Finally, the extreme view of *cognitive linguistics* (Langacker, 1987) dispenses with formal linguistic rules altogether, focusing instead upon the communicative conventions of particular languages.

CHAPTER SIX

# Semantics

## INTRODUCTION

How do we represent the meaning of words? How is our knowledge of the world organised? These are issues concerned with the study of meaning, or *semantics*. In the previous chapter we saw how the sentence processing mechanism parses sentences to construct a representation of the syntactic relationships between words. Important as this stage might be, it is only an intermediate step towards the real goal of comprehension, constructing a representation of the meaning of the sentence. In this chapter we shall look at how the meaning of individual words are represented, and in the next how we combine these meanings to form a representation of the *meaning* of the whole sentence and beyond.

Our discussion of non-semantic reading in Chapter 4 showed that words and their meanings can be dissociated. There is further intuitive evidence to support this dissociation (Hirsh-Pasek, Reeves, & Golinkoff, 1993). First, we can translate words from one language to another. Furthermore, not every word meaning is represented by a simple, single word in every language (see Chapter 10 for further discussion of this). Second, there is an imperfect mapping between words and their meanings such that some words have more than one meaning (ambiguity) while some words have the same meaning (synonymy). Third, the meaning of words to some extent depends upon the context.

The word "big" means different things in the phrases "the big ant" and "the big rocket".

Tulving (1972) drew what is now regarded as a fundamental distinction between *episodic* and *semantic* memory. Episodic memory is our memory for events and particular episodes; semantic memory is, in simple terms, our general knowledge. Hence my knowledge that the capital of France is Paris is stored in semantic memory; my memory of learning as a child in a geography lesson at school when the Eiffel Tower was built is an instance of an episodic memory. Semantic memory develops from or is abstracted from episodes which may be repeated many times. I cannot now recall when I learnt the name of the capital of France, but clearly I must have been exposed to it at least once. We have already seen that our mental dictionary has been given the name *lexicon*, and similarly our store of semantic knowledge can be called our mental *encyclopaedia*. Clearly there is a close relationship between the two, both in development and in the developed system, but they must also be separable for the reasons given above. Neuropsychology reveals important dissociations in this respect. We have seen that words and their meanings can be dissociated; but we must be wary of confusing a loss of semantic information with the inability to access or use that information. This problem is particularly important when we consider semantic neuropsychological deficits.

The notion of meaning is closely bound to that of *categorisation*. Concepts are very closely related to meaning. A concept refers to a mental representation that determines how things are related or categorised. It enables us to group things together, so that instances of a category all have something in common. Thus concepts somehow specify category membership. All words have an underlying concept, but not all concepts are labelled by a word. We have a word "dog" which we can use about certain things in the world, but not others. There are two fundamental questions about this. We can say that the philosophical question is: how does the concept of "dog" relate to the members of the category dog? The psychological question is: how is the meaning of "dog" represented, and how do we pick out instances of dogs in the environment? As just hinted, we could in principle have a word, say "brog", to refer to brown dogs. That we do not have such a term is probably because this is not a particularly useful concept in this domain. Rosch (1978) pointed out that categorisation is not arbitrary, but determined by two important features of our cognitive system. First, the categories we form are determined in part by the way in which we perceive the structure of the world. Perceptual features are tied together because they form objects and have a shared function. How the categories we form are determined by the biological factors is an important topic, about which little is known. We shall return to this with the example of colour in our discussion of the relationship between language and thought in Chapter 10. Second, the structure of categories might be determined by a principle known as *cognitive economy*. This states that semantic memory is organised so as to avoid excessive duplication. Of course we cannot be too economical, as we often need to make distinctions between members of some categories more than others. Rosch proposed that this compromise resulted in a *basic level* of categorisation which tends to be the default level at which we categorise and think unless there is particular reason to do otherwise. Unless given reason to do otherwise, we use the basic level of "chairs", rather than the lower level of "armchairs" or the higher level of "furniture".



It should be obvious that the study of meaning therefore necessitates capturing the way in which words refer to things that are all members of the same category and have something in common, yet are different from non-members. (Of course something can belong to two categories at once: we can have a category labelled by the word "ghost", and another by the word "invisible", and indeed we can join the two to form the category of invisible ghosts labelled by the words "invisible ghosts".) There are two issues here. First, what distinguishes items of one category from items of another? Second, how are hierarchical relationships between categories to be captured? There are category relationships between words. For example, the basic level category "dog" has a large number of category *superordinate* levels above it (such as "mammal", "animal", "animate thing", and "object") and *subordinates* (such as "terrier", "rottweiler", and "alsatian"—these are said to be category co-ordinates of each other).

Hierarchical relationships between categories are one clear way in which words can be related in meaning, but there are other ways that are equally important. Some words refer to properties of things referred to by other words (e.g. "dog" and "paw"). Some words (*antonyms*) are opposites in meaning (e.g. "hot" and "cold"). We can attempt to define many words: for example, we might offer the definition "unmarried man" for "bachelor". Another fundamental issue for semantics concerns how we should capture all these relationships.

It should by now be clear that semantics is in many ways the interface between language and the rest of perception and cognition. This relationship is made explicit in the work of Jackendoff (1983), who proposed a theory of the connection between semantics and other cognitive, perceptual, and motor processes. From these considerations, he proposed two constraints on a general theory of semantics. The grammatical constraint says that we should prefer a semantic theory that explains otherwise arbitrary generalisations about syntax and the

lexicon. Both some AI theories and theories based on logic (in particular, a form of logic known as predicate calculus) fail this constraint as to work at all they both have to make up entities that do not correspond to anything real. The cognitive constraint says that there is a level of representation where semantics must interface with other psychological representations such as those derived from perception.

In this chapter we will focus upon two main topics. First, how do we represent the meaning of words? In particular, how does a model of meaning deal with the issues we have just raised? Second, what does the neuropsychology of meaning tell us about its representation and its relationship with the encyclopaedia? We will consider the development of meaning, and how it is related to exposure to specific episodes, in Chapter 12.

## CLASSICAL APPROACHES

It is useful to distinguish immediately between a word's *denotation* and its *connotation*. The denotation of a word is its core, essential meaning. The connotations of a word are all of its secondary implications, or emotional or evaluative associations. For example, the denotation of the word "dog" is its core meaning: it is the relation between a word and a class of objects to which it can refer. The connotations of "dog" might be "nice", "frightening", or "smelly". Put another way, everyone agrees on the denotation, but the connotations differ from person to person. Here we are primarily concerned with denotation, although the distinction can become quite hazy.

Ask a person on the street what the meaning of "dog" is, and they might well point to one. This theory of meaning that words mean what they refer to is one of the oldest, and is called the *referential theory*. There are two major problems with this lay theory, however. First, it is not at all clear how such a theory treats abstract concepts. How can you point to "truth", yet alone point to the meaning of a word such as "whomsoever"? Second, there is a dissociation between a word and the things to which it can refer. Consider the words "Hesperus" (Greek for "The Evening Star") and "Phosphorus" (Greek for "The Morning Star"). They have the same reference or extension in our universe, namely the planet Venus, but they have different senses or *intensions*. Indeed, the ancients did not know that they were the same thing, so even though the words "Hesperus" and "Phosphorus" actually refer to the same thing (the planet Venus), the words have different senses. "Hesperus" could only be used to refer to a planet in the evening sky, and "Phosphorus" could only be used to refer to a planet in the morning sky. This distinction



was made explicit in the work of Frege (1892/1952), who distinguished between the *sense* (often called the *intension*) of a word and its *reference* (often called its *extension*). The intension is its abstract specification or meaning, determining how a word is related in meaning to other words, and which specifies the properties an object must have to be a member of the class, while the extension is what it stands for in the world, that is the objects picked out by that intension. These notions can be extended from words or descriptive phrases to expressions or sentences. Frege took the extension of a sentence to be its truth value (which is simply whether it is true or not), and its intension to be its underlying sense that was derived by combining the intensions of the component words. This formal semantics approach of building logical models of meaning has been developed by logicians into complex systems of meaning known as *model-theoretic semantics*. (Because of the importance of truth in these theories, it is sometimes known as truth-theoretic semantics.) Although the original idea was to provide an account of logic, mathematics, and computing languages, its application has been extended to natural language.

