

Chapter 1

Introduction

In this chapter I provide an outline of the role of semantics in linguistic theory. I begin with a discussion of the central role of truth and reference. In this book I develop a truth-conditional model of semantics, the purpose of which is to explain the conditions under which a sentence can be uttered truthfully. I carry this out by appealing to a theory of reference, which tells us how linguistic expressions can be linked to objects in the world. In the next two sections I discuss some of the primary empirical issues in semantics and pragmatics. In the final section I discuss my methodology which involves building logical models of the meanings of expressions in terms of the meanings of their parts.

1.1 TRUTH AND REFERENCE

In this section we study the philosophical and psychological issues that motivate my decision to center our study of semantics around the notions of truth and reference. In the first subsection we consider the fundamental role of truth in developing a theory of meaning in natural language. I also lay out the general structure of a theory of truth. In the next subsection I discuss the referential nature of language, or in other words, how language relates to the world. I also discuss how a theory of reference can be employed as the foundation of a theory of truth. In the last subsection I discuss some of the philosophical and psychological positions that have been taken with respect to truth and reference. I concentrate on the conventional nature of linguistic meaning, contrasting it with conceptualist and realist theories.

Before beginning this all too brief section, I would like to point out that truth and reference are two of the most widely studied areas not only in linguistic semantics but also in twentieth-century philosophy as a whole.

This being the case, I can only provide a taste of the nature of the debates that surround these contentious issues and pointers for those interested in learning more.

1.1.1 Truth

At least since the time of Plato, the concept of truth has been seen as fundamental in understanding the nature of human language. This preoccupation with truth is reflected in Plato's *Sophist*, *Phaedo*, and *Thaetetus*, and is further developed in Aristotle's *Metaphysics*, *Categories*, and *De Interpretatione*.

One of the principal goals of linguistic semantics, as initiated in its present form by its founding fathers, Frege and Russell, has been to develop a workable theory of the conditions under which a sentence can be uttered truthfully. To *utter* a sentence is to use it in a specific context. Given the crucial role of context in natural language, it should be clear that utterances, rather than sentences, can be either true or false. Of course, some sentences are *eternal* in that their truth conditions do not depend on context, but these are few and far between in ordinary discourse. In this book we will be focusing primarily on what is known as *truth-conditional semantics*. There are certainly a wide range of uses of sentences for purposes other than conveying a truth value, and we will consider some of these in sections 1.3.1 and 1.3.2. Nevertheless, the primary focus of semantics has been on formulating truth conditions for literal uses of declarative sentences.

While the notion of truth has been central since the advent of Western philosophy, it was not until the work of Tarski (1935, 1944) that a mathematically sound truth-conditional semantics was formulated. Before we consider Tarski's semantic approach in section 1.4.3, we consider his general approach to a theory of truth. Tarski introduced the following paradigm, instances of which have come to be known as *T-sentences* (Davidson 1967c).

(1) *s* is true if and only if *P*.

The variable *s* is to be filled with a sentence, whereas the variable *P* is to be instantiated with the conditions for the truth of *s*. Ideally, we would like a general theory that gives us truth conditions for every sentence in the language. Developing truth conditions for a large fragment of English is the primary goal of this book, and of a great deal of contemporary semantic investigation.

Tarski provided the following kind of instance of a *T*-sentence.

(2) *The cat is on the mat* is true if and only if the cat is on the mat.

Such a construal of truth conditions may at first appear to be circular, tautological, or nonsensical, although this is not the case. First note that on the left-hand side of the biconditional is a claim about the truth of a particular sentence, namely $s = \textit{The cat is on the mat}$. In technical terms, we have what is known as a *mention* or *name* of the sentence. In the text I will italicize words that are mentioned (in the literature, quotes are often used). On the right-hand side of the biconditional in (2), however, we interpret the sentence as we would any other sentence of English. In other words, we have what is known as a *use* of the sentence. To clarify the use/mention distinction, consider the following pair of examples.

- (3) a. *Chris* is a word with five letters.
 b. #*Chris* is a word with five letters.

The first case, (3a), mentions the word *Chris*; it is true because the word *Chris* does in fact have five letters. The second example, (3b), involves a use of the word *Chris*; it is semantically infelicitous (a condition we mark with the symbol “#”) because *Chris* is an individual, not a word. Note that it is uses of sentences that refer to truth values, and uses of names that (sometimes) refer to individuals.

With this distinction in hand, notice that the left-hand side of the *T*-sentence contains a statement that a particular sentence is true, whereas the right-hand side is just a statement of English. The primary reason that *T*-sentences such as (2) seem circular is that we are using English as a metalanguage in which to develop a theory of English itself. A *metalanguage* is a language for theorizing about another language, known as the *object language*. Of course, this relationship is relative. A language \mathcal{L}_1 may have a metalanguage \mathcal{L}_2 , which in turn has a metalanguage \mathcal{L}_3 , leaving \mathcal{L}_2 in the position of being a metalanguage relative to \mathcal{L}_1 and an object language relative to \mathcal{L}_3 . It is usual for a metalanguage to include terms that pick out elements of the object language. What is confusing about (2) is that English is acting as both the metalanguage and the object language. An alternative instance of the *T*-sentence in (2) might involve a shift in object language, say to Italian.

(4) *Il gatto é sul tappeto* is true if and only if the cat is on the mat.

In this case there is no apparent circularity because the two languages are distinct (despite some common historical antecedents); Italian is the

object language, and English is the metalanguage. Certain paradoxes arise if one is not careful when using the same language for both object language and metalanguage, and we consider some of these in section 11.1.5.

Of course, for a metalanguage to be of use in explicating an object language, the metalanguage must itself be well understood. The reason that natural language is often used as a metalanguage is that it is usually assumed that readers of English understand the language. Thus it is typical to find English used as the metalanguage in discussions of the English language. English is often used in fields other than linguistics, such as history and biology. But natural languages, though quite expressive, have a number of drawbacks when employed as metalanguages. Specifically, natural-language expressions are often vague, ambiguous, prone to mis-construal, and so on. Often this problem can be rectified by employing additional words and phrases, whose meaning is carefully spelled out (usually in English). For instance, this is the strategy employed in mathematics, chemistry, philosophy, the law, and almost every other field of inquiry.

For my purposes, I will find it convenient to employ the mathematical constructions of set theory to supplement English as a metalanguage. A firm mathematical foundation was once considered at the heart of what was known as *generative grammar*. Chomsky (1957, 5) had the following to say.

Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy, and consequently, gain a deeper understanding of the linguistic data. . . . I think that some of those linguists who have questioned the value of precise and technical development of linguistic theory have failed to recognize the productive potential in the method of rigorously stating a proposed theory and applying it to strictly linguistic material with no attempt to avoid unacceptable conclusions by *ad hoc* adjustments or loose formulation.

Such sentiments engendered a major paradigm shift away from imprecise theories formulated loosely in English, to more exacting mathematical theories. Ironically, Chomsky (1990b), in a reply to Pullum (1989), has recently reversed his position. Chomsky now argues against formal methodology, claiming that it is premature to formulate explicit and mathematically precise theories about domains of which we have such a limited understanding. He claims that to make a theory fully explicit would

require arbitrary decisions at fine-grained levels of detail. In the end, this debate comes down to one about whether or not to use mathematics in the metalanguage. I strongly disagree with Chomsky's present belief, apparently shared by the majority of his followers, that the state of linguistic theorizing is not precise enough so that "inquiry can proceed in a constructive way" (Chomsky 1990, 146). In this book, we will be following Chomsky's original injunction and endeavor to construct precise, mathematical theories of the structure of language.

1.1.2 Reference

A longstanding approach to formulating the meaning of a sentence is by specifying the conditions in the world under which the sentence could be truthfully uttered. In other words, a sentence is true if it corresponds to the way the world is, and it is false otherwise. Such theories have come to be known as *referential* or *correspondence theories of truth*. One of the foundations of referential theories is their assumption that the truth of an utterance depends on facts about the world, independently of the cognitive agents uttering and hearing the sentence. The goal is to provide an objective measure of the truth or falsehood of utterances based on the objects in the real world to which they refer. Such a view of language is obviously attractive when language is used in domains such as science or even politics. I consider some of the philosophical ramifications of referential theories and their rivals in the next section.

