

3

The paradigmatic and syntactic delimitation of lexical units

3.1 Introductory

In the previous chapter a number of important decisions were taken which enable us to establish the location of lexical elements within sentences, that is to say, to delimit them syntagmatically. We must now confront the rather more daunting problems of differentiating lexical units paradigmatically.

It will be necessary to introduce a distinction, which has up to now not been needed, between two kinds of element relevant to lexical semantics. The two types will be called **lexical units** and **lexemes**.¹ In this book our main, although by no means exclusive, concern is with the former. Lexical units are those form-meaning complexes with (relatively) stable and discrete semantic properties which stand in meaning relations such as antonymy (e.g. *long* : *short*) and hyponymy (e.g. *dog* : *animal*), and which interact syntagmatically with contexts in various ways to produce, for instance, the different sorts of anomaly discussed in chapter 1. A particular lexical unit, of course, expresses its semantic identity through such relations, but its essence cannot be exhaustively characterised in terms of any determinate set of such relations. The meaning aspect of a lexical unit will be termed a **sense**. Lexemes, on the other hand, are the items listed in the lexicon, or 'ideal dictionary', of a language; these will be discussed in 3.10.

It may be wondered why it is necessary, or even advantageous, to have two sorts of unit for lexical semantics. The reason is that they have different functions, which impose different constraints on their nature. Senses need to represent unitary 'quanta' of meaning, but they do not need to be finite in number. There is nothing in the notion of oppositeness, for instance, which dictates that there should necessarily be only a finite number of opposite pairs in a language. A lexeme, on the other hand, may well be associated with indefinitely many senses, but the set of lexemes must be finitely enumerable. Consider, by way of illustration, the example of *topless*. We may speak of (a) *a topless dress* or (b) *a topless dancer*.

Each of these is lexically distinct, in that it has, for instance, different typical contrasts (e.g. *long-sleeved* for (a) and *nude* for (b)), and the two readings are called forth by different types of context. They are also relatively stable across contexts: for instance, *a topless swim-suit* would seem to exemplify the same sense of *topless* as (a), and *a topless barmaid* the same as (b). Why then can we not simply say that *topless* (a) and *topless* (b) are different (although perhaps related) lexemes? One important reason is that the number of possible distinct uses of *topless* seems to be, in principle, open; so any attempt to draw up a determinate closed list would be of questionable validity. In addition to the examples mentioned above, one might also encounter *a topless bar*, *I hear Torquay has gone topless*. And can we entirely rule out *topless by-laws*, or *the topless watchdog committee* (with the function of monitoring the behavioural effects of *toplessness*)? Another example of such openness is provided by what we shall call the **unit-type** ambiguity. For instance, *jacket* in *I like this jacket* may be understood to refer to a particular individual jacket (the unit), or to a type of jacket. However, there seems, in principle, to be no limit to the number of possible type readings in such cases. Suppose someone in a greengrocery picks up an apple and says: *Is this the fruit you mean?* Besides the unit reading, the speaker may be intending to refer, among other possibilities, to: that variety of apple (e.g. Cox's Orange Pippin); apples in general; fruit from a particular supplier; home-grown apples; etc., etc. While in particular contexts some readings may well be much more likely than others (if this were not the case, many utterances would be harder to understand than they are), the number of possible readings is clearly limited only by imagination. (Except, of course, that the type cannot be more general than the lexical item used to refer to it: *this apple* cannot refer to fruit in general.) It seems that there is a high degree of creativity in the lexicon which we must take account of. The creativity inherent in the grammar of a language has often been pointed out: an unlimited number of sentences may be produced from a finite set of elements together with rules for their combination.² Lexical creativity is probably of a similar order and, like syntactic creativity, must have a finite aspect. It will be assumed in this book that a (relatively) closed set of lexical units is stored in the mental lexicon, together with rules or principles of some kind³ which permit the production of a possibly unlimited number of new (i.e. not specifically stored) units.

3.2 Selection and modulation of senses

One of the basic problems of lexical semantics is the apparent

multiplicity of semantic uses of a single word form (without grammatical difference). There seems little doubt that such variation is the rule rather than the exception: the meaning of any word form is in some sense different in every distinct context in which it occurs.

However, that does not mean that the 'word-form-in-context' is the appropriate unit for lexicological purposes. There are two distinct types of variation in the semantic contribution that a word form makes to different sentences – or, to look at it from a different point of view, two ways in which the sentential context of a word form may affect its semantic contribution to the sentence. It will be argued that one of these types of variation involves the selection, by the context, of different units of sense, while the other type is a matter of contextual modification of a single sense.

The difference between the two contextual effects can perhaps be approached initially by considering two corresponding ways in which a word form, in a single context, may be open to more than one interpretation. *Cousin* and *bank* in 1 and 2, respectively, illustrate the difference:

1. Sue is visiting her cousin.
2. We finally reached the bank.

Cousin in 1 can, of course, refer to either a male or a female cousin. But the sentence can function as a satisfactory communication without either the hearer perceiving, or the speaker intending to convey, anything concerning the sex of the person referred to. This is because *cousin* has a general meaning which covers all the more specific possibilities (not only with regard to sex, but also with regard to an indefinitely large number of other matters, such as height, age, eye-colour, etc.). *Bank* in 2 can also be interpreted in more than one way (e.g. "margin of river" or "establishment for the custody of money"); but it has no general meaning covering these possibilities. Furthermore, the interpretation cannot be left undecided: both speaker and hearer must select a reading (the same reading) if the sentence is to play its part in a normal conversational exchange. We shall say that the word form *cousin* is **general** with respect to the distinction "male cousin"/"female cousin";⁴ *bank*, on the other hand, will be said to be **ambiguous** with respect to the sense distinction "financial institution"/"side of river". In other words, the two meanings "male cousin" and "female cousin" are both associated with the same lexical unit *cousin*, whose meaning is more general than either; they therefore do not represent distinct senses of *cousin*. The meanings "financial institution" and "side of river", on the other hand, do represent two distinct

senses, so there are two lexical units *bank* corresponding to these senses. (Every word form is general with respect to some semantic distinctions, and (at least potentially) ambiguous with respect to others.) Let us now examine in greater detail the different ways in which contexts exert a restrictive influence on the meanings associated with word forms which occur within them.

There are two fundamental ways⁵ in which the effective semantic contribution of a word form may vary under the influence of different contexts. First, a single sense can be modified in an unlimited number of ways by different contexts, each context emphasising certain semantic traits, and obscuring or suppressing others; just as a dirty window-pane will allow some parts of the scene beyond it to be seen clearly, and will partially or completely obscure other parts – and a different pane will affect the same scene differently. This effect of a context on an included lexical unit will be termed **†modulation**; the variation within a sense caused by modulation is largely continuous and fluid in nature. The second manner of semantic variation concerns the activation by different contexts of different senses associated with ambiguous word forms. This will be termed **†contextual selection** (of senses); in the nature of things, this sort of variation proceeds in discrete jumps rather than continuously. The two types of variability are normally operative together; that is, a selected sense is also subject to modulation by the context which forced its selection. Let us first look a little more closely at modulation.

We shall discuss sense modulation under two main headings: first, changes in the status of semantic traits along the dimension of necessity – which will be termed **†promotion** and **†demotion**; and second, the highlighting and background of traits. As an example of promotion and demotion, consider the semantic traits associated with *nurse* in 3 and 4:

3. A nurse attended us.
 4. A pregnant nurse attended us.
- In 3, the trait “female” is expected, and the trait “male” unexpected; but in 4, although *nurse* represents the same lexical unit as in 3, “female” is at the very least canonical (if not criterial), while “male” is denoted to anomalous or impossible status.⁶ As a further example, consider 5:
5. Arthur poured the butter into a dish.

Out of context, or in a neutral context, “liquid” is either a possible or unexpected trait of *butter*. But in 5 it is at least canonical. Sentence 5 also illustrates another aspect of modulation, which we shall call **†linkage**

of traits. It is clear that the butter referred to in 5, if it is normal, is not only liquid, but also hot; “hot” is therefore a canonical trait. Now, “hot” is merely a possible trait of *butter* in, for instance, *Arthur put the butter into a dish*; and it is certainly not the case that any lexical unit functioning as the direct object of *pour* has “hot” as a canonical trait – in *Arthur poured the milk into a dish*, for instance, “hot” is, again, merely possible. It is the combination of *pour* with *butter* (in direct object position) – or, more directly, the interaction of the traits “butter” and “liquid” – which promotes “hot” from possible to canonical status. This is a very simple example; it is easy to conceive of extremely varied and complex patterns of linkage appearing in various contexts. This will not, however, be elaborated on here; we shall merely note that it is an important aspect of modulation.⁷

Another effect of contextual modulation on the sense of a lexical unit involves the relative **†highlighting** or **†backgrounding** of semantic traits. Different sorts of trait can be affected in this way. Two examples will suffice. First, some part of an object (or process, etc.) may be thrown into relief relative to other parts. For instance, *The car needs servicing* and *The car needs washing* highlight different parts of the car. (This is not to say that *car* refers to something different in each of these sentences – in both cases it is the whole car which is referred to.⁸) Second, it is commonly the case that what is highlighted or backgrounded is an attribute, or range of attributes, of the entity referred to. For instance, *We can't afford that car* highlights the price of the car, *Our car couldn't keep up with his highlights* its performance, and *The car crushed Arthur's foot* its weight. It is in respect of ‘contextually modulated sense’ that a lexical unit may be justifiably said to have a different meaning in every distinct context in which it occurs.

