning of dog is the set of dogs in the world. Actually, the theories are somewhat more complex than this, because the set of dogs in the world is changing all the time: New dogs are born and old ones die every minute. Furthermore, we may talk about hypothetical or fictional dogs, which would not be found in the world. Thus, the meaning of dog must be something more complex, like the set of all the dogs that ever existed and that will exist, or the set of all possible dogs (“If I had a dog, it would be brown”). In this case, as in many others, the meaning would be an infinitely large set of objects, since there is an infinite number of possible dogs.

Referential semantics is extremely popular in linguistics, but it is not acceptable as a psychological theory. The reason is that people do not know or have access to these sets of objects or events. Fortunately, I do not know all the dogs in the world, much less all dogs that ever existed or will exist, or all possible dogs. If understanding the meaning of “dogs have four legs” requires reference to the set of all possible dogs, then I can’t possibly comprehend this sentence, because I can’t know all possible dogs. This approach to meaning, then, is not suitable as a psychological theory (see Chierchia and McConnell-Ginet 1990; or Dowty, Wall, and Peters 1981;
for more detailed explanation and defense of referential semantics). It is possible that a psychological theory will turn out to be related to the referential theory. For example, whatever my meaning of *dog* is, perhaps it allows me to pick out any possible dog, even though my meaning is not actually the set of all possible dogs. If such a correspondence could be worked out, then separate linguistic and psychological theories might be compatible, even if they involve very different kinds of entities (infinite sets and relations in one case, some kind of psychological representation in the other).

The psychological approach assumes that people do not know about every example of each word they know. Instead, it assumes that people have some sort of mental description that allows them to pick out examples of the word and to understand it when they hear it. For example, I do not know all possible dogs, but I have some description of what *dog* means in my mental *lexicon* (or dictionary) that can be used to identify dogs. Furthermore, when people tell me things like “my dog bit me,” I can understand them, because I can retrieve the mental descriptions of these words and combine them to result in a proposition that describes the meaning of the entire sentence. This description will be accurate enough in most cases for me to make contact between words and the world, as referential semantics assumed. That is, my mental descriptions of *mad* and *dog* are such that I can combine them to identify actual mad dogs and then talk about them. So, the goal of any psychological approach to word meaning is to specify that mental description associated with each word, which allows people to use it.

I want to remind the reader that my taking this perspective is due to our ultimate goal of explaining how people understand and produce words. If one has a different goal, such as explaining how words got the meanings they have, or how meaning relates to logical validity, one might take a very different perspective. In providing a conceptual theory, I am not saying that it exhausts all that needs to be said about meaning. In too many discussions of meaning, there are arguments in which writers show that X cannot fully account for some semantic phenomenon, and so they conclude that X is not part of a theory of meaning. My own guess is that many different kinds of perspectives will be required to explain all of meaning. The present perspective is the most appropriate one for addressing psychological issues.

**Summary of the Conceptual View**

My claim in this chapter is that word meanings are psychologically represented by mapping words onto conceptual structures. That is, a word gets its significance by
Word 1 \(\iff\) Concept 1
Word 2 \(\iff\) Concept 2
Word 3 \(\iff\) Concept 3
Word 4 \(\iff\) Concept 4
...
Word N \(\iff\) Concept N

Figure 11.1
The "word = concept" idea. There is a one-to-one mapping between words and concepts.

being connected to a concept or a coherent structure in our conceptual representation of the world. To put it another way, the meaning is built out of concepts. The way that this works can be somewhat complex, and specifying it is difficult, because (understating wildly) we do not know everything we would like to know about conceptual structure. As a result, whatever limitations there are in our understanding of concepts will be carried over onto our understanding of how concepts represent meaning. In order to clarify how the two are related, let me work through some very simple, wrong ideas, which will illustrate why the answer must be more complex.

One very simple idea about the relation of word meanings to concepts could be called the "words = concepts" idea, illustrated in figure 11.1. As the figure shows, every concept is connected to exactly one word, and every word has exactly one concept as its meaning. Clearly, this is oversimplified: There are synonyms and ambiguous words in which the mapping is not one-to-one (i.e., ambiguous words must be connected to two different concepts, and synonyms must be connected to the same concept). However, it would be fairly simple to allow these in our system, by listing ambiguous words twice, for example. A more serious problem with this version of the concept-meaning relationship is that there are many concepts that do not have a word to go with them. In fact, people have published books of concepts that do not have words but which ought to. Some examples that were distributed over an electronic bulletin board include:

**ELBONICS:** n. The actions of two people maneuvering for one armrest in a movie theater or airplane seat.

**PUPKUS:** n. The moist residue left on a window after a dog presses its nose to it.²

I, for one, was familiar with these concepts, though I never had a word for them before reading these suggestions. Another example I have used in my courses is the
clumps of dust that accumulate under beds or in closets of rooms that have wood floors. I typically find that about half the class has a name for these things (dust bunnies or dust monsters being the most popular), but about half does not. So, the mapping from concepts to words is incomplete. Although every content word must have a meaning (I will ignore primarily syntactic words like articles), this does not mean that every concept is connected to a word. So, words ≠ concepts.

A second view is illustrated in figure 11.2, in which every word is connected to exactly one concept, even though some concepts are not labeled by words. Again, ambiguous words are a problem here, as they have two meanings, and so must be connected to two concepts. There is no point in pretending that the meanings of bank are really a single concept. But if we agree to list ambiguous words twice, maybe this view is a reasonable one: Each word has one concept that represents its meaning. Here our ignorance of conceptual structure is something of a problem, since it is not entirely clear what “one concept” is. If our conceptual system is a highly interconnected set of facts and beliefs (as the knowledge view suggests, for example), then picking out a single concept could be difficult, since chopping the concept away from its connections would not correctly represent how it works within the conceptual system. Furthermore, couldn’t it be that a word picks out part of a concept, rather than the whole thing? For example, it seems likely to me that the word leap refers to one part of an event concept (it talks about the jumping off but not the landing). A word like toenail probably refers to a subsection of a con-
cept—a toenail is part of a toe and is probably represented within the toe or foot concept rather than being a fully independent concept. A final concern is that even unambiguous words often have a number of different, related senses. For example, theater can refer to the institution which puts on plays and the building in which one views the plays (e.g., “American theater is in the doldrums”; “I met her outside the theater”). These are clearly not the same thing—one is an institution and one is a building—even though they are highly related. We will be discussing this phenomenon in some detail below.

In short, we can’t stick with the view shown in figure 11.2, even for unambiguous words. Taking a conceptual view does not require that there be exactly one concept that represents a word’s meaning. Instead, we will have to be quite flexible about how concepts make up the meaning of a word, and much of the remainder of this chapter consists of a discussion of just this. For the moment, I would suggest the following principles for a conceptual approach to word meaning. First, word meanings are made up of pieces of conceptual structure. Second, an unambiguous word must pick out a coherent substructure within conceptual knowledge. (Ambiguous words pick out n coherent structures, one for each of their n meanings.) Third, when an unambiguous word has multiple related senses, the meanings are overlapping or related conceptual structures. For example, the two senses of theater discussed above are related by both containing information about plays.

Although it may seem that not much can follow from these principles, in fact, important parts of the psychology of word meaning can now be easily explained by referring to the psychology of concepts that we discussed earlier in the book. That is, principles of concept use will carry over to become principles of word meaning. Furthermore, there are a couple of corollaries that follow from these principles that are important, even if somewhat obvious. First, semantic content entails conceptual content. That is, if a word you know means something, that something must be part of your conceptual structure. Second, no semantic distinctions can be made that are not distinguished in conceptual structure. For example, you couldn’t distinguish the words chair and stool if you didn’t perceive the difference between these kinds of things and have that difference represented in your concepts of furniture.

One concern that people sometimes have about this psychological approach is the idea that conceptual representations are necessarily private—they are inside your head, rather than being publicly available. This is a concern because word meaning appears to be public, in that there is a meaning for the word dog in the language, which must be shared by all speakers in order to be understood. If everyone has his or her own private concept, then on what basis could we communicate (J. D. Fodor
1977)? This argument is to some degree correct, but it ignores the fact that people do not associate any old concept to a word. Instead, they learn through socialization which concepts go with which words. So, as a child, you learned that dog refers to a certain kind of animal. If you first developed the hypothesis that dog refers to any four-legged mammal, you would soon find yourself miscommunicating with people. They would not understand you when you referred to a sheep as dog, and you would not understand them when they said that all dogs bark, and so on. Thus, there is a social process of converging on meaning that is an important (and neglected) aspect of language (Clark 1996). The result of that process is that different people within a community relate words to very similar concepts. Even though my concept of dogs is an internal mental representation, it has been shaped to be the concept for the word dog by many years' interaction with other English speakers. This process of convergence can be seen experimentally when speakers are placed into a novel environment and need to make up words or descriptions for the new objects (Clark and Wilkes-Gibbs 1986; Garrod and Doherty 1994; Markman and Makin 1998). So the private nature of concepts does not prevent them from being the basis of communication.

**Theoretical Arguments for Conceptual Representations**

Why should we say that concepts are our representation of word meaning? There are empirical reasons based on experiments that I will discuss in the next section. However, there are also theoretical reasons for why this is the most likely representation. To understand why this is, we need to take a step back and consider why people have concepts. Concepts represent our knowledge of the kinds of things in the world. They allow us to identify a new object as a dog or a shoe, and they then allow us to infer unseen properties of the object, such as its likely behavior or function. They also help to explain the properties or actions of different objects. Thus, concepts provide critical information for our interactions with objects and our participation in events.

When we communicate, we use what we hear to control our interactions with the world. For example, if someone warns us “There is a mad dog in that yard,” we can infer various actions that would be appropriate and inappropriate based on this information. We don’t need to be actually bitten by the mad dog in order to take action; the linguistic warning obviates that necessity. It has often been argued that this aspect of language, communicating information critical to survival, is the reason that humans have evolved a highly sophisticated capacity for communication.
Since we represent our information about the world in terms of concepts, language needs to make contact with those concepts in order to be useful. For example, if someone tells you “Never approach a dog with your hand raised over its head” (good advice), this will only be useful to you if the word dog makes contact with your concept of dogs. That is, in a different situation, perhaps weeks or months later, you need to be able to categorize an object as a dog and then recall the advice that was given linguistically, and turn that information into behavior. You need to make a connection between the word dog and that dog, between the word hand and your hand, between the word approach and your present action of approaching, and so on. Since the way we represent the world is via our concepts (as just defined), the way to use linguistic information is to have it make a change in our conceptual structures. That is, the information in such a sentence would need to be added to our concept of dogs in order to be put into use.

For linguistic meanings to alter our knowledge of the world, it would be most efficient if they were directly connected to our concepts. In contrast, suppose that linguistic meanings were purely linguistic, that meanings were represented in “semantic features” that were not part of our conceptual knowledge, much like syntactic features probably are. Now sentences would not influence our conceptual structures and therefore influence our interaction with the world, because they would only be representations in these linguistic features, and not in the concepts we use to interface with the world. In short, you would now go into the yard with the mad dog. (One might naturally attempt to augment this view by suggesting that meaning is purely linguistic but that those linguistic entities are connected to our conceptual structure as well, which allows language to influence our understanding of the world. However, the semantic features in this theory turn out to be completely redundant with the conceptual structures. They could be eliminated without any loss; Murphy 1991.)

Empirical Evidence for Conceptual Representations

We are not limited to purely theoretical arguments for this position: There is overwhelming empirical evidence for the conceptual basis of word meaning. The evidence discussed below follows a consistent pattern. I will identify a phenomenon that is found in the concept literature and then show that it is also found with words and/or in linguistic tasks. (Indeed, many of the experiments discussed up until this chapter have used words as their stimuli, although I did not draw attention to this aspect of them.) The argument, then, is that if a general property of concepts is
also found in the use of words, this is evidence that the words are represented via those concepts. Although one such property might arise through coincidence, there is evidence for many such properties throughout the psycholinguistic literature. Indeed, I do not know of any phenomenon in the psychology of concepts that could conceivably be found in words that has not been found. If word meanings are not represented in terms of concepts, then they must be represented in terms of something else that just happens to have the exact same properties as concepts. By Occam’s razor, I will conclude that word meanings are represented in terms of concepts.

Category Effects
This evidence directly reflects the use of categories in sentence comprehension. Federmeier and Kutas (1999) performed a reading study in which they sampled brain waves (Event-Related Potentials, or ERPs) while subjects read sentences. The materials consisted of sentence pairs that created an expectation for the final word. For example:

They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of . . .

Here subjects would have a strong expectation of the word palms. In one condition, the expected word was actually shown. In another, an unexpected word from the same category was shown, for example, pines for the above example. In the final condition, an unexpected word from a different category appeared, like tulips. When unexpected words appeared, the ERPs revealed a component known as the N400 (a negative wave occurring about 400 ms after the onset of the word), which often indicates semantic anomaly or an unexpected event. Importantly, the N400 was significantly less when the presented word was in the same category as the word that was expected; for example, the N400 was smaller for pines than for tulips, even though both were inconsistent with the context.