Referential theories are as venerable as theories of truth and have been part of the philosophical discourse at least since Plato's Socratic dialogues, in which he states a referential theory of truth more or less identical to the one I have just given. Aristotle reinforced this choice in his *Metaphysics*, clearing away many of the ancillary assumptions and digressions of Plato (Prior 1967b).

For an example of how a referential theory of truth may be formulated, we return to the example I introduced in (2).

(5) The cat is on the mat.

To determine whether or not an utterance of this sentence is true, we need to know the objects to which its salient constituent expressions refer. For instance, we need to know the referents of the definite noun phrases *the cat* and *the mat*. Furthermore, we need to know under what conditions one object can be said to be on another object. With these three referents from (5), two individuals and a relation, we are in a position to determine

the truth of the sentence as a whole by investigating whether the referent of the subject, *the cat*, stands in the relation specified by *on* to the referent of the prepositional object, *the mat*. To determine the referents of the various expressions, we need to know how the meanings of the various expressions connect to the world, such as *cat*, *mat*, and *on*, as well as the auxiliary *is* and the definite article *the*. With such a theory, it does not really matter whether the utterer of (5) knows whether the cat is on the mat, or even whether he believes that the cat is on the mat; the truth or falsehood of the sentence only depends on the referents. Certain subtleties arise concerning the determination of the referents, which may depend on the context of utterance, a point to which we return in section 1.2.4.

A general framework for the construction of referential theories of truth was provided by Tarski (1935). In fact, this work was fundamental in establishing the very consistency of the referential enterprise. Tarski accomplished this by developing a theory of truth for first-order logic. The objects of reference in this theory were individuals and relations. The key insight provided by Tarski was a recursive procedure for combining the meanings of constituents to determine the truth of arbitrarily complex sentences. I provide an overview of Tarski's general method in section 1.4.2. The rest of this book will be devoted to developing a Tarskian theory of truth for natural language.

While it may at first appear obvious that some notion of reference to objects in the world is a necessary step in determining the truth of a sentence, this theory has not gone unopposed. A principal contender is the *coherence theory of truth*. The coherence theory determines the truth of a statement not by its correspondence with the world but rather by how it coheres within a larger collection of statements. Historically, coherence theories of truth have been closely allied to rationalist metaphysical positions, as put forward by Leibniz, Spinoza, and later Hegel (White 1967).

The other prominent position opposed to a direct, referential theory of language is *conceptualism*. Under a conceptualist theory, language derives its meaning through its relation to our mental states. Some conceptualist theories, most notably *solipsism*, go so far as to deny the very existence of a reality beyond our minds. Other, more moderate conceptualists instead assume that language refers to mental entities or *representations*, with or without elaborating on the relation between our mental representations and the world. Representationalism has been a popular theory among some linguists and philosophers, most notably Katz and Fodor (1963; Katz 1966, 1972; Fodor 1975), Jackendoff (1972), and the *generative*

semanticists, as represented by Lakoff (1970a). Recent versions of conceptualism have been argued for by the proponents of *cognitive semantics* (Fauconnier 1985, Lakoff 1987, Langacker 1987). Fodor (1975) and Chomsky (1986b, 1990a) even go so far as to assume that such representations of concepts are genetically innate.

1.1.3 Conventionalism and Realism

Although I have adopted a referential theory of meaning, the specific nature of individuals, concepts, and relations among them remains in need of explanation. Furthermore, how individuals, concepts, and relations enter into a language must also be explained. In this section we consider some popular theories of reference, as well as some philosophical arguments against the possibility of developing a truly referential theory.

The approach I favor to the thorny issue of how reference is established is known under the general heading of *conventionalism*. According to such a theory, expressions derive their referential force by means of conventions established by linguistic communities. In other words, it is up to speakers of the language, either individually or collectively, to legislate the referents of expressions in their language. Conventionalism provides a natural explanation of how different communities are able to specialize their conceptual and individual categories. Under a conventionalist theory, any number of factors may influence the development of a language's basic vocabulary and structure. Specific motivations for choice may include cultural ideals, scientific knowledge, and psychological considerations. Such linguistic communities need not be separated by a great cultural divide; physicists and lawyers, for instance, have their own distinct, specialized vocabularies and usages. In a conventionalist theory, the way in which a language subdivides the world into objects and relations has come to be known as a *conceptual scheme*.

To fully specify a conventionalist theory of language, it would be necessary to understand how conventions arise and, just as important, how they can be shared among different speakers. The most important aspect of conventionalist theories is that individuals, concepts, and relations come about because we in some sense create them as cognitive agents. Recently, the linguistic-semantics community has paid a great deal of attention to the role of our psychological and physical constitutions in the determination of our conceptual schemes (Fauconnier 1985; Herskovits 1986; Jackendoff 1983; Lakoff 1987; Langacker 1987; Rosch, Gray, Johnson, and Boys-Braem 1976). To what degree our psychological makeups

determine our conceptual schemes is a matter of some debate; as we noted above, Fodor and Chomsky believe that such conceptual schemes are *innate*. Proponents of the *innateness hypothesis* argue that there is no other way in which such concepts could be acquired. The innateness hypothesis is also one of the most prevalent assumptions made by the theoretical-syntax community, ever since Chomsky (1965) identified the notion of an explanatory linguistic theory with one that could adequately characterize language acquisition and speculated that language must be innate.

An alternative to conventionalism is *realism*, a theory in which individuals, concepts, and relations among them are assumed to be independent of cognitive agents. A historically significant form of realism, known as *platonism*, has its roots in the writings of Plato. Platonism is the theory that there is an extant realm of ideas, sometimes known as (platonic) *ideals* or *universals*, to which our words refer. Platonists argue that these ideas exist independently of our minds but may be discovered by cognitive agents. Only indirectly, through the connection of ideas to the world, is our language able to refer to the world. Frege's (1892) notion of senses, to which we return in section 1.2.5, is the most prominent platonic theory in linguistic semantics. More recently, Katz (1981) has advocated a platonist view of concepts.

Although thoroughgoing platonists are scarce these days, there are many who argue that at least some individuals, concepts, and relations exist in the world. This brand of realism shares with platonism the view that we discover, rather than create, concepts. For instance, many realists assume that there is a circumscribed class of real concepts, known as *natural kinds*, that form the foundation of the conceptual system of our language. Putnam (1970, 1973, 1975) has been one of the most prominent supporters of a notion of natural kinds in interpreting human language. Kripke (1972) put forward a theory of referential uses of names that directly connects them to real individuals, rather than by assuming that a name is just an abbreviation for a cluster of concepts with which to pick out an individual. We return to theories of individual reference in sections 7.8 and 7.10. Natural kinds, like platonic ideals, provide a potential explanation of how different individuals can come to share similar, or even identical, concepts, both within and across cultural and historical boundaries. An extreme brand of realism has been championed by Lewis (1973, 1986), who argues that so-called *possible worlds*, that is, other ways in which the world might have been, are just as real as any other objects (we return to the topic of possible worlds in section 1.2.5). But if a realist

theory admits too many natural kinds, for instance, by positing one for every social and individual conceptual class, it becomes difficult to distinguish the theory from a conceptualist one. Some realists have actually attempted to bring realism in line with conventionalism, by stating that one way in which concepts can be real is if they are conceptualized by cognitive agents (Barwise and Perry 1981, 1983).

There is a great deal of controversy surrounding the nature of conceptual schemes and their connection to reality. The *empiricist* approach assumed that language was connected to the world through primitive observations (Russell 1940, Ayer 1946, Carnap 1947, Hempel 1950). Such a connection was taken to provide a foundation for scientific inquiry. One of the foundations of empiricism was a division between so-called *analytic* and *synthetic* truths. Analytic truths are ones that follow from logic, whereas synthetic truths require empirical verification. Mainstream philosophers of science abandoned the strict empiricism of the logical positivists, primarily due to the persuasive arguments of Nagel (1944) and Quine (1953b). Quine, for instance, argues that the notion of analyticity and attempts to explain it are inherently circular. Instead, he abandons the empiricist distinction between analytic and synthetic truths in favor of a more holistic approach, the result of which is a “blurring of the supposed boundary between speculative metaphysics and natural science” (1953b, 20). Specifically, Quine denies the objective realism of concepts, stating, “in point of epistemological footing the physical objects and the gods differ only in degree and not in kind” (1953b, 44).