We have been speaking, so far, of the effects of context on the meaning of a single lexical unit. But a context normally also acts in such a way as to cause a single sense, from among those associated with any ambiguous word form, to become operative. When a sentence is uttered, it is rarely the utterer's intention that it should be interpreted in two (or more) different ways simultaneously. It is probable that deliberate equivocation in respect of the intended sense of word forms is always to some extent odd. This means that, for the vast majority of utterances, hearers are expected to identify specific intended senses for every ambiguous word form that they contain. The process of sense selection is, of course, extremely complex, with many interacting factors. However, in general, one can say that a hearer selects that combination of lexical readings which leads to

the most normal possible utterance-in-context. In other words, a hearer will generally assume that the producer of an utterance wants to communicate something, and has chosen the linguistic context of his utterance with a view to furthering this aim.⁹ Broadly speaking, we can identify two types of normality – sentence-internal normality and contextual normality (it is probably the case that the latter is the stronger determinant of sense selection). Very often a sentence contains more than one ambiguous word form; in such cases, there will occur a kind of mutual negotiation between the various options so as to achieve the most normal combination. This process is illustrated in 6:

6. Several rare ferns grow on the steep banks of the burn where it runs into the lake.

It is highly unlikely that any reader of this sentence will interpret *rare* in the sense of “undercooked” (as in *a rare steak*), or *steep* in the sense of “unjustifiably high” (as in *steep charges*), or *bank* in the sense of “financial institution”, or *burn* in the sense of “injury caused by fire”, or *run* in the sense of “progress by advancing each foot alternately never having both feet on the ground simultaneously”, etc. There is only one selection of senses here which yields a normal sentence (i.e. the sentence form is not ambiguous). Contextual normality involves such matters as relevance, informativeness and consistency. Consider 7:

7. A: It's dark in here, isn't it?
B: Yes. Aren't there any lights?

B's utterance (in the context of A's) is normal if *lights* is interpreted to mean “sources of illumination”, but would be of, at best, obscure relevance if interpreted to mean “lungs of sheep”. (Notice, however, that B's utterance does not display internal abnormality on either interpretation.)

So far we have taken it for granted that the distinction between ambiguity and generality is intuitively obvious. In some cases it is, but in others it is not; this judgement certainly does not figure amongst the basic set of intuitive judgements on which we base our analyses. We must now, therefore, consider in some detail the question of explicit diagnostic tests for ambiguity and generality.

3.3 Indirect tests for ambiguity

One approach to the diagnosis of ambiguity relies on finding, for two occurrences of a word form, different relations of meaning with other items. These relations may be of the paradigmatic variety (e.g.

oppositeness, synonymy, etc.) or they may be of the so-called **paronymic** sort (that is to say, involving identity of root, but difference of syntactic category, as, for instance, with *act:actor*, *race:racist*).¹⁰ We shall describe evidence of this type as ‘indirect’; arguments will be put forward that indirect evidence has severe drawbacks as a method of diagnosing ambiguity. The following three ‘tests’ for ambiguity will serve to illustrate the approach.

- I. If there exists a synonym of one occurrence of a word form which is not a synonym of a second, syntactically identical occurrence of the same word form in a different context, then that word form is ambiguous, and the two occurrences exemplify different senses.

Thus, for example, one might suggest *lucifer* as a synonym for *match* in 8 (but not in 9), and *contest* as a synonym in 9 (but not in 8):

8. Guy struck the match.
9. The match was a draw.

From this, the principle expressed in I would allow us to conclude (correctly, in this instance), that *match* was ambiguous, and in 8 and 9 represented different senses.

- II. If there exists a word or expression standing in a relation of oppositeness to one occurrence of a word form, which does not stand in the same relation to a second, syntactically identical occurrence of the same word form in a different context, then that word form is ambiguous, and the two occurrences exemplify different senses.

In 10, for instance, *dark* (but not *heavy*) stands in a relation of oppositeness to *light*, whereas in 11 *heavy* is a satisfactory opposite, but *dark* is not:

10. The room was painted in light colours.
11. Arthur has rather a light teaching load.

Light is therefore, according to the test, an ambiguous lexical form, and 10 and 11 manifest different senses.

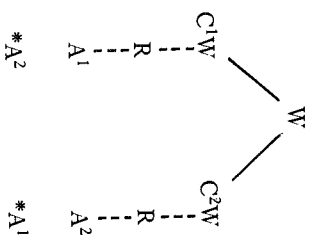
- III. If there exists a word which stands in a paronymic relation to one occurrence of a word form, but does not stand in the same relation to a second, syntactically identical occurrence of the same word form in a different context, then that word form is ambiguous, and the two occurrences exemplify different senses.

Consider *race* in 12 and 13:

12. The race was won by Arkle.
13. They are a war-like race.

The verb *to race* and the noun *racing* are paronymically related to the occurrence of *race* in 12, but not to that in 13; on the other hand, *racial* and *racist* are related to *race* in 13, but not in 12. Hence *race*, according to the test, is an ambiguous lexical form, and 12 and 13 manifest different senses. Once again, the diagnosis seems intuitively correct.

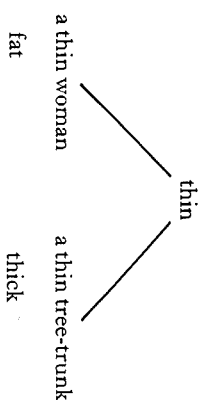
Other tests of the same general type may be proposed, but none bring anything radically new to the picture. They all suffer from a major weakness, which is that for every instance in which a word form possesses different synonyms, opposites, morphological derivatives, or whatever, in different contexts, there are several possible explanations, only one of which involves ambiguity of the word form; hence, further evidence of a different sort is required to determine which explanation is correct in any given instance. Suppose there exists a word form *W*, which in context *C*(1) stands in a particular meaning relation to another element *A*(1), but in context *C*(2) stands in the same meaning relation not to *A*(1), but to *A*(2):



There are at least three possible reasons why *W* should have different relational partners in *C*(1) and *C*(2). One is, of course, that *W* is ambiguous, and *C*(1) and *C*(2) select different senses. This is presumably what happens in the cases of *light*, *match* and *race* discussed above. Another possibility is that *C*(1) and *C*(2) modulate a single sense of *W* in mutually exclusive ways. Thus *monarch* in 14 has *queen* but not *king* as a synonym, whereas in 15 it has *king* but not *queen*:

14. The Ruritanian monarch is expecting her second baby.
15. The child's father is the reigning monarch.

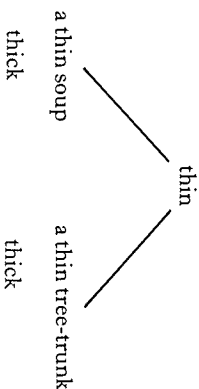
In such a case there would be no evidence of ambiguity. A third possibility is that *A*(1) and *A*(2) are sensitive to differences between *C*(1) and *C*(2) to which *W* is indifferent. Consider the following case of *thin* and its opposites:



It does not seem illuminating to say either that *thin* is ambiguous, or that these contexts restrict its meaning in mutually exclusive ways.¹¹

It is clear that nothing can be reliably inferred from the mere fact that a word form has different meaning relations in different contexts, and independent evidence concerning ambiguity or generality is required. But if such evidence is available, then it is superfluous to appeal to differential relations.

Indirect tests can be used in another way, which in some cases can seem more reliable. Instead of looking for different relations in different contexts to prove ambiguity, one may adduce sameness of relations as evidence of generality. Thus, the fact that *thin* in both of the contexts illustrated above has *slender* as a synonym could be cited as evidence that it is not ambiguous after all. However, using indirect criteria in this way is no more reliable: one simply falls into a different trap, because the item which stands in the 'same' meaning relation in different contexts to the lexical form being tested may itself be ambiguous. There is an intuitively clear example of this involving *thin*:



It is of course true that one of the main purposes of distinguishing discrete senses is to have available a unit which can stand in relations such as synonymy and oppositeness. However, it seems clear that these units must be established in some other way. Fortunately there are more successful and reliable ways of distinguishing ambiguity from generality, and to these we now turn.

