CHAPTER 12

Syntagmatic relations

CHAPTER 12

Syntagmatic relations

12.1 Normal and abnormal co-occurrence

It is an obvious fact that some combinations of words 'go together' naturally, and it is easy to imagine a situation in which they could function as part of a discourse. Other sets of words do not go together in this way: it is impossible, or at least very difficult, to imagine a situation in which they could be used (although we must not underestimate the flexibility and ingenuity of the human mind in this respect). This chapter is about the semantic relations between lexical units in the same discourse, string, sentence, or other syntactic structure, which govern their well-formedness. (There are, of course, important relations between larger discourse elements such as clauses, sentences, and larger units which are important for discourse cohesion and coherence. Here, however, we are concerned only with the lexical level.) All meanings co-present in a discourse affect one another to some degree and in one way or another. The interactions are complex and not yet fully understood; here only a sketchy outline can be offered. Before any details can be examined, it is necessary to make a distinction between two types of interaction between meaningful elements in a discourse. We shall distinguish the two types by the terms **discourse interaction** and **syntagmatic interaction.** We shall be eventually concerned mainly with the latter type.

Consider the following sentence:

(i) John and Mary will be joined in holy matrimony next week: who's going to get the spuds?

There are two sorts of oddness here. The first is the register clash between *holy matrimony* and *spuds.* This can easily be cured:

(2) John and Mary will be joined in holy matrimony next week: who is going to get the potatoes?

But we are still left with the difficulty of finding the relevance of potatoes to

John and Mary's marriage. (There would be no problem if *potatoes* was replaced by *confetti,* or even, *rice.)* These are both aspects of discourse interaction, as in each case the clash, as we may call it, is not between one item and its most intimate syntactic neighbour. The register clash is relatively superficial. Certain lexical items—they may have any syntactic function—serve as markers of degree of formality. Obviously contradictory markers are going to clash. The irreconcilability of *marriage* and *potatoes* has a much deeper origin, to do with the construction of plausible scenarios involving the two concepts, and drawing on cultural knowledge, which we cannot go into here, but is not obviously syntactically governed.

Another distinction needs to be made (it has already been adumbrated above) before the discussion can be advanced. There are two potential focuses of interest in studying syntagmatic semantic relations: one is whether, or to what extent, a particular combination makes sense, the other is whether, or to what extent, a combination is normal or abnormal. Although these two characteristics often coincide, they are by no means the same thing. For instance, *My geraniums kicked the bucket in the hot weather* is perfectly understandable, but is none the less somewhat odd; conversely, a difficult article on a topic in, say, formal semantics, may have no odd sentences in it. In this chapter we shall be concentrating mainly on whether syntagmatic combinations are normal or abnormal; this is both easier to control, and also more revealing of a wider range of facts.

It is undeniable that the normality of a particular string of words (even one which is close knit syntactically) can be affected by the wider context in which they are set. This means that we must be careful what we mean when we say that a particular string is abnormal. Take a case like *heavy on air.* This might strike the ear as odd, if no context is given. But suppose the conversation is about space travel and the need to develop ways of recycling vital materials like water and air. In this context it is not difficult to make sense of a statement to the effect that a particular device is *heavy on air.* The reason the original presentation was odd was that the default readings of the constituent items do not go together; the effect of the context was to enable a relevant selection of interpretations to be made. Similarly, Chomsky's *colourless green ideas* might not be so anomalous if used to describe a boring lecture on environmental issues. The moral of this is that we are not concerned with strings of words, but with strings of readings. Very often, a potential anomaly is a clue to the fact that either a different reading of some item in the string must be selected, or a new reading must be created.

In some cases, oddness can be alleviated by contextual manipulation even when there is no change of reading. This is particularly true of zeugma: it is possible to create a zeugma-like effect which has a purely discourse origin. However, this can be made to disappear with contextual elaboration. There are, on the other hand, cases where discourse manipulation does not serve to remove the zeugma, which is thus shown to be a lexical effect. What

we are interested in here is oddness which cannot be conjured away in this fashion.

Let us now look at different types of inherent combinatory abnormality.

12.2 Types of abnormality

Two basic types of abnormality resulting from the combination of two senses can be distinguished. The first is where meanings simply do not 'go together'; the second is when one meaning adds nothing new to another one with which it is combined and thus appears unnecessary, or redundant. We shall call these **clash** and **pleonasm,** respectively.