The importance of formal approaches to meaning for psycholinguistics is not clear. Although they help refine what meaning might be, they appear to say little about how we represent or compute it.

## SEMANTIC NETWORKS

One of the most influential of all processing approaches to meaning is based on the idea that the meaning of a word is embedded within a network of other meanings. In a semantic network, knowledge is given meaning only by the way in which it relates to other knowledge. Some of the earliest theories of meaning, from Aristotle to the Behaviourists, viewed meaning as deriving from a word's *association*. From infancy, we are exposed to many episodes involving the word "dog". For the behaviourists, the meaning of the word "dog" was simply the sum of all our associations to the word: it obtains its meaning by its place in a network of associations. The meaning of "dog" might involve an association with "barks", "four legs", "furry", and so on. It soon became apparent that association in itself was insufficiently powerful to be able to capture all aspects of meaning. There is no structure in an associative network, with no relationship between words, no hierarchy of information, and no cognitive economy. In a semantic network, this additional power is obtained by making the connections between items do something—they are not merely associations representing contiguity

of frequent co-occurrence, but themselves have a semantic value. That is, in a semantic network the links have meaning.

### The Collins and Quillian semantic network model

Perhaps the best known example of a semantic network is that of Collins and Quillian (1969). The idea arose from an attempt to develop a *teachable language comprehender* to assist machine translation between languages.

A semantic network is particularly useful for representing information about natural kind terms. These are words that refer to naturally occurring categories and their members—such as types of animal or metal or precious stone. The scheme attributes fundamental importance to their inherently hierarchical nature: for example, a bald eagle is a type of eagle, an eagle is a type of bird of prey, a bird of prey is a bird, and a bird is a type of animal. This hierarchical format suggests a straightforward way to implement cognitive economy. If you store the information that birds have wings at the level of bird, you do not need to repeat it at the level of particular instances (e.g. eagles, bald eagles, and robins). An example of a fragment of such a network is shown in Fig. 6.1. In the network, *nodes* are connected by links which specify the relationship between the linked nodes; the most common link is an ISA link which means that the lower level node "is a" type of the higher level node. Attributes are stored at the lowest possible node at which they are true of all lower nodes in the network.

*The sentence verification task.* One of the most commonly used tasks in semantic memory research is that of *sentence verification*. Subjects are presented with simple "facts" and have to press one button if the sentence is true, another if it is false. The reaction time is an index of how difficult the decision was. Collins and Quillian (1969) presented subjects with sentences such as (1) to (4).

1. A robin is a robin.
2. A robin is a bird.
3. A robin is an animal
4. A robin is a fish.

Sentence (4) is of course false. Sentence (1) is trivially true, but it obviously still takes subjects some time to respond "yes"; clearly they have to read the sentence and initiate a response, but it does provide a baseline measure. The response time to (1) is less than to (2), which in turn is less than that to (3). Furthermore, the difference between the reaction times is about the same—that is, there is a linear relationship.



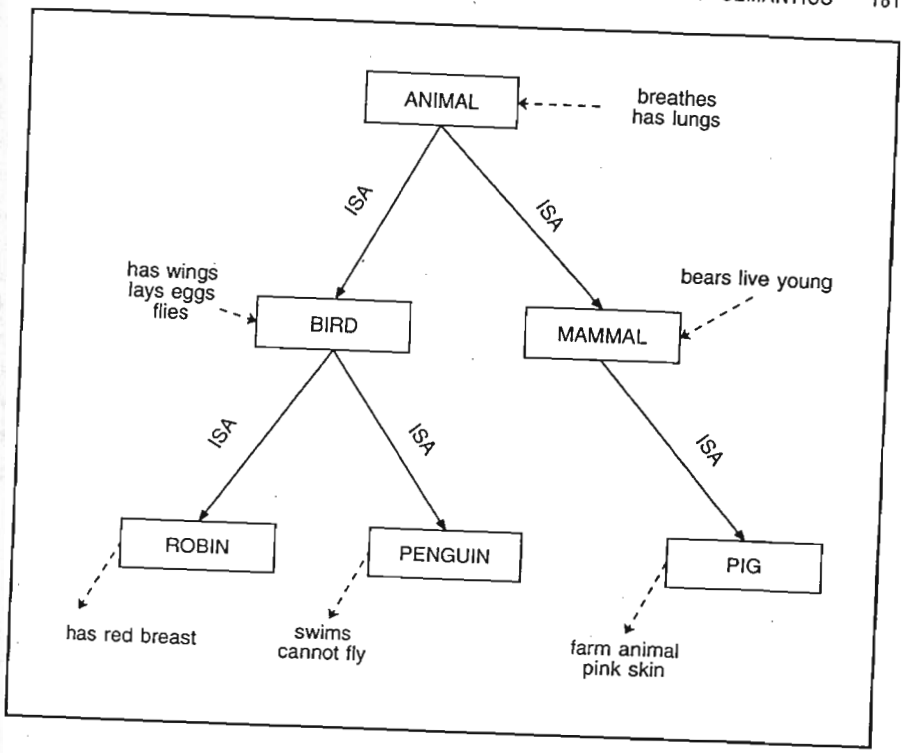


FIG. 6.1. Example of a hierarchical semantic network. (Based on Collins & Quillian, 1969, p. 241.)

Why do we get these results? According to this model, subjects produce responses by starting off from the node in the network that is the subject in the sentence (here "robin"), and travelling through the network until they find the information necessary. As this travelling takes a fixed amount of time for each link, the further away the information is, the slower the response time. To get from "robin" to "bird" involves travelling along only one link, but to get from "robin" to "animal" necessitates travelling along two links. That is, the semantic distance between "robin" and "animal" is greater than that between "robin" and "bird". If the information is not found the "no" response is made. The characteristic of property inheritance also shows the same pattern of response times, as we have to travel along links to retrieve the property from the appropriate level. Hence reaction times are fastest to (5), as the "red-breasted" attribute is stored at the "robin" node, slower to (6), as "has wings" is stored at the "bird" level above "robin", and slowest to (7), as this information is stored two levels above "robin" at the "animal" level.

5. A robin has a red breast.
6. A robin has wings.
7. A robin has lungs.

These data from early sentence verification experiments therefore supported the Collins and Quillian model.

*Problems with the Collins and Quillian model.* A number of problems for this model soon emerged. First, clearly not all information is easily represented in hierarchical form. What is the relationship between "truth", "justice" and "law", for example? A second problem is that the materials in the sentence verification task that appear to support the hierarchical model confound semantic distance with what is called *conjoint frequency*. This is exemplified by the example of the words "bird" and "robin"; these words appear together in the language—they are used in the same sentence for example—far more than do "bird" and "animal". Conjoint frequency is a measure of how frequently two words co-occur. When you control for conjoint frequency, the linear relationship between semantic distance and time disappears (Conrad, 1972; Wilkins, 1971); in particular hierarchical effects can no longer be found for verifying statements about attributes ("a canary has lungs") although they persist for class inclusion ("a canary is an animal"). This suggests that the original sentence verification results were found simply because the sentences that give the faster verification times contain words which are more closely associated. Another possible confound in the original sentence verification experiments is with category size; the class of "animals" is by definition bigger than the class of "birds", so perhaps this is why it takes longer to search (Landauer & Freedman, 1968). However, they did not properly control for typicality and semantic distance (see Rips, Shoben, & Smith, 1973; Smith, Shoben, & Rips, 1974).

Third, the hierarchical model makes some incorrect predictions. We find that a sentence such as (8) is verified much faster than (9), even though animal is higher in the hierarchy than mammal (Rips et al., 1973).

8. a cow is an animal
9. a cow is a mammal

We do not reject all untrue statements equally slowly. Sentence (10) is rejected faster than (11), even though both are equally untrue (Schaeffer & Wallace, 1969, 1970; Wilkins, 1971). This is called the *relatedness effect*: the more related two things are, the harder it is to disentangle them, even if they are not ultimately from the same class.