But Quine did not abandon empiricism entirely. He still believed that conceptual schemes could be judged for their efficacy in formulating predictive theories about the world. Such a view is often identified with the belief that the world comes not only through physical reality; also required is a true or proper way to conceptualize it, that is, to divide it up into individuals, relations, and the like. Recently this view has been challenged by philosophers such as Rorty (1979, 1989), Davidson (1989), and Putnam (1988), who deny not only that there is a fact of the matter concerning conceptual schemes but also that humans will ever be able to evaluate conceptual schemes themselves. This latter shortcoming arises because we are forever trapped in our current conceptual scheme. While they believe we can change our conceptual schemes, as a result, for instance, of a scientific paradigm shift (Kuhn 1962), there is no way to achieve a god’s-eye perspective from which to evaluate our conceptualizations, as Putnam maintains. Though these philosophers do not deny

the existence of a mind-independent reality, they believe there is no way that we, as human cognitive agents, will ever be able to disentangle the dichotomy of conceptual scheme and content.

1.2 TOPICS IN SEMANTICS

In this section I provide a brief survey of some of the topics that linguistic semantic theories address.

1.2.1 Synonymy, Entailment, and Contradiction

Traditionally, natural-language semantics has been devoted to explaining several relationships that hold among utterances of sentences and smaller phrases. In this section I discuss some of these relations and their role in semantics.

One of the primary semantic relations is that of *synonymy*. We say that two phrases are *synonymous* if they have the same meaning. For instance, the following two sentences are usually taken to have the same meaning.

- (6) a. Dana saw Kim.
 b. Kim was seen by Dana.

The second sentence, (6b), is merely the passive form of the first sentence, (6a). Any time that one sentence in (6) can be used truthfully, so can the other. Even though these sentences are synonymous from a semantic perspective, they do have functional differences in larger contexts. For instance, consider following each of the examples in (6) with the following sentence.

- (7) He heard her, too.

It seems that this continuation is only felicitous when following the active sentence, (6a). In particular, the resolution of pronouns, such as *he* and *her*, may depend on the active/passive distinction. But such putative constraints are rather subtle. Dick Oehrle (p.c.) suggests the following dialogue.

- (8) *Speaker A*: Dana saw Taylor. But he didn't see Kim.
Speaker B: Kim was seen by Dana. He heard her, too.

With the explicit contrast of the second speaker's contribution with the first speaker's, reversing the order of the pronouns is much more acceptable. It has become quite popular to integrate such dynamic aspects of a sentence's contribution to discourse into the meaning of a sentence (see Chierchia 1995 for an up-to-date overview). If such a dynamic approach

to meaning is adopted, then the examples in (6) would not be considered synonymous.

Some cases of semantic interactions with discourse are even more subtle. For instance, consider the clefted form of the active and passive sentences in (6).

(9) It was Kim who Dana saw.

It is not so clear that the clefted sentence (9) can be applied in the same situations as the active and passive versions in (6). The clefted form seems to have a uniqueness effect in that Kim was the only person that Dana saw.

We might assume that nouns can be synonymous. Consider the classic example of *bachelor* and *unmarried man*. It seems that one can be applied to an individual just in case the other can. We can often consider the problem of synonymy of expressions smaller than sentences in terms of their substitutability in larger contexts. For instance, consider the following pair of sentences.

- (10) a. Pat is not a bachelor.
b. Pat is not an unmarried man.

It seems that the sentences in (10) are just as synonymous as their component nouns. The terms *bachelor* and *unmarried man* are the canonical exemplars of terms that are said to be *analytically equivalent*. An example due to Plato is that of *featherless biped* and *human*. In Plato's experience, all featherless bipeds were human and vice-versa. If this held in the wider world, it would be an example of a so-called *synthetic equivalence*; whether the two terms refer to the same set of objects is a matter for empirical investigation to determine. But consider terms like *vampire* and *unicorn*, which apply to exactly the same set of individuals, namely none, on the assumption, of course, that there are no vampires or unicorns. In this case, we are not so hasty to conclude that *unicorn* and *vampire* are synonymous; they appear to have very different meanings. This case illustrates the difficulty of formulating definitions for terms like *synonymous*. I consider the distinction between *unicorn* and *vampire* and other, similar distinctions in section 1.2.5.

In a truth-conditional approach to semantics, we might say that two sentences are *synonymous* if they can be truthfully uttered in exactly the same situations. This reduces the problem to the nature of truth and the nature of situated utterances, topics we return to below.

Another important relationship between sentences is that of *entailment*. Informally, one sentence entails another if the truth of the first sentence guarantees the truth of the second. For instance, consider the following pairs of sentences.

- (11) a. Terry is an unmarried man.
b. Terry is a man.
- (12) a. Terry is tall and handsome.
b. Terry is tall.
- (13) a. Terry ate a cake.
b. Terry ate.

In all of these cases, (11) through (13), the first sentence entails the second. In each case, the second sentence provides strictly less information than the first. In general, two synonymous sentences will entail one another. But as with synonymy, the notion of entailment is extremely subtle, and there are borderline cases that are difficult to untangle.

Two sentences are said to be *contradictory* if they could not both be truthfully uttered in the same context. Consider the following canonical examples.

- (14) a. Lee is asleep.
b. Lee is awake.
- (15) a. Lee is asleep.
b. Lee is not asleep.

In both (14) and (15), it is not possible for both sentences to be uttered truthfully. Of course, their contradictory nature depends on their being uttered at the same time (we return to the notion of time dependence in section 1.2.4).

A sentence may also be self-contradictory, in which case we say that it is *inconsistent* or simply a *contradiction*. Consider the following examples.

- (16) a. Terry is 30 years old, and Terry is not 30 years old.
b. Pythagoras constructed a round square.
c. My computer danced a jig.

The first example, (16a), is logically inconsistent. The second example, (16b), is inconsistent because an object cannot be both round and a square. The last example, (16c), is an instance of what is often referred to as an *anomaly*; computers simply are not the kinds of objects that dance jigs.

There is a dual notion to contradictory sentences: sentences that are true whenever they are uttered are said to be *tautologous* or *eternal truths*. Consider the following examples.

- (17) a. Dana ran or didn't run.
 b. Every bachelor is unmarried.

Assuming that the time at which both clauses are interpreted remains constant, the first case, (17a), derives its truth from logic alone. The second example, (17b), is true because bachelorhood entails not being married.

1.2.2 Presupposition

Another key way in which sentences convey information is by what is known as *presupposition*. Presupposition is related to entailment, but is not expressed as a relation between sentences. Rather, a sentence is said to presuppose a piece of information when the sentence may be felicitously uttered only if the information is valid. For instance, consider the following common examples.

- (18) a. My opponent supports the plan that will hurt the people.
 b. Terry has (not) stopped attending lodge meetings.
 c. The present king of France is (not) bald.

For the first sentence, (18a), to be used felicitously, there must be a plan that will hurt the people. For the sentence to be true, the opponent of the speaker must support that plan. Here the fact that the plan hurts the people is presupposed. Note that if we negate the sentence, the presupposition that there is a plan that hurts the people does not go away (unlike ordinary entailments). The second sentence, (18b), has similar behavior in that it entails that Terry was attending lodge meetings. Notice that the presupposition is not canceled when the sentence is negated. For Russell's (1905) famous example, (18c), there are conflicting opinions as to whether this sentence presupposes the existence of a present king of France, an issue to which we return in section 7.8. Presuppositions are even more insidious when embedded in questions:

- (19) Has Sandy stopped drinking?

There is no way to answer this apparently simple yes/no question with a simple *yes* or *no* without committing to the proposition that Sandy drank in the past. Either answer admits the fact that Sandy was once drinking. If Sandy never was drinking, the respondent must attempt to address the falsehood of the presupposition. Often presuppositions are quite subtle to

address; it is common to find presuppositions in political debates, advertisements, and other manipulative language.

1.2.3 Ambiguity and Vagueness

Another important semantic aspect of sentences is their potential *ambiguity*. Ambiguity arises from many different sources, as the following examples illustrate.

- (20) a. I went to the bank.
 b. Put the block in the box on the table.
 c. At least one referee read every paper.

The first example, (20a), involves *lexical ambiguity*. The word *bank* can mean either a river bank or a savings bank (among other things, like a bank shot in pool). The second example, (20b), involves what is known as *structural ambiguity*. It can either be read as a request to take a particular block and put it in the box on the table, or a request to take the block in the box and put it on the table. The last example, (20c), involves what is known as *scope ambiguity*; it can mean that there is one referee who happened to read every paper, or that for every paper, there was at least one referee who read it.

