

10 The Lexicon

D. A. CRUSE

1 Introduction

To take a simplistic view, the bare essentials for a language are, first, a set of basic units, and second, a set of rules for combining them into larger, more complex units like phrases and sentences. A list of the basic units constitutes the lexicon of the language; a specification of the combinatory rules constitutes the grammar.

The basic units must have both a form and a meaning (in the broadest sense); the entries in the lexicon must specify these, together with information necessary for the proper application of the grammatical rules. The combinatory rules will tell us not only what complex forms are allowed in the language, but also how their meanings are to be computed.

What are the units that are listed in the lexicon? The obvious answer is that they are words, and that is what we shall take them to be (although the matter is not quite so straightforward as it might at first seem). To the layperson, probably the most important thing about a word is what it means; this chapter has a similar bias, being chiefly about words and their meanings. We begin by looking at what sort of things words are, as a linguist sees them.

2 Words

It is notoriously difficult to frame a definition of a word which is satisfactory for all languages, and even for everything word-like in a particular language. We shall assume that, as in Wittgenstein's famous example of *game*, no succinct definition applicable to all cases is possible, and that the best approach is to look for features characteristic of central examples of the class.

2.1 Lexical forms, lexical units and lexemes

The word *word* is used in different senses, and it will be as well to clarify the most important of these right from the start. Suppose we are doing a crossword puzzle. It is quite possible that, say, *walk* is the correct answer to a clue, but *walked* is not: from this perspective *walk* and *walked* are different words. Now imagine someone who encounters sentence (1):

(1) I have disconfirmed the doctor's prognosis

and asks: "Is there such a word as *disconfirm* in English? Look it up in the dictionary." What does *word* mean here? Clearly not what it means in the crossword context, since the dubious sentence contained *disconfirmed*, but the question asks about *disconfirm*. Let us at this point make a terminological distinction between *word forms* and *lexemes*. We shall say that *walk* and *walked*, and *disconfirm* and *disconfirmed* are different word forms, but whereas *walk* and *disconfirm* represent different lexemes, *walk* and *walked* are different word forms belonging to the same lexeme.

What, then, is a lexeme? As a first step let us say that they are the units listed in a dictionary. A dictionary provides a list of the lexemes of a language, each indexed by one of its word forms. (Which word form a dictionary uses to indicate a lexeme is at least partly a matter of convention. For instance, in English, for verbs, we use the bare stem: *run*, *walk*; in French it is the infinitive: *courir*, *marcher*; in Latin, the first person singular of the present indicative: *curro*, *ambulo*.)

A more technical characterization is that a lexeme is a set of related meanings associated with a set of related word forms. Sometimes meanings associated with a single word form are clearly unrelated, as in the case of *bank* (financial) and *bank* (river); these would therefore be assigned to different lexemes. In other cases a relationship can easily be intuited, as with *position* (location), *position* (opinion), and *position* (job), and these will be considered to belong to the same lexeme. Most dictionaries give separate main entries to distinct lexemes, even if they share the same forms, but group related meanings under a single main entry. What we shall call "a set of related word forms" is a set of forms which differ only in respect of inflectional affixes (such as the singular and plural forms of nouns, or the past, present, and future forms of verbs).

This is fine, but how do we then designate the three individual items *position*? It is usual to call the distinct meanings *senses*, but what is the sound-meaning complex? I shall call them *lexical units*. Actually, in many (perhaps most) contexts it is perfectly clear what *word* means: the expressions *word form*, *lexeme*, and *lexical unit* will therefore only be used when there is a danger of confusion.

2.2 *Individuating word-forms: graphic and phonetic clues*

Most modern writing systems (English is no exception) indicate word (here, obviously, “word form”) boundaries by means of spaces. This makes reading a lot easier. However, there is usually no analog of written spaces in spoken language, although this usually comes as a surprise to the layperson, because spoken words are clearly demarcated perceptually. There may, nonetheless, be signals of other types which indicate the positions of the boundaries of spoken words. For instance, many languages have a regular stress pattern for words, as in Czech, where words are always stressed on the first syllable. Other signs may be more complex or subtle. For instance, to take a venerable example, English speakers can discern purely from the sound the different positions of the word boundaries in *night rate*, and *Nye trait*, and between *parks treat* and *Park Street*, at least when they are carefully pronounced, even though there is no silence between the words. This is because, for instance, the variety of /r/ which occurs at the beginning of a word is different from that which appears when it is preceded by /t/, and is different again if the /t/ is preceded by /s/.

2.3 *Grammatical properties of words*

Prototypically (we shall not explore the exceptions) the stretches of a sentence that constitute word(form)s can be recognized by the fact that they are the largest chunks of the sentence which cannot be interrupted by the insertion of new material. Take the sentence *The government is strongly opposed to denationalization*. The possibilities of inserting new material are as follows:

The (present) government, (apparently), is (very) strongly (and implacably) opposed (not only) to (creeping) denationalization, but . . . etc.

It will be noticed that the insertions all respect word boundaries, and all sequences of two words can be interrupted.

The parts of a word also cannot be reordered (**ationizaldenation*), although, at least in languages with some freedom of word-order, the words themselves can be rearranged (obviously to varying degrees).

2.4 *The semantic properties of words*

There are several constraints on what a word can possibly mean. First, though, a non-constraint. It might be thought that there could not be a word meaning, for instance, “to eat corn flakes while standing facing south on a Sunday morning.” However, a brief period of reflection should convince the reader

that such a meaning is not really impossible, merely unlikely in our culture: in a society where corn flakes were ritually dedicated to the god of the south, it would not be at all surprising if such an action received a lexical label. We shall look at two more serious constraints on possible word meanings, conceptual unity and internal completeness.

Whatever attracts a lexical label must have some degree of conceptual coherence. Let us confine our attention to what can be referred to by a noun: in the broadest sense, these are "things." Prototypical things are characterized by spatial continuity and persistence through time. Non-prototypical things must have something which confers unity on them. In front of me as I write, I can see, among other things, a bottle of Buxton mineral water, a photograph of Commander Data from *Star Trek*, and a ball of string. Is there any chance that these could be designated collectively by a noun? In a sense, yes: they could constitute the whole of my worldly possessions, and there could be a name for this (on the lines of my "estate," when I die). But that would not be a name for that particular set of things. Alternatively, they could be the requisites for, say, the Klingon Ceremony of Nga (or whatever). But in the absence of some such extrinsic unifying factor, the items mentioned would not be (collectively) nameable.