Federmeier and Kutas argue that the results are due to subjects forming an expectation for what word would appear at the end. The category effect reveals that this semantic expectation involves the category structure of semantic memory. That is, items in the same category are more related than are items in different categories. What is surprising about this effect is that it is found even though no categorization test was involved—only reading—and that the category relation was between the presented word and a word that was not actually shown. Category-level information, then, seems to be involved in sentence comprehension.
Typicality Effects

As chapter 2 reviewed in detail, there have been many demonstrations that not all category members are equal. This is revealed both in unclear category members (items that are not clearly in or out of the category) and in typicality structure. This kind of result is found with every type of stimulus that has been investigated. For example, it is found with dot patterns (Posner and Keele 1970), alphanumeric strings (Rosch and Mervis 1975), patches of color (Noshofsky 1988), geometric shapes (Medin and Shaffer 1978), and so on. Thus, it is a phenomenon that is not specifically tied to linguistic materials by any means. Nonetheless, typicality gradients have been found in innumerable tasks using linguistic stimuli.

For example, Rosch’s first studies of typicality were of color names (see Rosch 1973b), where she found that there is a small range of hues that is prototypical of a color term, and other hues are perceived as being less typical even if they are clearly in the category. There are also unclear members that are not definitely in or out of the category (is teal a kind of blue or green?). Such findings led her to find similar results for superordinate categories like vehicles, weapons, and clothing (Rosch 1975). At the same time, research in semantic memory was finding that words like bird and animal had typicality structure that greatly influenced their use. Rips, Shoben, and Smith (1973) found that it took longer to verify sentences like “An eagle is a bird” than “A robin is a bird,” where eagle is less typical than robin. This result has been replicated many, many times since then. In a clever task, Rosch (1977) asked people to generate sentences that included category names (like bird). She then took those sentences and replaced them with typical or atypical category members (like robin and goose). She found that the sentences replaced with typical members were rated as much more natural than the ones with the atypical members.

These experiments mostly involved various kinds of judgments, such as sentence verification or ratings. However, typicality also influences naturalistic sentence processing tasks. For example, in language production, people are more likely to produce a typical item before an atypical item. Kelly, Bock, and Keil (1986) asked subjects to repeat sentences, after a delay. Often subjects made slight errors in repeating the sentence. When a sentence contained two category names, the errors were more likely to involve placing the more typical item first. So, subjects were more likely to say “The child’s errand was to buy an apple and lemon at the fruit stand” than “The child’s errand was to buy a lemon and apple at the fruit stand.”

The authors argued that the sentence planning process was sensitive to the typicality of concepts, which in turn influenced the accessibility of words. That is, apple comes to mind before lemon, because it is a more typical fruit.
Garrod and Sanford (1977) examined the comprehension of anaphoric noun phrases based on category relations. For example, if I am talking about a thief, I can later say “The criminal escaped,” and you would understand me to be talking about the same thief. This is because a thief is a kind of criminal, and so it is likely that I am talking about the same person. Garrod and Sanford compared examples like the following:

(1) A tank came roaring around the corner.
   The vehicle nearly flattened a pedestrian.

(2) A bus came roaring around the corner.
   The vehicle nearly flattened a pedestrian.

Here, the *vehicle* in the second sentence refers to either the tank or bus mentioned in the first sentence. Clearly, a bus is a much more typical vehicle than a tank. Garrod and Sanford predicted that it would be more difficult to understand the anaphoric relationship between these terms when the antecedent was atypical. They gave sentence pairs of this sort to subjects to read, measuring the amount of time they spent on each sentence. Garrod and Sanford found that subjects took longer to read the second sentence of (1), which has the atypical category member as the antecedent. Furthermore, if the sentences exchanged the category terms, as in (3),

(3) A vehicle came roaring around the corner.
   The tank/bus nearly flattened a pedestrian.

an analogous result was found: When the second sentence included *tank*, it was more difficult to understand than when it included *bus*. (The first demonstration is more convincing, because it is the exact same sentence that is being compared across conditions.)

In this experiment, then, normal comprehension of sentences was sensitive to the typicality of category items, suggesting that in understanding what the *vehicle* referred to, readers were accessing conceptual information about vehicles and then comparing it to information about tanks or buses. Because more typical things are more similar to their category representations, the anaphoric connection is easier to establish for them.

Studies of word learning have long found that children tend to learn typical referents of words before atypical ones. For example, Anglin (1977) found that children could correctly label typical animals, but not atypical ones, as *animal*. In fact, they were able to classify unfamiliar, yet typical animals like wombats and anteaters when they were not able to classify familiar but atypical animals (from the child’s perspective) like butterflies or ants as being animals.
In sum, the findings of typicality throughout the concept literature have parallels in studies of language comprehension, learning, and production. Indeed, many of the first demonstrations of typicality effects, by Rosch, Smith, and their colleagues, were done with linguistic stimuli. Although not focusing on experimental evidence, Taylor (1995) has written an entire book on typicality effects in language (mostly in semantics, but also in phonology and syntax). (This is in fact a very useful introduction to category structure from a linguistic perspective. It is especially good at discussing cases other than noun concepts, which are almost the only topic in the psychological literature.) All these examples provide evidence for the view that word meaning is represented in the conceptual system.

**Basic Level Effects**

Another important phenomenon of conceptual structure is that of the basic level (see chapter 7). Although objects can usually be described at a number of levels of specificity, there is usually one that is preferred, for example, table, rather than more general categories (superordinates) like furniture or object, and rather than more specific categories (subordinates) like work table, dining room table, or wooden dining room table. Even artificial categories can have a basic level (see Murphy 1991; Murphy and Smith 1982), and so it is not a linguistic phenomenon. Nonetheless, it is found in a number of tasks involving words for categories.

The most fundamental finding is that in object labeling, people prefer to use the basic-level name to other names. That is, when shown a table, they prefer to call it *table* than to use a more specific or general name (see Rosch et al. 1976; and also Cruse 1977; Lin, Murphy, and Shoben 1997; Morris and Murphy 1990). In text, basic-level names are by far the most likely label for a single object, whereas superordinates are typically used only to refer to multiple objects (Wszewski and Murphy 1989). One can also see basic-level effects in the language as a whole. Linguistic anthropologists have noted that there is usually a preferred level of names in the language, such that basic-level categories have a single-word name (like *table*), whereas more specific categories often involve multiple morphemes (like *dining-room table* or *speckled trout*). (Berlin, Brecklove, and Raven 1973). Another difference is that basic-level names are usually count nouns whereas superordinates are often mass nouns (Markman 1985). For example, I can say “I have two couches,” but not “I have two furnitures,” because *furniture* is a mass noun, like *sand or water*. So, I can say “I have some furnitures” (with no plural) just as I can say “I have some water” (with no plural). But I can’t say “I have some table.” Markman (1985) showed that this difference is found in a number of different languages and is not just a quirk of English.
Finally, children find it much easier to acquire basic-level categories than more general or more specific categories. This is true both in novel category-learning experiments (Horton and Markman 1980) and in everyday vocabulary acquisition (Anglin 1977; Rosch et al. 1976, Experiments 8 and 9; and many other studies—see chapter 10). Thus this aspect of conceptual structure influences vocabulary acquisition. However, it is also true that parents tend to use basic-level names in talking to their children (Brown 1958a)—perhaps that accounts for the acquisition data. It is likely that this is the cart rather than the horse (see chapter 10 for discussion). That is, children would find it difficult to understand other terms, and so parents speak in a way that allows them to be understood. As a general rule, the concepts that are most useful and easiest to learn are the ones that children learn the words for first. This provides further evidence that meaning is based on conceptual structure.

Conceptual Combination
The main question of conceptual combination is how people understand phrases like apartment dog, mountain stream, peeled apple, and sport that is also a game. The very definition of conceptual combination, then, is in linguistic terms—how people understand phrases. As I will discuss in chapter 12, it is difficult to tell whether conceptual combination is a question of concepts or of language understanding. The two perspectives melt together here, suggesting that there is a very close connection between them. Rather than repeat the information in that chapter, I will simply point out how it relates to the issue at hand.

Typicality raises its head again here, as one of the main topics of study in conceptual combination is in predicting the typicality of items in combined categories (Hampton 1988a,b; Smith et al. 1988). Murphy (1990) found that it was more difficult to understand an adjective-noun phrase when the adjective was rated as atypical of the noun than when it was rated as typical: For example, loud museum was more difficult than quiet museum. So, typicality is found in this linguistic domain as well. I argue in chapter 12 that an understanding of conceptual combination would require background knowledge of the sort referred to in the knowledge approach to categorization. This is another similarity. Wisniewski and Love (1998) have argued that the alignment of conceptual structures is important in understanding conceptual combinations, just as it is in other conceptual tasks.

Summary of Evidence
When one starts looking for it, there is a very large body of evidence showing that conceptual properties influence linguistic tasks just as much as they do nonlinguistic
thought or artificial categories. Indeed, in the concept literature, little theoretical distinction is made between experiments using known words and those using artificial materials. There isn’t one theory of concept learning or representation for dot patterns and a qualitatively different one for nouns, say. The reason is that, as just documented, the same phenomena occur in nonlinguistic concepts and in word use, and so any theory of one will serve to a large degree as a theory of the other.

For any one of these cases, it is possible to devise a counterargument, and occasionally one will read such counterarguments about a single such issue in the literature. For example, perhaps linguistic categories don’t really reflect typicality structure, but instead typicality differences are due to associations (e.g., robin is more associated to bird than chicken is). Perhaps the early acquisition of basic-level categories is a result of word frequency (parents using basic-level names more than others). And so on. However, to put it briefly, these arguments are unconvincing. Unless there is a competing theory that says how word meanings are represented and that also can independently find evidence for an explanation of each of these effects, a much more powerful explanation is that word meanings are represented by concepts. In this way, every conceptual effect that is also found with linguistic materials is automatically explained. If word meaning is not based on concepts, the findings of many parallels between the two is an unbelievable coincidence. So, I would suggest that we not believe it.

Learning Word Meanings

Word learning and conceptual development are discussed at some length in chapter 10. What I would like to do here is to discuss some theoretical questions that arise about word learning when one takes the conceptual approach. This discussion is largely speculative, rather than covering a well-specified area of empirical research. However, I believe that discussion is warranted, given that there are some potentially confusing problems relating word meanings and concepts in early childhood, when both are undergoing rapid development. Those who want to know facts should consult the earlier chapter.

From the perspective of a conceptual model of word meaning, there are two very general (oversimplified) ways that we can think of a word’s meaning as being learned. The first is that one could already have a concept and then learn its a name. Mervis (1987), for example, argues that children already have concepts for a number of words that they learn. The word learning corresponds primarily to associating the verbal label to this concept. For example, when her son learned the word duck,
Mervis believed that he already knew what ducks were, because of their distinctive shapes and behaviors and the constrained context in which he encountered them (at the lake in the park). When he learned the word, it was just a matter of acquiring the label for this kind of thing. Furthermore, Ari's initial use of the word *duck* did not correspond to what he heard from adults, but to his own (immature) understanding of what ducks are. There is experimental evidence that children will acquire categories without labels and then associate a new label to them when it is provided (Merriman, Schuster, and Hager 1991). On this view, then, the name is associated to an existing concept. E. Clark (1983) argued that children often identify lexical gaps—things in the world that they do not know the name for—which they then try to learn the names for or even make up their own name, if necessary. When this occurs, the concept is driving the word-learning process.

The second general way that a word meaning could be learned is that a new concept could be formed as a result of initial word learning. For example, if you went to a baseball game for the first time, you might not have the concept of a home run. However, in the context of hearing people talk about home runs, you might start to look for the distinctive attributes of home runs. By virtue of trying to understand what people meant by *home run*, you would be forced to form a new concept. In so doing, the usual processes of concept formation would be likely to hold. For example, you would find it easier to learn the new word's meaning if you viewed typical examples rather than atypical examples; if you were exposed to contrast categories, you would also learn how to distinguish that word's use from another's; if you saw a number of examples, you would learn the meaning better than if you only viewed one example; and so on. So, on this second story, you don't already know the concept of home runs and then learn the label. Instead, you hear the label a number of times and then try to work out what the concept behind it must be.

I would argue that the typical word-learning situation combines aspects of both ways of learning a new word. It is probably relatively rare that we have the correct already-formed concept ready to attach to a word; it is also probably rare that we really have no inkling of the concept before we hear the word. For example, imagine that you are a young child living in a big city. Further suppose that you are taking a trip to the country with your parents, and that you are going to learn the words *cow* and *horse* as a result of this trip. It is unlikely that you have no concepts at all corresponding to these things, because no doubt you have seen pictures of them, heard the words used, seen them on TV, and so on. On the other hand, you may not know much about them besides the fact that both are large four-legged mammals. Some of
these animals appear to be fatter and slower than others, but that's true of many kinds of animals—it doesn't necessarily put them in different categories.

On your trip to the country, then, you'll see cows and horses up close, and you may now be better able to spot some of the intrinsic differences between them. You can start to notice important differences based solely on exposure to the objects. (As we know from chapter 9, even infants can do this.) At the same time, your parents will name some of these animals for you. Since the labels cow and horse that you hear probably do not map perfectly onto any pre-existing classes you may have formed, you will look for differences between the objects so labeled (Clark 1983). That is, the use of distinctive names may force you to distinguish these somewhat similar concepts, whereas you might not have done so or might have taken longer to do so if you had not been learning the names. So, name-learning can influence or cause category acquisition, based in part on differences you were already noticing in any case: The fat animals with the udders seem somewhat different from the thinner ones without the udders; the first sometimes moo and the second neigh. The influence of the word does not in this case cause you to form an entirely new concept but may help you learn which of the apparent distinctions are important. For example, you can't distinguish which things are called cow and which horse based on their color, but you can based (partly) on the presence of horns. You can't distinguish them based on location, but you can (partly) based on whether they are being ridden. So, the names focus your attention on the features that are most useful for distinguishing the things being named and thereby influence the categories you form.