Ambiguity can be distinguished from the closely related concept of *vagueness*. Ambiguity reflects the possibility of interpreting an expression in more than one way. For instance, the term *nut* is ambiguous among a fastener for a bolt, a kind of food that grows on trees, and more colloquially, an insane person. An utterance of *nut* means one of these things but not all of them or even the disjunction of them. Vagueness, on the other hand, arises when one expression is intrinsically underspecified but cannot be used with only one meaning among several. The term *brother-in-law* is an example of a vague term. Although it can be applied to the brother of a spouse or to the male spouse of a sibling, it cannot be uttered to include one possibility and not the other. One way to see this distinction is with the standard linguistic test of sensitivity to negation, as in the following two examples.

- (21) a. Leslie is not a nut.
 b. Leslie is not Pat's brother-in-law.

If an expression is ambiguous, a use of it picks out one of its ambiguous senses, which is then negated. Negating a vague term, on the other hand, negates all of the possible meanings. In the first example, (21a), there is an ambiguity involved. An utterance of it could be false in the situation in

which Leslie is insane but is not a hard-shelled fruit. But contrast this with (21b); the term *brother-in-law* is *vague* in the sense that it is equivalent in meaning to the disjunction of being a brother of a spouse or the husband of a sister. If Leslie is the husband of Pat's sister or a brother of Pat's wife, then (21b) is going to be false.

Another standard test for distinguishing vagueness from ambiguity arises in cases of verb-phrase ellipsis, wherein an elided verb phrase must be interpreted in the same way as its antecedent.

- (22) a. Robin has a brother-in-law, and Sandy does too.
 b. #The broker went to the bank, and the riverboat did too.

The first example, (22a), allows Robin to have one kind of brother-in-law and Sandy the other. On the other hand, the second case, (22b), cannot be interpreted as having the broker visit a savings bank and the riverboat a river bank (but it could, of course, be true if the broker and riverboat both went to the bank of a river or both went to a savings bank). Similarly, coordination requires only one sense of an expression to be used.

- (23) a. #Terry is in Chicago and a good mood.
 b. #Terry has and will score a goal.

These are examples of a phenomenon known as *zeugma*, in which two different senses of the same word are juxtaposed. Such usages are considered extragrammatical.

Unfortunately, the term *vague* is itself ambiguous. Besides meaning something like disjunctively underspecified, a term is also said to be *vague* if membership in the class denoted by the term is not precisely specified. Adjectives like *red* and *tall* are common instances of expressions that fit the second definition of vagueness. It is not at all clear at which point along the spectrum from canonical red to canonical orange an object stops being red and begins to be orange. Similarly, there is no precise height above which someone is considered to be tall. A further complication arises because so-called *comparative* terms like *tall* involve comparison to some extrinsically specified class. Consider being tall for a basketball player versus being tall for a jockey. Even with an explicit comparison class, say linguistics professors, the boundary between tall and nontall remains underdetermined. Of course, many nouns, verbs, and probably all prepositions are vague. In fact, it is rather rare to encounter terms that are not vague outside of certain scientific or mathematical sub-languages. Such considerations sparked Wittgenstein's (1953) theory of *family resemblances*, in which he assumed there were canonical exemplars

or *prototypes* of terms, and other objects were gauged for inclusion in a class on the basis of their similarity to the prototype. Such theories have also found their way into psychological work on categorization (Rosch et al. 1976). Others have claimed that the classes determined by (at least some) natural-language terms are determined by a core set of simpler concepts, together with some peripheral concepts (Smith and Medin 1981, Collins and Quillian 1969). For instance, canaries' being yellow is central to canaries; their having skin is a less centrally associated concept. A taxonomic organization of terms that reflects the central and peripheral properties is clearly reflected in most dictionary definitions of concepts, although this may be simply an artifact of dictionaries being forced to explain the meanings of words in terms of other words. Others have argued for the centrality of *natural kinds* in settling issues of vagueness. For instance, Putnam (1975) argues that being composed of H_2O is central to the concept of water, which is for him a canonical instance of a natural kind, even if those who used the term *water* in 1750 had no grasp of the chemical structure of water. But even terms like *water*, at least in common usage, are vague in the sense that it is not clear at what level of impurity or pollution a body of liquid turns from water into something else.

1.2.4 Indexicality and the Situation of an Utterance

Often the meaning of an expression in a natural language depends on the context in which it is uttered. Canonical examples include first- and second-person pronouns such as *I* and *you*, temporal adverbials such as *now* and *yesterday*, and the tenses of verbs such as *ran* and *will run*. The reference of these expressions depend on facts about the *utterance situation*, such as the speaker and hearer in the case of the pronouns and the time of utterance in the temporal cases. For instance, *I* always refers to the speaker, and *you* to the hearer. As is standard in linguistics, I use the term *speaker* and *hearer* generically to refer not only to spoken communication but also to the initiator and recipient of written language, signed language, semaphores, and so on. The temporal adverbial *now* always refers to the time of utterance, and *yesterday* to the day before the time of utterance. These terms are known as *indexicals*, due to the logical theories that were proposed to explain them (Bar-Hillel 1954). These theories provided a number of *indices*, which are parameters representing the context of utterance on which expressions depend to determine their interpretations. For instance, the speaker and hearer and the time and

location of an utterance are often supplied as indices. A more general class of context-determined expressions are known as *demonstratives*, which include terms such as *this*, *that*, and *those*. The term *deixis* has traditionally been applied to cover the whole spectrum of contextually referential phenomena.

The indexical view of meaning is slightly more sophisticated than the simple referential theory. Rather than establishing their referents once and for all, the interpretation of an expression such as *I* will depend on an index, which must be supplied by the context of utterance. Of course, the issue immediately arises as to just how many indices are necessary for the interpretation of natural language and just where the line is to be drawn, if anywhere, between indexicality and other discourse phenomena such as the resolution of other anaphoric dependencies (see chapter 9). Lewis (1970) proposed a finite set of indices, but these included several specialized indices to deal with the discourse dependency of expressions such as *aforementioned*. Cresswell (1973) then extended the notion of index to allow an arbitrary number of properties determined by the context of utterance, construed very broadly. Kaplan (1977, 1978) is usually credited with formulating the general theory of indexicality, as later put to use by Montague (1973), in which the meaning of an expression is a mapping from a context of utterance to an interpretation. The basic notion of indexicality, though, is clearly defined by Bar-Hillel (1954). In its most general form, an indexical theory of *meaning* provides a general characterization of how *interpretations* may be derived from contexts.

Barwise and Perry (1981, 1983) generalized the functional notion of meaning to a relational one in which a context does not uniquely determine an interpretation. One motivation for Barwise and Perry's relational approach to context and reference is its ability to account for many of the ways in which information can be extracted from an utterance. For instance, if someone utters a noun phrase such as *my wife* and you know from the context who the intended referent is, you will be able to extract the information that the speaker is married, and in fact married to the person in question (on the assumption that the utterance was truthful). The notion of reference is quite subtle though, and we will return to it in section 7.8.

1.2.5 Sense and Reference

Any referential semantic theory that attempts to explain the use of verbs like *believe* is immediately faced with the following puzzling examples (Frege 1892).

- (24) a. The morning star is the evening star.
 b. The evening star is the evening star.
 c. The ancients believed the morning star is the evening star.
 d. The ancients believed the evening star is the evening star.

Because the terms *the morning star* and *the evening star* both refer to the planet Venus, which can be seen on the horizon both in the evening and the morning, a naive referential theory would assign *the morning star* and *the evening star* the same meaning, namely Venus. If the referent of a sentence is taken to be a truth value, as is commonly assumed in truth-conditional semantics, then (24a) and (24b) have the same meaning, because both are true. The further assumption that belief sentences express relationships between their subjects and the meanings of their sentential objects brings the theory up against the obvious counterexamples in (24c) and (24d). Even though the ancients did not realize that the morning star and the evening star were both Venus, (24c) and (24d) will have the same interpretation according to the naive theory.