What I am calling "internal completeness" is more easily illustrated than explained. Take the phrase *a very large man*. The notion that there should exist a word meaning "large man" is not at all exotic (think of *giant*); nor is the idea of a word meaning "very large" (e.g. *enormous*); there could well be a word meaning "very large man" (perhaps *colossus*), too. But what about a word meaning "very . . . man," i.e. a noun such that any adjective modifying it is automatically intensified? This, surely, offends against our deepest semantic intuitions: it is an impossible meaning for a word. The same would be true of a putative "word" *beve* meaning "drink chilled . . .," such that *beve wine* would mean "drink chilled wine" (words meaning "drink wine" or "chilled wine," or even "drink chilled wine" could not be ruled out). The explanation seems to be on the following lines. We first need to distinguish *dependent* and *independent* components of a semantic combination. The independent component is the one which governs the external relations of the combination as a whole. So, for instance, in *very large*, it is *large* which governs the combinability of the phrase *very large* with other items. Thus the oddness of, say, *?very large air* is due to a clash between *large* and *air* – there is no clash between *very* and *air* (think of *very cold air*). By similar reasoning, the independent item in *chilled wine* is *wine*, and in *drink chilled wine* is *drink*. This process of reasoning allows us to establish chains of (semantic) dependencies (it does not matter whether the elements in a chain are expressed as different words, or are incorporated into the meaning of a single word). For instance, the chain for *very large man* is:

"very" → "large" → "man"

and that for *drink chilled wine* is:

“chilled” → “wine” → “drink”

The constraint that we are looking at says that the elements that constitute the meaning of a word must form a continuous dependency chain, with no gaps needing to be filled by elements from outside the word.

3 Lexical Semantics

The study of the meanings of words within linguistics is called *lexical semantics*. Under this banner a variety of spheres of interest, theoretical orientations and methods of study flourish.

3.1 Theoretical approaches

To a large extent, how one goes about the business of studying meaning depends on what picture one has of the sort of thing meaning is. Some grasp of the major options will be useful as a background to the more detailed discussions which follow. We shall concentrate on two issues, holism vs. localism, and the relation between linguistic meaning and concepts. Let us begin with the holism / localism debate. Essentially, a holist believes that the meaning of a word is fundamentally relational, that is to say, it is a matter of relations with other words in the language. A localist believes that a word’s meaning is self-contained, and describable independently of the meanings of other words.

3.1.1 The contextual / holistic approach

Within linguistics, what philosophers of language call *holistic* theories of meaning are usually called *contextual* theories. These come in several varieties: two will be briefly illustrated here.

The first type falls under the heading of *structural semantics*. The basic notion of the interdependence of meanings can be illustrated as follows. Think of a child learning the names of the animals. The fact that a child can say *It’s a dog* every time s/he is given a dog to identify, does not prove s/he has grasped what *dog* means; just as important, is that s/he should avoid saying *It’s a dog* when faced with a cat, or fox, or whatever. In other words, the meaning of *dog* (or any other word) cannot be learnt in isolation. A structuralist such as Lyons (the seminal reference is Lyons (1963)) builds on this basic insight, and characterizes the meaning of a word as its position in a network of relationships. Let us consider what that would mean in the case of *dog*. First, *dog* belongs to a set of words with which it has an exclusion relationship, that is to say, *It’s a dog* implies *It’s not a cat / lion / camel / cow / squirrel / etc.*; furthermore, all these fall into the denotation of a more inclusive term *animal*. *Animal* (at least

on one reading) also belongs to a set whose members are mutually exclusive (including *insect*, *fish*, *bird*, etc.); these in turn are included in *living thing*, and so on. But *dog* has many other relations, for instance, with *tail*, *paw*, *head*; with *pack*; with *bark*, *howl*; with *kennel*, which itself has relations with other structures such as *hut*, *cabin*, *house*, and so on. Ultimately, every word is linked directly or indirectly, by means of specific links such as “is a,” “is not a,” “has a,” “is part of,” “lives in a,” etc., with virtually every other word in the lexicon: on the holist view, the meaning of a word is not fully comprehended until all these links are known (although, obviously, some links are more central than others).

An alternative version of a contextual theory takes its origin from Wittgenstein’s dictum: *Don’t ask for the meaning, ask for the use*. This is suggestive, but lacking in precision as a basis for a theory of meaning: what, precisely, do we mean by *use*? J. R. Firth (quoted in Mackin (1978)) gave the notion a useful twist when he said: *Words shall be known by the company they keep*. This line of thinking was developed into a holistic theory of meaning by W. Haas. (Haas’s ideas are not readily accessible in published form; a summary can be found in Cruse (1986: ch. 1).) Haas started out from the idea that every grammatically well-formed sequence of words was either fully normal semantically, like *The dog barked*, or to some degree abnormal, like *?The cat barked* or *?The dog evaporated*. He then argued that if two words differ in meaning, this fact will inevitably be reflected in a difference of normality in some context or other. For instance, that there is a difference in meaning between *illness* and *disease* follows from the fact that *during my illness* is more normal than *?during my disease*. Haas went on to characterize the meaning of a word as its normality profile across all its grammatically well-formed contexts, actual or potential: absolute synonyms, on this view, are words which have the same normality in all contexts.

3.1.2 *The componential / localistic approach*

A localist believes that the meaning of a word is a self-sufficient entity which in principle is finitely describable. Whereas holists tend to see the meaning of a word as a set of relations, either with other words, or with possible contexts, a localist will typically say that these relations are a consequence of the word’s meaning.

The most popular varieties of localism portray the meaning of a word as a finite assemblage of elementary bits of meaning, each of which accounts for some aspect of the semantic behavior of the whole. These “semantic atoms” (variously known as *semantic components*, *semantic features*, *semantic markers*) are drawn from a finite inventory, and in the strongest versions of the theory are psychologically real (in the sense that if we knew enough about the brain we would be able to identify a distinctive neuronal structure corresponding to each feature), and they are universal (in the sense that they form part of the language capacity that each human being is born with). It is impossible to give

a satisfactory picture of any of the existing systems in a short space, but the following examples will give the flavor of such analyses:

- filly = [HORSE] [FEMALE] [YOUNG]
- boy = [HUMAN] [MALE] [YOUNG]
- kill = [CAUSE] [BECOME] [NOT] [ALIVE]
- chair = [OBJECT] [FURNITURE] [FOR SITTING]
[FOR ONE PERSON] [WITH BACK]

3.1.3 *The conceptual approach*

Much debate centers on the relation, if any, (but surely there must be some) between linguistic meaning and concepts, or, as far as we are concerned in this chapter, between word meanings and concepts. Earlier semanticists (including Lyons and Haas) did not believe that anything solid was known about concepts; they therefore preferred to pursue their semantic studies without reference to such entities. The rise of cognitive psychology has made concepts more respectable, and few would now deny their significance. The debate now concerns whether, or to what extent, meaning can be identified with concepts: do words map directly onto concepts, or is there an intermediate level of semantic structure where word meaning is to be located, and the connection with concepts indirect? The present author's sympathies lie with the conceptual approach. A conceptual (or "cognitive") semanticist would argue that there is no theoretical work for an autonomous linguistic semantic level to do that cannot be performed at the conceptual level. He would also argue that the connection between words and the outside world is mediated through concepts, and that therefore examining world-word relations is not the most profitable approach to word meaning.

4 **How Many Meanings? Contextual Variability of Word Meaning**

Most words are capable of presenting different semantic faces in different contexts. Sometimes the differences are major and clear cut, as in:

- (2) The boat was moored to the bank.
- (3) She works in a bank.