3.4 Direct criteria for ambiguity

Three different types of criteria for ambiguity will be proposed. It may ultimately be possible to show that they all reduce to a single basic criterion, but here they will be presented separately. Generally speaking, unless there are specific reasons why one or other of the criteria should be inapplicable (some of these reasons will be discussed below), we shall expect an ambiguous item to satisfy all the criteria.

The first criterion is frequently difficult to apply in practice, but it is conceptually important. It is that the senses of an ambiguous word form should not in every case be totally conditioned by their contexts, unlike the interpretations which arise as a result of contextual modulation. This means that an ambiguous word form set in a disambiguating context may well carry more information than can be accounted for in terms of interaction between the context-independent meaning of the word form, and the semantic properties of the context. In cases of contextual modulation, on the other hand, ALL information is derived from these sources. Consider sentences 16 and 17:

16. Arthur washed and polished the car.

17. John lubricated the car.

The most likely interpretation of 16 is that not every part of the car underwent washing and polishing, but the exterior surface only. What is the basis for this conclusion? It is derived entirely from the general meaning of *car*, together with the semantic properties of the context (remember that general knowledge concerning cars and operations carried out on them is, on the view of meaning adopted in this book, embedded in the meanings of *car*, *wash*, *polish*, etc.). A similar account can be given of the most likely interpretation of *car* in 17. Or take the case of *monarch* in 14 (repeated here for convenience):

14. The Ruritanian monarch is expecting her second baby.

We can be virtually certain that the monarch in question is a queen, because of the restricting effect of the context on the general meaning of *monarch*. Notice that a similar interpretation would arise, and no loss of information would result, if *monarch* were replaced by a synonym or paraphrase such as *sovereign*, or *crowned head* (and *automobile* would interact in the same way with the context if it were substituted for *car* in 16 and 17). Contrast these, however, with *bank* in 18 and 19:

18. Her husband is the manager of a local bank.

19. At this point, the bank was covered with brambles.

Let us try to account for the (most probable) different interpretations of *bank* in the way that we did for *car*. It is first necessary to decide on a synonym or paraphrase of the context-invariant meaning of *bank*. This already poses problems, but let us say, for the sake of argument, that it is equivalent to *place*. We can then observe the effect of substituting *place* for *bank* in 18 and 19:

20. Her husband is the manager of a local place.

21. At this point, the place was covered with brambles.

There is quite clearly a loss of information, so we have failed to show that the interpretations of *bank* are the result of contextual modulation of a general meaning. It may be concluded, therefore, that the different contexts are selecting discrete senses of *bank*. Another instance of incomplete contextual determination is to be observed with *dog*. Let us for the moment take it as established that *dog* has a general sense, denoting the whole species, irrespective of sex. In sentences such as 22, however, *dog* has a more specific meaning, and refers only to males:

22. John prefers bitches to dogs.

Now it might be argued that the resultant sense of *dog* here is caused by contextual modulation of the general sense: *dog* cannot in this context refer to females if logical consistency is to be preserved, which leaves only males as possible referents. Consider now, however, 23:

23. Incredibly, John prefers an aged, half-blind bitch to a dog,
as his canine companion.

If the interpretation of *dog* in this sentence were the result of contextual modulation of the general sense, it ought to include reference to, for instance, young females with good eyesight. But once again, it refers to male dogs only. This reading cannot be explained by contextual modulation, so it must be the result of selection from a set of discrete possibilities. In fact, the same is true of 22. That contextual modulation of the general sense of *dog* cannot explain the specific interpretation in 22 is shown by the lack of a parallel specific interpretation of *canine* (in its jocular use as a noun) when it is substituted for *dog*:

24. ? John prefers bitches to canines.

(We shall consider below why 24 should be anomalous.)

Some understanding of the way the semantic effects of selection may be independent of, and indeed may transcend, those properties of the

context which are responsible for the selection can be gleaned from the following analogy. Suppose that it is known that a certain event is to occur on a certain day, but may take place at only one of two possible times, namely, 12.00 noon or 12.00 midnight. If one were subsequently to receive a report that when the event occurred, the sun had set, one would be able to infer that it had taken place at exactly 12.00 midnight. The precision of this inference goes well beyond what is explicitly present in the report, which acts rather like a trigger setting off one of two pre-existing possibilities. In a similar manner, the context of *dog* in 22 and 23 acts like a trigger which activates one of a set of pre-existing bundles of semantic properties, each having a precision and richness not directly sanctioned by the context. In principle all ambiguous items should be capable of manifesting these characteristics.

Our second criterion for ambiguity is that separate senses should be **independently maximisable**. Under certain conditions, the application of certain terms must be maximised within the current universe of discourse, even at the expense of oddness. Consider 25 (which resembles 24):

25. ? Mary likes mares better than horses.

One might have thought that the context makes it clear that *horses* is to be interpreted as "stallions"; however, such an interpretation is not available for this type of sentence. The reason is that since mares have been mentioned, they fall within the current universe of discourse, and by the rule of maximisation (the details of which are not entirely clear) must be included in the reference of *horses*. This, of course, leads to logical inconsistency, and hence oddness. (Notice, however, that there is no anomaly if the reference of *horses* is explicitly restricted: *Mary prefers mares to horses which can sire foals* or *Mary prefers mares to these horses* uttered in a situation where only stallions are present.) On the other hand, 26, unlike 25, is perfectly normal:

26. John prefers bitches to dogs.

The general sense of *dog* would of course give rise to anomaly in 26, because of the rule of maximisation. The reason 26 is not odd is that *dog* has another sense, which even when maximised excludes bitches, and this is automatically selected by the context. By contrast, 27 selects the general reading of *dog* (the specific reading would be odd here, but not for reasons connected with maximisation):

27. Arthur breeds dogs.

Thus 26 and 27 taken together constitute strong evidence that *dog* is ambiguous.

The existence of two independent senses of *dog*, each independently maximisable, is responsible for the fact that A's question in 28, if the dog in question is female, can be truthfully answered either 'Yes' or 'No' (depending on which sense the respondent believes the questioner to be intending):

28. A: Is that a dog?
B: (i) Yes, it's a spaniel.
(ii) No, it's a bitch.

There is no parallel set of circumstances in which the question in 29 can be truthfully answered 'Yes' or 'No':

29. A: Is the subject of this poem a monarch?
B: (i) Yes, it is a queen.
(ii) ? No, it is a king.

Because there is only one sense of monarch, namely, the general one, and because its reference must be maximised, if the subject of the poem was a king or a queen, then 'Yes' is the only truthful answer. As with 28, situations can be imagined in which the questions in 30 and 31 can be truthfully answered either negatively or positively:

30. A: Has Charles changed his position?
B: (i) Yes, he's now sitting next to the chairman.
(ii) No, he still supports corporal punishment.
31. A: Did Arthur make it to the bank?
B: (i) Yes, he's a strong swimmer.
(ii) No, he was arrested as soon as he came out of the water.

The same should be true, in principle, of any truly ambiguous expression.¹² Ambiguity tests of the third kind utilise the fact that independent senses of a lexical form are antagonistic to one another; that is to say, they cannot be brought into play simultaneously without oddness. Contexts which do activate more than one sense at a time give rise to the variety of oddness we have labelled *zeugma*:

32. ? John and his driving licence expired last Thursday.

The simultaneous bringing into play of two senses can be effected either

by coordination, as in 32, where *John* and *his driving licence* select different senses of the verb *expire*, or by anaphora, as in 33:

33. ? John's driving licence expired last Thursday; so did John.

So did is an anaphoric verb phrase; that is to say, its referential properties operate not directly, but indirectly; through a previously mentioned verb phrase, in this case *expired last Thursday*, which must be re-applied, this time with *John* as subject. But since this demands a different sense from the one appropriate to its first occurrence, the result is zeugma. A general term cannot give rise to zeugma in this way:

34. My cousin, who is pregnant, was born on the same day as Arthur's, who is the father.

Arthur's refers anaphorically through *cousin*. The context makes it clear that the two cousins are of different sexes; however, the sentence is not zeugmatic, so we may conclude that *cousin* does not have two senses "male cousin" and "female cousin".