In this way, category learning and name learning can go hand in hand, sometimes over a very extended period. There is no simple relationship between the learning of the name and learning of the concept. In some cases, the categories are probably learned as a response to hearing different names. In other cases, children may notice an interesting difference between categories and then ask about it or realize that it corresponds to a naming difference. For example, a child may notice that a chihuahua looks different from other dogs and therefore become very receptive to learning a new name for it, or may even ask "What's that?" The linguistic input now might just be a label for the concept he or she has already formed. In most cases, the full meaning of the word is probably not acquired immediately, as some categories can take much time to fully grasp (e.g., the distinction between geese, ducks, and swans), especially if their members are not encountered that often. All this is due to the fact that children are learning about the world—and in particular, the classes of entities in the world and their properties—at the same time they are learning language. What
they learn in one feeds back to the other in an interactive way. This story would be much simpler if only children would first learn all about the world without learning any words. Perhaps after passing a standardized test, they would be deemed ready to learn the words for the things they know. Then language acquisition would be a simple case of paired-associate learning, in which a new label would have to be learned as the name of a known concept. However, language is so useful as a tool of learning that it would be very difficult to learn all one’s concepts without it. But this makes our story much more complex because when children first acquire the meaning of a word, they are likely not to have the full concept yet.

In short, both vocabulary learning and concept acquisition are moving targets. As conceptual structure develops, word meanings have to change to reflect that development. But as word learning progresses, this also creates changes in conceptual structure.

A final complication of the relation between concepts and word learning involves certain peculiar errors children make in using words, especially at the very beginning of word use. It is well known that children make overextensions and underextensions, or applying a word too widely and too narrowly, respectively. In fact, for any word, a child is likely to do both—include some objects that are not correctly labeled by the word, and exclude some objects that should be included (Anglin 1977; Clark 1973). Some of these errors are perfectly understandable as limitations of the child’s concepts (Mervis 1987). For example, if a European child does not recognize an ostrich as a bird, this would be completely consistent with the concept of birds that the child has formed based on his or her everyday experience.

In fact, such developments point out one advantage of the conceptual view, namely that as concepts change, word use will automatically change along with them. For example, as a child learns what the critical properties are to being a bird (or to being any biological kind), his or her use of the word bird will change along with it. So, when the child begins to understand the importance of feathers to birds, he or she now knows not to call planes or bats birds. Learning new facts will be reflected in language use, since the facts are incorporated into conceptual structure, which is the basis of word meaning.

Harder to explain are cases in which children use words to refer to objects that are of drastically different types. Clark (1973) provides a summary of many diary studies of early word use, and some of the children’s overextensions are worrisome. For example, a child’s word for the moon was also applied to cakes, round marks on a window, writing on windows and books, round shapes in books, and the letter O. Another child used the word fly to refer to flies, specks of dirt, dust, small insects,
his own toes, crumbs, and a toad. If we take these referents to indicate the child's categories, this would suggest that children's concepts are profoundly and qualitatively different from those of adults, a conclusion that is at odds with most studies of children's concepts (see chapter 10). So, one could argue, this means that the word meanings are not truly picking out the child's concepts.

There are two problems with this conclusion. The first is that the concept the child may be using could be perfectly normal but simply the wrong one. For the case of moon described above, the child is probably using the word to refer to an adjective concept of roundness rather than as an object concept referring to a type of celestial body. The child does not really think that the letter O, cakes, and the moon are all the same thing, but she does think that they are all the same shape. Similarly, the fly example suggests that the child is using this word to refer to some category of small, dirty things. The fact that the child used the word fly, which normally refers to a category of insects, makes us look for a similar object category behind his uses of this word. But there is no reason to assume that the child is using the word to refer to an object category, as it could be an adjective concept (presence of dirt?) instead.

The second problem with this conclusion is that children's intentions often go beyond mere labeling of an object, but this is often difficult to tell when they are capable of saying only one or two words in an utterance. For example, if a child points at a bare foot and says "shoe," she might be indicating that she has a strange word meaning that consists of feet and shoes. However, the child also might not be labeling the foot but could be making a comment, such as "You don't have your shoe on," or "You put shoes on your feet." The child could intend a directive, such as "Put your shoe on!" On the basis of a single short utterance, it is sometimes hard to tell what the child thinks the word really means. For example, the child who used fly to refer to a toad could have been talking about the toad eating a fly. There is a general problem that young children with few words must adapt the words they know to refer to a much wider range of meanings that they would like to communicate. So, even a child who can plainly see that a cow is not a dog may call the cow doggie, because that is the only word he or she has for any mammal. However, the same child may show good discrimination of cows and dogs in a comprehension task (see Anglin 1977; Clark 1983; Gelman et al. 1998).

In short, children's spontaneous labeling should not be taken as direct evidence of conceptual structure, especially when their vocabulary is very small. When they seem to be evincing some peculiar category structure, a more careful test needs to be done to discover whether the labeling is due to a reasonable, but incorrect meaning,
to communicative pressure, or to the intention to make a comment about something rather than label it.

Elaborations of Word Meaning

Now that the reader is fully convinced that word meanings are constructed from conceptual structures (or is at least tired of reading about it), I will turn to more specific phenomena and models of the psychology of word meanings. Much of this discussion is an attempt to bridge the gap between the static representation of words in the head and the dynamic process of comprehension, which colors our interpretations of words in actual discourse and text.

For the most part, I have been talking here as if each word is related to a single concept. If only this were true! I am not worried here about ambiguous words like bank. These are not really a problem, because their meanings are so distinct that there is little chance of confusing them. Most linguists would argue that the different uses of bank (financial institution and side of a river) are simply different words that happen to have the same phonology, for historical reasons. Therefore, its different meanings do not have to be coordinated in the lexicon. What is a problem for us to explain is unambiguous words that have complex meanings.

Much of this complexity falls under the rubric of polysemy, which means "many meanings," and which refers to a word having a number of senses that seem to be related but that are also distinguishable. Polysemy should not be used to indicate true ambiguity of the bank sort, where the meanings are unrelated. Virtually every content word of English is polysemous to some degree. Let's consider some examples. The word table appears to be a straightforward, simple word. However, if we start to think about it (and, I confess, if we consult a dictionary), we can come up with a number of distinct senses of this word, including:

- a four-legged piece of furniture: "Put it on the table."
- the class of all such items: "The table is America's most popular furniture."
- a painting, sculpture, or photograph of the piece of furniture: "The table is painted very soulfully."
- the stuff that makes up a table: "After the explosion, there was table all over the ceiling."
- the setting of a dinner table: "The table is not suitable for company; get the good silverware."
- food served at the table: "Count Drago keeps the best table in the country."
- the company of people eating at a table: "The entire table shared the paté."
• a flat or level area: “The ranch sits on a table of land south of the mountains.”
• an arrangement of words or numbers, usually in columns: “Table 3 shows the results of the experiment.”
• a synopsis: “The table of contents is incomplete.”
• (verb) to propose something and (verb) to put something aside (depending on dialect): “That motion has been tabled.”

In addition to these common senses, there are also a number of technical ones (listed in the dictionary), which are presumably known to some speakers if not to all.
• the leaves of a backgammon board
• a facet of a cut precious stone
• the front of the body of a musical instrument
• a raised or sunken rectangular panel on the wall

Imagine, then, that we have a subject who is a backgammon-playing jeweler who is also a native speaker of English. What is her meaning for the word table? It must include the central, specific uses (as in kitchen table), uses related to the central ones (people sitting at the table), representations (pictures of tables, presentations of results), technical uses (parts of diamonds), verb meanings, and so on. This meaning is a mess. It is not a single concept by any stretch of the imagination. Furthermore, among common words in English, this is a very typical situation. For example, Caramazza and Grober (1976) analyzed the word line and found 26 senses. They reported finding 40 senses for run. Perhaps surprisingly, it is the most common words that tend to be the most polysemous (Zipf 1945). (You can check this by looking up common and uncommon words in the dictionary and just counting the number of entries they have. The difference is obvious.) So, this is not a phenomenon that can be swept under the carpet, because it must be solved by language users every time they encounter a polysemous word. That is, if I say “I’m putting this on the table,” you must identify which sense of table I intend here, as it is a very different thing to add data to an array of numbers than to put an object on a backgammon board or to submit something for consideration.

There are two main questions to be asked here, and it is important not to confuse them. One is a purely psychological question, namely, how you represent the meaning of table so that you can understand its different uses. The second is why and how it is that word meanings get extended to have these different senses. The latter is not a purely psychological question but a question about the process of meaning change in a language—that is, part of linguistics (see Sweetser 1990, for an interesting discussion). Although I think that the second is also subject to psycho-
logical constraints (Murphy 1997a), we will focus on the first question here. However, the fact of polysemy reveals that it is apparently easier for people to take old words and extend them to new meanings than to invent new words (see also Markman and Makin 1998). Over the course of the history of English, it was found to be efficient to extend the word *table* from its usual meanings to the meaning of a face of a jewel and a display of information. Speakers could have invented a new word, or borrowed one from another language, or used a longer expression ("half of a backgammon board"), and so on. But given the very high level of polysemy in the language, using the old words is apparently the preferred route even if it results in very complex word meanings.

One thing to notice about the different senses of a polysemous word is that although they are related, they can end up picking out very different things (Nunberg 1979; Taylor 1995). There is an obvious relation between the piece of furniture and the people seated around that piece of furniture, but these are also very different kinds of entities. One is a single object, and the other is a set of individuals; one is inanimate, the other is animate; one is defined by its shape and function, the other by its present location. To put it another way, these senses pick out entirely different categories. The same is true of the representation sense (a painting of a table would not make a good table), the table-setting sense, the jewelry sense, and the verb senses. Although it is not entirely clear whether we should consider the noun and verb forms of *table* as belonging to the same word (Caramazza and Grober 1976 argue that we should, but the forms must be separated at some level for syntactic processing), one can see that they are related meanings, even though they are very different categories (objects vs. actions).

Given this difference in the kinds of things picked out by different senses, it is inescapable that the word *table* must pick out a number of different concepts. Rather than there being some constant meaning across all uses of *table*, there must be at least a few different concepts being used. This process has sometimes been called *chaining* (Lakoff 1987; Taylor 1995). The idea is that a word with a certain use might then be extended to refer to a related kind of thing; this new use can then be extended to something that is related to that; and so on. With enough links in the chain, one word can get extremely different senses, even though the adjacent senses are related (Nunberg 1979). For example, Cruse (1986, p. 73) pointed out:

In the case of *mouth*, if one knew what an animal's mouth was, and one were to hear, for the first time, a reference to the *mouth of a river*, I surmise that there would be little difficulty in construing the meaning; but suppose one were familiar only with *mouth* used to refer to the mouth of a river, and one heard a reference to the horse's *mouth*, it is by no means certain that one's attention would be directed to the appropriate end of the horse!
Similarly, a word like paper can refer to a substance made of wood pulp ("This shirt is actually made of paper") and also to the editorial policy of a newspaper ("That paper is biased against the president"). It is easy to reconstruct the chain of meanings that leads to these two senses (i.e., from the wood pulp substance to sheets printed from that substance, to a publication printed on those sheets, to the management that produces that publication, to their editorial policy), but it is striking that two extremely different concepts get picked out by the same word. In anyone's concepts, wood pulp is not much like editorial policy. It is perhaps not surprising, then, that subjects rarely treat different senses of the same word as being in the same category in a standard category sorting task (Klein and Murphy, in press).

A few linguists have argued that different senses of a word really reflect a common core meaning, which is found in every use of the word (most notably Ruhl 1989). However, examples of the sort given above cast strong doubt on this theory (Rice 1992). In fact, these kinds of examples make one wonder if there is any limit on what senses could be stuck together in a single word. Generally, we can assume that a word will only be extended to something that is related to its other uses. But words that have very distant senses like line or paper show that all of the meanings do not need to be similar. Rather, it seems necessary only that any sense be predictable from the other senses you know (Nunberg 1989). That is, if you knew only the meaning of chicken referring to a barnyard animal, could you understand what someone meant when they said that there was chicken in the stew? Even though the stew does not appear to contain a whole chicken of the sort you are familiar with, you could infer that it is the meat of the chicken that has been incorporated into the stew, rather than the whole animal.

The process of chaining has been supported by the findings of Malt et al. (1999), who looked at the names given to a set of objects they collected from supermarkets, kitchens, and magazines. All of these were containers of some sort, but they received a variety of different names, such as bottle, container, jar, jug, and box. In many cases, these items were variations on older technology, and so they did not exist when these words first came into use. That is, a plastic juice container in the shape of a cartoon character (labeled a juice box by subjects) is something that did not exist ten years ago, much less when the words bottle or box first came into use. Thus, using one of these words to refer to this new object is an example of chaining. The invention of an object leads to a change of meaning when an old word is extended to cover this new thing.