At other times the difference is more subtle:

- (4) John's maths teacher is on maternity leave.
- (5) Bill's maths teacher is on paternity leave.

Here we can infer from the context that John's maths teacher is a woman whereas Bill's is a man.

It is important to be able to decide whether two interpretations of a word in different contexts represent one semantic unit or two. This is not a purely theoretical concern: for instance, a lexicographer will have to decide how many definitions to give for *bank* and *teacher*. We shall take the position that the basic unit of word meaning is the *sense*, and we shall say that a word has X senses if and only if it is X-ways ambiguous. We now need to be more explicit about what it means for a word to be ambiguous.

4.1 Ambiguity

Consider sentence (6):

(6) We managed to get to the bank just in time.

In the absence of a biasing context, the two readings of *bank* are in competition with one another: like the two visual construals of a Necker cube, only one can be at the focus of attention at any given moment. In a particular context, a speaker will "intend" only one of the meanings and will expect the hearer to make the same selection. There is no general meaning of *bank* which subsumes the two alternatives, and the options of remaining uncommitted or of taking both meanings on board are not available (outside of deliberate word play). Contrast this with the following case:

(7) We shall talk to Mary's teachers.

Of course, the individual teachers referred to in (7) must be either male or female, but (a) the speaker may not even know the sex of the teachers involved and will not expect the hearer to select a particular gender; (b) there is a general meaning of *teacher* which covers both possibilities; (c) the sex of the teachers can be left unspecified; furthermore, sentence (7) may well refer to a mixed group. By the criteria suggested, then, *teacher* is not ambiguous, and does not have two senses corresponding to "male teacher" and "female teacher"; a lexicographer would not need to give two definitions for *teacher*.

Ambiguous words typically pass the traditional *ambiguity tests*.

4.1.1 The identity test

In *John has changed his position; so has Mary*, the word *position* must be interpreted the same way in both halves of the sentence: if John has changed his mind on some political issue, then that's what Mary did, too; likewise if John has changed his location. This shows that *position* is ambiguous. In contrast, in

I spoke to a teacher; so did Mary, there is no pressure to interpret *teacher* in the same way (gender-wise) in each conjunct, hence *teacher* fails this test.

4.1.2 *The independent truth-conditions test*

It is easy to think of a situation where one could truthfully answer the following question both in the negative and the affirmative:

(8) Have you had a drink in the last six hours?

This shows that the readings “take alcoholic beverage” and “imbibe liquid” are distinct senses. There is no comparable possibility for simultaneously true *Yes* and *No* answers to (9):

(9) Have you spoken to a teacher?

4.1.3 *The zeugma test*

A context which activates more than one reading of an ambiguous word gives rise to a sense of punning:

(10) The old man expired on the same day as his driving license.

(11) When the chair became vacant, the University Appointments Committee sat on it for six months.

(The effect in (11) hinges on the ambiguity of both *chair* and *sat on*.)

4.2 *Polysemy and homonymy*

The alternative readings of an ambiguous word may be totally unrelated, as in the case of *bank*, or they may be related, as in the case of *position* (see below for some discussion of possible types of relatedness). An ambiguous word with unrelated readings is described as *homonymous*; if the readings are related, the word is said to be *polysemous*. Homonymous words are usually given two main entries in a dictionary; polysemous variants are normally listed under a single main heading.

5 Sense Relations

Sense relations are relations between word meanings. Of course, every word has a semantic relation of some kind with every other word, but not all such relations have any intrinsic interest. To be interesting, a relation must recur

with significant frequency throughout the vocabulary, and must be capable of supporting significant generalizations. (A much fuller treatment of sense relations than can be accommodated here may be found in Cruse 1986.)

There are two major classes of sense relation, depending on the grammatical relation between the words bearing the senses, namely, *paradigmatic*, and *syntagmatic* relations. Paradigmatic sense relations are relations between the meanings of words which can occupy the same syntactic slot, and serve to unite the range of lexical meanings available at a particular point in a sentence into a more or less coherent structure. Take, for instance, the (incomplete) sentence: *John grows a number of ----- in his garden*. There is a structured set of choices of words to fill the gap. One may choose very general words like *trees*, *flowers*, or *vegetables*, or something more specific, falling under one of the general terms, for instance, *conifers*, *cabbages*, *carnations*. We shall look at the structuring in such a set in more detail in a moment, but it can already be appreciated that the words provide an articulation of the experienced world.

Syntagmatic sense relations hold between words in the same phrase or sentence. Intuitively, some words “go together” semantically, while others “clash”: consider *drink wine* and *drink water*, compared with *drink rock* or *drink sound*. There is a relation of cohesiveness between *drink* and *wine* which is absent from *drink* and *rock*. Syntagmatic sense relations are thus involved with the semantic coherence of grammatical strings.

5.1 Paradigmatic sense relations

It is paradigmatic relations which have received the most attention from linguists. For convenience, they may be divided into two sorts, relations of identity and inclusion, and relations of opposition and exclusion.

5.1.1 Relations of inclusion and identity I: hyponymy

We begin with relations of inclusion. There are two basic types of these. In the first type, the inclusion is of one class in another, as in the case of *car* and *vehicle*, where cars constitute a subclass included in the larger class of vehicles; in the second type, the inclusion is observable at the level of individual entities, as in the case of *finger* and *hand*, where every individual hand includes a number of fingers as parts.

The class-inclusion relation, called *hyponymy*, is exemplified by *dog:animal*, *apple:fruit*, *tulip:flower*, *cathedral:building*, *beer:beverage*, *copper:metal*, *kitten:cat*, *mare:horse*, *actress:woman*, and so on; of the two related items the more specific is called the *hyponym* (e.g. *dog*, *apple*), and the more general is called the *superordinate* (less commonly, the *hyperonym*), e.g. *animal*, *fruit*. Notice that although *dog* is a hyponym of *animal*, it is a superordinate of, say, *spaniel*.

Hyponymy can be thought of as the “-- is a --” relation which guarantees the truth of general statements such as *An apple is a fruit* and *An actress is a*

woman. For a lexical item X to be a hyponym of another item Y, the truth of *An X is a Y* must follow logically from the meanings of X and Y. An expectation that if something is an X, it is likely to be also a Y, is not enough. For instance, if someone talks about a cat, most people will assume that the cat in question is somebody's pet. However, this does not entitle us to say that *cat* is a hyponym of *pet*, because there are cats which are not pets, and so *Cats are pets* is not automatically true by virtue of its meaning.

5.1.2 Relations of identity and inclusion II: meronymy

The part-whole relation, in its lexical aspect, is called *meronymy* (sometimes *partonymy*); for instance, *finger* is a *meronym* of *hand*, and *hand* is the *immediate holonym* of *finger*. The notion of meronymy, like hyponymy, is relational rather than absolute: *hand*, for instance, is the holonym of *finger*, but it is at the same time a meronym of *arm*, which in turn is a meronym of *body*. The chain of relations stops at *body*, which may be termed the *global holonym*. Other examples of meronymy are as follows: *arm:body*, *petal:flower*, *engine:car*, *blade:knife*. Prototypical meronymous pairs (where X is a meronym of Y) are normal in frames such as: *X is a part of Y*; *A Y has an X*; *The parts of a Y are A, B, C . . . and so on*. Meronymy must be clearly distinguished from hyponymy, although both involve a species of inclusion. An easy way to highlight the difference is to note that a finger is not a kind of hand (meronymy), nor is a dog a part of an animal (hyponymy).