10. a pine is a chair
11. a pine is a flower

Neither are all true statements involving the same semantic distance responded to equally quickly. Sentence (12) is verified faster than (13), even though both involve only one semantic link (Rips et al., 1973), and a "robin" is judged to be a more typical bird than a "penguin" or an "ostrich" (Rosch, 1973). This is called the *prototypicality effect*.

12. a robin is a bird
13. a penguin is a bird

In summary there are too many problematical findings from sentence verification experiments to accept the hierarchical network model in its original form. We shall see that of these troublesome findings, the prototypicality effect is particularly important.

*Revisions to the semantic network model.* Collins and Loftus (1975) proposed a revision of the model based upon the idea of spreading activation. The structure of the network became more complex, with the links between nodes varying in strength or distance (see Fig. 6.2). Hence "penguin" is more distant from "bird" than is "robin". The structure is no longer primarily hierarchical, although hierarchical relationships still form parts of the network. Access and priming in the network occur through a mechanism of spreading activation. The concepts of activation travelling along links of different strengths and many simple units connected together in complex ways are of course important concepts in connectionist models.

## SEMANTIC FEATURES

A different approach to semantic memory views the meaning of a word as encoded not by the position of a word in a network of meaning, but by its decomposition into smaller units of meaning called *semantic features*. This works very well for some simple domains where there is a clear relationship between the terms: one such domain much studied by anthropologists is that of kinship terms. A simplified example is shown in Table 6.1. Here the meanings of the four words "mother", "father", "son" and "daughter" can be captured by combinations of the three features "human", "male" or "female", and "older" or "younger". We *could* provide hierarchical arrangement of these features (e.g. human → young and old; young → male or female, and old → male or female) but this would either be totally unprincipled (there is no reason

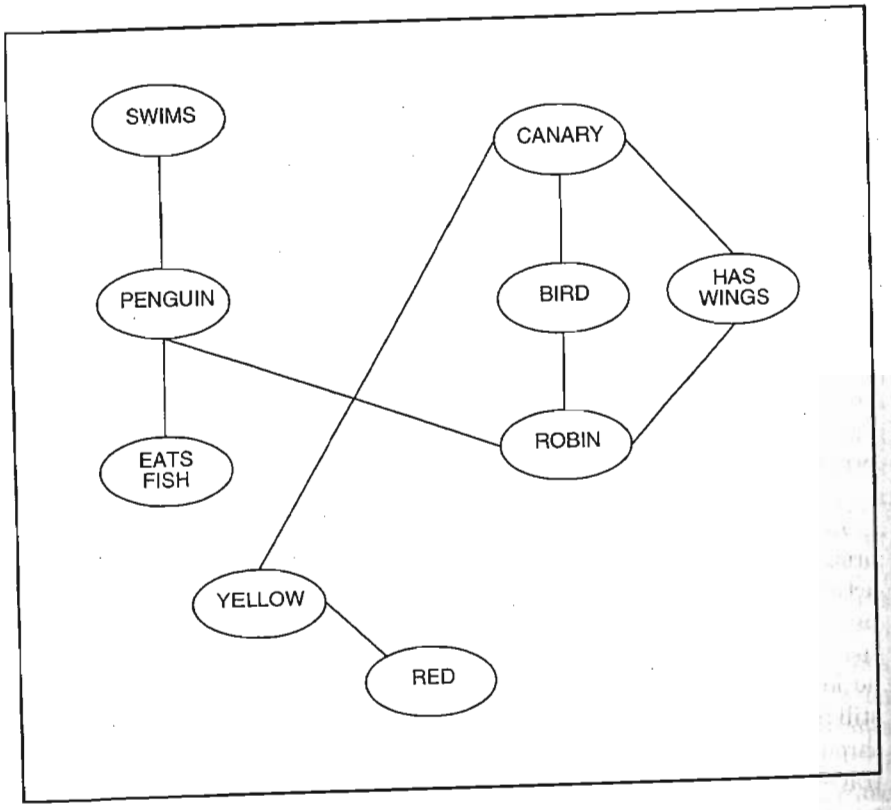


FIG. 6.2. Example of a spreading activation semantic network. (Based on Collins & Loftus, 1975, p. 412.) It should be noted that two dimensions cannot do justice to the necessary complexity of the network.

why adult/young should come before male/female, or vice versa) or involve duplication (if we store both hierarchical forms). Instead, we can list the meaning in terms of a list of features, so that father is (+ human, + male, + older).

**TABLE 6.1**
**Decomposition of kinship terms into semantic features**

| *FEATURE* | *Father* | *Mother* | *Daughter* | *Son* |
|---|---|---|---|---|
| Human | + | + | + | + |
| Older | + | + | – | – |
| Female | – | + | + | – |



We can take the idea of semantic features to the limit by trying to represent the meanings of all words in terms of combinations of as few semantic features as possible. When we use features in this way it is as though they become atoms of meaning, and are called *semantic primitives*. This approach has been particularly influential in AI. For example, Schank (1972, 1975) argued that the meaning of sentences could be represented by the *conceptual dependencies* between the semantic primitives underlying the words in the sentence. All common verbs can be analysed in terms of 12 primitive actions which concern the movement of objects, ideas, and abstract relations. For example, there are five physical actions (called "expel", "grasp", "ingest", "move", and "propel"), and two abstract ("attend", and "speak"). Their names are fairly self-explanatory, but it is not necessary to go into detail of their meanings here. Wilks (1976) described a semantic system where the meaning of 600 words in the simulation can be reduced to combinations of only 80 primitives. In this system the action sense of "beat" is denoted by ("strike" [subject—human] [object—animate] [instrument—thing]). The semantic representation and syntactic roles in which the word can partake are intimately linked.

### Early decompositional theories: Katz and Fodor

Because this approach is based on decomposing the meaning of a word into simpler constituents, such theories are sometimes said to be *decompositional* theories of meaning. One of the earliest of these was that of Katz and Fodor (1963). The emphasis of this theory was to show the meanings of sentences could be derived by combining the semantic features for each individual word in the sentence, with particular emphasis on showing how we understand ambiguous words in different contexts. Consider examples (14) and (15). A different sense of "ball" is used in each sentence. Consider further (16), which is semantically anomalous.

14. The witches put on their football strips and kicked the ball.
15. The witches put on their party frocks and went to the ball.
16. ? The rock kicked the ball.

There are no syntactic cues to be made use of here, so how do the meanings of the words in the sentence combine to resolve the ambiguity in (14) and (15) and identify the anomaly in (16)? First, Katz and Fodor postulated a decompositional theory of meaning so that the meanings of individual words in the sentence are broken down into their component semantic features (called *semantic markers* by Katz and Fodor). Second, the combination of features across words is governed by

particular constraints called *selection restrictions*. There is a selection restriction on the meaning of "kick" such that it must take an animate subject and an optional object, but if there is an object then it must be a physical object. An ambiguous word such as "ball" has two sets of semantic feature, one of which will be specified as something like (sphere, small, used in games, physical object ...), the other more like (dance, event ...). Only one of these contains the "physical object" feature, so kick picks out that sense. Similarly there is a selection restriction on the verb "went" such that it picks out locations and events, which is contradictory with the "physical object" sense of "ball". Finally, we have already mentioned that "kick" has a selection restriction specifying an animate subject, and the underlying semantic features of "rock" are incompatible with this. As there are no other possible subjects in this sentence we consider it anomalous.

### Feature list theories and semantic verification: Feature comparison

We have seen that decompositional theories of meaning enable us to list the meanings of words as lists of semantic features. What account does such a model give of performance on the sentence verification task, and in particular what account does it give of the problem to which hierarchical network models fall prey? Rips et al. (1973) proposed that there are two types of semantic feature. *Defining features* are essential to the underlying meaning of a word and relate to properties that things must have to be a member of that category (for example, a bird is living, it is feathered, lays eggs, and so forth), whereas *characteristic features* are usually true of instances of a category but are not necessarily true (for example, most birds can fly but penguins and ostriches cannot).