As a solution to this puzzle, Frege (1892) proposed that the meaning of an embedded sentence was not its referent (*Bedeutung* in German), but its *sense* (*Sinn* in German). In other words, the two sentences (24a) and (24b) have different senses, and hence different meanings. Unfortunately, Frege never formally articulated his notion of sense. But he did clarify one aspect of the theory, namely the relation of senses to referents, by saying “the sense of a proper name is grasped by everyone who knows the language or the totality of designations of which the proper name is a part” (1892, 26). Frege was also clear to separate the notion of sense from that of our ideas of such senses, which are necessarily incomplete and can often be erroneous.

Since Frege, the notion of senses containing the complete set of their denotations has been elaborated in a number of ways. For instance, Church (1951), although he only provides an axiomatic (rather than referential) treatment of Frege’s logical notions, states that an adequate model must provide a referent for each sense. Furthermore, he claims that “there is no difference in principle between this case [formal languages] and that of one of the natural languages” (1951, 110). Carnap (1947), though he disagrees with Church’s approach along a number of dimensions, proposes a semantic approach to sentence meanings in which a proposition is modeled as a function from “states of affairs” to referents. He applies this perspective to *modal adverbs* such *necessarily* and *possibly*. For instance, Carnap assumes a sentence is necessarily true if and only if

it is true under every interpretation of the nonlogical terms it contains (he assumed that logical terms, such as conjunction, negation, and quantification, are always interpreted the same way).

Kaplan (1964) developed Carnap's ideas and applied them to Church's language by modeling states of affairs as logical models of the underlying nonintensional language. For instance, a proposition is necessarily true if it is true in every model. Kaplan's approach suffers from a number of logical drawbacks, most notably its inability to deal with iterated attitudes such as the following.

(25) Terry knows that Chris believes that Dana runs.

The problem is that the intensional model is only one level deep, being built on top of a nonintensional model.

The philosophical and logical treatment of senses was revolutionized by the advent of Kripke's (1959, 1963a) possible-world semantics for modal logics, which cleared up a number of the more technical and empirical problems with previous approaches such as Church's and Carnap's. In Kripke's models, a *possible world* provides complete information about a possible way the world could be. It is assumed that the real world is simply one such possible world. Under such a theory, the sense of a sentence is just a function from possible worlds to truth values. In other words, a sense must determine the truth or falsity of a proposition in each of the possible worlds. Looked at the other way around, a possible world is just an abstract entity that determines the truth or falsity of every proposition. To determine whether a proposition is true in the actual world, it is applied as a function to the actual world, the result being its truth value in the actual world. The added power of Kripke's model derives from the richness of propositions in providing information in different worlds. Thus we could distinguish (24a) and (24b) if there were two possible worlds, one in which the morning star and the evening star are both the planet Venus, and one in which they are distinct. Under such a conception, the senses of (24a) and (24b) would be distinguished. If we then follow Frege in the assumption that belief is a relation between the individual denoted by the subject and the sense of its sentential object, we can also account for the different truth values of (24c) and (24d).

While this discussion has concentrated on sentences, almost all expressions of natural language display subtle distinctions in meanings beyond their immediate referents. For instance, a noun phrase such as *the president of the United States* may denote different individuals in different

possible worlds or at different times in the real world. Common nouns such as *vampire* and *unicorn* might actually apply to some individuals in some possible world. Often the synthetic/analytic division is recast in possible-world semantics as one of being true in the actual world versus being true in all possible worlds.

1.2.6 Lexical Semantics

A significant amount of effort in semantics has concentrated on the classification of the words and phrases of a language. Not surprisingly, several approaches to lexical semantics have been developed. Although lexical semantics will play only a minor role in this book, in this section I will describe a few of the more popular avenues of research in this area.

Traditionally, words have been categorized by their meanings. For instance, in the construction of a thesaurus, words are grouped according to similarity of meaning. Of course, it is difficult, perhaps even impossible, to find pairs of words that are truly synonymous. Often studies are made of the different *connotations* of roughly synonymous words, such as *tall* and *overgrown*. A similar undertaking involves *antonymy*, or words with meanings that vary along a single dimension but are otherwise similar, such as *large*, *medium*, and *small*, or *hot*, *warm*, *tepid*, and *cold*. A significant portion of the basic lexicon of a language can be organized into antonym classes. In addition to thesauruses, there are antonym dictionaries.

Another way in which words are classified involves *taxonomical* relationships. For instance, we often have words for broad categories, with further words reserved for successively smaller subclasses, e.g., *furniture*, *chair*, and *armchair*, or *vehicle*, *car*, and *sedan*.

Often there are groups of words that do not stand in a taxonomic relationship but nevertheless share contexts of use and a broad similarity of meaning. Examples are the pair *husband* and *wife* and the pair *hit* and *kick*. Other examples include *he*, *she* and *it*. Such relationships are known in linguistics as *paradigmatic*. Paradigmatic relationships are roughly defined as those between elements that can occur in the same contexts. A closely related notion is that of *syntagmatic* relations, which hold among words that occur together, but not in variation with one another. For instance, consider the pair *coffee* and *beans* or the pair *cat* and *meow*. A similar pattern is that of *collocational restrictions*. For instance, it is much more natural to use the expression *strong coffee* than the seemingly synonymous *powerful coffee*. A grasp of collocation patterns is essential in

generating natural-sounding utterances and is becoming increasingly important in the mechanical generation of natural-language by computers. Many researchers have attempted to link such classifications with human cognitive organization (Jackendoff 1983, Lakoff 1987). Standard word-association tests can often reveal closely related terms.

Because semantic relationships are notoriously tricky and subtle, much recent research in lexical semantics has focused on how semantic relationships are expressed in the syntactic patterns of language. One instance of this work is in the so-called *linking* of thematic roles and grammatical roles. Roughly, *thematic roles* indicate the semantic relationships between an event and its participants. For instance, in an event of Lee's hitting Pat, Lee is the instigator of the action, and Pat is the recipient. *Grammatical roles*, on the other hand, are marked by subjecthood and objecthood in a sentence, for instance. In English, the subject precedes the verb and the object follows in standard sentences, whereas in languages like Japanese, grammatical roles are reflected in morphological suffixes (or perhaps postpositional particles). In the typological pattern across the verbs of the world's languages we find instigators, often called *agents*, realized grammatically as subjects and recipients, often called *patients*, as objects. This observation formed the core of Fillmore's (1968) highly influential theory of *case grammar* and is the basis of the lexical *linking theory* of lexical functional grammar (L. Levin 1987). A semantic foundation for thematic roles was put forward in Dowty 1991, which also contains an extensive bibliography on the subject. B. Levin (1993) provided an extremely detailed survey and classification of the lexical patterns of English verbs from both a syntactic and a semantic perspective.

The relation between thematic roles and grammatical roles is only one kind of pattern that has been found to be *universal* across languages. Linguists, especially syntacticians of a typological persuasion, are usually much more concerned with universal patterns than language-specific ones. Some universal patterns simply concern the organization of language. For instance, all languages allow reference to individuals, actions, and relations, as well as methods for combining these, such as conjunction and disjunction. Universals can also be found in the possible meanings assigned to certain kinds of categories. We return to this topic in section 3.3, where we study constraints on the meanings of quantificational determiners, such as *every*, *no*, and *most*.

Another avenue of exploration in lexical semantics has been the *decompositional* approach. In a decompositional theory, primitive meaning units

are presupposed, along with limited means of combining them. Such studies were at the heart of the *structuralist* theory of semantics and later were incorporated into the theory known as *generative semantics*. Such theories often attempted to link their primitives to cognition. While most of the specific assumptions of the structuralists and generative semanticists have been abandoned, to a large extent their basic approach is reflected in contemporary research on lexical semantics. Decompositional approaches commonly resurface in computational approaches to natural-language understanding, the most widely known of which are Schank and Abelson's (1977) conceptual-dependency theory and knowledge-representation schemes based on semantic networks (Brachman 1979).

1.3 TOPICS IN PRAGMATICS

The branch of linguistics devoted to studying the effects of context on utterance interpretation has traditionally been known as *pragmatics*. Semantics, on the other hand, is usually taken to refer to the literal, context-independent aspects of meaning. Of course, the location of the dividing line between literal and context-sensitive aspects of interpretations is notoriously difficult to pin down. For instance, indexicality has traditionally been considered a pragmatic phenomenon, because the interpretation of indexical elements is clearly context-sensitive. More recently, the term *pragmatics* has been narrowed to encompass just those aspects of meaning that involve more than the literal interpretation of utterances, and this is how I will use the term. In this section I briefly survey three of the main areas of pragmatics: conversational implicature, speech acts, and discourse analysis. A thorough survey of the field, including the traditional issues of indexicality, deixis, and presupposition, can be found in Levinson 1983. The survey of discourse by Brown and Yule (1983) is also a good source for further reading. I will also discuss in this section the role of metaphor in language interpretation. While the study of metaphor is generally not thought to be pragmatic in nature, I include it here because of its nonliteral nature.