Not all portions of an object qualify as parts: a glass jug dropped on a stone floor does not break up into parts, but into pieces. The things we habitually call parts typically have a distinctive function or they are separated from sister parts by a formal discontinuity of some sort (or both). For instance, the wheels of a car have the function of allowing it to move smoothly over the ground, and transmit the motive power; the steering-wheel allows the direction of movement to be controlled; the door handles allow the doors to be opened and shut manually. Discontinuity manifests itself in a number of ways. For example, the wheels of a car are detachable and can move relative to the chassis; the fingers of a hand are not detachable, but have a certain freedom of movement; discontinuity may also be visual, like the cuff of a sleeve, or the iris of the eye.

Parts may be necessary or optional. The necessity in question is not a logical necessity, but a well-formedness condition: a hand with a finger missing is still a hand, but it is not a well-formed hand. In this sense, *finger* is a necessary (or *canonical*) part of *hand*, as is *prong* of *fork*. On the other hand, faces may be perfectly well-formed without beards, and doors without handles – here we are dealing with optional (or *facultative*) parts. Some parts are more tightly integrated into their wholes than others. An indication of less-than-full integration is the possibility of describing the part as “attached to” its whole; this is typically not normal with fully integrated parts. Contrast *The handle is attached to the door* (not fully integrated) and *?The handle is attached to the spoon* (fully integrated).

5.1.3 Relations of identity and inclusion III: synonymy

Dictionaries typically define synonyms on the lines of “words with the same or a similar meaning.” This description undoubtedly applies to all words that we would intuitively call synonyms: *begin* and *commence*, *death* and *demise*, *wedding* and *marriage*, *motor* and *engine*. However, it is not restrictive enough, as it surely also applies to, for instance, *mare* and *stallion*, which both refer to horses, but which are not synonyms. It would seem useful, therefore, to examine more closely the notion of “same or similar meaning.”

Synonym pairs or groups can be categorized according to how close the meanings of the words are. Three degrees of closeness can be recognized: *absolute synonymy*, *propositional synonymy*, and *near synonymy*.

The greatest possible resemblance between two senses is identity, in other words, absolute synonymy. A characterization of absolute synonyms based on Haas’s contextual approach was offered earlier, namely, that they are equinormal in all (grammatically well-formed) contexts. This is based on the assumption that any difference of meaning will reveal itself as a difference in co-occurrence possibilities, hence the discovery of a context where one of the putative synonyms is more normal than the other rules out the pair as absolute synonyms. This is an extremely strict criterion, and a rigorous testing of candidate pairs leads rapidly to the conviction that absolute synonyms are hard to come by. From the semiotic point of view this should probably not be surprising: there is no obvious reason why a language should have two forms with absolutely identical meanings. Let us look at a few possible examples of absolute synonymy:

(i) nearly / almost:

These are shown to be not absolute synonyms by the differences in normality between (15) and (16), and between (17) and (18):

(15) We’re very nearly home now.

(16) ?We’re very almost home now.

(17) He looks almost Chinese.

(18) ?He looks nearly Chinese.

(ii) big / large:

The difference in normality between (19) and (20) is enough to disqualify these:

(19) You’re making a big mistake.

(20) ?You’re making a large mistake.

(iii) begin / commence:

These, too, are disqualified:

- (21) Are you sitting comfortably, children? Then I'll begin.
(22) ?Are you sitting comfortably, children? Then I'll commence.

Absolute synonymy presumably approximates to what those people have in mind who maintain that true synonyms do not occur in natural languages.

There is perhaps a case for saying that absolute identity of meaning can occur between forms belonging to different varieties, especially dialects, of a language. An obvious example would be *fall* and *autumn* in American and British English, respectively. These are no different in principle to translational equivalents in different languages. Notice, however, that these would not come out as absolute synonyms by the Haasian test, since *fall* would be less normal than *autumn* in a sentential context that was otherwise lexically marked as British English. Saying that *fall* and *autumn* are identical in meaning presupposes a non-Haasian notion of what meaning is.

Propositional synonymy is less strict than absolute synonymy, and examples of this variety are consequently more numerous. It can be defined in logical terms: propositional synonyms can be substituted in any declarative sentence *salva veritate*, that is, without changing its truth value. By this criterion, *begin* and *commence* are propositional synonyms, because if *The lecture began at nine o'clock* is true, then so is *The lecture commenced at nine o'clock*, and vice versa.

There are too few absolute and propositional synonyms in any language to justify the existence of a dictionary of synonyms; the majority of what lexicographers call synonyms are, in our terms, near synonyms. The following illustrate sets of near-synonyms:

- (i) kill, murder, execute, assassinate
- (ii) laugh, chuckle, giggle, guffaw, snigger, titter
- (iii) walk, stroll, saunter, stride, amble
- (iv) anxious, nervous, worried, apprehensive, fearful
- (v) brave, courageous, plucky, bold, heroic
- (vi) calm, placid, tranquil, peaceful, serene

The words in these sets are not necessarily propositionally identical, so for at least some pairs it is not anomalous to assert one member and simultaneously deny the other:

- (23) He wasn't murdered, he was executed.
(24) They didn't chuckle, they tittered.
(25) He was plucky, but not heroic.

Near-synonyms often occur normally in the test-frame *X, or rather Y*, which signals first, that *Y* conveys propositional information not present in *X*, and second, that the difference is relatively minor. Thus, (26) is normal, but (27) is odd, because the difference in meaning is too great; (28) is odd because there is no propositional difference:

- (26) He was murdered, or rather, executed.
 (27) ?He was murdered, or rather, beaten up.
 (28) ?He was killed, or rather, deprived of life.

Near-synonyms, then, are words which share a salient common core of meaning, but differ in relatively minor respects. There is at present no more precise characterization of “minor” in this context.

Synonyms (of all kinds) often occur in clusters, and it is common for the cluster to be centered round a neutral word which subsumes all the rest, and of which the others are a semantic elaboration. For instance, *kill*, *laugh*, *walk*, *anxious*, *brave*, and *calm* are the central items, respectively, in the sets detailed above.

5.1.4 *Relations of opposition and exclusion I: incompatibility and co-meronymy*

We have looked at relations of inclusion; equally important are relations of exclusion, especially those that hold between sister items under a common inclusive term. Just as there are two sorts of inclusion, there are also two corresponding sorts of exclusion, which receive the labels *incompatibility* and *co-meronymy*.