*Sentence verification in feature list theories.* According to the characteristic-defining feature list theory, sentence verification involves making comparisons of the feature lists representing the meaning of the words involved in two stages. For this reason this is called the *feature comparison* theory. In the first stage, the overall featural similarity of the two words is compared, including both the defining and characteristic features. If there is very high overlap, we respond "true"; if there is very low overlap, we respond "false". If we compare "robin" and "bird", there is much overlap and no conflict in the complete list of features, so we can respond "true" very easily; with "robin" and "pig" there is very little overlap and a great deal of conflict, so we can respond "false" very quickly. However, if the amount of overlap is neither very high or low, we then have to go on to a second stage of comparison (which obviously takes additional time) where we consider only the defining features. An exact match on these is then necessary to respond "true".



For example, when we compare "penguin" and "bird", there is a moderate amount of overlap and some conflict (on flying, for example). An examination of the defining features of "penguin" then reveals that it is, after all, a type of bird.

### Evaluation of decompositional theories

The appeal of decompositional theories is mixed. This is a difficult area in which to carry out experiments, and we must consider other lines of evidence. Indeed, Hollan (1975) argued that it is impossible to devise an experiment to distinguish between feature list and semantic network theories because they are formally equivalent in that it is impossible to find a prediction that will distinguish between them (but see Rips, Smith, & Shoben, 1975, for a reply).

On the one hand, such theories have an intuitive appeal, and they make explicit how we make inferences based upon the meaning of words in sentence verification task. In reducing meaning to a small number of primitives, it is very economical. On the other hand, it is difficult to construct decompositional representations for even some of the most common words. Some categories do not have any obvious defining features that are common to all their members. The most famous example of this was provided by Wittgenstein (1953), who asked what all games have in common, and therefore how "game" should be defined—that is, how it should be decomposed into its semantic primitives. There is no clear complete definition; instead, it is as though there are many different "games" which have in common a *family resemblance*. So if we cannot define an apparently simple concept such as this, how are we going to cope with more complex examples? A glance at the examples we mentioned above should reveal another problem: even when we can apparently define words, the features which we come up with are not particularly appealing or intuitively obvious; one suspects that an alternative set could be generated with equal facility. It is not even clear that our definitions are complete: often it is as though we have to anticipate all possible important aspects of meaning in advance. Bolinger (1965) criticised Katz and Fodor's theory with the example of (17).

17. He became a bachelor.

The word "bachelor" is ambiguous between the senses of "unmarried man who has never been married" and "a person with a degree". Why do we select the second interpretation in the case of (17)? You might say that it is because we know that you cannot become an unmarried man who has never been married. So does that mean that "impossible to

become" is part of the underlying meaning of this sense of bachelor—that is, that this is one of its semantic features? This seems very implausible. Generally, interpretation of word meaning is very sensitive to world knowledge. Is it part of the meaning of "pig" that it does not have a trunk? This also seems most unlikely. If we consider that such problems are solved by making an inference rather than simple access to semantic memory, then the problem becomes more much more complex.

The feature-comparison theory has additional problems. First, it is very specific to the sentence verification task. Second, there are some methodological problems with the Smith et al. (1974) experiments. Semantic relatedness and stimulus familiarity were confounded in the original experimental materials (McCloskey, 1980). Moreover, Loftus (1973) showed that if you reverse the order of the nouns in the sentences used in sentence verification, you find effects not predicted by the theory. If we only compare lists of features for the instance and class nouns, their order should not matter. Hence (18) "Is a robin a bird?" should be verified in the same time as (19) "Is a bird a robin?".

18. Is a robin a bird?
19. Is a bird a robin?

She found that order is important and the verification times were a function of how often the category was given for a particular instance for sentences such as (18), but were a function of how often the instance was given for the category in sentences such as (19). However, the task involved in verifying sentences such as (19) seems unnatural compared with that of (18). Fourth, Holyoak and Glass (1975) showed that people may have specific strategies for disconfirming sentences, such as thinking of a specific counter-example, rather than carrying out extensive computation. Finally, and most tellingly, it is not easy to distinguish empirically between defining and characteristic features. Hampton (1979) showed that in practice defining features do not always define category membership. The model still cannot easily account for the finding that some categories have unclear or *fuzzy* boundaries. McCloskey and Glucksberg (1978) showed that although subjects agree on many items as members of categories, they also disagree on many. For example, although all subjects agree that "cancer" is a disease and "happiness" is not, half think that "stroke" is disease and about half think that it is not. Similarly, about half the subjects think that "pumpkin" is a type of fruit and half do not. Labov (1973) showed that there is no clear boundary between membership and non-membership of a category for a simple physical object like a "cup": "cup" and "bowl"



vary along a continuum, and different subjects put the cut-off point in different places. Furthermore, this point can be altered by asking subjects to focus on different aspects of the object; if they are asked to imagine an object which is otherwise half way between a cup and a bowl as containing mashed potato, subjects are more likely to think of it as a bowl.

*Is semantic decomposition obligatory?* Another important issue for feature theories is whether the decomposition of a word into its component semantic features is obligatory. That is, when we see a word like "bachelor", is the retrieval of its features an automatic process? In featural terms, the meaning of the unmarried man sense of "bachelor" must clearly contain features that correspond to (+ unmarried, + man), although these in turn might summarise decomposition into more primitive features, or there might be others (see above). In any case, on the decompositional account, when you see or hear or think the word "bachelor", you automatically have to decompose it. Therefore you will draw automatically all the valid inferences that are implied by its featural representation—for example, the feature (+unmarried) automatically becomes available in all circumstances.

Automatic decomposition is a very difficult theory to test experimentally. However, Fodor, Fodor and Garrett (1975) observed that some words have a negative implicit in their definition. They called these *partial definitional negatives* (PDNs for short). For example, the word "bachelor" has such an implicit negative in (+unmarried), which is equivalent to (not married). It is well-known that double negatives, two negatives together, are harder to process than one alone. Fodor et al. compared sentences (20), (21), and (22).

20. The bachelor married Sybil.
21. The bachelor did not marry Sybil.
22. The widow did not marry Sybil.

According to decompositional theories, (20) contains an implicit negative in the form of the PDN in "bachelor". If this is correct, and such features are accessed automatically, then (21) is implicitly a double negative and should be harder to understand than a control sentence such as (22), which contains only an explicit negative and no PDN. Fodor, Fodor, and Garrett could find no processing difference between sentences of the types of (21) and (22). They concluded that features are not accessed automatically, and instead proposed a non-decompositional account in which the meaning of words is represented as a whole. To draw an inference such as "a bachelor is unmarried", you have to make

a special type of inference called a *meaning postulate*. We do this only when required. However, it is difficult to make up good controls (for example, sentences matched for length and syntactic complexity) for this type of experiment (see Katz, 1977).

Fodor, Garrett, Walker, and Parkes (1980) examined the representation of *lexical causatives*. These are verbs that bring about or cause new states of affairs. In a decompositional analysis such verbs would contain this in their semantic representation. For example, "kill" would be represented as something like (cause to die). In Fig. 6.3, (a) shows the surface structure for the two sentences with the apparently similar verbs "kiss" and "kill". For the control verb "kiss", the deep structure analysis is the same, but if "kill" is indeed decomposed into (cause to die), its deep structure should be like that of (b). Fodor et al. asked subjects to rate the perceived relatedness between words in these sentences. In (b), "Vlad" and "Agnes" are further apart than they are in the deep structure of "kissed", as there are more intervening nodes. Therefore "Vlad" and "Agnes" should be rated as less related in the sentence with the causative verb "Vlad killed Agnes" than with a non-causative verb as in "Vlad kissed Agnes". However, they found no difference in the perceived relatedness ratings in these sentences, and therefore no evidence that subjects decompose lexical causatives.

There are a number of assumptions in this analysis. The finding was questioned by Gergely and Bever (1986). In particular, they questioned whether perceived relatedness between words truly is a function of their structural distance. They provided experimental evidence to support their contention, concluding that the technique of intuitions about the relatedness of words cannot be used to test the relative underlying complexity of semantic representations. The conclusion also depends on a failure to show a difference rather than on obtaining a difference, which is always less satisfactory.