1.3.1 Conversational Implicature

Entailment and presupposition are only two of the ways in which language can convey information. Another important means of conveying information by language, originally articulated by Grice (1975), is through what he calls *conversational implicature*. Roughly, Grice noted that utter-

ances may convey information by means of their use in context and that such information goes beyond the literal content and entailments of the utterances. As a typical example, Grice provides the following dialogue.

(26) *Speaker A*: I am out of petrol.

Speaker B: There is a garage around the corner.

Grice notes that the response of speaker *B* to speaker *A*'s request, namely that there is a garage around the corner, has the conversational implicature that speaker *B* believes the garage is open for business and sells petrol. While this latter piece of information is not explicitly entailed by the statement that there is a garage around the corner, a cooperative use of the response implicates it.

In contrast to true entailment, implicatures are *cancelable*. That is, the response of speaker *B* in (26) does not fail to be true if the garage around the corner is closed. On the other hand, if around the corner there is nothing but a convenience store, then speaker *B*'s statement is simply false. Furthermore, implicatures can be explicitly overridden. For instance, speaker *B* may have continued his response in (26) with *but it's not open, so you'll have to go further down the road*.

The cornerstone of Grice's account of conversational implicature is the view of language as just one form of cooperative behavior. Along these lines, Grice formulated his *cooperative principle* for language as follows.

Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

Grice then observed that conversational implicatures arise due to the interaction of the cooperative principle with discourse. It should be clear how the implicature derived from the response of speaker *B* in (26) follows from the assumption that speaker *B* is being cooperative.

Under Grice's analysis, implicatures are derived from the meaning of an utterance and its interaction with the cooperative principle. The syntactic form of the utterance is irrelevant, and thus Grice claimed that implicatures are *nondetachable*. True implications are obviously nondetachable by definition. Presuppositions, on the other hand, are detachable. For instance, *The winner of the race is Sandy* presupposes the existence of a winner, whereas *Sandy won the race* does not, even though both sentences would be true in exactly the same set of circumstances.

Given his analysis, Grice drew the related conclusion that implicatures are *nonconventional*. That is, implicatures are derived from the meaning of

sentences and their situations of use, rather than from conventions of language per se. Support for this claim can be found by completing exercise 2.

An equally important means of conveying information that crucially involves the cooperative principle arises when the principle has apparently been disregarded. For instance, consider Grice's example of a tutor writing a reference letter for a pupil who is a candidate for a job in philosophy, who reports only the following.

- (27) Dear Sir,
 Mr. *X*'s command of English is excellent, and his attendance at tutorials has been regular.
 Yours, etc.

Grice reasonably suggests that the implicature here is that *X* is no good at philosophy. Given the context, if *X* had been an excellent student, the tutor would have indicated as much. The omission of this information forms the basis of the implicature.

Grice breaks the cooperative principle down into four broad categories, which he referred to as *conversational maxims*.

- (28) a. *Quantity*
 i. Make your contribution as informative as is required.
 ii. Do not make your contribution more informative than is required.
- b. *Quality*
 Try to make your contribution one that is true.
 i. Do not say what you believe to be false.
 ii. Do not say what you lack adequate evidence for.
- c. *Relation*
 Be relevant.
- d. *Manner*
 Be perspicuous.
 i. Avoid obscurity of expression.
 ii. Avoid ambiguity.
 iii. Be brief.
 iv. Be orderly.

Any of these Gricean maxims may be used to implicate information that is not entailed or presupposed by the literal meaning of an utterance. Grice provides examples of uses of each, as well as a more precise defini-

tion of conversational implicature and its relation to the maxims. Further examples and discussion may be found in Levinson 1983. Grice realistically admits that this list is likely to be incomplete, and in fact there have been many who have reclassified the cooperative principle into coarser, finer, or simply differently organized maxims. A currently popular approach has been to take the notion of *relevance* as central (Sperber and Wilson 1986).

1.3.2 Speech Acts

Another way in which language may be used to convey information that goes beyond its literal meaning is by the use of *speech acts*. The theory of speech acts, originally developed by Austin (1961) and further elaborated by Searle (1965, 1975), also places a great deal of importance on the role of language as cooperative behavior.

Austin (1961) introduced the notion of a *performative utterance*, and as typical instances he provided the following:

- (29) a. I name this ship the *Queen Elizabeth*.
b. I now pronounce you man and wife.
c. I give and bequeath my watch to my brother.

Utterances of the sentences in (29) can be used to christen a ship, marry a couple, or create a will and testament. As Austin claims, “If a person makes an utterance of this sort we should say that he is *doing* something rather than merely *saying* something” (1961, 251). Thus the key fact about performative utterances is that their users are performing some sort of action with language, rather than making a true or false claim about the world. This places an analysis of performatives outside of the scope of this book, which focuses on truth-conditional semantics.

Austin noted that performative sentences, such as those in (29), derive their meaning from the conventions of a given language community. Thus they could fail to be meaningful in many of the same ways as simple declarative sentences. For instance, an utterance of (29b) will be infelicitous if the speaker does not have the authority to perform a marriage or, in some cultures, if one member of the couple in question is already married. Performatives may also fail for more mundane semantic reasons. For instance, (29c) will fail if the speaker does not own a watch or does not have a brother.

Austin (1962) generalized his approach to language as action, introducing the general term *illocutionary act* for the use of language to

perform actions. The most well-known exposition of a theory of illocutionary acts is due to Searle (1965), although he referred to them using the now standard term *speech acts*. Searle provided examples of verbs that were typically used in speech acts, such as *warn*, *remark*, *command*, *criticize*, *apologize*, *promise*, and *welcome*. Austin claimed there were thousands of such speech-act verbs in English.

Searle stressed the following point.

It is not, as has generally been supposed, the symbol or word or sentence, or even the token of the symbol or word or sentence, which is the unit of linguistic communication, but rather it is the *production* of the token in the performance of the speech act that constitutes the basic unit of linguistic communication. To put this point more precisely, the production of the sentence token under certain conditions is the illocutionary act, and the illocutionary act is the minimal unit of linguistic communication. (1965, 222)

In other words, Searle emphasized the speech context, including the intentions of the speaker, in the interpretation of an utterance.

Perhaps the most influential aspect of Searle's analysis of speech acts was his account of *indirect speech acts* (Searle 1975). A typical indirect speech act would be an utterance of the following sentence as a request to pass the salt.

(30) Could you pass the salt?

Interpreted literally, (30) is a yes/no question. Yet one common usage of such a polar interrogative is as a request to pass the salt. While this may seem to be an almost completely conventionalized, polite usage, consider another example due to Searle.

(31) It is cold in here.

Such a statement could also be used as a request, say to turn up the heat or close a window, depending on the context.

One of the key aspects of Searle's analysis of speech acts was his emphasis on the component illocutionary acts that make up the larger utterance. Searle argued that to understand utterances such as (30) and (31), it is necessary to understand their component speech acts, including determining the reference of definites (such as *the salt*) and indexicals (such as *you* and *here*), as well as to understand the literal meanings of the expressions (such as *in* and *pass*). Interpreting the speech act itself is a matter of cooperation among cognitive agents, rather than a process determined by the linguistic conventions adopted by a linguistic community, and is well beyond the scope of this book. I will concentrate almost

exclusively on the referential and truth-conditional aspects of meaning and interpretation.

As Grice points out, in his analysis of meaning, a crucial feature of *A*'s meaning something by *x* is that "*A* intended the utterance of *x* to produce some effect in an audience by means of the recognition of this intention" (1975, 51). Searle (1965) distinguishes the effects of utterances that come about by the hearer's recognition of the speaker's intentions from those that do not. The former case simply involves ordinary uses of language. As an example of the latter case, he imagines a situation where an American soldier is captured by the Italian army during World War II. The American remembers one line of German from a poem, namely *Kennst du das Land, wo die Zitronen blühen?* (literally, 'Knowest thou the land where the lemon trees bloom?'), and utters this sentence in an effort to convince his Italian captors that he is German. For the American's ploy to succeed, it is crucial that the hearers not recognize the speaker's intention in uttering the line from a German poem. The American's utterance is used for what Austin terms its *perlocutionary effect*, rather than its illocutionary effect. As another example of perlocutionary effects, he cites the case of a speaker who uses abstruse language in an effort to impress the hearer with his erudition.