Incompatibility is the relation which holds between, for instance, *cat* and *dog*, *apple* and *banana*, *rose* and *tulip*, *man* and *woman*, *church* and *supermarket*, *bus* and *tractor*. The essence of this relation is mutual exclusion of classes: if something is a cat, then it follows ineluctably that it is not a dog, and vice versa – there is nothing that is simultaneously a cat and a dog. The same is true for the members of the other pairs mentioned. Note that this is not simple difference of meaning. Take the case of *novel* and *paperback*, which are both hyponyms of *book*. They clearly do not mean the same; on the other hand, they are not incompatibles, because something can be simultaneously a novel and a paperback. The same applies to *mother* and *doctor*, and *tall* and *blonde*.

A parallel relation of exclusion applies to sister meronyms of the same holonym, as in *nose*, *cheek*, *chin* of *face*, or *wheel*, *engine*, *chassis* of *car*, and so on. Here the exclusion is (at least prototypically) spatial: the sister parts of an individual whole do not have any material substance in common.

5.1.5 *Relations of opposition and exclusion II: opposites*

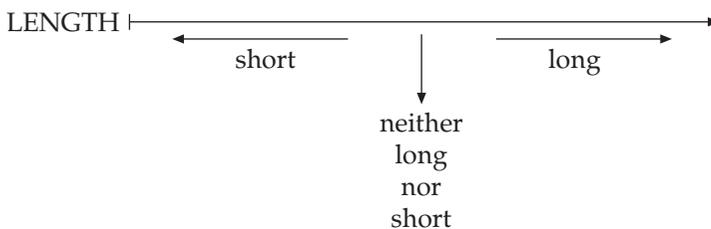
Oppositeness and synonymy are the only sense relations likely to be familiar to a layperson. Most languages have an everyday word for opposites; the relation is cognitively very basic and quite young children can grasp the notion. Opposites are incompatibles of a special type: they are inherently binary, that is to say, they belong together naturally and logically in pairs.

Opposites fall into a number of relatively clearly-defined types, the most important of which are *complementaries*, *antonyms*, *directional opposites* and *converses*.

Complementaries are probably the most basic sort. They can be distinguished from non-complementary incompatibles by the fact that negating either term logically implies the other. For instance, *Proposition P is not true* logically implies *Proposition P is false* and *Proposition P is not false* implies *Proposition P is true*; hence *true* and *false* are complementaries. They may be contrasted with ordinary incompatibles like *cat* and *dog*: *This is not a dog* does not imply *This is a cat*. Other complementary pairs are: *open:shut*, *dead:alive*, *stationary:moving*, *male:female*. A pair of complementaries bisects some conceptual domain, without allowing any “sitting on the fence”; whatever belongs in the domain must fall on one side of the divide or the other. (The negation test works only for items which belong in the domain presupposed by the test word: *This piece of chalk is not dead* does not imply *This piece of chalk is alive*, because chalk does not belong to the domain of things to which *dead* and *alive* properly apply.) The relation between complementaries can be portrayed as follows:

true	false
------	-------

Antonyms (in the narrow sense – the term is also often used to refer to opposites in general) are gradable adjectives (i.e. ones which can be modified without oddness by intensifiers such as *very*, *rather*, *extremely*, *a little*, and so on). Typical examples are *long:short*, *fast:slow*, *heavy:light*, *difficult:easy*, *thick:thin*, *good:bad*, *hot:cold*, *clean:dirty*. They indicate degrees of some property such as speed, weight, or length, one term denoting a value on the scale above some implicit standard appropriate to the context, and the other term denoting a value lower than the standard. Unlike complementaries they do not exhaustively bisect their domain – there is a neutral area between them, which can be described as, for instance, *neither good nor bad*, *neither long nor short*, *neither hot nor cold*, etc. The relation between antonyms can be portrayed as follows:



The comparative forms of antonyms vary according to whether they presuppose the positive forms or not. For instance, for something to be *hotter* than something else, it has to be already *hot*: *X is cold, but it's hotter than Y* is therefore odd. On the other hand, something that is *longer* than something else does not need to be *long*. *Hotter* is known as a *committed* comparative; *longer* is *impartial*. Committedness can be used to define three classes of antonyms: *polar*, *overlapping*, and *equipollent*.

Polar antonyms: both members of a pair are impartial in the comparative:

- (29) X is heavy, but it's lighter than Y.
 (30) X is light, but it's heavier than Y.

Other examples are: *long:short; high:low; wide:narrow; thick:thin; fast:slow; hard:soft*. Polar antonyms indicate degrees of objective, usually measurable, properties.

Overlapping antonyms: one member of a pair is committed in the comparative, the other is impartial:

- (31) ?X is good, but it's worse than Y.
 (32) X is bad, but it's better than Y.

Other examples are: *kind:cruel; clean:dirty; polite:rude*. The members of this class all have an evaluative polarity, one member being commendatory, the other derogatory.

Equipollent antonyms: both members are committed in the comparative:

- (33) ?X is hot, but it's colder than Y.
 (34) ?X is cold, but it's hotter than Y.

Other examples are: *happy:sad; proud of:ashamed of*. The members of this group typically denote sensations or emotions.

Membership of one of the groups described above correlates with other properties, of which the following are worth noting. One important feature of antonyms is the possibility of degree questions. There are two principal forms, (a) those using a noun related to the adjective, as in *What is the length / weight / thickness of X?*, and (b) *how*-questions, such as *How long / thick / heavy is it?* The characteristics of the degree questions in each group are as follows:

- (i) Polar antonyms: One antonym yields a neutral (impartial) *how*-question, the other (for most speakers) a somewhat abnormal question:
- (35) How long is the piece of wood?
 (normal and impartial)
 (36) How short is the piece of wood?
 (a bit odd, but if we have to interpret it, it is not impartial, but committed)

Polar antonyms also typically allow a *what*-question, but only with one of the terms of the opposition:

- (37) What is the length of the piece of wood? (impartial)
 (38) *What is the shortness of the piece of wood?

For both types of degree-question, the term which produces an impartial question is the one which indicates "more of" the gradable property (e.g. *long* = "more than average length," *thick* = "more than average thickness," and so on).

- (ii) Overlapping antonyms: One antonym yields a normal impartial *how*-question (e.g. *How good was the film?*) and its partner gives a normal, but committed *how*-question (e.g. *How bad were the exam results this year?*). In this case, it is the positively evaluative term which occurs in impartial questions, the other term being committed. Generally speaking, *what*-questions do not appear with antonyms from this group (*How clean was the room when you moved in?* / ?*What was the cleanness of the room when you moved in?*; *How polite was John when he came to see you?* / ?*What was John's politeness when he came to see you?*).
- (iii) Equipollent antonyms: Normal *how*-questions are possible with both terms, but both are committed: *How cold was it?*; *How hot was it?* A *what*-question is possible for *hot:cold* (*What is its temperature?*), but this pair seems unusual in this respect.