Malt et al. found that the names for these objects could be best predicted by the names of their neighbors via a chaining process. It was not generally the case that all
boxes were similar to a prototypical box. Rather, the plastic cartoon-character juice container was thought to be similar to other plastic juice boxes (without the cartoon character but containing juice and a straw), which were similar to cardboard juice boxes, which were similar to the most typical box (e.g., an open, squarish cardboard box). Malt et al. suggested that when a new form of packaging was developed, it took its name from the most closely related item. In some cases, it is not clear whether this is an on-going psychological process, or a historical process. That is, do people looking at a novel shampoo container think (in essence) “It's very much like other shampoo bottles, so I'll call it a bottle”? Alternatively, perhaps the manufacturer, which has been putting shampoo into bottles for 25 years, decides to call its new containers bottles as well, which therefore becomes the conventional name for it. Perhaps a small number of individuals decide on the name, which is then adopted by the community at large, without considerable thought on their part.

What does all this mean in terms of psychological representation? First, as noted, the word meaning is not a single concept or conceptual core. At least, no such core has been suggested for words like table, paper or line that would cover all their meanings. For such words there must be a one-to-many mapping from words to concepts. Given that, the learning problem is not exactly as we have been assuming earlier in the chapter, in which one simply (ha!) had to figure out which concept a word corresponded to. Now, one has to distinguish different senses and identify the concepts for each of them. So, after learning the word table as referring to kitchen, dining room, and coffee tables, the child also has to learn the sense in which it refers to people seated at the table. And in some cases, these senses must be kept separate, because they are fairly distinct. That is, one usually doesn’t use the word table to simultaneously refer to the conglomeration of the people and the object (a restaurant server wouldn’t say “The table is wobbly and wants to share the appetizers”; see Cruse 1986).

It is likely, however, that learning each specific sense is not necessary for all sorts of polysemy. Some patterns of polysemy are very common across the language (Nunberg 1979), and they can be immediately extended to other words. Imagine that you just saw an unusual animal at the zoo, about which you overheard someone say “Look at the zork.” Although you have only heard this word used to refer to a single animal, you would no doubt immediately understand uses that refer to the entire class of animals (“Zorks have only one kidney”), to representations of the animal (“My zork is the best sculpture in the show”), or to the meat of the animal (“Zork tastes kind of like chicken”). This is because these types of polysemy are found over and over in the language.
Indeed, Murphy (1997) found that people are very willing to extend new words in familiar ways of this sort, even though they had never heard them used in such uses—alogous to the zork example. These extensions may be driven by a kind of meaning-extension schema, for example, one that says that the name of an individual object can be used to refer to the whole class of such objects. In Murphy’s experiment these senses could not have been learned from observation, because the words had never been observed in these uses. So, some of our meaning extensions are not directly learned and encoded in memory but are derived on the spot. To give a real-life example, it seems extremely likely that once DVDs (Digital Video Discs) become popular (as of this writing, they are not quite yet), people will spontaneously say things like “I saw a great DVD last night,” referring to the content of the DVD. That is, they will use the word DVD polysemously to refer to the object and its content, just as they do with CD, book, or film.

Confusing matters somewhat are cases in which one cannot extend a word in the way one might expect to be able to. For example, we often extend the meaning of a name for a tool to mean the normal action of using that tool. You can hammer a nail with your shoe; or you can truck a package across town in your car. Such extensions are productive and predictable, yet some such extensions are not possible. We do not say that we ovened the bread or that we carred the passengers. The reason for this is the phenomenon of preemption (Clark and Clark 1979; Lehrer 1990), namely that one cannot give a word a meaning that another word already has. The word bake already means to cook something in an oven; therefore, we cannot use oven as a verb with this meaning. We cannot say that we carred the passengers, because the verb drive would do perfectly well in this situation. So, one constraint on extending a word to new senses is that the sense cannot match the meaning of another word.

The psychological question that arises from this discussion is how stored and derived meanings combine to produce the correct sense of a word when it is used. That is, do people have a listing in their mental lexicon for chicken of the sense “meat of a chicken” or do they derive it from the possibly more basic meaning of the animal? In the zork example (hearing the word only once in a zoo), we can be sure that people derive the meat sense from the animal sense, because we know they have never heard the word used in the meat sense. The new sense could not have been stored in the lexicon, and so it must have been derived from a more general principle. Perhaps this approach is quite general. Even very conventional meanings might be derived from more general reasoning about what the word must mean. For example, Clark (1991) discusses the example of names for skin color. Typically,
people of different races are referred to as having *white*, *black*, *yellow*, or *red* skin in spite of the fact that actual skin color is generally not at all close to these colors. How is it that *white* can be the name of a skin color (in addition to its normal usage), when in fact the people described are often pink or beige or a motley collection of colors? Clark argues that of the skin colors we are familiar with, one color is closest to being white, and so we can identify it as being the one talked about. Furthermore, the word *white* is probably more familiar and shorter than a more accurate color description would be. Finally, the interpretation of *white* as a skin color depends in part on the other familiar skin color terms, so that it includes fairly dark-skinned Mediterranean peoples, for example, because they are not covered by *black* or *red*.

Clark does not turn this point into a specific proposal for processing of these words, but if we were to do so, it would suggest that most polysemy is not stored in the lexicon but is figured out on the fly, using the kind of reasoning process just described. However, I suspect that Clark’s analysis is more likely to apply to the question of how conventional word senses develop historically in the language rather than to explain how people understand the words in fast-moving conversations. That is, people have heard the words *white* and *black* used to refer to skin colors hundreds or even thousands of times. It would be peculiar if they did not at some point develop a memory structure that stored this particular word sense but instead kept deriving it time after time. It seems unlikely that people will think, even implicitly, when they hear someone described as *white*, “Although his skin is not actually white, the light skin of Caucasians is closer to being white than it is to other basic color terms, and *white* is a more familiar word than....” Above I noted that we could use the word *zork* to refer to the meat of a zork even if we had never heard the word used that way. But with very common uses, we may bypass this derivation process and simply store the result. Since we talk about chicken meat more often than we talk about the animal (in my house, anyway), it is likely that we have formed the specific representation of *chicken* meaning chicken meat, rather than drawing an inference something like “She could not mean for me to chop up a whole live chicken here, so perhaps she is just referring to the meat of the breast of a chicken, which I happen to have in the refrigerator,” every time *chicken* is used to refer to meat. Again, after a few hundred uses of the word to refer to meat, people probably have learned this convention, even if it is derivable from general principles (Kawamoto 1993).

Experiments by Klein and Murphy (2001) suggest that some senses of a polysemous word are stored separately, rather than derived. They found priming when a
word was used twice in the same sense and interference when the word was switched from one sense to another. These results seem most consistent with the idea that the senses are represented and accessed separately rather than being derived each time they are encountered. In fact, the priming pattern for polysemous words like paper was very similar to that of homonyms like bank (though cf. Pickering and Frisson 2001).

The linguistic literature has generally focused on two possibilities: whether the different senses of a polysemous word are stored in the lexicon or are derived (e.g., Rice 1992; Ruhl 1989). However, it is possible or even likely that there is a continuum between these situations. Clearly, if you extend the new words DVD or zork to have a different meaning than you have ever heard before, you are deriving the new meaning. If you daily use the word chicken to indicate a kind of meat, you have probably stored this sense in your lexicon. But in between these cases may be words that have been extended only a few times and so are somewhat familiar in a different sense but are not very strongly represented in that sense. For example, perhaps you have only heard a few times the word elm to refer to a type of wood, but have often used it to refer to a type of tree. If I said to you “This table is made from elm,” you would have two ways of understanding this: the general principle that a name for a plant can also be used to refer to the substance that makes up that plant and retrieving the lexical entry saying that elm can refer to a kind of wood. Because the use of elm to refer to a kind of wood is not very well established in your lexicon, it may not dominate the comprehension process. That is, it may be that both of these processes are used in understanding the word in this way. There is as yet no experimental evidence that illuminates how these different sources of information are combined during comprehension.

It should be pointed out that at least some uses of words must be derived, namely what are often called nonce uses, in which a word or expression is used to have a meaning that is peculiar to that particular occasion (Clark 1983). For example, consider the case of a man who mainly paints still-lifes as his hobby. His wife mainly paints gardens and farms. If the man’s wife was trying to get him to change, he might say, “She wants us to paint the same things. She’s trying to turn me into a farmer” (Gerrig 1989, p. 199). Here, farmer doesn’t mean anything that is stored in our lexicon: It means “someone who paints farms.” This, however, is a use that can only be understood within that particular context, rather than being derivable from the meaning of farm on general principles. Not surprisingly, such innovations are usually somewhat harder to understand than using the word in its usual sense (for farmer, “someone who runs a farm”) (Gerrig 1989). However, they can nonetheless
be understood, and in some situations are very common indeed. For example, it is common (in English) to take a noun and turn it into a verb, as in “The boy has been porchng the newspaper recently” (Clark and Clark 1979; Clark 1983). The ease with which such innovations get uttered and understood reflects the constructive aspects of meaning, in which information from the discourse itself (say, a discussion about newspaper delivery) and from general world knowledge (that boys often deliver newspapers by tossing them near the front door) are accessed to construct a specific meaning that is not already in the lexicon (toss onto the porch). Such constructive processes are probably the same sort that are used in contextual modulation more generally, as will be discussed in the next section.

In summary, then, polysemy creates a real problem for any theory of word meaning, because it can result in radically different senses being picked out by the same word. Even though the senses are usually related to one another, perhaps through a chain of similarities, the categories of things being referred to can be very different. Polysemy has not received the attention it deserves in psychological theories of word meaning, and this is why this section of the chapter is heavy on linguistic examples and general pronouncements, and light on evidence. My analysis suggests that both main proposals for how polysemy could be handled are probably true. In some cases, the sense must be derived, and in others it is most likely retrieved from the lexicon. But the variables influencing each and the interaction of the two processes are simply unknown. The field will need to address this problem, because theories of word meaning that neglect polysemy cannot be complete. Furthermore, any attempt to build a computational language-comprehension device obviously cannot be successful if it ignores this issue. Perhaps it is from this direction that more concrete suggestions for handling the problem will come.

Indeed, Kawamoto (1993) touched on polysemy in his connectionist model of ambiguity resolution. The model represents semantics as a kind of space of features and feature combinations, and a given concept is a location in space corresponding to its particular set of features. This location is not a single point, however, but an area that encompasses all the related uses of the word. Although *dog* might refer to animals of a single species with many common properties, the individuals referred to also differ in size, color, personalities, behaviors, and so on, and so different uses of the word *dog* might pick out slightly different locations in space. According to Kawamoto, different senses of a polysemous word will become established in the lexicon when they are distinct enough and frequent enough to establish their own “local minimum” in that space. If there is a clump of uses of a word that is noticeably distinct from another clump of uses (e.g., *chicken* referring to an animal vs. the
meat), the model will act as if there are two different senses. If the uses are infrequent and evenly spread throughout the space, then the model will not distinguish different senses. Although Kawamoto (1993) did not implement his ideas about polysemy in his running model, this is an interesting proposal that has the prospect of dissolving some of the fuzzy distinctions of senses that arise in this topic. That is, rather than worrying about whether two similar uses of a word are the same or different senses and therefore whether they are represented separately, the model would directly represent the continuum of word uses.

A final issue is how the phenomenon of polysemy speaks to the question of the conceptual basis of word meaning. Clearly, it complicates matters considerably. However, the complication is not due to representing meanings by concepts but is in the actual structure of the lexicon. The fact that people use words like paper to refer to a substance, a copy of a daily publication, and an editorial policy means that the word's representation must be complex. The conceptual theory provides some basis for the complexity of the meaning, because one can see the relation between the senses in terms of their content. That is, it is because a publishing company establishes its editorial policy, which it prints in individual copies of a newspaper, which are made up of that substance, that the same word can be used to refer to all these things (see Pustejovsky 1995). In contrast, imagine we had a referential theory in which words were defined by the objects they picked out. There is no overlap between newspaper companies and copies of a newspaper or between editorial policy (indeed, it is not clear what its reference would even be) and a paper presented at a conference. Thus, the referential theory cannot explain why it is that the same word is used to refer to these different concepts, whereas the conceptual basis of word meaning at least has some basis for explaining the phenomenon of polysemy. This is not to say that it is completely understood, as linguistic study of the topic continues, and psychological study is just beginning. However, the conceptual basis of meaning at least provides a way of addressing the problem.

Modulation of Meaning

As we have already seen, words do not mean the same thing in different situations. Usually, it is the context, the setting and speakers or the rest of the sentence, that tells us which sense is the right one. If I say "The dog is in the yard," it is easy to infer that I must be intending dog to refer to an individual dog rather than to the class of all dogs, because it is very likely that an individual dog but not the entire class of dogs is in the yard. Furthermore, if I own a dog, and the person I'm talking
to knows this, then the person can further infer that I am talking about my particular dog. Similarly, a reference to chicken made in the kitchen probably picks out the meat sense, but outside on a farm, the animal sense is more likely. So, context can act to select the intended word sense. However, context can also specify in much more detail what is meant by a word, as I will discuss.