12.2.1 Semantic clash

The sorts of clash we are interested in here are those which resist contextual manipulation and can reasonably be considered to be lexical in nature. It is a feature of units of meaning that they impose semantic conditions of some sort on their syntagmatic partners: if these conditions are satisfied, the result is semantically well formed, and the combination is readily interpretable; if the conditions are not satisfied, some sort of clash results, which may trigger off a semantic transformation of some kind, which produces a reading that ***does*** satisfy the conditions. (For this reason, virtually no combination of ***words*** can be ruled out as anomalous.) We shall call the conditions **co-occurrence preferences** (rather than, as they are often designated, **co-occurrence restrictions,** which suggests a more yes/no, law-like condition than we actually find); they can also be thought of as presuppositions of the unit which imposes the conditions. Clashes come in varying degrees of severity. Presumably this property varies continuously, but as a first approximation, some distinctions can be made.

The first distinction is between clashes which result from the non-satisfaction of **collocational preferences,** and those which result from the non-satisfaction of **selectional preferences.** This latter distinction—between collocational and selectional preferences—depends on whether the preferences in question are an inherent consequence of propositional content or not. Take the case of *My geraniums have kicked the bucket.* There is here a semantic clash between *geraniums* and *kicked the bucket',* for full normality, *kick the bucket* requires a human subject. But the propositional content of *kick the bucket* is the same as that of *die:* it would not be honest to answer the question *Did my geraniums kick the bucket while I was away?* in the negative, if the geraniums in question had died, on the grounds that only humans can kick the bucket. The point is that kicking the bucket is not a special way of dying that only humans can suffer; it is more correct to say that the expression *kick the bucket* can only be used without oddness to refer to dying if certain contextual

conditions are satisfied, one of them being that the 'patient' should be human, another being that the situation should be informal. The conditions (preferences) do not arise ineluctably from the propositional meaning, but are, as it were, tagged on independently and somewhat arbitrarily. Contrast this case with the oddness of *My letter to Mary kicked the bucket.* Here the clash is not just, or even principally, between *letter* and the 'tagged on' meaning present in *kicked the bucket,* since the oddness is not significantly improved by putting *died* in place of *kicked the bucket.* There is a much more radical clash between the propositional meaning of *kick the bucket* and *my letter,* in that the concept of dying is only applicable to things/entities that at some time were alive. "Living subject" can be thought of as a logical presupposition of the default meaning of *die;* "human subject" is merely a stylistic presupposition of *kick the bucket.* If a collocational preference is contravened, we may say that **inappropriateness** results: inappropriateness is then the lowest degree of clash.

If what is here called a selectional preference is contravened, the clash is more serious. Two degrees of clash can just about be distinguished here. Consider *The cat barked,* or *a tiny giant. Bark* means "to make a noise" and is characteristic of dogs. But notice the difference between this case and *kick the bucket* in relation to humans. Whereas humans do not have a special way of dying (at least, this is not what *kick the bucket* denotes), dogs *do* have a special way of making a noise. So *bark* is not adequately glossed as "make a noise" (applied to dogs): it must be "make the characteristically canine noise". And it would not be misleading to answer the question *Did I hear the cat bark?* in the negative, if the cat had, in fact, miaowed (or, indeed, if it had been the dog which had made the noise). On the other hand, *bark* and *miaow* are in a sense the same kind of thing, both animal noises, so the clash is at a sort of intermediate level. In Cruse (1986) this was called **paradox.** Paradox is also involved when the 'wrong' value on a dimension is indicated: *It's too small to fit into this box, Rain falls upwards, usually, If you walk any faster, you'll be standing still.* Paradoxes are typically 'correctable'.

The most serious degree of clash is **incongruity.** This is when the ontological discrepancy is so large that no sense can be extracted at all, without radical reinterpretation. Since there is not even an inkling of sense, in the worst cases, there is no feeling that the utterance could be corrected. Examples are:

purple gestures of rat milk
the sky's nipple is a dictionary
crystalline miasmas of safety-pins
in phonemic toe-buckets

This is reminiscent of the worst sort of avant-garde poetry. A way of firming up these distinctions will be offered below, but it must be re-emphasized that degree of clash varies continuously, and the divisions are only first approximations.