1.3.3 Discourse Structure

It is evident that sentences cannot simply be randomly organized in the presentation of complicated information. Consider what would happen if all of the sentences in this chapter were put into a bag, and then extracted at random to form the text (a technique actually advocated by the beat author William S. Burroughs). Theories of discourse structure are generally centered around the organizations of written or spoken language as it occurs in narrative fiction, interpersonal dialogues, newspaper articles, television commercials, instruction manuals, classroom interaction, and just about any other conceivable occurrence of language in context. Usually such analyses focus on multisentence discourses and intersentential relationships. Brown and Yule (1983) provide a thorough survey of the field.

Of the more linguistically oriented studies, a great deal of attention has been paid to notions such as the *topic* and *focus* of a text, the role of stress and intonation on discourse structure, the theme of an exposition and its development, and the computation of the antecedents of pronouns and elided elements based on discourse structure. Various theories have been

proposed, including the development of grammars for texts. Most computational implementations of discourse structuring revolve around speaker and hearer intentions (see Grosz and Sidner 1986) and around the knowledge of prototypical events and their structure (Schank and Abelson 1977). Like other aspects of pragmatics, theories of discourse structure are usually constructed using semantic analyses of single phrases, clauses, and sentences as a basis. A popular technique, embodied in recent theories of *dynamic semantics* (Groenendijk and Stokhof 1990, 1991; Chierchia 1995), has been to model the meaning of a sentence as a mapping from discourse contexts to discourse contexts. The effect of a sentence with a definite noun phrase might be to introduce an individual into the domain of discourse, whereas a sentence with a pronoun might require that the input discourse context contain an antecedent for it. This allows the effects of longer stretches of text to be modeled by composing the meanings of sentences.

1.3.4 Metaphors, Idioms, and Language Change

One of the most subtle aspects of the semantics of natural language arises in the use of *metaphor*. Intuitively, a metaphorical usage of a term involves some shift in meaning away from how the term is conventionally understood. Thus metaphorical uses of terms or phrases are often contrasted with their literal or conventional uses.

First, let us consider some examples of metaphor, drawn from Searle 1979.

- (32) a. Sandy is a gorilla.
 b. I have climbed to the top of the greasy pole. (Benjamin Disraeli)
 c. Juliet is the sun.

In each of these cases, it is clear that the intended meaning is not the literal one. In each case there is a rough, literal paraphrase. For example, (32a), according to Searle, means that Sandy is fierce and nasty. The second example, (32b), means that Disraeli struggled to become prime minister of Great Britain. In the last example, (32c), there may be any number of plausible interpretations, or the speaker, Romeo, may have meant a great many things with the utterance, including that he could not live without Juliet or that she brought sunshine into his life (another metaphor).

A theory of metaphor must provide answers to a number of questions. Searle (1979) isolates the following primary issues. First, we must under-

stand how metaphorical meaning differs from literal meaning and from other figurative uses of language. Second, a theory of metaphor should explain how it is possible for speakers to formulate and for hearers to understand metaphors. The final piece of a theory of metaphors is actually extrinsic to the understanding of metaphors per se but is of perhaps of the greatest importance, both psychologically and semantically; the issue is why speakers use metaphors rather than expressing themselves literally.

Searle (1979) argues that the key aspect of metaphorical interpretation is the divergence between *speaker meaning* and *literal meaning*. He concludes that metaphorical meaning is a kind of speaker's meaning and that it crucially depends on understanding the literal meaning of an utterance. But rather than trying to provide a unified theory of metaphorical interpretation, Searle concludes that the issue of how metaphors function reduces to the issue of how one thing or concept reminds us of another. Thus the reason that some metaphors fail is that some things simply cannot be used to remind us of other things. This intentionally vague "theory" of the interpretation of metaphors does provide an explanation of their utility. A common view is that we depend on metaphors when there is no literal usage that will convey the meaning we intend, or at least not with the same force, economy, or perspicuity.

Metaphors have often been argued to be one of the driving forces both in scientific change and in language change. An excellent discussion of their utility for the sciences and the humanities can be found in Rorty 1989. Lakoff and his colleagues (1987, 1989; Lakoff and Johnson 1980) have greatly revived interest in metaphorical interpretations of natural language, focusing on those derived from our physical makeup as humans. Such theories have been especially important in the understanding of prepositions, such as *in*, as can be witnessed by the extensive study of prepositions in Herskovits 1986. For instance, consider the following examples.

- (33) a. The block is in the box.
b. Robin is in the army.
c. Robin is in charge.
d. Robin is in a good mood.

All of these sentences use the preposition *in* in a different sense, ranging from physical containment to disposition. These examples are only the tip of the iceberg; there are literally hundreds of common uses of most of the simple prepositions. Most of these uses have some metaphorical

relation to the notion of physical inclusion, though the connection is often tenuous.

Most theories of *idioms* treat them simply as dead metaphors. That is, an idiom is thought of as nothing more than a metaphor whose meaning has become frozen. For instance, consider the following sentences with idiomatic verb phrases.

- (34) a. Sandy kicked the bucket.
 b. Sandy let the cat out of the bag.

The first example, (34a), means that Sandy died. The second, (34b), means that Sandy gave away a secret. Note the degree to which these expressions are frozen. Variants such as *boot the pail* and *allow the feline to escape*, while having very similar literal meanings to *kick the bucket* and *let the cat out of the bag*, simply do not convey the same meanings as their idiomatic counterparts (except insofar as they invoke the idiom itself).

If we follow conventional wisdom and suppose that idioms are like any other expressions with conventionalized meanings, then there is only one problem left to solve, and that is the nature of their syntactic realization. This is complicated, though, because of the phrasal nature of idioms. The pattern of allowable modifiers and inflections, the discontinuity of idioms themselves, and their seeming disregard for normal lexical constituency require special care. Recently researchers have proposed two theories of the syntax and semantics of idioms that are consistent with the semantic approach I adopt in this book (Nunberg, Sag, and Wasow 1994; van der Linden 1993).

1.4 METHODOLOGY

In this section I discuss the methodology that I will employ throughout the rest of the book to develop a theory of natural-language semantics. I begin by motivating a compositional approach to semantics. I then discuss the use of mathematical models in linguistics.

1.4.1 Productivity

One of the most fundamental aspects of natural language lies in our ability to produce and understand novel utterances. This ability is often referred to as linguistic *creativity* or *productivity*. Perhaps even more surprising is that the set of grammatical sentences in English to which a semantic theory must assign meanings is unbounded (though arguably

countable). To see that this is so, simply consider the following sequence of expressions.

- (35) a. The kid ran.
b. I know that the kid ran.
c. You know that I know that the kid ran.
d. I know that you know that I know that the kid ran.

Significantly, each of these sentences has a distinct meaning. There are, of course, many other ways to construct unbounded sequences of sentences with distinct meanings.

It might be argued that there should be some upper bound to the length of a sentence, and any longer sentence should be classified as ungrammatical. After all, with only a relatively short lifespan, it would be impossible to produce a sentence with a trillion words. It is also clear that sentences of more than a few hundred words will be incomprehensible if they have a structure that is more complex than a simple conjunction of simpler sentences. With the examples in (35), it is not at all clear where to draw the line. It seems that if any sentence in (35) is grammatical, then so is the next sentence. There seems to be no principled way in which to decide at what point sentences become too long or complex to be considered grammatical. In every theory of grammar proposed to date, if *s* is admitted as a sentence, then *I know that s* is also admitted. Furthermore, it is clear how to semantically interpret each of the sentences in the sequence. Rather than choosing a length beyond which sentences are considered ungrammatical, linguists endeavor to create more general, productive theories.

The primary goal of theoretical linguistics, as opposed to psycholinguistics, is to formulate a theory of language itself, rather than the human ability to process it. Chomsky (1965) drew a distinction between linguistic *competence*, on the one hand, and *performance* on the other. Chomsky believed that our competence comprised a system of rules for the construction of utterances and their meanings, such as the one stated above, namely that if *s* is a sentence, so is *I know s*. Chomsky further assumed that the knowledge of some such rules is innate, as is the ability to use such rules. Linguistic performance is subject to the vagaries of all human cognitive activities, such as attention span, alertness, distractions, and so on. Even relatively short sentences can quickly overcome human

processing limitations if their structure is too complex. For instance, consider the case of center-embedded relative clauses.