An interesting property of overlapping antonyms is the feature of *inherentness*. Take the case of *bad:good*. Of two bad things, it is always possible to describe one as *worse* than the other: *The exam results this year were bad, but they were worse last year; This year's famine was worse than last year's*. However, the use of *better* is curiously restricted: *The exam results were bad this year, but they were better last year; ?This year's famine was better than last year's*. The general principle is that only things that are contingently bad (i.e. where good examples are conceivable) can be described using *better*: inherently bad things can only be qualified as *worse* (and, incidentally, cannot be questioned using *How good* . . . ? : **How good was John's accident?*).

Directional opposites: directional opposites are of two main types,

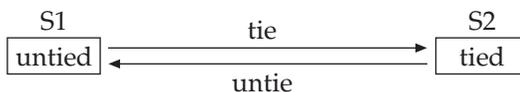
- (i) static directions, like *up:down, backwards:forwards, north:south*:

west ←————→ east

and

- (ii) dynamic directional opposites (usually called *reversives*) such as *rise:fall, advance:retreat, increase:decrease, lengthen:shorten, dress:undress, tie:untie, mount:dismount, enter:leave, damage:repair*, and so on.

We shall concentrate here on reversives. It will be noticed from the examples given that the notion of reversivity is extended from purely spatial domains to any sort of change of state. In general terms, a verb which is a member of a reversive pair denotes a change from an initial state (say S(1)) to a final state (S(2)); its partner will then denote a change from S(2) to S(1):



An important feature of such verbs is that the path of change is irrelevant. For instance, a train entering then leaving a station may well travel in only one direction: what is important for entering is to start out “not in” something, and to end up “in” it, and the reverse is the case for leaving. Or take the case of tying and untying one’s shoes: a film of someone untying their shoes is not identical to one of someone tying them run backwards: the nature of the process of change is not specified by reversible verbs, only the initial and final states.

Reversible verbs have another curious property. Consider the following sentence: *Mary tied the knot, then untied it again five minutes later*. Assuming that *again* is unstressed, what is asserted to have been repeated, by the use of *again*? It is not, in fact, the act of untying – this may be the first time that Mary has ever untied a knot; what is said to recur is the state of being untied. This is presumably a further reflection of the importance of the initial and final states in the semantics of reversible verbs.

Converses: Converses are pairs like *above:below* and *parent:offspring*. Unlike most opposites, both terms can be used to describe the same state of affairs: for instance, *A is above B* means the same as *B is below A*, except in respect of which term serves as the reference point; similarly, *A is B’s parent* designates the same relationship between A and B as *B is A’s offspring*. For this reason, some linguists consider converses to be a variety of synonym.

Converses may be 2-, 3-, or 4-place relations, according to the number of arguments involved. *Above:below* are 2-place converses; *bequeath:inherit* are 3-place converses (“*John bequeathed a fortune to Mary*” designates the same event as “*Mary inherited a fortune from John*”); *buy:sell* are 4-place (“*John bought the car for £1,000 from Bill*” describes the same transaction as “*Bill sold the car to John for £1,000*”).

5.2 Syntagmatic sense relations

We turn now to semantic relations between words which occur together in the same text. These can be roughly divided into two types. First, there are those which hold over relatively long stretches of text, between grammatically unrelated items, and which typically do not involve propositional meaning or directional properties. For instance, in *The Prime Minister attended the White House reception accompanied by his Dad*, there is a register clash between *Dad* and such formal items as *attended*, *reception* and *accompanied* (*Tony Blair went to the White House party with his Dad* sounds less weird); the clash would be resolved by replacing *Dad* with the propositionally synonymous *father*. Notice that in this case, the clashing items are some distance from one another, and are not directly related grammatically (*Dad* does not clash with *his*, nor *his Dad* with *by*). Second, there are relations which hold between closely related elements in the same grammatical construction and which do frequently involve propositional meaning and directional properties. For instance, the clashes in

?*a male aunt*, ?*a highly strong man* and ?*John drank a filing cabinet* involve the second type of relation. Take the case of the latter example: there is no clash between *John* and *drank*, or *John* and *filing cabinet*, the clash involves specifically *drank* and its direct object *filing cabinet*; the clash involves propositional meaning, since it can only be resolved by substituting either *drank* or *filing cabinet* with something propositionally different (e.g. *bought* or *wine*, respectively); *drank* imposes semantic restrictions on its direct objects (e.g. they must be liquids).

There are three possible effects of putting words together in a grammatically well-formed construction: either the result is normal, as in *John drank the wine*, or there is a semantic clash, as in *John drank the filing cabinet* or *a highly strong man*, or the result is pleonastic (or redundant) as in *a female aunt*.

Generally speaking, for a combination of words to be semantically normal, two conditions must be satisfied. If two words are joined together in a construction, it is usually possible to identify a selector, which imposes semantic conditions on possible partners, and a selectee, which satisfies (or does not satisfy) the conditions; the first requirement for a normal combination is that these conditions must be satisfied. This will avoid semantic clash. In the case of an adjective-noun combination, it is the adjective which is the selector: compare the ease with which semantic conditions (usually called selectional restrictions), can be specified for normality in the following:

A ---- woman.

There is no semantic generalization which covers, for instance, *intelligent*, *tall*, *pregnant*, *kind*, *highly paid*, *left-handed*, all of which combine normally with *woman*.

A pregnant ----.

Here, the restriction is much easier to capture: *pregnant* requires a head noun which (a) denotes a mammal, and (b) is not specifically marked as "not female" (like, for instance, *bull*).

It is generally the case that in a modifier-head construction, such as adjective-noun or intensifier-adjective, the modifier is the selector; in a head-complement construction, such as *drinks beer* in *John drinks beer*, it is the head of the construction, i.e. the verb, which is the selector.

In ?*a female aunt* there is no semantic clash, but the combination is still odd, which indicates that a further condition must be satisfied. This is that a grammatically dependent item (modifier or complement) must contribute semantic information not already present in the head. Clearly, the notion FEMALE is part of the meaning of *aunt*, so the word *female* adds nothing, and consequently ?*a female aunt* is pleonastic, or redundant; *a lesbian aunt*, on the other hand, is semantically normal, because although *lesbian* incorporates the notion FEMALE, it also brings new information not predictable from *aunt* alone.

6 Meaning Extensions and Change

6.1 *Established readings and nonce readings*

Some of the alternative senses of a word are permanent and established features of the language, and we would expect them to be recorded in any dictionary worthy of the name. We may also assume that they are laid down in neural form in the mental lexicons of competent speakers of the language. This is the case with the two readings of *bank* discussed earlier. In a sentence like *Mary works in a bank*, we can say that the context “selects” a reading from among those that are permanently laid down, in the sense that only one of them yields a normal combination.

But take another case. Imagine a reception with a large number of guests who will later proceed to a dinner. There is no dining room large enough to accommodate all the guests, so they are divided into two groups. On arrival, each guest is presented with either a rose or a carnation. When it is time for dinner, the head waiter announces: *Will all roses proceed to Dining Room A, and carnations to Room B, please.* These uses of *rose* and *carnation* are perfectly comprehensible in context, but they are not an established part of the language, nor would we expect to find them in any dictionary, however complete. These are said to be nonce readings. How do they arise? Not by selection, but by coercion: if none of the established readings fits the context, then some process of sense-generation is triggered off, which produces a new reading.