12.2.2 Pleonasm

A pleonastic relation between two elements occurs when one of them seems redundant, and appears not to add any semantic information not already given by the other element. So, for instance:

(3) John kicked the ball with his foot.

Here *with his foot* adds nothing, since we know from *kick* what the instrument of striking was. Pleonasm can be avoided either by omitting *with his foot:*

(4) John kicked the ball.

or by replacing *kick* with *strike:*

(5) John struck the ball with his foot.

Notice, however, that (6) is not pleonastic:

(6) John kicked the ball with his left foot.

This is because the phrase *with his left foot* now contains new information: the repetition involved in *foot* is unavoidable as otherwise *left* could not be incorporated. Similarly (7) is pleonastic, because *male* gives no information that is not already conveyed by *uncle:*

(7) One of my male uncles told me.

On the other hand, *my gay uncle* is not pleonastic, although *gay* (nowadays) incorporates the notion "male", since *gay* also brings new information not present in *uncle.*

It is important to realize that repetition does not automatically bring about pleonasm. In some cases it is required by the grammar. For instance, in the phrase *two books,* one might argue that plurality is signalled twice, once by the numeral *two,* and then by the -5 of *books.* In some languages, for instance, Turkish, although a plural affix exists, the noun would have no plural marker in such circumstances: *kitaplar* ("books", *-lar* is the plural affix); *iki kitap* ("two books"). In some cases, the repeated item simply applies twice, sometimes with dramatic effects:

(8) I don't not want it.

Here the negative acts on itself, cancelling itself out. (In many languages there is obligatory duplication of negative marking, without the above semantic effect.) In some cases, repetition has an intensifying, rather than a pleonastic effect:

(9) That is very, very good.
(10) Mary rushed quickly to the window.

Notice that the idea of "quickly" is part of the meaning of *rush,* which is why we get a paradox if we qualify an act of rushing with the opposite term:

(n) ?Mary rushed slowly to the window.

Sometimes the interpretation is not clear, as in *Will you repeat it again, please,* which some speakers will interpret simply as an intensification, while others require a previous repetition for well-formedness. The underlying rules are not clear, but it seems that repetition causes intensification most frequently when a graded property is involved.

12.3 Syntagmatic sense relations

If we try to set up syntagmatic sense relations on the pattern of paradigmatic relations we find right at the outset that there are certain differences. The main one is that there are no relations of a syntagmatic nature that have the generality and context independence of paradigmatic relations such as hyponymy and meronymy. All relations are tied to particular grammatical constructions, or at least to families of constructions. To take a simple example, the following exemplifies a clash between *chair* and *saw.*

(12) The chair saw John.

But these two words do not necessarily clash:

(13) John saw the chair.

The clash only occurs when the words are in a particular grammatical relationship. Bearing this fact in mind, we can set up three basic relations, according to whether the words in question go together normally, clash, or yield pleonasm:

philonyms: go together normally
SAW the CHAIR
xenonyms: clash
HEARTFELT INSOMNIA
tautonyms: produce pleonasm
an ACADEMIC UNIVERSITY

Remember that in each case the grammatical relation between the terms must be specified, and that we are assuming the combinations to be fully grammatical. The grammatical relations can be specified in a semantically neutral way: for instance, *chair* and *see* are subject-verb xenonyms, and *heartfelt* and *thanks* are modifier-head philonyms; or they can be specified in a semantically more concrete way: for instance, *man* and *see* are experiencer-verb philonyms, *snap* and *pleasure* are verb-patient xenonyms.

12.4 The directionality of syntagmatic constraints

Constraints on co-occurrence between lexical items usually have directional properties. Two aspects of this are of particular interest. The first concerns which item does the selecting (the **selector),** and which gets selected (the **selectee).** It is necessary to separate two notions of selection here. If we are thinking of the selection from a set of polysemous or homonymous readings, then in a sense the process is obviously at least potentially bidirectional and there is no clear distinction between selector and selectee. In the combination *a hard match,* for example, *hard* rules out the reading "device for producing a flame" for *match,* and *match* rules out the reading "not soft" for *hard,* and we are left with the interpretation "difficult contest". Here we have a combination of two words, each with more than one reading, but there is only one philonymous combination of readings, and this emerges as the preferred interpretation. However, if we look closely at the relations between the meanings of items in a grammatical construction, we usually find another species of directionality, in that it is much easier to specify the restrictions imposed by one of the items than the other. Suppose we set ourselves the task of specifying the semantic nature of the adjectives which form philonymous modifier-head pairs with a noun such as *match* ("contest"). Think of the range of possibilities:

(14) home, ill-tempered, exciting, hard-fought, postponed, three-day, all-ticket, important, decisive, qualifying

There is no cover term, or superordinate notion which encompasses all these, even approximately. The only thing they have in common is that they go normally with *match.* But look now at *hard* ("difficult") and perform the same exercise:

(15) game, exercise, problem, journey, climb,job, crossword, exam

In this case we can roughly define the qualifying head nouns as falling under the general heading of "human activity requiring effort". The same can be done with the other readings of both *hard* and *match',* in each case, the philonym partners of the *hard* readings can be given a general specification, but those of the *match* readings cannot. In fact it becomes clear that the mechanism of selection for *match* readings is as follows: *match* readings select those adjectives whose co-occurrence preferences they satisfy. In other words, the apparent bidirectional selection has a unidirectional basis.

The direction in which selection operates, is correlated with grammar. The relevant generalization is that adjectives select their head nouns and verbs select their complements; nouns, in general, are always selectees. This can be made into a more satisfying generalization in logical terms: predicates select, and arguments are selected. Why this should be so is an interesting question.

Perhaps it is something to do with the fact that arguments are so intimately connected via reference to things in the outside, non-linguistic world, whose nature is not limited to a finite set of properties. A set of properties may serve to identify a referent as belonging to this or that class, but the referent itself goes beyond those properties. For instance, *teacher* represents a set of properties, knowledge of which enables us to distinguish teachers from non-teachers in the world. But once we have identified an extralinguistic teacher, we can predicate things of them which have nothing to do with the identifying properties: *sad, angry, tall, rich,* etc. Predicates, on the other hand, inhabit a different, conceptual world, whose denizens ARE mutually limiting.

The second aspect of directionality concerns the phenomenon of pleonasm. Generally speaking, if a combination of words is to be normal (i.e. non-pleonastic), the combination must yield more information (in a broad sense) than either of the combined items on its own. This must be pretty obvious. What is slightly less obvious is that the burden of providing extra information falls asymmetrically on the combined items. The categories used above, of predicate and argument, are of no help in formulating a regularity here. What we need instead are the categories of **(semantic) head** and **(semantic) dependant.**

Roughly speaking, the semantic head of a combination (construction) is the element which governs the semantic relations of the combination, viewed as a unit, with other elements or combinations. Take the case of an adjective-noun combination: this combination may in turn combine with a verb, but it is only the semantic properties of the noun which determine whether the combination is normal or philonymic. Take the combination *The small table sneezed,* which we can all agree is odd. Suppose we hold *sneezed* constant and ask ourselves what is the minimal change which will restore normality. The answer is that we must change *table {The small boy sneezed)',* no fiddling about with the adjective will produce any effect. Of course, *small* semantically interacts with *table {*the small phoneme (meaning),* but once the combination is effected, *small* has no further combinatory role to play. Similar arguments show that it is the verb which governs the combinatorial properties of a verb phrase.

Now that we have a notion of semantic head and its dependants, we are in a position to state a generalization regarding pleonasm: it is the duty of a non-head to bring information not available in the head; the head is under no such compulsion. This conforms with the observation of pleonasm in:

(16) a female aunt
a new innovation
Please repeat it for me again.
He kicked it with his foot.
She chewed it in her mouth.
I heard it by listening.
etc.

(The reader may consider me pedantic on some of these examples.)

12.5 Syntagmatic and paradigmatic relations

There are certain systematic connections between syntagmatic and paradigmatic sense relations which are worth signalling.