Antagonism of senses also lies behind the so-called **identity test** for ambiguity.¹³ In 35, each part of the sentence contains an occurrence, either direct, or indirect via anaphora, of the ambiguous adjective *light*, and can therefore in theory be interpreted in two ways:

35. Mary is wearing a light coat; so is Sue.

However, the whole sentence does not have four (i.e. 2×2) interpretations, but two only. This is because the same reading of *light* must be selected in each part: either both ladies are wearing "undark" coats, or both are wearing "unheavy" coats. What is termed the **crossed interpretation**, with each part of the sentence manifesting a different sense, is prohibited. This prohibition is not a mysterious property of the grammatical process of anaphora; it is simply a consequence of the fact that *light* resists, as it were, the simultaneous activation of more than one of its senses. General terms allow crossed interpretations:

36. Mary has adopted a child; so has Sue.

There are four possible distributions of sexes compatible with this sentence, since there is no requirement that the two children should be of the same sex.

3.5 Some difficult cases

In this section the operation of ambiguity tests will be illustrated

by applying them to a selection of difficult cases. The difficulties mostly concern tests based on the antagonism of sister-senses (i.e. senses associated with a single lexical form). It is not possible simply to dispense with such tests, because there are occasions, especially when dealing with highly context-bound readings which do not appear in ambiguous sentences, when they are the only practicable way of diagnosing ambiguity.

I The first example involves the unit-type ambiguity. This is quite easy to demonstrate by means of the *Yes/No*-test:

37. A: Is this the jacket you want?

B: (i) Yes. (it's the type I want)

(ii) No. (this particular one is shop-soiled)

But it is much more difficult to show antagonism: many contexts which might be expected to manifest it do not:

38. This is our best-selling jacket: do try it on.

Jacket in the first clause clearly must have a type reading – one cannot repeatedly sell the same individual jacket. One might have thought that only a particular unit of the type could be 'tried on', but that seems not to be the case. One must beware of drawing hasty conclusions in this area. As it happens, it is possible to find contexts which isolate the two readings, and when these are yoked together, zeugma results. Sentence 39 allows only the 'unit' reading for *skirt* (this seems to be a property of *belong*):

39. That skirt belongs to Mary.

Sentence 40 can only bear a type reading:

40. My sister has the skirt Sue is wearing now.

Try to link these two readings together anaphorically, and the antagonism becomes plain:

41. ? The skirt Sue is wearing belongs to Mary; my sister has it, too.

It not infrequently happens that ambiguous readings are related in such a way that in certain contexts one reading entails the other. Such cases are a common cause of apparent failure of the zeugma-test (often called the 'pun-test') or the identity test. The two readings of *dog* are a case in point.¹⁴ In 42, for example, it appears that a crossed interpretation is possible, in that Mary's dog could well be male, and Bill's female:

42. Mary bought a dog; so did Bill.

Does this contradict the evidence presented above that *dog* is ambiguous? The answer is that it does not. When *dog* occurs in a sentential context in which the specific interpretation entails the general interpretation, we cannot be sure which sense is operative when reference is made to a male *dog*: the two senses under these circumstances are effectively inseparable. Hence the normality of 42 when the dogs referred to are of opposite sexes cannot be used as evidence against the existence of two senses of *dog*, since it can be fully accounted for by claiming that only the general sense is operative. However, the situation is much clearer when *dog* occurs in a context where neither sense entails the other, as in 43:

43. Arthur wants to know if that is a dog; so does Mike.

A moment's thought will convince the reader that the crossed reading is prohibited here: this sentence cannot be used to describe a situation where Arthur knows that the animal in question is an alsatian, but is unsure of its sex, while Mike knows that it is female, but thinks it might be a wolf. The pun-test, too, demands non-entailing contexts:

- 44a. Dogs can become pregnant at 12 months. (general sense only)
- b. Dogs mature later than birches. (specific sense only)
- c. ? Dogs can become pregnant at 12 months, but mature later than birches.

III Entailment between readings also bedevils attempts to demonstrate antagonism between the "exactly" and "at least" interpretations of numerals and other expressions of quantity.¹⁵ The Yes/No-test suggests that this is a genuine ambiguity:

- 45. A: Have you got £10 in your wallet?
- B: (i) Yes. In fact, I've got £12.
- (ii) No, I've got £12.

However, *John has (exactly) £10* entails *John has (at least) £10*, which perhaps explains why 46 is not zeugmatic:

- 46. You need £100 in your account to qualify for free banking.
- Arthur has it, now that he has added £50 to the £50 that was already there.

The first mention of £100 clearly demands an "at least" interpretation; what Arthur has is "exactly" £100; one might therefore not expect the *it* of the second sentence to be able to refer anaphorically to £100 in the first sentence without antagonism. However, because of the entailment

referred to above, the original and anaphoric occurrences of £100 can both be given the "at least" interpretation, thus avoiding antagonism. It is possible to construct isolating contexts which reveal antagonism, but they are extremely cumbersome:

- 47a. John, with £11, and Bill, with £12, both have the £10 necessary to open a savings account. ("at least")
- b. Tom, too, now has £10, having spent £2 out of his original £12. ("exactly" reading forced by *now*)
- c. ? John, with £11, has the £10 necessary to open a savings account; Tom, too, now has it, having spent £2 out of his original £12.

IV The case of *door* is interesting (a group of related words such as *window*, *hatch*, *sky-light*, etc. behave similarly). Two senses of *door* may be observed in 48, which can be truthfully answered either 'Yes' or 'No' in the following situation: the door in question has a 'cat-flap', and is standing open; the cat goes through the cat-flap, but not through the doorway:

- 48. Did the cat go through the door?

Once again, difficulties arise with the antagonism criteria. It might be predicted, for instance, that 49 would be zeugmatic, since what is smashed (the door-panel) is different from what is bricked up (the doorway):¹⁶

- 49. The door was smashed in so often that it had to be bricked up.

But there is no anomaly of any kind. Again, it appears that contexts of a particular kind must be avoided if the test is to succeed. In this case it is the part-whole relationship which is to blame. For certain predicates, applicability to parts entails applicability to wholes corresponding to the parts. Thus, if I touch the *table-leg*, by doing so I necessarily touch the *table*; if the *tea-pot handle* is broken, so is the *tea-pot*, and so on. It seems likely that this entailment is interfering with antagonism in 49 – both events are interpreted as happening to the 'global door', of which the door-panel is a part. The remedy, as before, is to avoid such contexts, and to use, to isolate the senses, only those contexts in which part does not entail whole (or, better still, contexts where part entails not-whole). When this is done, the antagonism of the senses is easily seen:

- 50. ? We took the door off its hinges and then walked through it.

The moral to be drawn from these examples is that apparent compatibility of readings must not be too hastily accepted as proof of generality: each case must be examined carefully to determine whether there are special factors preventing the appearance of zeugma. It may be reasonably confidently assumed that the different criteria for ambiguity which have been described in fact are sensitive to the same underlying semantic property, and that in the absence of 'special factors' will provide identical diagnoses.

3.6 Non-lexical sources of ambiguity

It is important to realise that not all sentence ambiguity originates in lexical ambiguity; furthermore, our tests for ambiguity are not, in general, capable of discriminating between lexical and non-lexical varieties. Usually this is not a serious source of practical difficulty, since most cases are intuitively clear; but it is unfortunately not easy to formulate explicit criteria for recognising lexical ambiguity. We shall adopt a 'default' definition and characterise as lexical all ambiguities for which there is no convincing non-lexical explanation. This means that something at least must be said about alternative types of ambiguity, although a detailed treatment would be well beyond the scope of this book.

We can crudely classify the sorts of ambiguity found in sentences as follows:

1. Pure syntactic ambiguity:
old men and women
French silk underwear
2. Quasi-syntactic ambiguity:
The astronaut entered the atmosphere again
a red pencil
3. Lexico-syntactic ambiguity:
We saw her duck.
I saw the door open.
4. Pure lexical ambiguity:
He reached the bank
What is his position?

Types 3 and 4 are of direct relevance to us, and are discussed in some detail in the present chapter; types 1 and 2, on the other hand, are irrelevant, and we need to know how to exclude them.

By 'pure syntactic ambiguity' is meant ambiguity in which the variant readings of a sentence involve identical lexical units; the ambiguity is thus necessarily a matter merely of the way the elements are grouped together.