A third possible effect of context is to enrich an existing reading, without producing a new sense. This is what happens to *teacher* in *John’s maths teacher is on maternity leave.*

6.2 *Literal and non-literal readings*

A distinction is often made between literal and non-literal meanings of (polysemous) words, the assumption being that only one of the readings is literal. While at first sight this distinction seems intuitively clear, on closer examination it is not so straightforward. One thing is clear; a literal meaning must at least be an established one; the criteria for privileging one out of the set of established readings, however, are less clear. Dictionaries often order their entries in terms of chronological order of earliest attestation in the language. However, the earliest recorded meaning of a word does not necessarily strike speakers’ intuitions as the literal meaning. There is no doubt, for instance, that the “die” meaning of *expire* predates the “cease to be valid” meaning, but (to my surprise) current British undergraduates, when asked to pick out the reading they intuitively feel to be the literal one, are virtually unanimous in

selecting the “cease to be valid” reading. Another possible criterion is frequency in a language: one might reasonably expect the literal meaning to be the most frequent. Once again, however, this does not always accord with (strong) native speaker intuitions. For instance, few would dispute that the “have a visual experience” meaning of *see* is the literal meaning and “understand” an extended meaning, yet evidence from one very large corpus of English usage indicates that the latter is the more frequent. Two possibly more valid criteria are, first, the default meaning, that is, the one which comes first to mind when one is confronted by the word out of context, and, second, the reading from which the others can most plausibly be derived by standard semantic processes. The latter criterion is most reliable when there are more than two readings. Take the case of *position*, and three readings:

- (i) “location in space”
- (ii) “job in a large firm, etc.”
- (iii) “opinion”

Starting from (i), it is easy to see how (ii) and (iii) may have arisen by means of a process of metaphorical extension; but starting from either (ii) or (iii), there is no obvious way of deriving the other two.

6.3 *Metaphor*

There are a number of strategies for deriving one reading from another. Three will be illustrated here. The first of these is metaphor. Metaphor is essentially the projection of the conceptual structure appropriate to a familiar field onto a different and less familiar field, and depends for its effectiveness on a sufficient resemblance between the two fields for the projection to be intelligible. For instance, in the case of *expire*, projecting the notion of dying onto the life-cycle of, say, a credit card, allows an immediately intelligible parallel to be drawn between death and the end of the period of usability of the card.

6.4 *Metonymy*

The second strategy of meaning extension is metonymy, which is based not on resemblances or analogies between items in different conceptual domains, but associations within a single conceptual domain. Referring to people wearing roses as *roses* is intelligible, not because of any structural parallels between the concept of a rose and the people designated, but because of the close association between the latter and roses in a particular situation. This usage is unlikely to become established. An example of metonymy that is so

well-established that we are hardly aware that it is non-literal is *The kettle's boiling*.

6.5 Specialization and generalization

A third process which produces new meanings from old ones is a change of inclusiveness, widening or narrowing down the meaning. The meaning "take alcoholic beverage" of *drink* is derived from the meaning "imbibe liquid" by specialization; *meat* has become specialized from meaning food of any kind to "animal flesh used as food"; *handsome* in Jane Austen seems to apply indifferently to men and women, but became specialized later to apply primarily to men, acquiring a particular nuance in its application to women; *interfere* in Jane Austen seems rather like current *intervene*, acquiring its disapproving connotations at a later date. The meaning of *cat* which includes lions, tigers, ocelots, and jaguars is derived by generalization from the meaning which includes only *felis domesticus*.

The loss of a whole sense may be considered a type of specialization: an example of this, again from Jane Austen, is *direction*, which has lost the reading "address" (as on a letter).

6.6 Amelioration and pejoration

Amelioration is when a neutral or pejorative term becomes commendatory, and *pejoration* is when the reverse movement occurs. The latter seems to be by far the more frequent, and is particularly prone to happen to words referring to women. Examples of words undergoing pejoration are: *madam*, *mistress*, *courtesan*, *wench*, *tart*, and so on. (The only two words referring to women that Jane Miller in *Womanwords* signals as having ameliorated are *jilt* which originally meant a prostitute, and *bat*, which "has lost its negative sexual connotations.")

7 Larger Groupings of Words

7.1 Word fields

We have already seen that the vocabulary of a language is not just a collection of words scattered randomly through semantic space: it is at least partially structured by recurrent sense relations. In some areas of the vocabulary the sense relations unite groups of words into larger structures, known as *lexical fields* (or *word fields*). We shall look briefly at examples of one type of larger structure, namely, *lexical hierarchies*. Hierarchies may be *non-branching*, as in figure 10.1, or *branching*, as in figure 10.2:



Figure 10.1

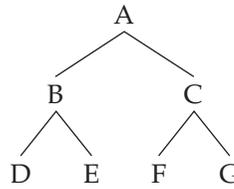


Figure 10.2

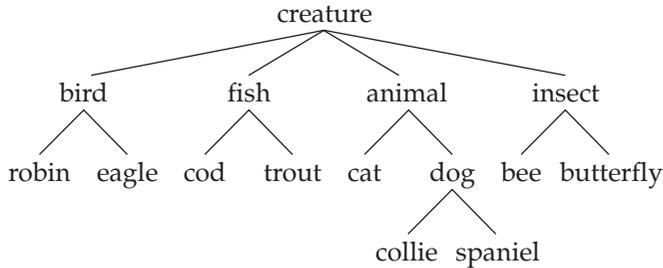


Figure 10.3

Note: A word of explanation is perhaps needed regarding the position of the word “animal” in the above hierarchy, which for some non-British speakers of English, who feel that “animal” and “creature” are synonyms, and that “animal” subsumes “bird,” “fish,” “insect,” and so on, may seem anomalous. This is not so in British usage: the *Collins Handguide to the Wild Animals of Britain and Europe*, Ovenden et al. 1979, a field guide to identification, includes mammals, amphibians (such as frogs and newts), and reptiles (such as snakes and lizards), but no birds, fish or insects (which are covered by sister volumes). I am not aware of any other language which has a term with exactly this denotation.

If these are lexical hierarchies, then A, B, C, etc. represent word senses, and the lines joining them represent sense relations. The following are examples of non-branching hierarchies (turning them on their side for convenience):

- (i) general – colonel – major – captain – lieutenant – etc.
- (ii) ocean – sea – lake – pond – puddle
- (iii) scorching – hot – warm – lukewarm – cool – cold – etc.
- (iv) tertiary (education) – secondary – primary – pre-school

A structural necessity for a branching hierarchy is that the branches must never converge. There are two main sorts of branching lexical hierarchy, which are called *taxonomies*, which are structured by hyponymy and incompatibility and *meronomies*, which are structured by meronymy and co-meronymy. Figure 10.3 illustrates a fragment of a taxonomy (valid for British English).