Whereas some studies have concluded that complex sentences which are hypothesised to contain more semantic primitives are no less memorable or harder to process than simpler sentences which presumably contain fewer primitives (Carpenter & Just, 1977; Kintsch, 1974), there are counter-arguments that these experiments confounded the number of primitives with other factors (Gentner, 1981). Sentences that contain primitives that facilitate interconnections between elements of the other sentence are recalled better than those which do not. For example, "give" decomposes into the notion of transferring the ownership of objects between participants in the sentence, while "sold" decomposes into the notion of transferring the ownership of objects plus an exchange of money between participants. Hence sentences of the type "Vlad sold the wand to Agnes" are remembered more accurately than



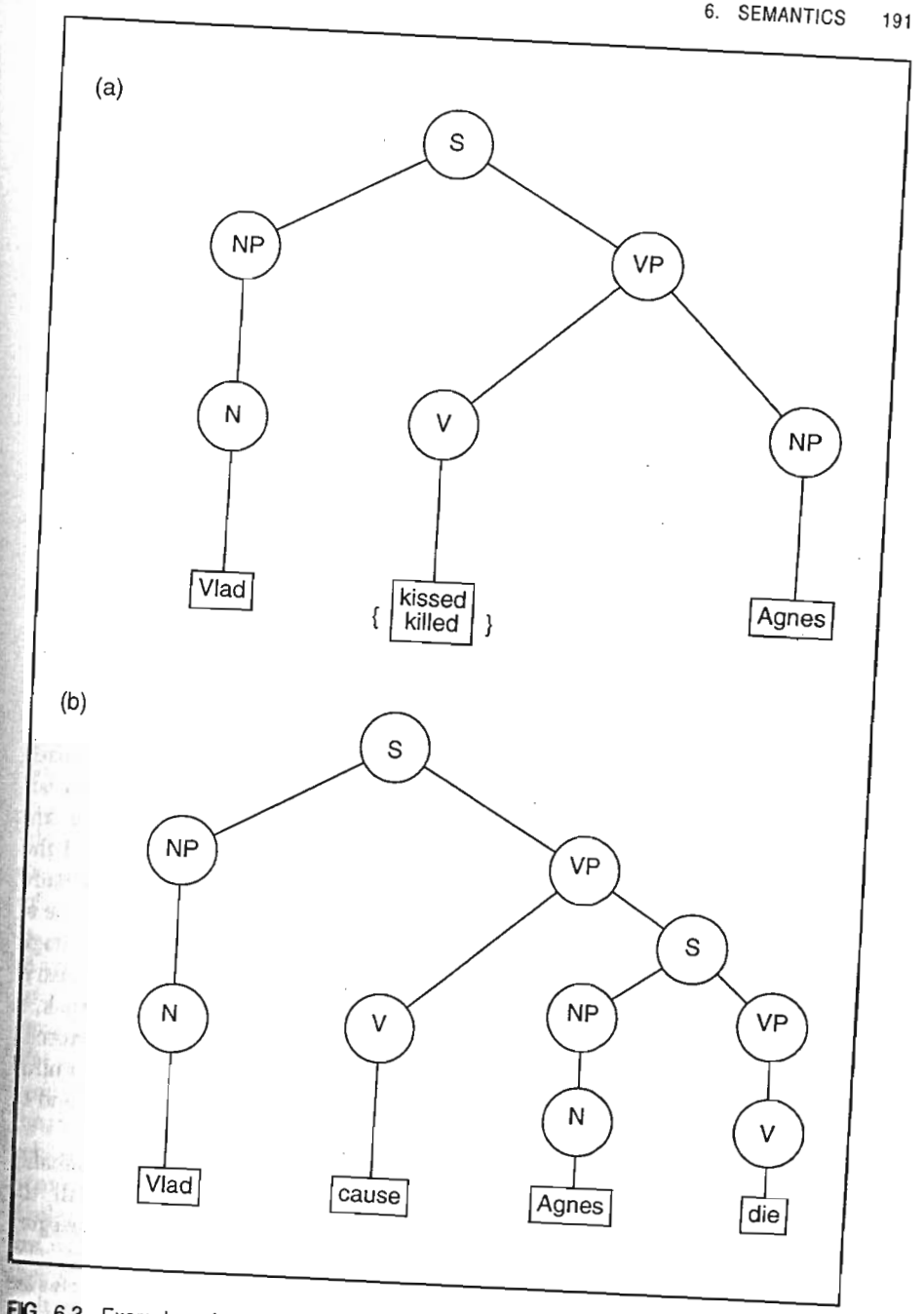


FIG. 6.3. Examples of analysis of semantics of causative verbs showing different deep structure distances. (Based on Fodor, Garrett, Walker, & Parkes, 1980, pp. 289–292.)

sentences of the type "Vlad gave the wand to Agnes" (Gentner, 1981; see also Coleman & Kay, 1981). However, memory tasks do not always provide an accurate reflection of what is actually happening at the time of processing.

This leaves us in a most unclear position regarding whether or not we do in fact decompose words into semantic features. The experiments so far have been concerned with showing that we do not decompose rather than showing that we do, and claims on the basis of "no significant difference" are never wholly satisfactory. On the other hand, automatic decomposition is fiendishly difficult to test. The status of definitional theories of meaning has proved very controversial. (See, for example, the debate between J.A. Fodor, 1978, 1979, and Johnson-Laird; Miller & Johnson-Laird, 1976).

*The probabilistic feature model.* Even though the status of decompositional theories in general is unclear, it is clear that the feature-comparison theory is untenable in its original form. Smith and Medin (1981) extended and modernised the feature theory in what is called the *probabilistic feature model*. In this approach there is an important distinction between the *core descriptions* and *identification procedures* of a concept. The core description includes the essential defining features of the concepts and captures the relationships between concepts, while identification procedures concern those aspects of meaning, as the name implies, that are related to identifying instances of the concept. For physical objects, perceptual features form an important part of the identification procedure. Semantic features are weighted according to a combination of how salient they are and the probability of their being true of a category. For example, the feature "has four limbs" has a large weighting because it is salient and true of all mammals; "bears live young" has a lower weighting because although true of almost all it is less salient; "eats meat" is even lower because it is not even true of most mammals. In a sentence verification task, a candidate instance is accepted as an instance of the category if it exceeds some critical weighted sum of features. For example, "a robin is a bird" is accepted quickly because the features of "robin" that correspond to "bird" easily exceed "bird's" threshold.

The revised model has the advantage of stressing the relationship between meaning and identification, and can account for all the verification time data. Because identifying an exemplar of a category only involves passing a threshold rather than examining the possession of defining features, classes that have "fuzzy" or unclear boundaries are no longer problematic. At this point it becomes difficult to distinguish empirically between this model and the prototype model to be discussed next. The probabilistic feature model is still developing and it is at present premature to evaluate its success.



## FAMILY RESEMBLANCE AND PROTOTYPES

We have seen that one of the problems of the decompositional theory of semantics is that for many words it is surprisingly difficult to come up with an intuitively appealing list of semantic features. Furthermore, many categories seem to be defined by a *family resemblance* between their members rather than the specification of defining features which all members must possess. On this view, a *prototype* is an average family member (Rosch, 1978). Potential members of the category are identified by how closely they resemble the prototype or category average. Some instances of a category are judged better exemplars than others. The prototype is the "best example" of a concept, and is often a non-existent, composite example. For example, a robin is very close to being a prototypical bird; it is of average size, has wings, feathers, can fly, and has average features in every respect. A penguin is a long way from being a prototypical bird, and hence we take longer to verify that it is indeed a member of the bird category. The idea of a prototype has arisen from many different areas of psychology. Posner and Keel (1968) showed subjects abstract patterns of dots. Unknown to the subjects, the patterns were distortions of just one underlying pattern of dots which the subjects did not see in the first instance. The underlying pattern of dots corresponds to the category prototype. Even though subjects never saw this pattern, they later treated it as the best example, responding to it better than the patterns they did see. We shall discuss the related work of Rosch on prototypes and colour naming in Chapter 10.