**Interactional Effects on Meaning**

Before getting to that, however, these issues must be put into a broader context. There is a tradition within experimental psychology of studying language in highly decontextualized settings. Typically, words or sentences appear on a computer screen with no apparent speaker, and the subject must make a judgment about the stimuli rather than responding as if the text were part of a normal conversation. Such a situation removes many critical aspects of language use, which some researchers have argued distorts our understanding of language processing. Clark (1996) describes many of these aspects. One concern that is relevant to our present discussion has to do with the differences in word meanings across communities. People who share experiences, specialized knowledge, hobbies, religion, and so on, may use words in a different way than those who do not share those things. If you and I are both jewelers, then you may understand my use of the word table to refer to a facet of a gem rather than a piece of furniture. If you didn’t think I was a jeweler, you might not understand my using the word in this way. People choose the vocabulary appropriate to the situation and their audience, and listeners understand words in the appropriate senses for those situations. In the most specific case, a word’s meaning is only fully known to the people involved in the conversation. For example, Garrod and Anderson (1987) had pairs of subjects solve a maze task. Different pairs used different and incompatible terms to describe the parts of the maze. If a sentence said by a speaker in one pair had been heard by a different pair, they would have misunderstood what was being described. More generally, during referential communication tasks, speakers start with long descriptions but soon shorten them to abbreviated versions that could not be understood without having heard the earlier, complete version (Clark and Wilkes-Gibbs 1986; Krauss and Weinheimer 1966). It is the earlier history of the conversation that allows the shortened version to be understood.

Exactly how people manage to keep track of all these different influences is a matter of some controversy (see Keysar et al. 2000; Polichak and Gerrig 1998). The controversy surrounds how fast and accurate people are at keeping track of all the information they share with people they are speaking with (e.g., Paul is a psycholo-
gist, soccer fan, and father, like me, but isn’t at my university or town, doesn’t share my religion or interest in beer, etc.). Does the jeweler who hears someone refer to a table mistakenly think that a jewel face is being referred to before realizing that the speaker is not a jeweler? Such psychological issues surrounding the production and comprehension of words are not yet settled. However, it is nonetheless clear that speakers and listeners use words differently depending on their shared communities and prior discourse.

Having said all that, I must point out that the psychology of concepts is of limited help in explaining some of these phenomena. If word meaning is built out of conceptual structures, this tells us that the senses of table must all be represented in conceptual units, for universally known senses as well as novel meanings that have arisen in a discourse. However, the psychology of concepts simply is little help in explaining the nature of personal interactions, episodic memories, and the prior discourse that influences meaning, because these are not conceptual factors. As a result, the remainder of this section focuses on factors more directly related to one’s conceptual and world knowledge. These influences arise not from interactions with a specific person but from more general knowledge of words in a sentence. Ultimately, both this sort of knowledge and information about one’s interlocutor will need to be integrated in a complete model of language use.

Contextual Modulation

In many cases, a word that has a general meaning is actually understood in a more specific way when it occurs in a conversation. For example, Anderson and Ortony (1975) contrasted the sentences, “The accountant pounded the stake” and “The accountant pounded the desk.” The action of pounding that we get from these sentences is somewhat different. And indeed, Anderson and Ortony found that a word like hammer was a better recall cue for the first sentence than for the second in a later memory test.10 This suggests that when the subjects read these sentences, they elaborated the meaning of pounded so that it had a more detailed representation, based on the situation that the sentence evoked. This process is called contextual modulation (Cruse 1986) or modification. In the first case, pounded is modulated to include hitting with a mallet or hammer, whereas in the second case, the same word is understood to involve hitting with a fist, though neither of these is explicitly said.

Roth and Shoben (1983) used measures of category structure to study contextual modulation. They presented subjects with sentences containing a category name and then asked them to rate the typicality of different category members. For example:
(4) During the midmorning break, the two secretaries gossiped as they drank the beverage.

(5) Before starting his day, the truck driver had the beverage and a donut at the truck stop.

Roth and Shoben predicted that tea would be a more likely beverage than milk for sentence (4) but that milk would be more typical than tea for sentence (5). Indeed, when subjects were asked to rate the typicality of items as being in the category named in the sentence, this is just what was found. More importantly, when subjects read a subsequent sentence referring to tea or milk, the tea sentence was easier following sentence (5) above, but the milk sentence was easier following sentence (4). That is, even though the same category name, beverage, is used in (4) and (5), readers apparently interpret this word more specifically, based on what particular beverages they think are more likely to be drunk in the described situation. This particular result is therefore directly related to category structures, suggesting that the Garrod and Sanford (1977) result described earlier can be modulated by context. That is, people will not simply read the word tank more slowly than bus when they are thinking of a vehicle; this result can be changed depending on which kind of vehicle seems more likely in the described situation. (See also Barsalou 1991, for more on contextual changes in category structure, though not specifically related to language.)

This process of contextual modulation of word meaning is not confined to general category terms like beverage. In fact, Cruse (1986) suggests that “the meaning of any word form is in some sense different in every distinct context in which it occurs” (p. 51). We should not overplay this point, as some of the differences in meanings can be very subtle indeed. However, we should not ignore the effect, either, which is more often done. Let’s just consider a few examples to get a sense of the differences in meaning:

(6) Audrey is visiting her mother in San Diego.

(7) Audrey is visiting her friend this evening.

(8) Audrey is visiting her bank.

(9) Audrey is visiting Madagascar.

Although the same verb visiting is used in each case, the act of visiting is somewhat different, due to the differences in the object of the verb and in the time specified (if any). So, Audrey’s visiting her mother in San Diego (said in Illinois) probably involves staying at her mother’s house and spending considerable time on and off
with her mother over the course of several days. But in visiting her friend this evening, Audrey is probably spending a continuous, but much shorter time with her friend, perhaps engaged in a specific activity (dinner, a video). Both (6) and (7) suggest that visiting is primarily a social occasion, but when Audrey visits her bank, it may be for a business meeting, for a brief period of time. Finally, to visit Madagascar, Audrey must have been gone for a fairly long time, and must have been continually in Madagascar during the visit. Furthermore, she obviously did not socially interact with “Madagascar,” but viewed the countryside, met the people, ate the food (or tried to avoid it), and so on. There is nothing particularly insightful or unusual in these observations. They simply point out that a word placed into different sentences can yield very different interpretations. So, it is a mistake to think that one can store a single word meaning that will exhaust the word’s significance in most contexts.

Cruse (1986, p. 52) describes the process of contextual modulation as “… A single sense can be modified in an unlimited number of ways by different contexts, each context emphasising certain semantic traits, and obscuring or suppressing others.” Consider the following example (based on one of Cruse’s):

(10) Carl poured the butter over the vegetables.

Here, we clearly understand the butter to be liquid, and therefore hot. As Cruse points out, however, it is unlikely that we have distinguished these different aspects of butter in our lexical entry for the word butter. We could make innumerable distinctions between different forms of butter, but we don’t have one lexical entry for butter that is cold and solid, another that is hot and melted, one that is in a big stick, one that is molded into little pieces shaped like leaves, and so on. Similarly, we may not have different entries in our mental dictionaries for visit in which one is a continuous, short social occasion, another is a long-term, discontinuous social occasion, another is a short, continuous business occasion, and so on. At least some of these contextual modulations are probably derived on the fly.

Murphy and Andrew (1993; see also chapter 12) asked subjects to provide synonyms or antonyms for adjectives. The words, however, were presented as parts of phrases, such as fresh water, fresh air, fresh bread, and so on. The results showed that the antonyms and synonyms supplied changed depending on the other word in the phrase. For example, a synonym for fresh water might be spring water, yet clearly spring bread is not a synonym for fresh bread. The adjective fresh seems to take on somewhat more specific meanings depending on the noun it modifies. The specific meanings led subjects to provide different synonyms in different contexts.
The effects of context, as noted in Cruse’s examples, or in Roth and Shoben’s experiment, give evidence for a knowledge-based interpretation process. Roth and Shoben’s subjects apparently believe that secretaries are more likely to drink tea than milk during their breaks; and they also believe that people are more likely to have milk with donuts than tea. This kind of information is part of our general knowledge of beverages, donuts, social activities in an office, and so on. Such information must become readily available from hearing the word during sentence comprehension. Similarly, Murphy and Andrew (1993) discussed examples that suggest that the antonyms and synonyms that subjects provided were not based on simple lexical associations but rather depended on a complex understanding of the situation being described. For example, one subject gave as a synonym for dry cake the phrase overcooked cake. Although dry and overcooked are clearly not synonyms in general (most dry things have never been cooked, much less overcooked), a cake could become dry through overcooking. So, this person accessed knowledge about the possible causes of a cake being dry in this apparently purely linguistic task of providing a synonym. (See also Hörmann’s 1983 experiments on the interpretations of quantified phrases like a few crumbs and a few children.)

Modulation is also found in verbs. Gentner and France (1988) presented subjects with simple subject-verb sentences that were somewhat anomalous, such as “The lizard worshipped” or “The responsibility softened.” Often such sentences can be understood if one or more parts are interpreted metaphorically. Gentner and France found, using a number of different measures, that the verb was much more likely to be interpreted metaphorically or undergo some other kind of meaning change than the noun was. For example, one paraphrase for “The lizard worshipped” was “The chameleon stared at the sun in a daze.” Here, lizard is interpreted literally, but worship has been changed so that it is something that a lizard might do. The authors propose that nouns have more cohesive meanings, at least when they refer to objects like lizards and cars. Verbs are often more relational and so are more susceptible to their meanings being stretched by the objects that they relate. Another reason, which the authors didn’t consider, has to do with word order. Since the word lizard appears first, with no context to suggest that it has anything other than its expected meaning, readers may be less likely to try to reinterpret it. In contrast, worship appears later and so may have to adjust itself to the meanings that have already been evoked.

Cruse (1986, p. 152) also suggests that contextual modification is greater for verbs and adjectives, because their meanings depend on the nouns that they are associated with. Saying that John crawled across the floor does not describe the
same motion as saying that a centipede did. However, whether there is more or less such modification for nouns vs. adjectives and verbs is very hard to say. For example, Cruse (p. 74) also points out that nouns can have different interpretations in contexts, as in (I have added some examples to his):

(11) the handle of...
   a door  
   a drawer  
   a suitcase  
   an umbrella  
   a sword  
   a saw  
   a shopping bag  
   a car door

As can be seen, the precise shape, function, and material making up these handles differs greatly, but this information can only be determined by the subsequent prepositional phrase. Thus, although verbs and adjectives may change their precise meaning depending on what they describe, nouns can undergo considerable contextual modulation as well.¹¹

Throughout the linguistic literature, there has been a debate on whether the lexicon is more like a dictionary or an encyclopedia. That is, does the meaning associated with dog include only the most critical (perhaps "defining") information relevant to the use of this word, or does it include the entire store of what we know about dogs? The concept view I have been proposing is closer to the encyclopedia view, because the concept of dog is what we know about dogs in the world, rather than some specifically linguistic information. However, this does not mean that every piece of information that one knows about dogs is actually retrieved and used every time one hears the word dog. One reason for taking an encyclopedic view, however, is to account for contextual modulation of the sort described here. For example, one can only know that the handle of a sword is a long thing, continuous with the sword's blade, by using one's knowledge of what swords are usually like. Similarly, the handle of a shopping bag is likely to be made of string or plastic, and to be an enclosed semicircular shape. We know this through our world knowledge of shopping bags. (Not very long ago, the expression "handle of a shopping bag" would have described something made of string, whereas now it would probably be made of plastic. This is a change in the world, which is reflected in our concepts of shopping bags, which thereby changes our interpretation of the phrase.)
Through the years, many linguists have become very exercised about the distinction between "real" (dictionary) meaning and other aspects of meaning that they thought were not truly linguistic. Such researchers might claim that the difference between the different kinds of handles is not one of linguistic meaning, but of something else (reference, discourse models, or interpretation). However, there is no psychological evidence supporting this distinction. If you eliminate the encyclopedic knowledge, pure dictionary meaning has great trouble in explaining the phenomena discussed here: how polysemy is resolved, how contextual modulation occurs, and how people come to more specific meanings for sentences than the linguistic meanings of individual words include. These are all important processes that must be dealt with in almost every utterance. (Keep in mind that conceptual variables are only part of the answer to these problems, as conversational interaction also plays a part.) Thus, it is not the case that "real" meaning is somehow more important than the encyclopedic meaning, because both must be invoked to explain comprehension. To my mind, this debate has become a dated terminological dispute, in which writers try to argue that some information is, but other information is not, part of the linguistic meaning of a word based on their prior theories of semantics instead of any empirical motivation. Nonetheless, even writers who believe that meaning is (or should be) like a dictionary must admit that real-world knowledge has enormous effects on how sentences in actual utterances are interpreted. Until there is clear evidence that the dictionary meaning is accessed and used separately from the encyclopedic meaning, there is little psychological point in distinguishing them. More detailed reviews of this debate may be found in many standard discussions of semantics and psycholinguistics, including Clark and Clark (1977), Cruse (1986), J. D. Fodor (1977), and Taylor (1995).