For instance, the meaning of *old men and women* differs according to whether *old* goes with *men* only:

(old men) and women

or with *men and women*:

old (men and women)

Likewise, *French silk underwear* may be underwear made of French silk ((*French silk underwear*) or French underwear made of silk (*French silk underwear*)). Such cases are characteristically very insensitive to the semantic properties of the constituent lexical items: *melodious trills and scales*; *porcelain egg container*. The so-called 'ambiguities of scope' can be included in this category; although they are often lexically restricted, they can be fully accounted for in terms of 'what goes with what'. Take, for example, sentence 51:

51. I don't like him.

Imnocuous though it may seem at first sight, this can be interpreted (at least in the written form) in two ways: either "I dislike him" (the most usual reading), or, in suitable contexts, "It's not true that I like him" (for instance, in *I don't dislike him, but I don't like him either*). There is no need to postulate different negative elements, or different meanings of *like*: it is enough to allow the negative element either to take the whole of the rest of the sentence as its scope (*Neg (I like him)*), in which case the meaning will be "It's not true that I like him," or the single element *like (I Neg-like him)*, in which case the meaning will be "I dislike him."¹⁷

'Quasi-syntactic' ambiguities require careful consideration because there may be a temptation to diagnose them as cases of lexical ambiguity. This is because there is no straightforward syntactic explanation of the ambiguity: not only are the lexical units identical for the two interpretations, but they are identically grouped, too. And yet this type of ambiguity bears a striking resemblance to the scope ambiguities described above. Consider the case of *The astronaut entered the atmosphere again*. The two meanings are (i) "the astronaut entered the atmosphere for (at least) the second time" and (ii) "the astronaut returned to the atmosphere (after what could have been his/her first trip into space)". This ambiguity can be accounted for without the need either for two different elements *enter*, or two different elements *again*, if we regard the meaning of *enter* as being constituted out of more elementary semantic entities which are related quasi-syntactically:

"enter" = [COME TO BE] [IN]

The two readings can then be represented as follows:

- (i) ([COME TO BE] [IN]) [AGAIN]
 (ii) [COME TO BE] ([IN] [AGAIN])

The availability of an explanation along these lines (however it might be formalised in relation to the syntax) renders a lexical solution unnecessary.¹⁸ Another example is a *red pencil*, which has the two readings (i) "a pencil painted red" and (ii) "a pencil which writes red". It may be thought that in reading (ii), *pencil* should be taken to refer only to the core of the pencil. This is not so, however: there is little doubt that in both interpretations *pencil* refers to the whole object (or at least potentially does so). Notice that *I have a red pencil and a blue one* has no crossed interpretation, which is what we expect from a genuine ambiguity. Yet *The red pencil is the chewed one* is quite normal on both readings, which would not be expected if on one of the readings *pencil* referred only to the core. It seems that the adjective *red* can apply either to the whole of the referent of the noun that it accompanies, or to a salient, or major functional, part of it. The same potential ambiguity is present in *a stainless steel hammer*, and even (although pragmatically less likely) *a felt pen*. It is not clear at present exactly what the rules are in such cases, nor whether the choices of readings are as clear-cut as they at first seem. What is clear, however, is that we are not dealing with lexical ambiguity.

3.7 Establishment of senses

A lexical form may well be associated with an unlimited number of possible senses, but these are not all of equal status. If we take seriously the notion of 'unlimited number', there must be, for any lexical form, potential senses which have never been realised in use: equally, every lexical form has at least one relatively well-utilised sense. We may thus envisage a gradient of what we shall term **establishment of senses**. (Individual speakers may, of course, differ markedly in respect of the degree of establishment of different senses, but a substantial measure of consensus may be assumed.)

The difference between established senses and potential senses is not merely one of frequency of use, although this is undoubtedly an important component of the difference: established senses are presumably represented differently in the mind's lexicon. It seems appropriate to distinguish two kinds of contextual selection, according to whether the selected sense is established or not. In the former case, where selection is from among pre-established senses, the context acts merely as a kind of filter: we shall

refer to this as **passive selection**. Where, on the other hand, the selected sense is not established, the context acts rather as a stimulus for a productive process, namely, the activation of a set of rules or principles which 'generate' the sense in question. The latter type of selection will be called **productive**. The difference between the two types of selection may be assumed to be of psycholinguistic importance.

There is a possible test for the establishment of a sense, which has consequences for the second family of tests for ambiguity described earlier. It appears that it is possible to assert one of the senses of a lexical form, using the bare form, while at the same time denying (explicitly or implicitly) another of the senses, only if the asserted sense is fully established. A few examples will make this clear. Take the case of *novel*, which can have the readings (i) "narrative text" or (ii) "physical object (embodying a narrative text)". The two readings may be observed in 52 and 53 respectively:

52. His new novel will be published next spring.
 53. Why is your desk always piled high with novels?
 Now consider 54 and 55:

54. I'm not interested in the cover design, or the binding – I'm interested in the novel.
 55. ? I'm not interested in the plot, or the characterisation, or anything of that nature – I'm interested in the novel.

Notice that 54 is more or less normal: the "physical object" reading is explicitly denied, and *novel* is consequently understood with the "text" interpretation. Sentence 55, on the other hand, is uninterpretable: since the "text" reading has been excluded, it appears that there is no other possible reading, so the sentence is anomalous. It would make sense if we were free to take *novel* to refer to the physical object; but in this sentence such an interpretation is not available. It seems reasonable to conclude that only the "text" reading is fully established. In the case of the numerals, it is the "exactly X" reading which is fully established according to this test. Thus the final *£10* in 56, if it carries the main sentence stress, can only mean "exactly *£10*":

56. A: I would earn at least *£10* an hour there.
 B: Well, here you'll earn *£10*.

However, the bare mention of *£10* cannot carry the "at least" interpretation in contrast to an explicitly expressed "exactly *£10*":

57. A: I would earn just £10 an hour there.
B: ? Well, here you will earn £10.

If the "at least" reading had been available, the sentence would not be odd. From this we may conclude that the "at least" sense of numerals is not established. In the case of the unit-type ambiguity, it is the unit readings which pass this test (cf. 58), while the 'type' readings fail (cf. 59):

58. I don't want that type of jacket, I want that jacket.
59. ? I don't mean that individual dog, I mean that dog.

In all the above cases, the lexical form in question has only one established sense. This, however, is not a rule: more than one sense may be established, as the normal interpretability of all the following examples shows:

- 60a. I'm not only interested in male dogs, I'm interested in dogs.
b. I'm not interested in all members of the canine race
irrespective of sex – I'm interested in dogs.
61a. I didn't put my money in the side of a river, I put it in the
bank.
b. I didn't moor the boat to a financial institution, I moored it
to the bank.
62a. Charles has moved to another seat in the conference hall, but
he has not changed his position.
b. Charles hasn't changed his mind on EEC membership, but
he has changed his position.

These examples point to a limitation on one of the tests for ambiguity elaborated earlier. It appears that certain ways of applying the criterion of independent maximisability are valid only for established senses. Sentences of the form of 26, for example, require established senses. Negative results in such cases must therefore be checked either against other criteria, or against other ways – such as the *Yes/No*-test – of applying the same criterion. (Positive results, of course, present no problems.)

The number of fully established senses is presumably finite at any one time (though it may differ for different members of the language community, and at different times for the same speaker). It might therefore be thought advantageous to limit the class of lexical units to these. However, although our attention will naturally be more strongly drawn to established senses, to limit the discussion in principle to these would lead to a distorted picture of word-meaning. This is because less-than-fully-established senses

are lexicologically almost indistinguishable from fully established ones, in that they enter largely the same range of syntagmatic and paradigmatic relations of meaning (the sentences cited above, of course, show that they are not absolutely identical). We shall therefore not limit our investigations in any principled way to established senses; whether a sense is established or not is, however, of significance for lexicography.

3.8 Sense-spectra

It has been argued up to now that although word-meaning is in a sense infinitely variable, nonetheless discrete units – 'atoms' or 'quanta' of sense – can be identified which at least in some respects are stable across contexts, and which are the appropriate basic units for lexical semantics. Certain aspects of word-meaning, however, are difficult to reconcile with this view: particularly awkward are what we shall term **'sense-spectra'**.