More formally, we can say that a prototype is a special type of *schema*. A schema is a frame for organising knowledge which can be structured as a series of slots plus fillers (see next chapter for more detail). A prototype is a schema with all the slots filled in with average values. For example, the schema for "bird" comprises a series of slots such as "can fly?" ("yes" for robin and wren, "no" for penguin and emu), "bill length" ("short" for robin, "long" for curlew), and "length" ("short" for robin, "long" for stork). The bird prototype will have the most common or average values for all these slots (can fly, short bill, short legs). Hence a robin will be closer to the prototype than an emu. Category boundaries are unclear or "fuzzy". For some items, it is not clear which category they should belong in; and in some extreme cases, some instances may be in two categories (for example, a tomato may be categorised as both a vegetable and a fruit).

There is a wealth of evidence supporting prototype theory. Rosch and Mervis (1975) measured family resemblance among instances of concepts such as fruit, furniture, and vehicles by asking subjects to list their features. Although some features were given by all subjects for particular concepts, these were not technically defining features, as they did not distinguish the concept from other concepts. For example, all subjects might say of "birds" that "they're alive", but then so are all other animals. The more specific features that were listed were not shared by all instances of a concept—for example, not all birds fly.

A number of results demonstrate the processing advantage of a prototype over particular instances (see for example Mervis, Catlin, & Rosch, 1975). Sentence verification time is faster for prototypical members of a category. Prototypical members can swap for category names in sentences whereas non-prototypical members cannot. Words for typical objects are learnt before words for atypical ones. In a free recall task, adults retrieve typical members before atypical ones (Kail & Nippold, 1984). Prototypes share more features with other instances of the category but minimise the featural overlap with related categories (Rosch & Mervis, 1975). Hence "apple" is a prototypical "fruit" (Battig & Montague, 1969), and is similar to other fruit and dissimilar to "vegetables", but "tomato" is a peripheral member and indeed overlaps with "vegetable". There are prototypes which possess an advantage over other members of the category even when they are all formally identical. Subjects consider the number "13" to be a better "odd number" than "23" or "501" (Armstrong, Gleitman, & Gleitman, 1983), and "mother" is a better example of "female" than "waitress". These typicality effects can also be found in sentence verification times.

*Basic levels.* In the introduction to this chapter we saw that there is a basic level of categorisation which is particularly psychologically salient. The basic level is the level that has the most distinctive attributes and provides the most economical arrangement of semantic memory. There is a large gain in distinctiveness from the basic level to levels above, but only a small one to levels below. For example, there seems to be a large jump from "chairs" to "furniture" and other types of furniture such as "tables", but a less obvious difference between different types of chair. Objects at the basic level are readily distinguished from each other, but objects in levels beneath the basic level are not so easily distinguished from each other. It is the level at which we think in the sense that those are the labels we choose in the absence of any particular need to do otherwise. The basic level is the most general category for which a concrete image of the whole category can be formed (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).



Rosch et al. (1976) showed that basic levels have a number of advantages over other categories: subjects can easily list most of the attributes of the basic level; it is the level of description most likely to be spontaneously used by adults; sentence verification time is faster for basic level terms; and it is typically acquired first by children. We can also name objects at the basic level faster than at the superordinate or subordinate levels (Jolicoeur, Gluck, & Kosslyn, 1984).

*Problems with the prototype model.* Although the prototype is now generally considered to be the best approach to the representation of meaning that we have, it is not without its criticisms. Hampton (1981) pointed out that not all concepts have prototype characteristics: abstract concepts in particular are difficult to fit into this scheme. What does it mean, for example, to talk about the prototype for "truth"? The prototype model does not explain why categories cohere: Lakoff (1987) points to some examples of very complex concepts for which it is far from obvious how there could be a prototype—the Australian Aboriginal language Dyirbal has a coherent category of "women, fire, and dangerous things" marked by the word "balan". It cannot explain why typicality judgements vary systematically depending on the context (Barsalou, 1985). Finally, the characterisation of the basic level as that of the most psychologically fundamental is not quite as clear cut as at first sight. The amount of information we can retrieve about subordinate levels varies with our expertise (Tanaka & Taylor, 1991). Birdwatchers, for example, know nearly as much about subordinate members such as blackbirds, jays, and dartford warblers, as they do about the basic level. Komatsu (1992) describes these problems in detail.

## NEW APPROACHES TO SEMANTICS

Recent work questions the necessity of whether *abstraction* is an essential component of conceptual representation. An alternative view is that of representing exemplars without abstraction: each concept is representing a particular, previously encountered instance. This has been called the *instance approach* (Komatsu, 1992). There are different varieties of the instance approach depending on how many instances are stored, and on the quality of these instances. (See, for example, Hintzman, 1986; Murphy & Medin, 1985; Nosofsky, 1991; Smith & Medin, 1981; and Whittlesea, 1987). The instance approach provides greater informational richness at the expense of cognitive economy.

Connectionism has not yet had as much impact on high level processes such as semantic memory as it has had on lower level

processes such as word recognition. Nevertheless there has already been some interesting work in this area. We saw in Chapter 4 how Hinton and Shallice (1991) and Plaut and Shallice (1993) incorporated the semantic representation of words into a model of the semantic route of meaning. In this model meaning is represented as a pattern of activation across a number of semantic feature units, or *sememes*, such as "hard", "soft", "maximum-size-less-foot", "made-of-metal", and "used-for-recreation". No-one is claiming that the semantic features that they use in their simulations are necessarily those which humans use, but there is some evidence for this sort of approach from recent data on word naming by Masson (1995). Indeed, the semantic features that underlie word meaning may not have any simple linguistic encoding. (If there are many features corresponding to quite low-level information they are sometimes called *micro-features*.) The semantic level is more complex than this, however, because the semantic units do not stand in isolation. First, they are grouped together so that features that are mutually excluded inhibit each other so that only one can be active at any one time. For example, an object cannot be both "hard" and "soft", or "maximum-size-less-foot" and "maximum-size-greater-two-yards" at the same time. Second, they are operated upon by another set of units called "clean-up" units that modify their activation. These allow combinations of sememes to influence each other. Although their description is couched mathematically in terms of multidimensional vector spaces, as we saw in Chapter 4, it can be thought of as a landscape with many hills and valleys. The bottom of each valley corresponds to a particular word meaning. Words that are similar in meaning will be in valleys that are close together. The initial pattern of activation produced by a word when it first activates the network might be very different from its ultimate semantic representation, but as long as you start somewhere along the sides of the right valley, you will eventually find its bottom. The valley bottoms which correspond to particular word meanings are called *attractors*, and for this reason this type of network is called an *attractor network*.

Note that this approach is not necessarily a competitor to the prototype theory; one instance of a category might cause one pattern of activation across the semantic units, another instance will cause another, similar pattern, and so on. We can talk of the prototype that defines that category as the average pattern of activation of all the instances.

At this stage it is too early to say exactly how such an approach would deal with particular semantic processing tasks, such as sentence verification and categorisation, although it is fairly straightforward to see how it might be done in terms of comparing patterns of activation.



Such comparisons are a move away from viewing all processing as occurring through the interaction of many simple units. Systems that combine connectionist networks with complex, rule-based operations are called *hybrid* models. It is possible that the future of semantic memory lies with these.

## THE NEUROPSYCHOLOGY OF SEMANTICS

Can we learn anything about the representation of word meaning by looking at how it breaks down as a result of brain damage? It is a recurring theme of this book that brain damage can selectively impair a range of language functions. Of course, just because a subject cannot name a word or object does not mean that the semantic representation of that word has been lost or damaged. We have seen in our discussions of both dyslexia and anomia that subjects can fail to access the phonology of a word while they still have access to its semantic representation. We know this for a number of reasons: they may have access to part of the phonological form, they might be able to comprehend the word in speech if they cannot read it, they might be able to produce it spontaneously, they know what to do with the objects in real life, and they can group appropriate pictures together. Nevertheless there are some instances where the semantic representation is disrupted. In Chapter 8 we shall look in detail at Howard and Orchard-Lisle's (1984) patient JCU who has a general semantic disorder such that she was unable to distinguish object names from close semantic relatives. But can neuropsychology tell us anything about the organisation of semantic memory?