The encyclopedic aspect of meaning can be seen in the claim that many words' meanings only make sense when they are taken in the context of a background of knowledge and practice that is shared within a community. This point was probably first made most explicitly in the writing of Fillmore (1982) and was later taken up by Lakoff (1987). Fillmore pointed out that even when verbs seem to be well defined by a listing of simple semantic features, they often can only be used properly when it is understood what kind of situation they were intended to apply in. So, verbs like buy and sell only make sense in a society that has a conventional notion of economic transactions. For example, if I happened to give you a pen at the same time you gave me a gold coin, I could not describe this as "I sold my pen to you for a gold coin." To truly buy and sell something, the actors need to have certain intentions within a system of exchange, and these intentions are difficult to specify, be-
cause they take place in the context of an elaborate economic system. In another example, J. A. Fodor (1981) pointed out that the verb *paint* does not simply mean to cause paint to be applied to a surface. For example, if I put a small bomb in a jar of paint, and it blew up all over Jerry Fodor, I should not say "The bomb painted Jerry Fodor." This is not the notion of painting one has in mind. Rather, painting takes place within a general idea of decoration and protection of surfaces, in which the paint is carefully applied by an agent to a given surface in order to achieve certain results. Other physical actions that might happen to result in the same effect would not really be called painting, because they are not done within the context of decoration that is assumed when talking about painting.\(^{12}\)

In a final example, the word *bachelor* applies within a societal system in which men and women normally marry at a certain age. In this frame, a bachelor is a man who is eligible to marry but never has. But the Pope is an eligible adult male, and yet he does not seem to be a very good bachelor. Similarly, a gay male couple that has been together for 15 years does not seem to consist of two bachelors, even though they fit the definition. In both cases, the primary features of the word are fulfilled, but the assumptions and background underlying the word are not, as the people have removed themselves from the normal assumptions of marriage as it has traditionally occurred in our society—the Pope, by virtue of having taken a vow of chastity, the gay couple by their sexual orientation.

Thus, there is an argument to be made that word meanings cannot fully be represented by typical feature lists used to describe a single concept. These lists presuppose a certain understanding of the domain, and they only apply within that framework. For those of us who live within that framework, it can be very difficult to see its importance to the word's meaning, because its assumptions are not explicitly represented in each word, but within our overall conceptual structure. In short, word meanings may rely on more complex kinds of assumptions and knowledge than one might think on first glance.

The same thing can be found in concepts more generally. For example, we might tell a child that a dog is something that barks and has four legs. This is sufficient to pick out many dogs. However, we realize that if we were to take a cow and modify it so that it barked, or if we were to build a four-legged mechanical robot, we would not actually have a dog. Instead, a number of deep biological properties are involved in making something a dog or cat—biological properties responsible in part for generating the four legs and barking. People are sensitive to these properties even if they can't explicitly state them (Keil 1989). Such background knowledge is exactly the sort of thing emphasized by the knowledge approach to concepts and reviewed
in chapter 6. So, the claims made by Fillmore and others about word meaning are consistent with this approach to concepts.

**Associative-Computational Approaches to Word Meaning**

I have been focusing on a conceptual approach to word meaning here, but this has not always been the major approach to word meaning in linguistics and psychology. Given the nature of this book, it is not possible to do justice to the history of psychological models of meaning or to fully discuss alternative views. However, one general approach that has been important in the history of psychology has been based on word associations. That is, the meaning of a word is the set of other words (or perhaps words plus other mental entities) it is associated to. This approach has become more prominent again, due to some high-profile computational models that use text association to specify the meaning of a word. These approaches may seem at odds with the conceptual view proposed in the present chapter, and so I will review and critique them here.

**Past Computational Models of Meaning**

Before discussing the newest associative model, it is useful to consider early computational models of meaning and why they were found to be inadequate. Charles Osgood and his colleagues (e.g., Osgood, Suci, and Tanenbaum 1957) made one of the first quantitative attempts to measure meaning by asking subjects to rate words on 50 different scales of adjective contrasts, such as happy-sad or slow-fast. This test was called the **semantic differential**. A quantitative analysis showed that the 50 scales could be effectively reduced (through a factor analysis) to three orthogonal factors, which Osgood et al. called **evaluative** (e.g., good-bad), **potency** (e.g., strong-weak), and **activity** (e.g., tense-relaxed). Many words could be distinguished based on their values on these scales, and experimental work showed that words with similar scores were behaviorally similar in some respects. Furthermore, once the three factors were identified, Osgood et al. were able to find specific scales that measured each one, so that each word would not have to be rated on all 50 scales. This procedure could not be a total representation of meaning simply because it does not provide the information necessary to tell one when to use a word and what the word means when it is used. For example, if you didn't know the word *zork* but did know its values on the three scales (slightly negative, somewhat active, very potent), you wouldn't have a clue what to call a zork or what someone was talking about when they mentioned zorks to you. The three factors are far too vague to represent the
many components of word meaning. In spite of this shortcoming, Osgood’s work was very well known, because it was one of the first attempts to quantitatively model semantic information.

A more general and potentially useful attempt to represent meaning came about with the development of scaling procedures in the 1960s, such as multidimensional scaling and various clustering techniques. These procedures require information about the similarity of different words, which was most often obtained by asking subjects to rate the similarity of pairs of words. For example, how similar are dog and cow? dog and cat? cow and cat? and so on. These similarities between all pairs of words being studied would be input into a program, which would output a structure. In the case of multidimensional scaling, this structure would be a similarity space. That is, it would create something like a map in which words were the points, and the distance between them represented how similar they are. An example is shown for a set of food names in figure 11.3 (from data reported by Ross and Murphy 1999).

There are two general ways to interpret such scaling solutions. The first is to look for interpretations of the dimensions. Is there some overall difference between the items on the left and the items on the right and between the items on the top and those on the bottom? Sometimes one can find a consistent difference, which would suggest that this dimension is an important one in people’s representations of the words. (One can also rotate the entire solution, since the model only determines the distances between the items, not the specific orientation of the whole solution. Rotating the solution can often reveal meaningful dimensions.) For example, Dimension 2 in figure 11.3 might represent the time when these foods are normally eaten. Breakfast foods appear at the top, and dinner entrees at the bottom, with desserts and snacks in between. Foods that might be eaten with any meal (like water, milk, and butter) are in the middle. In a famous use of multidimensional scaling, Rips, Shoben, and Smith (1973) found that mammals differed primarily in size and ferocity. On one dimension, mice and cats differed from deer and bear (size), but on another dimension, the cat appeared with the lion and bear but was distant from vegetarians like sheep and cows. When one can find interpretable dimensions (which is not always the case), it suggests that this kind of variation is important in the domain under investigation.

Another use of such solutions is to look for clusters of objects. In figure 11.3, one can clearly see groups of meats, sweets, fruit, vegetables, and breakfast foods that are fairly distinct from the other clusters. In some cases, these clusters correspond to obvious categories that might well have been guessed in advance (like fruit and
Figure 11.3
A multidimensional scaling solution for the data of Ross and Murphy (1999) on similarity of different foods. (Although it is based on the data in that article, this solution was not published there.) Thanks to Larry Hubert for providing the solution.
meats), but in others, there might be surprising groupings that would not have been expected (like breakfast foods).

When such techniques were developed, there was some hope that they could provide a more quantitative basis for representing word meaning. One could find underlying bases for word meaning that subjects might not overtly express, and one might also discover surprising clusters of words that one did not realize would go together. In fact, both of these things do happen with some of these techniques. However, they also have severe limitations as representations of word meaning.

One of the limitations is much like the one described for the semantic differential, namely that the solutions obtained simply do not represent much of the meaning. Typically, multidimensional scaling, for example, results in two to four dimensions. However, word meaning is extremely complex, and no small set of dimensions can account for how words are used. For example, even if the scaling solution suggests that muffins and cereal are thought of as breakfast foods, this does not explain what the differences are between them, what should be called a muffin as opposed to cereal, that cereal usually has milk poured on it, but muffins do not, and so on. All of this information is used in deciding when to use the words muffin and cereal, but it cannot all be represented in such scaling solutions. Similarly, although pigs and goats are similar on the size and predacity dimensions, they differ greatly on other dimensions, which was not revealed in Rips et al.’s solution. If people only knew what the solution shows, they couldn’t distinguish actual pigs and goats. In short, scaling solutions often do not provide sufficient detail to account for people’s semantic knowledge.

Another limitation is that the solutions obtained depend greatly on the particular set of items tested. For example, Rips et al. tested only mammals and found that size and predacity were important dimensions. However, if they had tested only plant-eating animals, it is likely that predacity would no longer be a dimension, and some other difference among the animals would have become more salient and would have appeared in the solution (e.g., domesticated vs. wild), even though it was not present in Rips et al.’s original solution. The scaling solutions are context sensitive for two reasons. First, people’s similarity judgments are context sensitive. If one is judging many different kinds of foods, muffins and cereal may seem quite similar. However, if one were judging only breakfast foods, muffins and cereal would now seem quite different, as the differences between baked goods and loose cereals would be much more salient. Second, the scaling procedures are also context sensitive in that they attempt to account for the largest differences in the data. That is, muffins and cereal are extremely far from lobster and hamburger in people’s judgments, and
the scaling solution must move these items as far apart as possible in order to account for these differences. Smaller differences within clusters (e.g., bagels are a bit more similar to bread than to muffins) are not accounted for as well, because they do not account for as much variance in the ratings. If one were simply to remove the meats from the data included in figure 11.3, one might find a better representation of the breakfast foods.

In any case, the result is that quantitative scaling solutions have been useful for some purposes, especially in exploring unfamiliar domains. However, as a representation of people’s knowledge of words, these techniques have not been as helpful as was originally hoped for when they were developed. Their main problem is that they do not provide enough detail to account for word use. For reviews of such scaling techniques, see Van Mechelen et al. (1993) and Shepard (1974). As we will now see, somewhat different, more powerful techniques have recently been developed that may overcome the problems of these scaling techniques.

Recent Associative Models of Semantic Memory
Fast processors and large computer memories have spurred the development of a new form of semantic analysis, called Latent Semantic Analysis (LSA, Landauer and Dumais 1997) or Hyperspace Analogue to Language (HAL, Lund and Burgess 1996). This computational technique is based on an associative approach to word meaning. Associationism has a very long history in psychology, most importantly dating from the work of British nineteenth-century philosophers who believed that all thought could be explained by strings of associations. In the early twentieth century, associationistic approaches to meaning were popular because they required reference to little internal cognitive processing. Within the context of American behaviorism, the free association task seemed to be an excellent measure of stimulus-response learning, and many psychologists were able to convince themselves that actual language production and comprehension were simply a matter of stringing together longer and longer series of associations. As a result, mental representations and complex mechanisms would not be needed to explain language use. The failure of such theories was a major event in the founding of cognitive psychology (Chomsky 1959), but associations between words is a fact of life that is still studied (perhaps over-studied) by many researchers in lexical processing.

The recent associative enterprises go beyond the simple idea that word 1 is connected to word 2 by an associative link. They suggest that a word can be characterized by its entire set of associations to all other words. That is, people are exposed to tens of thousands of word types, within millions of sentences that they read and
hear. Certain words tend to occur together and other words hardly ever co-occur in people’s experience. On this view, these patterns of co-occurrence represent meaning: Words with similar patterns have similar meanings.

Both the models mentioned above use very large text corpora as sources of information about how words are associated. For example, Landauer and Dumais (1997) input 4.6 million words from a student encyclopedia to their LSA model, finding 60,768 word types. They calculated how often each word type occurred with each other word type in the same article in the encyclopedia, creating a 60,768 × 60,768 matrix of co-occurrences. (Lund and Burgess’s technique weights co-occurrence scores by how close the two words are in the text.) A complicated mathematical technique was used to reduce this matrix to a more reasonable size of 300 dimensions. (Burgess and colleagues do not always reduce the matrix.) Each word had a value on each of these 300 dimensions, which represented the association information in the entire text. Using these dimensions, the semantic similarity between two words can be calculated as the cosine between the two vectors (a bit like a correlation, going from −1 to 1, showing how similar the two words are across these dimensions). Words with nearly identical patterns of co-occurrences would have cosines near 1, whereas those with orthogonal patterns would have cosines of 0.

Why should this work? That is, why should this procedure result in a measure of semantic similarity? Some words that occur together often are not semantically similar. For example, knife and butter might co-occur pretty often, but they are semantically very different. However, LSA does not use only the information about how often knife and butter occur together but also how often they occur with all the other words in the corpus. For example, butter might occur with a lot of other foods on shopping lists or in discussion of dairy products, and knife would not occur in such lists. Similarly, knife might occur with other words describing weapons or cutting, and butter would not occur with such words. Because LSA looks at the entire pattern of co-occurrences, it tends to find that words are similar if they are found in similar sentences, and, it is hoped, this will be a good measure of semantic similarity.

Landauer and Dumais (1997) and Lund and Burgess (1996) report a number of experiments and tests of their models, some of which are fairly impressive. Landauer and Dumais tested their system on a synonym test from the Test of English as a Foreign Language (TOEFL), which is used as a college admission test for nonnative speakers of English at American universities. The model got 64.4% correct, which was almost identical to the performance of a large sample of college applicants who took the test. Although these applicants did not all know English very well,
Landauer and Dumais report that such a score would allow admission to many American universities. Lund, Burgess, and Atchley (1995) showed that their model can predict results in word priming experiments of the sort that are often performed in the study of lexical access. In an extremely controversial move, Landauer and his colleagues are adapting LSA to be used as an automatic grading tool for student papers. Essentially, the more the pattern of lexical co-occurrences in the student paper is like that of the course materials, the higher the student's grade. There is evidence that the model may have some validity as a measure of the amount of learning (see Wolfe et al. 1998; and Rehder et al. 1998). A readable article describing the grading system and including outraged howls of critics who object to grading by machine may be found in Thompson (1999).