There are cases where variant readings of a single lexical form would seem to be more appropriately visualised as points on a continuum – a seamless fabric of meaning with no clear boundaries. This would not necessarily conflict with the picture of word-meaning developed so far if a single superordinate sense could be found which covered all the variants. However, there do appear to exist examples of gradual variation which cannot be made to share a superordinate; in such cases the absence of boundaries between senses is an embarrassment. The appearance which sense-spectra present can be compared with a so-called 'dialect continuum': speakers from village A can communicate with those from village B, who are able to converse with speakers from C; these, in turn, can communicate with speakers from village D. However, speakers from A cannot hold a conversation with speakers from D, and without the evidence of the intervening stages, one would be tempted to say that they spoke different languages. But it is impossible to say at what point along the continuum the change from one form of the language to another occurs, or to determine how many distinct forms there are. Another analogy is with the evolutionary biologist's notion of a 'ring-species': a population A, of some species, interbreeds with a neighbouring population B, B with C, C with D, and so on, round the world, until population X is reached, whose territory adjoins that of the original A. But A and X do not interbreed: they give every appearance of being distinct species. Again it is impossible to say where the change-over from one species to the next occurs, and how many species there are. The fact seems to be that in such cases it is inappropriate to think in terms of discrete variation. In the semantic analogues to these

continua, two readings which are close together on the continuum can be coordinated without zeugmatic incompatibility (this is the semantic parallel to mutual intelligibility and interbreeding), whereas readings which are far apart are incompatible. Examples of this are far from rare: on the contrary, this state of affairs would seem to be the norm, for example, for senses which have undergone 'metaphorical extension'.¹⁹

As an example of this sort of semantic continuum, which we shall call a **sense-spectrum**, consider the following use of *mouth*:

63. John keeps opening and shutting his mouth like a fish.
 64. This parasite attaches itself to the mouths of fishes, sea-squirts, etc.
 65. The mouth of the sea-squirt resembles that of a bottle.
 66. The mouth of the cave resembles that of a bottle.
 67. The mouth of the enormous cave was also that of the underground river.

Allowing for a degree of non-anomalous unusualness in the sentences (such sequences are, for various reasons, rather difficult to construct) it seems that we have got from *John's mouth to the mouth of the river* without encountering zeugmatic incompatibility. The normal conclusion from this would be that the readings of *mouth* in 63-67 were contextual modulations of a single superordinate sense. This is ruled out, however, not only by the difficulty of finding a paraphrase of the supposed superordinate sense, but also by the clearly zeugmatic nature of 68:

68. ? The poisoned chocolate entered the Contessa's mouth at the same instant that the yacht entered that of the river.

This is, of course, a simplified picture of a sense-spectrum: it should be thought of as having, at least potentially, many dimensions, and as continually growing, amoeba-like.²⁰

One of the points on the sense-spectrum presented above – and this is typical of the metaphorical variety – has a special status, which manifests itself in two principal ways. First, it is the only sense which can appear in a neutral, or minimal context, as in 69:

69. At school, we are doing a project on mouths.

It seems unlikely that 69 could be taken to include river mouths. All the other possibilities are highly context-bound, in that they can only appear in relatively explicit contexts – compare the a and b sentences in the following:

- 70a. ? The body was found near the mouth.
 b. the mouth of the cave.
 71a. ? This bird is often to be seen near mouths.
 b. the mouths of rivers.
 (cf. also: near estuaries.)
 72a. ? The candle was stuck in the mouth.
 b. the mouth of the bottle.

The independent sense is often also the 'literal' sense, in that it is the only one, or at any rate the most plausible one, from which all the others can be derived by metaphorical interpretation. (It may sometimes happen that of two senses, either one could plausibly be a metaphorical extension of the other, as with, for example, *expire* (driving licence, etc.) and *expire* (person).) In the case of *mouth*, if one knew what an animal's mouth was, and one were to hear, for the first time, a reference to *the mouth of a river*, I surmise that there would be little difficulty in construing the meaning; but suppose one were familiar only with *mouth* used to refer to the mouth of a river, and one heard a reference to *the horse's mouth*, it is by no means certain that one's attention would be directed to the appropriate end of the horse!

The proper descriptive treatment of sense-spectra, and points along them, is somewhat problematical. A full sense-spectrum is not a satisfactory lexical unit: it does not, for instance, enter into any recognised lexical relations. Individual points along a spectrum, on the other hand, seem at first sight to be insufficiently distinguished from one another. However, there are reasons for believing that these are the most appropriate lexicological units. Although when viewed as part of a spectrum their distinctness is questionable, they typically function in widely different semantic fields, and within these their discreteness and stability are not in question. Take the case of *mouth of river*: it participates in a significant number of meaning-relations:

mouth : source	(opposites)
mouth : river	(part-whole)
mouth : bed	(coordinate parts)
mouth : estuary	(superordinate-hyponym)

None of these relations are shared by, for instance, *mouth of bottle*. Furthermore, the sense of *mouth (of river)* is stable across a variety of contexts (i.e. subject only to modulation) provided that "of river" is understood.

So far, so good. But here we are faced with a dilemma. If we allow the existence of distinct sets of lexical relations to individuate senses along

a sense-spectrum, we are re-instating the indirect criteria dismissed earlier as being inadequate. If, on the other hand, we adopt a complex unit such as *mouth of river* as a basic lexical unit, this would be inconsistent with our earlier decision not to regard, for instance, *foot the bill* as a single unit. We shall adopt here the first of these solutions, as being the least objectionable of the two. That is to say, we shall recognise sense-units along a sense-spectrum – to be called **local senses** – by their participation in distinct lexical fields (here, to be understood merely as sets of lexical items interrelated by determinate meaning relations such as oppositeness, hyponymy, part-whole, etc.). This method of delimiting senses will be confined to sense-spectra.

The true extent of the phenomenon is not at present clear, but not all sense-spectra are of the metaphorical sort. It seems likely, for instance, that the senses of *handle* form a spectrum:

- handle of door
- of drawer
- of suitcase
- of umbrella
- of sword
- of knife
- of spoon

There is more than a suspicion of zeugmatic tension when the end-items are yoked together:

- ? He grasped the handle of the door in one hand, and that of the spoon in the other.

The different senses of *handle* can be delimited in the manner suggested above for *mouth*.

3.9 Syntactic delimitation

Lexicological units must not only be delimited paradigmatically, that is, within a constant syntactic frame: we want also to be able to say of two occurrences of a lexical form in different syntactic environments whether they are occurrences of the same lexical unit, or two different units. Consider the occurrences of *open* in the following:

- 73a. The open door.
- b. The door is open.
- c. The door won't open.
- d. John will open the door.

How many different items *open* are represented here? The sort of criteria which we used for paradigmatic delimitation are of no help here.

It would seem reasonable to adopt as a general principle that any two occurrences of a lexical form which represent two different grammatical elements should be regarded, *ipso facto*, as lexically distinct. However, there does not seem to exist an accepted notion of 'grammatically different element' which is sufficiently well-defined to carry the whole burden of distinguishing lexical units. Mere occurrence in syntactically different environments is not a sufficient criterion for the grammatical distinctness of two elements. For instance, the following two occurrences of *man* can be said to be in syntactically different environments:

- 74a. Arthur saw the man.
- b. The man's brother was here.

However, there are various reasons for saying that *man* is the same grammatical element in 74a and b. An important one is that the possible substitutes for *man* (preserving grammaticality, but not necessarily semantic normality) are virtually identical in the two positions. We might therefore demand difference of grammatical paradigm as a minimum requirement for distinctness. However, this is not sufficient, either, although it may well be necessary. Consider the following examples:

- 75a. The main customer
old
- b. The customer is asleep
*asleep
old
- 76a. Michael is eating his sandwiches.
preparing
- b. Michael is laughing
*laughing
eating
*preparing

It is extremely dubious, in spite of the differences in grammatical paradigm, whether anything would be gained by classifying the two occurrences of *old*, or those of *eat*, as grammatically, hence lexically, distinct. Other purely grammatical criteria may be suggested, but none seem capable of guaranteeing the desired results.

A more satisfactory way of delimiting lexical units is to look for grammatical

differences which correlate with differences of meaning. Take, for example, the occurrences of *open* cited above (73a-d). Grammatically distinctive traits can be found for each of these. Looking, for instance, at grammatically equivalent substitutions, *main* is possible only in a, *ajar* only in b, *disappear* only in c, and *hit* only in d. Other differences may be cited: only in c and d can *open* take -s as an affix, and only in a and b can *open* be modified by *wide*; c and d differ in that the noun phrases which form normal subjects of *open* in c are those which form normal objects of *open* in d (and similarly with odd subjects in c), so that, for example, the normality of *The book opened* is paralleled by that of *John opened the book*, and the oddness of ?*The page opened* by that of ?*John opened the page*. Most, but not all, of these grammatical differences are correlated with semantic differences. Taking the meaning of *open* in 73b as basic, we can paraphrase 73c (not exactly, but quite closely) as "the door came to be open", and 73d as "John caused the door to come to be open." In any sentence, the appropriate interpretation of *open* can be determined from its grammatical nature (i.e. whether it is adjective, transitive or intransitive verb, etc.). The fact that the occurrences of *open* in 73b, c and d exemplify a regular correlation between semantic and grammatical properties provides a justification for regarding them as lexically distinct. However, there is no similar way of differentiating 73a and b semantically, so, in spite of grammatical evidence of distinctness, they are to be considered lexically identical.