It will be appreciated that the hierarchy illustrated in figure 10.3 can be extended upwards, downwards, and sideways, ultimately producing a huge structure encompassing all living things. Most taxonomies are much more

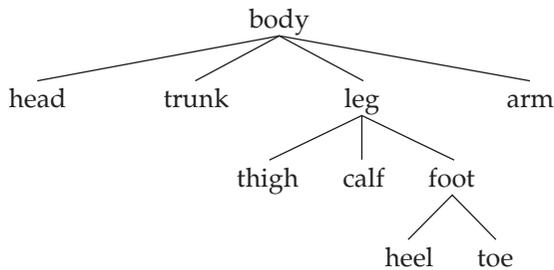


Figure 10.4

fragmentary than this, covering such areas as vehicles, buildings, clothes, and so on.

A taxonomy typically has well-defined levels: in figure 10.2, A is at Level 1, B and C are at Level 2, D, E, F, and G are at Level 3. One level of a taxonomy has a special status and is known as the *basic level*, as it seems that the structure is organized around this level. It is the level at which the “best” categories occur. This means that resemblance between fellow members and distinctiveness from members of sister categories are both maximized. Take the categories ANIMAL, DOG, and SPANIEL. Animals are distinct enough from non-animals, but their overall degree of mutual resemblance is relatively low; in the case of spaniels, they resemble one another highly, but differ in only minor ways from other types of dog; in the category DOG, on the other hand, mutual resemblance and distinctiveness are both high. Basic level items are also the default names for things, the names we use for simple, everyday reference. Suppose A hears a noise in the garden and asks B what it is. B looks out of the window and sees a spaniel. What does he say? Any of the following would be true:

- (i) Oh, it's just an animal.
- (ii) Oh, it's just a dog.
- (iii) Oh, it's just a spaniel.

In the absence of special circumstances, (ii), which contains the basic level item, is by far the most likely.

Some examples (in capitals) of basic level items are as follows:

- a. vehicle – CAR – hatchback
- b. fruit – APPLE – Granny Smith
- c. living thing – creature – animal – CAT – Manx cat (British)
- d. object – implement – SPOON – teaspoon

The other main type of branching lexical hierarchy is the part-whole variety. (In figure 10.4, as in figure 10.3, only some of the branches are shown.) Part-whole hierarchies are just as numerous in the vocabularies of natural languages as are taxonomies. They differ from taxonomies in a number of respects, but

perhaps the most significant difference concerns structural levels: meronomies tend to have no, or only weakly developed, levels, hence there is no equivalent to the basic level of a taxonomy.

7.2 Word families

Another type of grouping of associated words is the *word family*. Most complex lexemes are built up out of a *root* and one or more *derivational affixes*. For instance, the word *undress* is composed of the root *dress* and the prefix *un-*; the noun *re-entry* is composed of the root *ent(e)r*, the prefix *re-* and the suffix *-y*. A word family is composed of all the words derived from a given root. For instance, the following all belong to one word family:

nation	national (adj.)	national (n.)
nationally	nationalize	denationalize
nationality	nationalism	nationalist
international	transnational	nationhood
(etc.)		

As native speakers, we have a quite complex knowledge of which derivations are possible, and what they mean. We know, for instance, that although a *painter* is someone who paints, and a *painting* is the concrete end result of a painter's efforts, a *screwdriver* is not someone who inserts screws, but an appliance for doing this, and a *killing* is not the concrete end result of the process denoted by *kill* (i.e. a corpse), but an instance of the act itself. We also know that whereas a *diner* may be someone who is dining, it can also be a place where one dines, and we know that there is no parallel reading for *painter*.

7.3 Domain-specific vocabulary

Another type of word-grouping controlled by native speakers is the vocabulary appropriate to a particular situation, for instance, a race meeting:

1	2	3	4
horse	jockey	course	bet
favorite	owner	race	run
form	trainer	fence	win
odds	bookmaker	winner's enclosure	jump
colors	steward	starting-gate	take a fence
winner	stable-boy	heavy going	come up on the inside
handicap		reins	fall
yearling			disqualify
			scratch
			etc.

A broad grouping like this is composed of a number of nested sub-domains, such as weighing in, saddling, starting, running the race, laying bets, and so on.

7.4 *Layers of vocabulary*

We shall use the expression *layers of vocabulary* to refer to much larger groupings of words, each of which will incorporate many structures like those described, which are confined to certain areas of usage. For instance, there are technical vocabularies, such as those used by art historians, or doctors, when communicating with others of their kind. There are also collections of words which associate together at different levels of formality. These are effortlessly called up in appropriate contexts, and they must be somehow linked in storage.

7.5 *The mental lexicon*

Each of us has in our cognitive system some kind of inventory of all the words that we know, together with all the information – semantic, grammatical and phonetic / graphic – necessary for their correct use. Estimates of the number of words known by an average adult speaker vary from 150,000 to 250,000 (see, for instance, the discussion in Aitchison (1987: 5–7)). This represents a vast quantity of information.

The inventory is accessed via written or spoken forms every time we hear or read something in a language we know, and via some kind of semantic representation every time we produce language (recall that, because of widespread synonymy and polysemy, the mapping between forms and meanings is not one-to-one, but many-to-many). Although much is known, the details of representation and processes of use are still very imperfectly understood; nonetheless, the astonishing speed at which words are retrieved and identified – within about a fifth of a second from the start of the word, for spoken language – points to a highly efficient and organized storage system.

Every person's mental lexicon is different from everyone else's, yet by and large we manage to understand each other; this presumably indicates an adequate degree of overlap between individual lexicons.

7.6 *Vocabularies*

In addition to the mental lexicons of individual speakers of a language, it is possible to think of the total lexical stock of a language, which covers all its speakers, including those belonging to distinct speech communities, and including those who are now dead. Of course, the boundaries of such an entity

are very vague, and will differ according to the purposes of the compilers and users (how far back in time do we go? how many dialect, or specialized technical forms do we include?, etc.). The natural home for such a vocabulary is the dictionary, and the natural way of drawing it up is by studying corpuses. The contents of a dictionary do not correspond to the contents of the mental lexicon of any single speaker, nor do they represent what is common to all speakers; yet every entry must be justified by some degree of common ownership in one or other of the sub-communities using the language.

8 Conclusion

We have now surveyed, at least in broad outline, the whole domain of words in language, from the detailed properties of individual words, through relational properties along the major paradigmatic and syntagmatic axes, to communities of words, large and small, tightly or loosely structured.

It should be borne in mind, however, that many detailed aspects of the lexicon are still only imperfectly understood, and of these, a number are currently the object of intense research activity. Two recent major stimuli to research on the lexicon are worth mentioning. The first has been the advent of powerful computers, and the attempt by computational linguists to develop programs capable of “understanding” natural language texts. The syntactic problems have proved relatively tractable; the big problem has turned out to be the lexicon – deciding what a computer must “know” about word meanings and how they are to be represented. The second major stimulus has been the development of large-scale corpuses of spoken and written language (together with tools for processing them), which allow an accurate picture to be gained of how words are actually used. This has, among other things, revolutionized lexicography, and no doubt the full effects of both stimuli have yet to be seen.