The success of such models could be taken as undermining the conceptual approach to meaning taken in this chapter. If these models can truly account for how words are used, then word meaning may be a large associative network that merely encodes patterns of co-occurrence. Landauer and Dumais (1997) suggest that such a network would help to explain how it is that children are able to acquire so many words so quickly. When an unfamiliar word appears in a sentence context, children can assume that it is similar in meaning to other words that have appeared in similar contexts. Because the only input to the model is words, meaning would be represented through links to other words rather than through the knowledge underlying those words—that is, concepts. LSA's success in some tests may indicate that conceptual information is simply not that necessary to explaining meaning.

Before we get too carried away, however, we need to see more of the semantic structures that these models create. Landauer and Dumais do not provide examples of the semantic structures that their model arrives at (a fact that is itself worrisome). Lund and Burgess (1996, table 2) are more helpful. In their text base (consisting of 160 million words from newsgroup messages), the closest neighbors to jugs were the words juice, butter, vinegar, bottles, and cans. The closest neighbors to monopoly were threat, huge, moral, gun, and large. (A later paper by Burgess, Livesay, and Lund 1998, suggests that these kinds of neighborhoods are common.) Clearly, these words are related in some ways to their neighbors. However, the meaning of jugs is simply not similar to the meaning of juice, butter, or vinegar; monopoly is not semantically like threat or gun. Even though monopolies are large, and may be considered a threat, the fact that these two words are related to monopoly is not very informative. Is the monopoly a threat, or is someone threatening it? Who is it a threat to, and how does it threaten them? If one just knows that monopoly and
threat are “similar,” one does not really understand what monopolies are and do. Similarly, are monopolies large like elephants or buildings, or like something else? There is nothing in this set of neighbors specifying that a monopoly is an illegal activity by a company that controls a certain commodity in order to unfairly raise prices for it and stifle competition. Once one knows this, one can easily understand how monopolies are large (in the sense of controlling nearly the entire supply of the commodity), are threats (to other companies or to consumers, but in economic terms), might not be moral (why isn’t immoral the related term?), and so on. Having the concept explains why these words are related, rather than vice versa.

The general problem with nets of associations is that knowing what words are associated to one another does not specify what the meaning of an individual word is. First, there is an overall problem that one simply cannot understand the meanings of words by reference to other words (as discussed by Johnson-Laird 1983, for example). If one only knows dog by its similarity to cat and cow and bone … and cat by its similarity to dog and cow and bone … and cow by its similarity to cat and dog and bone … and so on, one is caught in a circle of similar words. One simply does not know enough from similarities to other words to know what should be called cat and what dog. Words must be connected to our knowledge of things in the world, not just other words.

A more specific problem is that the model does not specify the basis for each association. Although jugs may be related to both vinegar and bottles, these relations are extremely different, and an overall similarity score does not represent these differences. One wouldn’t know from the scores that “Put it in the jugs” is similar to “Put it in the bottles” but not “Put it in the vinegar.” Apparently the similarity score is good enough to allow LSA to pass a synonym test or even to grade student papers, but it will not be enough to represent the very detailed knowledge people have about the referents of words, which is used in speech and comprehension.

These shortcomings are exactly the things that the conceptual approach does well. Since concepts are our nonlinguistic representation of the world, by connecting words to these representations, we can explain how people can connect sentences and words to objects and events in the world. Concepts are just the things that are evoked by our perceptual systems and that control our actions. Thus, by hooking up words to concepts, we can break out of the circle of words connected to words and tie language to perception and action. Furthermore, since one’s concept of a jug, say, would include detailed information about its origins, parts, materials, functions, and so on, the concept is more than sufficient to distinguish the meaning of jugs
from that of vinegar and, for that matter, bottles, and it directly represents the information that allows one to see that jug and bottle are near synonyms, but jug and vinegar are not.

In short, even the more recent associationistic approaches do not cast doubt on the overall conceptual approach I have been proposing here. I think that it is likely that children can learn words by using the context that novel words appear in, as Landauer and Dumais (1997) argue. And it is certainly possible that a system that accumulates massive numbers of associations through a large text base will be able to do a number of useful tasks. However, as a psychological model of word use, the associationistic approaches do not seem likely to replace the conceptual approach.

These programs are still new (as of this writing), and it is possible that they will overcome what now appear to be shortcomings. For example, the accuracy of the models' behavior depends greatly on how many words they are exposed to. It is only with very large samples of text that they can accurately represent the meaning of individual words. Although they have in fact been run on large text corpora, the sizes of these corpora are still much less than the verbal experience of a college student, say. So, perhaps with a few more hundred exposures to real-life occurrences of jugs, HAL would no longer list it as being similar to vinegar. However, it is important for contrast to remember that children can figure out much of the meaning of a word from a single reference, which arguably provides better information than even thousands of word co-occurrences.

Another possibility is that (again, with sufficient input) the 300 dimensions that LSA uses could turn out to be meaningful semantic dimensions. That is, rather than just providing a global measure of similarity between cat and horse, LSA might allow us to examine the dimensions on which they are similar and dissimilar and thereby specify the properties of cats and horses. For example, dimension 129 might be size, and cat might have a moderately low value on that; and dimension 212 might be tail, and both cat and horse might have high values on that. Although current implementations of such systems do not generally result in many interpretable dimensions, it is again possible that future versions will. If so, then we might be able to use LSA representations to derive more familiar conceptual representations of feature lists or schemata. And unlike Osgood's semantic differential, or most scaling procedures, the large number of dimensions used in LSA could conceivably be enough to represent many different aspects of meaning. At that point, LSA would not necessarily be a different theory of word meaning but could be seen as a way of implementing conceptual knowledge. In the meantime, such models do not pose a threat to the conceptual approach to meaning.
Linguistic Determinism

One of the most popular issues in the psychology of thought and language is whether our concepts are determined by the language we speak. This notion of linguistic determinism is often called the Sapir-Whorf hypothesis (usually shortened to the Whorfian hypothesis). A complete account of the history of the psychology of meaning would give considerable coverage to this theory. However, this view has had little concrete support and as a result has had little influence on psychological theories of word meaning, regardless of its hold on the public’s imagination.

In everyday life, one often hears statements of the sort “Eskimos have 50 different words for snow,” along with the suggestion that this makes them see and think about snow differently than do, say, people living in New Jersey. Such a statement is a kind of arctic legend rather than a claim about a specific language (e.g., what Eskimos exactly are being talked about, what language, and what exactly are the words?), and it seems, insofar as one can tell with such legends, not to be true (Pullum 1991). But this claim can serve as a prototypical example of the idea that languages differ in ways that might influence thought. Similarly, when one travels, one may notice that people in different lands think about some things differently than we do. And if one learns the language, one sees that they also have a somewhat different way of talking about those things—often having more or fewer words, using different metaphoric expressions, or dividing up the world differently than English does. All this can lead to the conclusion that when learning one’s native language, one’s thoughts are largely determined. That is, when learning English, one learns only the single word snow, and so one sees snow as being one kind of thing; when learning Eskimo, one must now see 50 different things instead of just snow. The idea is that our concepts are fairly malleable relative to language, which is imposed by one’s culture, and which no single individual can change very easily. As a result, language acquisition forces a certain conceptual structure on the individual.

There is certainly a note of plausibility to this story. As pointed out earlier, during language acquisition, children are forced to learn distinctions that they might not have spontaneously noticed, because the word meanings depend on them. Although children would probably notice the difference between geese and ducks at some point, their parents’ use of two different labels for these birds can accelerate the process. However, behind such straightforward examples lie some much stronger claims that are more controversial. For example, in saying that language determines thought, is one saying that people cannot notice the distinction unless it is in the
language? Such a statement would clearly be too strong. Indeed, as I have just reviewed, most content words have different senses, which refer to things that are different, but that we perfectly well distinguish, such as table referring to a kind of furniture and to the people who are seated at such a piece of furniture. We are not fooled by the use of a common name to treat these things identically (e.g., try to polish the people or to feed the furniture). The importance of polysemy within a language is seldom recognized when considering differences between languages. If words had such a large effect on our concepts, one would not expect to have such drastically different senses referred to by the same word within a language.

Even to say that people would find it harder to notice distinctions that are not marked in the language is a strong claim. So far as I know, there is no verified result that shows this difference (see below). Sometimes the existence of different names is itself taken as evidence for cognitive or perceptual differences: How could the Eskimos use 50 different words for snow unless they had 50 different concepts, in contrast to our one concept? However, this is obviously a completely circular argument. To demonstrate an effect of language on cognition, there must be some evidence that Eskimos conceive of snow differently than those who have a more snow-deprived vocabulary, and that evidence should be a nonlinguistic task.

Furthermore, consider the question of why these linguistic differences come about. Why is it that the Eskimos happen to have 50 different names for snow? Why do the French have 100 different terms for the taste of wine (using another example heard on the street)? If linguistic determinism is true, then it would just be an accident of the language. That is, the Eskimo language happens to have a lot of names for snow, and therefore Eskimo children are forced to form many different concepts of snow, whereas French children are off the hook as regards snow, but are arbitrarily forced to learn the wine terms. (The French seem to have the better end of the deal.) However, these linguistic differences have very good nonlinguistic reasons. For example, if Eskimos did have 50 words for snow (again, see Pullum 1991), this is likely because they live in a very snowy area and need to make different distinctions in order to carry out their everyday activities. And France has many more varieties of wine than do the Northern Territories. Historically, then, it seems much more likely that their activities and thoughts determined the vocabulary than vice versa. Furthermore, even within English, there are experts who have many different words for snow, namely skiers. Clearly, this is not due to linguistic determinism either (given that standard English doesn’t include these terms) but arose through the need of skiers to communicate different qualities of snow to one another. They wanted to communicate these qualities because different types of snow have different implica-
tions for skiing, such as whether it is fast or slow and whether one can steer on it. It was the need to describe known distinctions that drove the development of a vocabulary, rather than vocabulary determining the concepts skiers have.

In spite of this discussion, one might argue that for a child born into a society where the vocabulary makes certain distinctions, the child is stuck with learning those distinctions and with not learning others that could be made. And so, even if historically the vocabulary came about through conceptual reasons rather than vice versa, the child must learn the language now, and, as a result, the child’s concepts are driven by the language. Nonetheless, there is very little evidence that language changes our perception of things in a very deep way, independently of experience and societal interests. This has been most studied within the domain of color. For example, people who come from a language with only two color names (essentially light and dark) appear to have color perception and memory similar to people who know many color names, like English speakers (Rosch 1973). The perceptual structure of color space and processing of colors in memory do not seem to rely on language. Linguistic determinism has also been studied in more complex nonperceptual domains such as counterfactuals, statements about events that did not occur, such as “If I had won the lottery, I would have paid my rent.” Although some languages distinctively mark counterfactuals and others do not, this does not turn out to influence their speakers’ counterfactual reasoning, in spite of claims to the contrary (Au 1983).

There is strong resistance to giving up the notion of linguistic determinism in some circles. For example, Lucy (1992) gave a book-long defense of the Whorfian hypothesis, which largely amounted to criticizing all the evidence that had been found against it. However, his defense could not cite actual, strong evidence in favor of the hypothesis—it could only attack the evidence contradicting it. Surely this is very weak support for a scientific hypothesis.

If there is any truth to linguistic determinism, I think that it is most likely to be in simpler domains than the ones typically studied in the psychology of concepts. Recent writers on linguistic determinism have emphasized the importance of grammaticalized features rather than lexical domains like words for snow and wine tastes (see Gumperz and Levinson 1996; Lucy 1992). For example, Slobin (1996) points out that verbs in Turkish are marked for whether the speaker directly observed the action or only knows about it second hand. Turkish speakers therefore must keep track of this information when speaking, and perhaps do so in general. I have heard Chinese speakers complain that they found it annoying to have to continually keep track of everyone’s gender, so that they could distinguish he and she when speaking
English. The fact that I don’t find this difficult may indicate that English has focused my attention on gender more than Chinese would have. Finally, Bowerman (1996) points out that speakers of Tzeltal, a Mayan language, must use different verbs to say that X is on Y, depending on X’s shape. Perhaps Tzeltal speakers focus on shape or even perceive it somewhat differently than English speakers.

All of these things may well be true, though most of them do not have direct, nonlinguistic evidence to support the notion that language has changed speakers’ concepts. These grammaticized distinctions are often based on broadly applied features (like gender, spatial relations, or plurality) instead of conceptual information relevant to the domains that we have been focusing on (animals, plants, artifacts), and they may result in greater attention or emphasis placed on some of the relevant features, rather than a change in conceptual organization. For example, Chinese speakers clearly categorize people in terms of gender, but perhaps English speakers do so faster and even in situations when it is not very relevant, because of their need to specify gender in personal pronouns. This would be a quantitative rather than a qualitative difference.

Another possibility, proposed by Slobin (1996), is that when people are planning to speak, they attend to the categories that their language requires in order to plan their utterances. For some languages, this will be plurality, gender, or tense; for other languages, it will be object shape, social distance, or the duration of the event. However, when simply observing events or in their general knowledge about the world, people may be little influenced by these same properties.