3.10 Lexemes

One of the most remarkable features of language is the fact that it makes infinite use of finite resources'. This dictum is more familiar in its application to grammar. But it is valid also for the lexical domain. We have already had glimpses of the indeterminate multiplicity of lexical senses: a lexicographer, however, needs a finitely enumerable set of lexical elements with which to work. The appropriate unit for this purpose is the **lexeme**: a dictionary contains (among other things) an alphabetical list of the lexemes of a language. We shall characterise a lexeme as a family of lexical units.

However, before outlining the principles governing the assignment of lexical units to lexemes, it is necessary to introduce a refinement into our conception of a lexical unit. We have so far assumed that it is a word form associated with a single sense, and that a difference of word form entails a difference of lexical unit. But this is not quite satisfactory. Strictly speaking, we would be obliged, on this view, to regard, for instance, *obey*,

obeys and *obeyed* as representing different lexical units. It would, however, be more advantageous for our purposes to be able to say that they were alternative manifestations of the same lexical unit *obey*. To characterise the form aspect of a lexical unit, therefore, we need to generalise across — or abstract from — a set of word forms. In order to characterise this more abstract notion of lexical unit more precisely, a distinction must be made between **inflectional** and **derivational** affixes. An affix is a grammatical element, belonging to a closed set, which can only function as a component of a word: *dis-*, *un-*, *-ment*, *-ise*, *-ed*, *-s* are all affixes. Each affix is obligatorily attached to a **stem**²¹ containing or consisting of an open set item: *dis-obey*, *un-popular*, *central-ise*, *dismount-ed*, *long-er*, etc. A stem may be simple (as *obey* in *dis-obey*), or complex (as *disobey* in *dis-obeyed*). Affixes are of two sorts — derivational and inflectional. Derivational affixes produce new lexical units: *true: untrue*, *kind: kindness*, *help: helpful*, *lion: lioness*, etc. They play no direct role in the syntax of a sentence, and can be recognised by the fact that words containing them (**derived words**) can typically be replaced in any sentence, without syntactic change, by a word which does not contain the affix:²²

Her kindness (voice) was overwhelming.
I found them extremely helpful (stupid).

Typically, derived words are listed as separate items in a dictionary. Inflectional affixes, on the other hand, do not produce new lexical units: *book: books*, *obey: obeyed*, *long: longer*. In principle for any word bearing an inflectional affix, it is possible to find contexts where all possible substitutes must contain either the same affix, or one belonging to the same closed set: consider the possible substitutes for *walked* in *Cedric walked home, longer in Mine is longer than yours* or *books in those books*.

We can now re-define a lexical unit. First, we may call the abstract unit of form which is realised in actual sentences as the appropriate member of a set of word forms differing only in respect of inflections a **lexical form**; and we can extend the notion of lexical form to cover an abstraction from the variously inflected manifestations of an idiom or dead metaphor. A lexical unit is then the union of a lexical form and a single sense. Let us now return to the question of assigning lexical units to lexemes.

For lexical units with identical grammatical properties, two alternative criteria for membership of the same lexeme will be proposed. The first is the most important. It is that two lexical units will be assigned to the same lexeme if there exists a lexical rule which permits the prediction of the existence of the sense of one of them from the existence of the

sense of the other. The existence of a rule presupposes that senses associated with more than one lexical form fall within its scope (otherwise there would be no rule). Hence, we shall accept as evidence of the presence of a rule a recurrent semantic contrast between senses, that is to say, a contrast which holds between senses associated with at least two different lexical forms. On this basis, the unit and type readings of *jacket* in *I like this jacket* belong to the same lexeme, because the same contrast recurs with *skirt*, *dress*, *coat*, *hat*, etc. (We shall not concern ourselves here with the exact formulation of the regularity: we shall merely note the evidence of its presence.) Similarly, the two lexical units represented by *brilliant* in *John is brilliant* and *This is a brilliant book* are to be assigned to the same lexeme, the evidence being the recurrence of the relation with *confused*, *angry*, *bitter*, etc. A parallel situation exists with *sad* in *John is sad* and *This is a sad poem*; this relationship, too, is recurrent (cf. *light-hearted*), but is different from that observed in the case of *brilliant*. A brilliant book is (roughly) the expression of a brilliant person, but a sad poem is rather one which induces sadness in the reader. Consider, too, the two readings of *flatten out* which occur in 77 and 78:

77. The surface of the mixture began to flatten out.
78. After Kendal, the countryside begins to flatten out.

The same difference of sense recurs in the following:

79. The soil began to dry up.
80. Once you leave the Bekaa Valley, the countryside begins to dry up.

Examples such as these can be multiplied indefinitely.²³

It is perhaps worth noting briefly at this point a special type of recurrent semantic relationship between lexical units sharing a lexical form, which is of particular significance in lexical semantics (it is discussed in greater detail in connection with markedness and neutralisation in chapter 11). The two senses²⁴ carried by the lexical form *dog* in 81a and b, and the two senses of *lion* in 82a and b, and of *heavy* in 84a and b stand in a relation of this type:

- 81a. Dogs, both male and female, make excellent pets.
b. Dogs are more aggressive than bitches.
82a. Lions breed well in captivity.
b. When fully grown, a lion is bigger than a lioness.

The two senses of *long* in 83a and b, and of *heavy* in 84a and b stand in a slightly different relation, but one of the same type:

- 83a. How long is it?
b. How long it is!
84a. How heavy is it?
b. How heavy it is!

The alternative criterion for assigning lexical units to a single lexeme is that their senses should be local senses belonging to the same sense-spectrum. Thus all the senses of *mouth* discussed earlier will represent lexical units belonging to a single lexeme. This criterion is quite strict, and does not allow the grouping together of all senses normally considered to be metaphorically related. For instance, there is no spectrum connecting the two senses of *expire*, so their lexical units would not be assigned to the same lexeme. The same is true of the readings of *position* that we have examined in connection with ambiguity. This differs from normal lexicographic practice, which is to group all metaphorically related senses together.

Among the lexical units which go to make up a lexeme it is possible to distinguish some that are more basic, or central, and others that are less so. It is clear that established units (i.e. those with established senses) are more central than unestablished ones: an ideal dictionary would be expected to define all the established senses within each lexeme. But even among established units we can distinguish grades of centrality. Most basic of all are lexical units which become operative in minimal, or neutral, contexts. These may be termed the **primary lexical units** of a lexeme — a category that would include, for instance, *dog* ("species"), *heavy* ("weight"), *novel* ("text"), etc. Some lexical units, even though established, are selected only in specific restricted contexts, or in contexts where the primary units would lead to abnormality. This is true of *dog* ("male"), *heavy* ("copious consumption"), etc. Such units may be termed **secondary** (the primary/secondary distinction here is not, of course, a strict dichotomy — the accessibility, or ease of activation, of lexical units may be assumed to vary continuously). There remain the unestablished units, generally indeterminate in number, and varying in the degree of contextual pressure required to activate them. Probably some degree of oddness is an inescapable penalty for calling an unestablished unit into service; this abnormality may be very slight, as in *A large novel fell on my head*, or it may be considerable, as in *I received a lot of kindness from him — would you like to try a bottle?*

The principle of recurrent relationships can also serve for the association of grammatically different lexical units. In such cases, the recurrent relationship must be simultaneously grammatical and semantic if the units

are to be assigned to the same lexeme. The following are examples of such recurrence:

- 85a. John moved the rock / The rock moved.
- b. John turned the key / The key turned.
- 86a. Have some apple / Have an apple.
- b. Have some potato / Have a potato.
- 87a. Put them in a can / Can them.
- b. Put them in a box / Box them.

Notice, however, that although the following exhibit a syntactic parallel with the cases cited above, the semantic relationship is not maintained, so the lexical units must be assigned to different lexemes:

- 88. Get him into a corner / Corner him.
- 89. Put his name in a book / Book him.

Again, this is not in accordance with normal lexicographic practice, which is, first, to regard differences of major syntactic category (e.g. noun, verb, adjective) as justifying a separate main entry, irrespective of the presence or absence of recurrent relationships. In respect of minor syntactic differences (e.g. transitive *v.* intransitive verbs; mass *v.* count nouns, etc.) dictionary makers are generally somewhat inconsistent. To summarise: a lexeme is a family of lexical units; a lexical unit is the union of a single sense with a lexical form; a lexical form is an abstraction from a set of word forms (or alternatively – it is a family of word forms) which differ only in respect of inflections.