### Category-specific semantic disorders

Perhaps the most intriguing and hotly debated phenomenon in this area is *category-specific disorders*. In these cases some semantic categories are disrupted, whereas other related ones are preserved. For example, Warrington and Shallice's (1984) patient JBR had a semantic category-specific impairment, such that he performed much better at naming inanimate objects than animate objects. He also had a relative comprehension deficit with living things. At first sight this suggests that semantic memory is organised in terms of categories, and that it is divided into animate and inanimate object categories. JBR's brain damage has caused the loss of the animate category. There are even more precise disorders. Hart, Berndt, and Caramazza (1985) reported a patient MD who also had specific difficulties in naming fruit and vegetables. Difficulties with a particular semantic category are not

restricted to naming pictures of its members: they arise across a range of tasks, also including picture–name matching and carrying out gestures appropriate to the object (Warrington & Shallice, 1984).

There are three possible explanations for category-specific disorders. The first is the obvious one I have just mentioned, that different types of semantic information are located at different sites in the brain, so that brain damage destroys some types and not others. On this view, information about fruit and vegetables is stored specifically in one part of the brain. If this explanation is correct then category-specific disorders are important because they reveal the structure of the categories as represented by the brain. Hence the distinction between living and non-living things would be a fundamental organising principle in semantic memory. Problems with this localisation theory are discussed in detail by Farah (1994), but in summary it is argued that it goes against everything we know about the organisation of the brain. The second possible explanation is that the categories that are disrupted share some incidental property that makes them susceptible to loss. Riddoch, Humphreys, Coltheart, and Funnell (1988) proposed that categories that tend to be lost tend to include many similar and confusable items. However, it is not clear that these patients have any perceptual disorder (Caplan, 1992). The third, which we shall examine shortly, is that members of the lost category share other characteristics other than being simply animate or inanimate.

The third possible explanation, then, is that the differences between the categories are mediated by some other variable. One particular proposal is that the underlying distinction between animate and inanimate objects might be one based on the observation that artefacts are distinguished from one another in terms of their functional properties, while items within biological categories tend to be differentiated primarily in terms of their physical properties (Warrington & McCarthy, 1987; Warrington & Shallice, 1984; and see below). Hence JBR, who generally showed a deficit for living things, also performed poorly on naming musical instruments, precious stones, and names of fabrics. What these things all have in common is that, like living things, they are recognised primarily in terms of their perceptual characteristics, rather than being distinguished from each other on largely functional terms.

There are a number of methodological problems in studying category-specific disorders. Funnell and Sheridan (1992) reported an apparent category-specific effect whereby their patient, SL, appeared to show a selective deficit in naming pictures and defining words for living versus non-living things. However, when they controlled for the familiarity of the stimulus, this effect disappeared. They went on to



make a more general observation about the materials used for these types of experiment. Most experiments use as stimulus materials a set of black-and-white line drawings from Snodgrass and Vanderwart (1980). Some examples are given in Fig. 6.4. Funnell and Sheridan (1992) showed that within this set there were more pictures of low frequency animate objects than there were low frequency inanimate objects. There were few low-familiarity non-living things and few high-familiarity living things. That is, a random sample of animate things is likely to be of overall lower frequency than a random sample of inanimate objects. Hence if frequency is important in brain-damaged naming, an artefactual effect will show up unless care is taken to control for frequency across the categories. Furthermore, there were two anomalous sub-categories: human body parts (high familiarity but a sub-category of living things) and musical instruments (low frequency but inanimate). These were the two anomalous categories mentioned by Warrington and Shallice (1984) in their description of JBR.

Stewart, Parkin, and Hunkin (1992) also argued that there had been a lack of control of name frequency, but pointed out in addition that the complexity and familiarity of the pictures used in these experiments varied between categories. In reply, Sartori, Miozzo, and Job (1993), analysing data from their patient "Michelangelo" concluded that Michelangelo had a real category-specific deficit for living things even when you control for these factors. The debate continued in Parkin and Stewart (1993), and Job, Miozzo, and Sartori (1993), particularly illustrating the need for better measures of visual featural complexity and similarity.

Many of the counter-arguments are weakened by the presence of double dissociations between the observed categories. Warrington and McCarthy (1983, 1987) describe patients who are the reverse of JBR in that they perform better on living objects than inanimate objects. It is

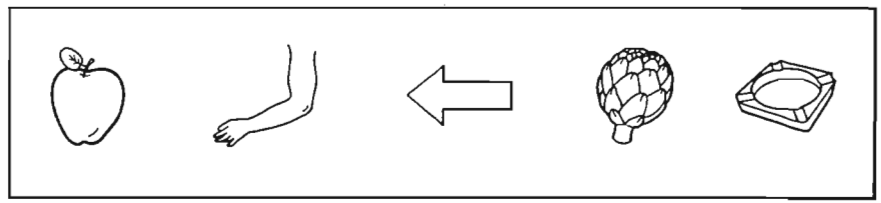

FIG. 6.4. Examples of line drawings from the Snodgrass and Vanderwart set. (Reprinted from Snodgrass, J. G., & Vanderwart, M., 1980. A standardised set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215, p. 197. © 1980 by the American Psychological Association. Reprinted by permission.)

necessary however to examine the detailed pattern of performance in these cases. Their patient YOT, for example, who generally had an impairment in naming inanimate objects relative to animate ones, on closer examination could identify large outdoor objects such as buildings and vehicles. There also appears to be a distinction between small and large man-made artefacts.

Finally, Farah, Hammond, Mehta, and Ratcliff (1989) showed that control subjects were poor at answering questions on the perceptual features of both living and non-living objects (e.g. "Are the hind legs of kangaroos larger than their front legs?"). This suggests we do not know much about the visual details of low frequency objects because, perhaps, they are never examined closely.

### Modality-specific disorders

How many semantic systems are there? So far there has been little discussion of how meanings are accessed across different modalities. We have discussed semantic information as though there is only one semantic store, regardless of how it is accessed. Paivio (1971) has instead proposed a *dual-code theory* of semantic representation, with a perceptual code for the perceptual characteristics of a concept, and a verbal code for the abstract, non-sensory aspects of a concept. Experimental tests of this hypothesis have produced mixed results (Snodgrass, 1984). For example, subjects are often faster to access abstract information from pictures than from words (see for example Banks & Flora, 1977).

The idea of modality-specific semantic stores, whereby verbal (words) and non-verbal (pictures) material are separated, has enjoyed something of a resurgence owing to data from brain-damaged subjects. There are three main reasons for this (Caplan, 1992). First, priming effects have been discovered which have been found to be limited to verbal material. Second, cases of anomia have been discovered where it is limited to one sensory modality rather than the more common case when it is found across all modalities. Third, there is a finding that patients with semantic deficits are not always equally impaired for verbal and visual material (e.g. Warrington, 1975). Warrington and Shallice's (1979) patient JBR showed a much larger benefit from cueing in reading a written word than in naming the corresponding picture. They interpreted this finding as evidence for separate verbal and visual conceptual systems. However, as Caplan (1992) points out, there is no demonstration that automatic retrieval as manifested by short onset priming is impaired.

Alternative explanations have been offered for these early studies. Riddoch et al. (1988) argued that patients who perform better on verbal



material may have a subtle impairment of complex visual processing. This idea is supported by the finding that this disturbance in processing pictures is worse for categories where there are many visually similar objects (for example, fruit and vegetables). The reverse dissociation of better performance on visual material may arise because of the abundance of indirect visual cues in pictures. For example, Caplan (1992) points out that the presence of a large gaping mouth and heavy paws in the picture of a lion is an excellent indirect cue to how to answer a comprehension question such as "Is it dangerous or not?" even if you do not know it is a picture of a lion.