- (36) a. The mouse ran.
 b. The mouse that the cat chased ran.
 c. The mouse that the cat that the dog bit chased ran.
 d. The mouse that the cat that the dog that the person owns bit chased ran.
 ⋮

Even the third example, (36c), with only two center embedded clauses, is pretty incomprehensible.

Abstracting away from performance issues, Chomsky took it to be the job of linguistics to construct a competence model. Even before Chomsky, most linguists distinguished human language itself from our ability to use it. Notable exceptions were those who took language to be determined by our use of it, such as the later Wittgenstein and the behavioral psychologists. In the last few decades there has been a tremendous amount of effort expended on the debate concerning whether humans are born with innate linguistic knowledge, or whether language, like many other aspects of our knowledge, is induced from the environment. Often these debates have been sparked by particular theories of the organization of the knowledge of language in our brains and of whether such knowledge can be separated from our presumably more procedural ability to use such knowledge. A detailed and relatively nonpartisan survey of the empirical data and various positions that have been taken with respect to language acquisition and processing can be found in Taylor and Taylor 1990.

1.4.2 Compositionality

The fundamental approach to modeling the productivity of human language has its roots in the writings of Frege (1892). He was attempting to explain how a finite system can lead to a productive grammar of language. Frege presented roughly the following theory of how such a system could be organized, which has come to be known as the principle of *compositionality*.

DEFINITION: COMPOSITIONALITY The meaning of an expression is determined by the meanings of its parts and the way in which they are combined.

On the assumption that there are only finitely many atomic parts involved in language, say those corresponding to words, and furthermore that there

are only finitely many grammatical constructions that can be employed to build larger structures, this definition shows how a finite number of building blocks can be used to generate an infinite set of sentences and interpretations. Typically, the smallest meaningful unit in language is assumed to be the *morpheme*. Morphemes can be *roots* or *stems*, such as *happy* or *run*; they can be *inflectional affixes*, such as the tense suffixes *-ed* and *-ing*, or they can be *derivational affixes*, such as the negative *un-* or the nominalizer *-tion*. The morphological structure of some languages is much less concatenative than in English. In some Bantu languages, morphological marking takes the form of tones overlaid on the stems. In other languages, such as Arabic and Sierra Miwok, stems consist of consonants and inflections of vowels, which are then merged into words. But in any case, there is typically a simple way to identify the underlying meaningful units and their compositional structure.

As Partee (1984b) points out, with a statement of Frege's principle as general as the one above, the notion of compositionality appears nearly trivial. There seems very little hope of building a noncompositional theory of meaning that would account for the productivity of language. In practice, with a rich enough collection of meanings and methods for combining them, just about any theory associating strings of symbols with meanings can be made compositional in the sense above. Thus rather than being an empirical claim, the compositionality principle has come to be a methodological principle. But strictly compositional theories are not, in fact, the only possibility; often a compositional formulation of a semantic theory may not be the most perspicuous one. For instance, consider discourse-representation theory, which has both compositional and noncompositional presentations (see section 7.10).

The theories of meaning I present will be compositional, but this statement requires some elaboration of what I mean by *meaning*. I will use the term *meaning* to denote the context-independent aspects of the semantics of an expression. I reserve the term *interpretation* for the semantic result of refining the meaning with contextual information. For instance, the meaning of the indexical expression *I* is the speaker, the interpretation of an utterance of *I* will be the speaker who uttered the expression. In what follows, I will adopt the methodology of Kaplan (1977, 1978), as discussed in section 1.2.4, and treat the meaning of an expression as a relation between contexts and interpretations. Furthermore, I follow Frege (1892) in treating interpretations as senses rather than referents, which provides an additional degree of freedom.

1.4.3 Model Theory and Grammar Fragments

The primary tool employed by contemporary semanticists is model theory, which is a branch of mathematical logic dealing with the interpretation of logical languages and theories expressed within them. Model-theoretic approaches to natural-language semantics were originally proposed by Tarski (1935). Following in the Tarskian tradition, most of the model-theoretic approaches to natural language have been devoted to truth-conditional semantics.

Logical model theory as applied to natural-language semantics has often been misconstrued as making stronger claims than its proponents are comfortable with. For instance, model theory is usually employed to produce models of a natural language itself, not our knowledge of or facility to use natural language. Logical models should rather be understood as idealizing away from speakers and thus should not be given epistemological interpretations. Along another dimension, models are rarely comprehensive, usually modeling only a small part of language. Furthermore, the objects employed in models are usually not intended to be provided with a metaphysical interpretation, though many proponents of model theory have done just that. These properties of model-theoretic semantics can be productively compared with the theory of differential equations as applied to the motion of a football. No one would claim that a football computes a differential equation, or even represents one, as it flies through the air. Nor would most physicists claim that differential equations were real in the metaphysical sense. Physicists also realize that most differential equations provide only an approximation of a football's movement, abstracting away from considerations such as turbulence, air temperature, and other details. Nevertheless, differential equations do provide useful approximations; the basic theory of differential equations and the classical theory of mechanics (as opposed to the more accurate relativistic one) formed the basis for the construction of the Apollo rockets that transported humans to the moon.

The most ambitious and influential application of model theory to natural-language semantics was carried out by Richard Montague in the late 1960s. He went so far as to write, "I reject the contention that an important theoretical difference exists between formal and natural language" (Montague 1970a, 188). While I believe this claim to be exaggerated, I draw many insights from both Montague's general grammatical framework (1970b) and his particular grammars (1973).

Chomsky (1957) pioneered the technique of building precise models of subsets of natural language, concentrating on syntactic and phonological phenomena. Montague employed logical model theory to develop semantic models in the same spirit. Since the early 1980s, Chomsky (1990b) has argued against the utility of building grammar fragments, and subsequently their use in syntax and phonology has gone out of vogue in “mainstream” linguistic theorizing. But there is still a strong community of syntacticians who employ the fragment-building technique, including those working in lexical functional grammar (LFG) (Kaplan and Bresnan 1982), generalized phrase-structure grammar (GPSG) (Gazdar, Klein, Pullum, and Sag 1985), head-driven phrase-structure grammar (HPSG) (Pollard and Sag 1994), and the formalism I adopt, categorial grammar (CG) (Morrill 1994a). The construction of grammar fragments remains prevalent in studies of natural-language semantics. One of the primary threads of this book is the construction of a semantic model of a relatively large fragment of English. In following a compositional approach to semantics, I will also need to formulate a precise syntactic model of English, though this will not be my principal concern, and I will refrain from introducing too much syntactic detail into my models.

In the next two chapters I will be developing the logical model theory I will later employ in extensional grammar fragments. This will involve developing a general theory of functions, known as the λ -calculus, which I do in the next chapter. In the chapter after that, I will build models of higher-order logic, which are built on the foundation of the λ -calculus.

Exercises

1. From a source such as a newspaper or transcribed speech, find several occurrences of metaphor. For each instance, explain what the intended interpretation is and what clues there are that it should be interpreted nonliterally. From a similar source, find several occurrences of presupposition. Why did the author or speaker choose to convey information by presupposition? From a similar source, find several instances of implicature and explain why implicature was chosen as a vehicle for conveying information.
2. Sketch a situation and an utterance in that situation where the utterance is true but its implicature is not. Next provide a situation and an utterance that is literally false but has a true implicature. Finally, provide a sentence and sketch two situations in which it would have different implicatures. (Levinson 1983)
3. Provide (the initial parts of) three or four sequences of grammatical sentences whose length is unbounded. Do not use the same grammatical pattern as in (35) or (36).

4. Choose a few contiguous sentences from some naturally occurring text and discuss the nonlinguistic knowledge necessary to interpret them. Consider both contextual information and encyclopedic information.
5. Jot down a few uses of indirect speech acts from discourses that you have recently participated in.
6. Provide an example of a linguistic community, and discuss both the motivation for and the kind of language used in the community.
7. Discuss the extent to which our language affects our culture and vice-versa.
8. Discuss the range of empirical facts that we have available to us as linguists. How are such facts interpreted?
9. Choose a short thesaurus entry and discuss how the synonyms listed vary in connotation.