Discussions of linguistic determinism in the past have given rise to more heat than light, and the topic is still emotionally charged for many. At this point, I think that some good ideas and good techniques for studying linguistic determinism have been developed. However, the core evidence that one would want for linguistic determinism—adequate description of linguistic differences, followed by independent demonstrations of parallel conceptual differences—is still largely lacking. The newer approaches represented in the Gumperz and Levinson (1996) book hold out some promise that such evidence will be found, if it exists. However, the domains being looked at by such researchers are not closely related to the main topics investigated in the psychology of concepts, and so it is still an open question about how much language determines adult concepts of objects, events, and so on.

In this context, it is interesting to consider a proposition that is essentially the opposite of the Whorfian hypothesis, that certain linguistic categories are due to universal conceptual and communicative constraints. For example, it is often said that nouns and verbs are found in every language, and adjectives in most languages.
But how does one identify the category of noun in different languages? One can attempt to answer this question syntactically: Nouns are things that take plural and case markings, whereas verbs take tense or aspect markings. However, even this is problematic, as different languages have different syntactic rules for the parts of speech. For example, Chinese does not normally inflect nouns for the plural, and English does not use case markings, whereas Latin does both. How can we say, then, that Chinese, English, and Latin all have the syntactic category of nouns? The traditional answer has been a semantic one, that nouns refer to persons, places, or things, verbs refer to actions or events, and adjectives refer to properties. However, this basis is not perfect either, as some nouns refer to events or properties, like *the party* or *whiteness*, and some adjectives refer to entities, like *corporate*.

Croft (1991) suggests an interesting resolution to this problem. He argues that nouns, adjectives, and verbs arise from a basic communicative need to refer to entities, to modify those references, and to predicate something about those entities, respectively. Therefore, there is a semantic core to nowness and verbhood, based on typical properties of objects and predicates. But rather than a well-defined category, these are prototype categories, with the typical objects and events at their center, and less typical things at the periphery. The particular syntactic properties of each category are related to these conceptual differences, Croft argues. For example, nouns are often inflected for number, topicality, and case, whereas verbs are often inflected for tense and aspect. This is because objects can be counted, are often the topics of conversation, and have different relations to one another, whereas verbs often refer to actions, which happen at a particular time and have other process distinctions (being complete or incomplete; being essentially instantaneous or lasting for some time) marked by tense and aspect. As a result, it is no coincidence that English and Spanish both mark nouns for plurality and have verbs that are marked for tense. It is our basic concepts of objects, event, and properties that lead languages to tend to have similar morpho-syntactic properties. Croft also notes that items that are semantically dissimilar to the category prototypes also often fail to have the typical syntactic properties. For example, although *whiteness* and *water* are both syntactically nouns, they cannot be pluralized, because they do not refer to objects.

In summary, Croft’s analysis suggests that language may reflect conceptual universals. Rather than concepts following language, some linguistic structures may arise from conceptual structures combined with communicative needs. I cannot do justice to Croft’s theory here, but it provides an interesting contrast to the Whorfian emphasis on the reverse direction of influence.
Figure 11.4
A final attempt to illustrate how word meanings are built out of concepts.

Final Summary of the Conceptual Approach

After all the discussions, arguments, and data presented in this chapter, where does this leave us with regard to the relation between concepts and word meaning? It is clear that conceptual structures are deeply involved in word meaning, given all the conceptual effects found in language use and learning. However, the simple relations depicted in figures 11.1 and 11.2 do not adequately represent the relation, as they do not account for phenomena such as polysemy and contextual modulation. An attempt at a more complete picture is made in figure 11.4, but it must be admitted
that a complete model of word meaning cannot be represented in a simple diagram, because the number of variables and complexity of the structures involved precludes a simple depiction. In the figure, the nodes and their connections represent conceptual structure, and words are connected to a node or substructure.

The notion here is that an individual (unambiguous) word usually has a number of different senses, like word 2 in the figure (the others don’t, for purposes of clarity). These senses are overlapping subparts of the conceptual structure. As it is likely that many of these senses (or at least the most common ones) are explicitly represented, instead of being derived from a core meaning (Klein and Murphy 2001), there must be a number of explicitly marked links between the word form and conceptual structures. The meaning of the word then consists in coherent subparts of conceptual knowledge that are picked out by the lexical item. However, note that the conceptual component also includes other information from the general domain, including superordinates, coordinates (nodes that share an immediate superordinate), and other related concepts. For example, sense 1 of word 2 picks out a single node, N2, in the conceptual structure. N2 has superordinate N1, coordinates N3 and N4, and subordinates N5 and N6 represented in the structure. Although such concepts are not directly picked out by the word, they are important for specifying the meaning by providing contrast, background knowledge, and underlying assumptions (e.g., Fillmore 1982). For example, the exact meaning of the word picking out N2 depends in part on distinguishing it from the related concepts N3 and N4. And since background knowledge is apparently critical for categorization and reasoning about concepts (see chapter 6), word use can draw on such resources as well. Figure 11.4 also attempts to illustrate the general pattern of chaining of word senses, as the parts of conceptual structure picked out by different senses overlap to a large degree. Sense 1 of word 2 is included in sense 2, which overlaps with sense 3. There are also some senses that are not themselves overlapping but are related by intermediate overlapping senses (like the material and editorial policy senses of the word paper), as in senses 1 and 3 of word 2 in the figure.

What cannot be represented in such a figure are the processes by which the senses are identified and represented, the way that related concepts and background knowledge constrain senses, and the processes by which an interpretation is constructed each time a word appears in a particular context. However, figure 11.4 certainly does serve to indicate the complexity of the picture relative to that envisioned in figures 11.1 and 11.2. And in fact, part of the reason concepts have been criticized as representations of word meaning is, I believe, because researchers
assumed the simple relations of the sort shown in those figures. Unfortunately, the relation is quite complex, which is hardly surprising when one considers the complexity of word meaning itself.

One particular process that the figure cannot display is how new senses are derived for a word. For example, if I have only heard the word DVD used to refer to the physical disks, when I hear a new sentence like “This DVD is very boring,” I must derive the sense meaning the disk’s content. Imagine that my original sense is like that of sense 2 of word 2 in the figure, a fairly complex meaning structure indicating a physical device of a certain shape that contains audiovisual content, if played by a decoder through a TV. However, the “boring DVD” refers only to the content, a substructure of the original sense, perhaps the node picked out in sense 1 in the figure. Thus, I must be able to shift from understanding the word as referring to the larger structure to understanding it as referring to a subpart of that structure. Pustejovsky (1995) formally describes how such operations could be carried out. No doubt other cases exist in which one must go in the other direction—extending a word from a narrow sense to a broader one. Understanding the meaning of a word in context, then involves not only retrieving prestored structures, but more constructive processes based on those structures.

Even a model at this level of generality does help us to understand some longstanding questions in the psychology of word meaning. One question has to do with whether word meanings are decomposed during use. Theories of semantic representation often suggested that word meaning was a list of semantic components (Katz and Fodor 1963), and therefore theories of comprehension suggested that understanding consisted of retrieving each word’s components when it was understood. One problem with this is that it is very difficult to provide a simple set of components that is a word’s meaning, as I have already suggested (in the paint and buy examples). Rather diffuse background knowledge can be a necessary part of interpretation even if it does not (and could not) appear in the list of features. An empirical problem is that words that appear to be more complex, in the sense of having more components, are not generally more difficult to understand as measured in a variety of experiments (Johnson-Laird 1981, chapter 10 provides a readable review of this literature), though one exception is that negative words are harder to understand than their positive counterparts (like absent-present, Clark 1974). Nonetheless, it is difficult to understand exactly what it would mean to comprehend a word if it is not to bring to mind its components. Although some have proposed that a word meaning is a holistic entity that is not analyzable (J. A.
Fodor 1981), every concrete theory of semantic representation does use components of some kind, because there is no other obvious way to represent the meaning.

By relating words to conceptual structure, we are now less tempted to talk about lexical “decomposition.” First, we have acknowledged that a word has a number of senses—not just one decomposition (see figure 11.4)—and coordinating and selecting among those senses is an important aspect of comprehension, which cannot be predicted by the number of components of the word in general. Second, as figure 11.4 also reveals, the substructure being picked out by the word is not a simple list of components, but a possibly elaborate structure. It is in general hard to know how to count up the number of nodes, links, and their relations in order to determine each word’s overall complexity. One would have to know the structure very precisely in order to obtain an accurate measure, more precisely than we typically do know. Third, the conceptual structure being picked out by the word is closely integrated with related structures, and these probably have considerable influence on the use of the word. For example, the word dog is embedded in biological knowledge of animals in general as well as knowledge of pets, the differences between dogs and cats, and so forth, and this knowledge might be activated to various degrees on some occasions in which one hears the word dog—even though it is not an “official” part of the meaning.

Finally, it is worth pointing out that we need not assume that every aspect of conceptual knowledge is retrieved when a word is understood. In fact, it would be very difficult to do so. It has long been known that when a sentence emphasizes one aspect of a word meaning, that aspect is more activated, and other aspects may not be activated (McKoon and Ratcliff 1988; Tabossi 1988). So, it is likely that not all conceptual knowledge associated with a word is retrieved into memory when the word is encountered, but some knowledge is picked out as being particularly relevant. Thus, one would not always expect more complex words to be harder to understand than simple words, as the amount of information explicitly encoded would likely vary across conditions.

Theoretical Implications

The literature on word meaning has not been directed towards distinguishing theories of concept representation. As a result, the phenomena that have been discussed here are generally not very probative in evaluating different accounts of how concepts are represented. For example, although typicality effects have been found
in linguistic data, in sentence comprehension, speech production, and rating studies, these studies have not investigated whether people's representations of word meanings are exemplar-based or some kind of prototype representation. (Of course, such phenomena do reinforce the severe problems with the classical view described throughout this book.) Often such studies appear to assume a prototype representation, but it seems likely that exemplar models could explain most of the data just as easily. For example, if reading the word vehicle activates memories of specific vehicles, it is more likely that a bus exemplar would be activated than a tank exemplar would be. Thus, the Garrod and Sanford (1977) anaphora data could possibly be accounted for by an exemplar-based representation. Of course, there are some phenomena that I have already identified as being somewhat problematic for exemplar models (e.g., hierarchical classification, see chapter 7), and they are equally problematic when they are applied to words and word meaning.

The phenomenon of chaining is particularly important here. That is, as emphasized by Taylor (1995; and Lakoff 1987), broadly applied words often do not have a simple prototype structure but can be understood as a chain of meanings that are pairwise related (e.g., the examples of paper and the juice box discussed above). This process has a very exemplarish quality to it (as noted by Malt et al. 1999), because it is similarity to a particular past use that is important, just as exemplar theory says that it is similarity to known exemplars that is important. However, it is not yet clear whether the chaining is taking place between exemplars or summary representations. The plastic cartoon-character juice box is not just similar to a particular plastic squarish juice box—it is similar to a whole class of such juice boxes. Perhaps it is because this novel item is similar to a subcategory of boxes (rather than to an exemplar) that it can be called a box. If plastic juice boxes form a subcategory of boxes, they might be represented by a prototype-like summary representation of that subcategory rather than as individual exemplars. So, chaining could involve either summary representations or exemplars.

Thus, it does not seem that the main phenomena of word meaning can be chalked up to either the prototype or exemplar model very clearly.

The knowledge approach comes off well in this analysis, however, from a number of sources. First, it has long been known that in order to understand language, one must make innumerable plausible inferences to fill in what the speaker or writer intended to say (e.g., Bransford, Barclay, and Franks 1972; Gibbs 1984; Grice 1975; Haviland and Clark 1974; Schank and Abelson 1977; among many others). For example, the sentence “Carl poured the butter” requires one to infer that the butter must be hot and in liquid form. Such inferences obviously involve our knowledge of
the world (e.g., only liquid can be poured, butter is liquid only when melted, and therefore is hot), which is apparently activated very quickly and continually during real language comprehension.

Knowledge effects can be seen more specifically in conceptual combination (see chapter 12), contextual modulation, typicality shifts (Roth and Shoben 1983), determining the intended sense in polysemy, and the use of background assumptions in determining word meaning (Fillmore 1982). That is, there is no way to tell from the word *table* by itself whether one is talking about the piece of furniture or the collection of people eating dinner. But when embedded in a sentence such as “The table left only a 10% tip,” there is only one plausible choice—only one choice compatible with our world knowledge. Although one can argue that such knowledge is not really part of the word meaning but is more general domain knowledge, it effectively becomes part of the meaning, because the knowledge is used to determine the specific interpretation of that word in that context. Indeed, it is in lacking such everyday knowledge that computerized translation and language comprehension devices have often had great difficulty, as well as their difficulty in using pragmatics and knowledge of the speaker. Thus, it can be very difficult to know where to draw the line between what is part of the word meaning per se and what is “background” knowledge. It is not clear to me that drawing this line will be theoretically useful.

In short, insofar as background knowledge and plausible reasoning are involved in determining which sense is intended for a polysemous word for contextual modulation and for drawing inferences in language comprehension, this provides further evidence that background knowledge is involved in the conceptual representations underlying word meaning.