Nevertheless, some more recent work is more difficult to explain away. Bub, Black, Hampson, and Kertesz (1988) describe the case of MP, who showed very poor comprehension of verbal material, and did not show automatic semantic priming, but much better comprehension of pictures. The nature of the detailed information MP was able to provide about the objects in the pictures, such as the colour of a banana from a black-and-white line drawing, could not easily be inferred from perceptual cues without access to semantic information about the object. Warrington and Shallice (1984) found high item consistency as long as modality was held constant, again suggesting different semantic systems were involved. Finally, supportive evidence comes from modality-specific anomia, in which the naming disorder is confined to one modality. For example, in *optic aphasia* (Beauvois, 1982) patients are unable to name objects when presented visually, but can if they are exposed to them through other modalities. The interpretation of these data is controversial. Riddoch and Humphreys (1987) argue that optic aphasia is a disorder of accessing a unitary semantic system through the visual system, rather than disruption to a visual modality-specific semantic system. Much hangs upon the interpretation of gestures made by the patient. Do they indeed reflect preserved visual semantics—so that patients understand the objects they see—with disruption of verbal semantics, or are they merely inferences from perceptual attributes of objects? Certainly Riddoch and Humphreys' patient JV produced only the most general of gestures to objects, and other experiments indicated a profound disturbance of comprehension of visual objects. Of course, we must remember the caveat that different patients display different behaviours, and one must be wary of drawing too general a conclusion from a single patient.

Caplan (1992) proposes a compromise position whereby only a subset of semantic information is dedicated to specific modalities. The perceptual information necessary to identify an object is certainly only a subset of the meaning of a concept, and this might well differ from the verbal semantic information elicited by that concept.

*Modality-specific and category-specific effects* Is there any relationship between these findings of modality-specific and category-specific effects? Farah and McClelland (1991) argued that there is, in that damage to a modality-specific semantic memory system can lead to category-specific deficits. The architecture of their connectionist model comprises three "pools" of units: verbal input and output units (corresponding to name units), visual input and output units (picture units), and semantic memory units (divided into visual and functional units). The network was then taught to associate the correct semantic and name pattern when presented with each picture pattern, and to produce the correct semantic and picture pattern when presented with each name pattern. They then lesioned this network. They found that damage to visual semantic units primarily impairs knowledge of living things, while damage to functional semantic units primarily impairs knowledge about non-living things. Hence it is possible that semantic category-specific anomia arises not directly because knowledge about animate and inanimate objects is stored in different parts of the brain and can therefore be selectively erased, but because knowledge of animate objects is derived primarily from visual information, whereas knowledge of inanimate objects is derived primarily from functional information.

### Access versus degradation

How can we tell when a concept is truly lost (or at least is degraded) from when there is difficulty in gaining access to it? Shallice (1988) and Warrington and Shallice (1979), among others, discuss five criteria to distinguish problems associated with the loss of a representation from problems of accessing it. First, performance should be consistent across trials. If an item is permanently lost, it should obviously never be possible to access it. If an item is available on some trials rather than others the difficulty is one of access. Second, superordinate information should be relatively well-preserved. Warrington (1975) found that superordinate information (e.g. that a lion is an animal, when "lion" is lost) may be preserved when more specific information is lost. She proposed that the destruction of semantic memory occurs hierarchically, from the bottom (more specific levels) up. Impaired access should affect all levels equally. Third, low frequency items should be lost first. Low frequency items should be more susceptible to loss, whereas problems of access should affect all levels equally. Fourth, priming should no longer be effective. If an item is lost it cannot be primed. Fifth, if the knowledge is lost then performance should be independent of the presentation rate, while disturbances of access should be sensitive to the rate of presentation of the material.



There has been considerable debate about how reliably these criteria do in fact distinguish access disorders from loss. It has been claimed that there are ways of presenting difficulties in access such that the same pattern of results could be found as if items were lost. See Shallice (1988) and Caplan (1992) for more detail, but in general to infer loss of items one needs to see at least consistent failure to access items across tasks. More recent work on this dichotomy is becoming mathematically sophisticated (see Faglioni & Botti, 1993).

### Dementia

Dementia is a generalised decay of cognitive functioning generally found in old age. Although its exact cause is as yet unknown, memory and semantic information are particularly prone to disruption. As such many of the characteristics we have just discussed are exemplified.

Chertkow and Bub (1990) showed that patients with a type of dementia known as Alzheimer's disease demonstrated no improvement in picture naming when cued. That is, saying of a picture of a lion "It's like a tiger" did not help performance. On the other hand, Chertkow, Bub, and Seidenberg (1989) found that the same group of patients showed a clear effect of semantic priming in a lexical decision task with an SOA (stimulus onset asynchrony; see Chapter 3) of 500 milliseconds. There are two conclusions to be drawn from this study. First, for short-term priming to be found suggests that the items could not have been lost, and therefore caution must be employed if one task (such as cueing) suggests that they are. Second, Caplan (1992) points out therefore that cueing, involving attentional processing, must be distinguished from what he refers to as true priming, which involves short-term automatic processing.

Research on the breakdown of semantic memory in dementia is a fairly recent development (see Nebes, 1989, for a review). It is likely that over the next few years this will contribute greatly to our understanding of the representation of meaning.

### Evaluation of the neuropsychology of semantic memory

At first sight this research seems promising. The types of dissociations found in category-specific disorders might reveal the fundamental organising principles of semantic memory. Unfortunately caution is needed on two counts. First, the area is plagued by methodological difficulties. Second, the organising principles might not be what they first seem to be. For example, the loss of the animate category might be really attributable to difficulty in accessing perceptual features, and the

loss of the inanimate category might be attributable to difficulty in accessing features to do with function. At this stage it would be premature to draw any firm conclusion: this has become a very complex area.

## CONCLUDING SUMMARY

Semantics is the study of meaning, and is inextricably bound up with the study of concepts and categorisation. We must distinguish between the intensional and extensional aspects of a word's meaning. We have looked at three main psycholinguistic approaches to word meaning. First, the idea that the meaning of a word is to be found in the way in which that item is embedded within a network representing the meaning of everything we know. Second, the meaning of a word is decomposed into a set of semantic features. There are a number of problems with both of these approaches. Probably the best approach to meaning that is currently available is that of prototypes. Connectionist modelling has yet to make much impact in this area.

We have also considered the neuropsychology of meaning. We have seen how a word and its meaning can be dissociated. Category-specific disorders are potentially a useful source of information about the organisation of semantic memory. Their study is however fraught with methodological problems. Neither is it clear whether at present they signify problems specifically concerning a particular semantic category, or whether instead they represent problems with some more fundamental organising principle. There is also considerable current debate as to whether there is a unitary semantic system, or whether there are different semantic stores for the visual and verbal systems. The study of dementia is also important, but here we must be careful to distinguish between the loss of semantic information from loss of the ability to access that information.

In the next chapter we shall see how we combine the meanings of individual words with the syntactic representation to form a representation of the meaning of the whole sentence.

## FURTHER READING

An good overview of cognitive psychology approaches to meaning and the representation of knowledge is provided by Eysenck and Keane (1995). The classic linguistics work on semantics is Lyons (1977b,c). Johnson-Laird (1983) provides an excellent review of a number of approaches to semantics, including the relevance of the more



philosophical approaches. Perhaps its most detailed and formalised philosophical system of meaning is that of *Montague grammar* (Montague, 1974). Montague grammar shows how the meaning of sentences can be formally derived by combining model-theoretic representations of the meanings or intensions of their component words. Although the theory is too advanced to be considered here, an introduction can be found in Johnson-Laird (1983, Chapter 8).

General problems with network models are discussed by Johnson-Laird, Herrman, and Chaffin (1984). Chang (1986) reviews the experimental support for psychological models of semantic memory. Komatsu (1992) and Medin (1989) review the literature on concepts. For further discussion of connectionist approaches to semantic memory, see Quinlan (1991).

Caplan (1992) provides and extensive review of the neuropsychology of semantic memory. Other cases have been reported by Sartori and Job (1988), who describe patient "Michelangelo", and de Renzi and Lucchelli (1994), who describe patient "Felicia", both of whom have a relative deficit on naming animals, fruit, and vegetables. A special issue of the journal *Cognitive Neuropsychology* (1988, volume 5, issue 1) was devoted to the neuropsychology of this subject, particularly upon modality-specific effects. Davidoff and de Bleser (1993) review the literature on optic aphasia, and Plaut and Shallice (1993b) describe a connectionist model of it. See Shallice (1988) and Caplan (1992) for further discussion of criteria for distinguishing between the loss of information and difficulty in accessing that information.