It is commonplace to describe a lexeme which has a number of senses as **polysemous** (or as manifesting the property of **polysemy**), and a lexical form which realises lexical units belonging to more than one lexeme as **homonymous**. These terms, especially *polysemous* and *polysemy*, although innocuous if used circumspectly, are not entirely ideal for our purposes, because they carry with them a view of lexical meaning in which there is a tendency to regard the lexeme as the primary semantic unit, and the different lexical units as 'merely variants'. Our approach, however, focusses on the individual lexical unit as the primary operational semantic unit, and consigns the lexeme to a secondary position.

Notes

Linguists who have worked in lexical semantics can be broadly divided into two categories: on the one hand, there are those who believe that a word form is associated with a number (perhaps finite, perhaps not) of discrete senses; and on the other, there are those who believe

that the discreteness of lexical senses is illusory. Advocates of a formal-theoretical approach to meaning not unnaturally favour the first alternative – for an exposition of this view, see Kempson (1977: 79–83). Protagonists of the other view include Matthews (1979: 67–75) and Moore and Carling (1982: ch. 5). Lyons, too, seems inclined to this position (1977: 550–69). I find myself highly sympathetic to the arguments of both sides; in this chapter I present what is in some respects a compromise view – I try to have my cake and eat it. (The views expressed here differ in some respects from those presented in Cruse (1982).)

3.1

1. Kempson (1977: 82–3) uses *lexeme* to mean something very close to our *lexical unit*. For me, a lexeme is a family of lexical units. I agree with Kempson in giving primacy to the lexical unit; but I agree with, for instance, Palmer (1976: 65–71), Lyons (1977: 550–69) and Cowie (1982) in assigning multiple semantic roles to a lexeme.
2. For the notion of 'creativity' in syntax, see Chomsky (1965: 3–9). To be able to produce indefinitely many sentences from a finite set of elements and rules, at least some of the rules must be **recursive** – that is, able to apply repeatedly (see Lyons 1968: 221–2). It is not unlikely that some of the sense-creating rules are also recursive.
3. See 3.10 for some examples.

3.2

4. Kempson (1977: 128–34) uses *vagueness* to refer to what we call *generality*. We shall use *vague* in more or less its everyday sense in opposition to *well-defined*. For us, generality and vagueness can vary independently. For instance, *vertebrate* is more general than *animal* (in its everyday sense) since birds and fish are vertebrates; but it is less vague – it is easier to specify qualifying characteristics for *vertebrate* than for *animal*. (It is characteristic of scientific terms to be relatively well-defined.) For a fuller discussion of vagueness, see Alston (1964: ch. 5).
5. This preliminary account will exaggerate the sharpness of the distinction between these two. See, however, section 8 of this chapter.
6. The mechanism underlying this change of status is discussed further in 4.12 and 12.2.
7. The phenomenon of linkage is one reason for treating the principle of compositionality with the greatest circumspection (indispensable though it is to any semantics – formal or informal).
8. The reader is reminded that, strictly, *car* in the sentence *The car needs washing* 'refers' only if the sentence is uttered in an appropriate situation.
9. These assumptions form part of the 'Cooperative Principle' governing conversational exchanges suggested in Grice (1975). See also Wilson and Sperber (1981), Leech (1983: 79–103), Levinson (1983: 97–106).

3.3

- 3.3 Criteria for lexical ambiguity of the sort which are here labelled 'indirect' are of considerable antiquity. Ross (1981: 40–7) attributes a number of them to Aristotle. See also Cruse (1982). I am informed by N. E. Collinge (private communication) that they occur even earlier, in Plato. For a modern example of the use of such criteria see Cowie (1982).
10. For parody see Ross (1981: 136–41).
11. *Fat* and *thick* have different collocational restrictions and these are not shared by *thin*. (See 12.2 for discussion of collocational restrictions.)

12. This criterion may be more or less equivalently (but more precisely) expressed as follows:

For any sentence form containing an ambiguous word form, there should exist, in principle, situations in which the sentence form can be properly used to express two distinct propositions, which are identical except for differences consequent on the choice of sense associated with the ambiguous word form, and which have opposite truth values.

- Both Lyons (1977: 404) and Kempson (1977: 128-9) deny – wrongly, in my opinion – that a successful test for ambiguity can be constructed along these lines. Kempson's argument runs roughly as follows. Suppose two linguists are in dispute as to whether *John killed Bill* is ambiguous between an intentional and an unintentional interpretation (intuitively, one is free to interpret it either way), and they decide to use the criterion of different truth values to settle the matter. Imagine, now, a situation in which John kills Bill unintentionally. Linguist A, who does not believe the sentence to be ambiguous, says that it is true relative to the situation described; linguist B, who believes *John killed Bill* to be ambiguous, says that it is true on the "unintentional" reading and false on the "intentional" reading. And they will have got no further forward – the test has resolved nothing. However, what this line of argument ignores is that we cannot properly use a particular sentence form to express whatever proposition comes into our heads. Linguist B is correct in saying that the proposition "John killed Bill intentionally" is false relative to the situation described, and the proposition "John killed Bill unintentionally" is true. He is wrong, however, in his implicit assumption that *John killed Bill* is a proper linguistic vehicle to express the proposition "John killed Bill intentionally" (or "... unintentionally"). The proper expression of this proposition would implicitly deny the proposition "John killed Bill unintentionally" – but there is no way that *John killed Bill* could be used to deny this proposition, whether implicitly or explicitly.
13. See Zwicky and Sadock (1975).

3-5

14. In Zwicky and Sadock (1975: 14) and Kempson (1977: 136) the possibility of using the identity test in such circumstances is denied. (Kempson, however, no longer subscribes to this view (private communication).)
15. See Kempson and Cornack (1981). Notice that numerals can also have an "at most" reading, as in *Can you run 100 metres in 10 seconds?* and *I'm aiming at 10 stones by Easter* (said by a swimmer). Yet another possibility is the "round number" interpretation, which is the most likely in, for instance, *I'll see you in 10 minutes* (see Wachtel (1980) and Charnell (1980)).
16. The normality of this sentence is cited by Nunberg (1979: 150) as evidence that *door* is not ambiguous in this way.

3-6

17. For fuller discussion of syntactic ambiguity see Kooij (1971) and Zwicky and Sadock (1975).
18. Not all conceivable differences of interpretation attributable to variation of scope represent true ambiguities. I agree with Kempson's arguments (1977: 132-5) that *It wasn't a woman that came to the door* is not ambiguous. (If it was a girl who came to the door, it might be argued that only the trait

"adult" is being negated, but not "human" or "female", whereas if it was a man, "female" is negated, but not "human" or "adult".) This sentence does not pass the Yes/No-test. But I disagree with her conclusions regarding *John almost killed the hostages* (Kempson (1977: 132)) and sentences exhibiting internal and external negation – like our example 51 – (Kempson (1977: 148-154)), both of which I believe represent true ambiguity.

- 3-8 The existence of semi-distinct local senses on sense-spectra blurs to some extent the distinction that we have up to now sharply maintained between ambiguity and generality, selection and modulation. It now seems probable that we are, in fact, dealing with yet another continuum. This does not, however, invalidate the original distinction: there are still innumerable clear instances of both ambiguity and generality.

19. It may well be that the meaning of every lexical unit should be regarded as at least potentially a sense-spectrum.
20. See also the discussion of the meaning of *pin* in Matthews (1979: 71-2). This looks like another typical example of a spectrum.

- 3-10 The definition of *lexeme* adopted here differs both from that of Kempson and from that of Lyons (which represent the main current alternatives), occupying, in a sense, an intermediate position. For Kempson (1977: 79-83), every distinguished sense represents a different lexeme, and she sees no theoretical justification for groupings of senses. Lyons adopts what I take to be a more traditional approach (1977: ch. 9). If I understand him correctly, for him each of the items which stand in a relation such as antonymy is a 'lexeme-in-a-particular-sense'. He thus regards what we call the lexeme as the basic lexical item. Our approach centres on a single-sense (univocal) unit (in this respect, therefore, agreeing with Kempson), but at the same time recognises groupings based on relatedness of sense (in this respect agreeing with Lyons).

21. The term *stem* is sometimes used in a narrower sense, to designate what an inflectional affix is attached to, *base* being used to refer to what a derivational affix is attached to. Our usage follows Allerton (1979: ch. 10).
22. This is the case in English, at any rate. The reader should be warned that this is not a comprehensive account of the differences between inflection and derivation. For a fuller discussion, see Matthews (1974: ch. 3). (Matthews speaks of 'lexical' – rather than 'derivational' – morphology.)
23. Cf. Leech's 'rules of semantic transfer' (1974: 216-17).
24. Lexical units contract semantic relations with other units by virtue of their senses. There is therefore no difference between saying that a certain semantic relation holds between two lexical units, and saying that it holds between the senses of those units.