12.5.1 Pleonasm

In cases of pleonasm, the oddness can in general be 'cured' by substituting one of the tautonyms by a hyponym or hyponymous expression, or the other by a superordinate. This gives us a way of identifying the head and dependent elements: the head is the item whose substitution by a superordinate cures the pleonasm. What the successful substitutions do, of course, is to restore the situation where the dependent item contributes new information. Some examples follow:

(17) male uncle (pleonastic)
gay/macho uncle (normal: *gay* and *macho* are hyponyms of *male)*
male relation (normal: *relation* is a superordinate of *uncle)*

(18) He kicked it with his foot, (pleonastic)
He kicked it with his left foot, (normal: *left foot* is hyponymous to *foot)*
He struck it with his foot, (normal: *struck* is superordinate to *kick)*

12.5.2 Clash

The severity of a clash can be roughly estimated by examining the minimal change required to cure it. This enables us to put a little more flesh on the notions of *inappropriateness, paradox,* and *incongruity*. Inappropriateness is a type of clash which can be cured by substitution of one of the xenonyms by a propositional synonym:

(19) The geranium passed away, (inappropriateness)
The geranium died, (normal: *died* is a propositional synonym of *pass away)*

Paradox is a more serious type of clash which can be cured by substituting one of the xenonyms by an incompatible or immediate superordinate:

(20) The cat barked, (paradox)
The dog barked, (normal: *dog* is an incompatible of *cat)*
The animal barked, (normal: *animal* is a superordinate of *cat)*
The cat emitted a noise, (normal: *emit a noise* is superordinate to *bark)*

Incongruity is an incurable clash:

(21) powdered thrills *(^.finely divided experiences)*

12.5.3 Normality/philonyms

It is not generally the case that if X is a philonym of Y, then any superordinate of X is also a philonym. (One can easily think of cases where the result is normal: *The dogtanimal barked,* to look no further.) This is because the result may be pleonastic: *He kicked it with his left foot!?foot.* Nor is it the case that if X is a philonym of Y, then any hyponym of X is also a philonym. Again drawing on the above examples, *The dog barked! The collie barked* is fine, but *The animal barked!?The cat barked* is not. However, it might be surmised that if X is a philonym of Y, no superordinate of X can be a xenonym of Y. Thinking of a hyponym as having 'more meaning' than its superordinate, and assuming that any clash must be attributable to some bit of the meaning of X, how can taking away a bit of meaning produce a clash? Well, what about *a homeopathic doctorl?a homeopathic human being?* The explanation for this seems to run as follows. If the meaning of X can be represented as [A] + [B], then an adjective modifying X may attach itself uniquely to [B]. Suppose, now, that Y contains only the component [A]; the adjective is then forced to attach itself to [A], with which it may clash. This is a plausible explanation of what happens with *homeopathic doctor:* if we analyse "doctor" into [HUMAN] + [PRACTICES MEDICINE], then *homeopathic* will modify only the second component, and when that is removed, it will be forced to modify [HUMAN], with which it clashes. Whether this can happen also with natural kinds is an interesting question.

12.6 Some puzzles

The effect of putting words together is not always what might be predicted on general grounds. A particular example of this is the failure of pleonasm to appear in certain circumstances. Consider the following examples:

(22) Mary rushed quickly to the door.
(23) John murmured softly in Bertha's ear.
(24) Some children were shouting loudly in the street.
(25) During last summer's scorching heat-wave . . .
(26) Jack gasped—a huge giant stood at the door.

Somehow, these are not as bad as they should be: after all, surely quickness is of the essence of rushing, softness of murmuring, loudness of shouting, and so on. Also, substituting antonyms for these epithets results in paradox:

(27) ?Mary rushed slowly to the door.
(28) ?John murmured loudly in Mary's ear.
(29) ?Some children were shouting softly in the street.
(30) ?Jack gasped—a small giant stood at the door.

In examples like (22)-(26), instead of pleonasm, we seem to get either reinforcement, or something like semantic agreement. It is difficult to say under what circumstances pleonasm does not appear. All the examples mentioned here involve some gradable (adverbial) property which is incorporated into the meaning of a verb: expressing the same idea with a separate adverb has the effect of reinforcing the notion. The same effect appears with *Johnny was very, very, very naughty,* where every extra *very* adds intensity; on the other hand, in *Johnny was extremely, extremely, extremely naughty,* the extra *extremely's* come across (to me at least) as merely redundant.

Another type of situation where pleonasm fails to appear occurs with certain verbs of bodily motion. Consider the following:

(31) Mary shrugged her shoulders.
(32) Mary stamped her foot in annoyance.
(33) Mary pouted her lips.

Why are these not pleonastic? What else can one shrug with except one's shoulders, or pout with, except one's lips? Also, *What Mary pouted was her lips* and *What Mary shrugged were her shoulders* are pleonastic, and, of course, *What Mary shrugged were her thighs* and *What Mary pouted were her ears* are paradoxical. The generalization here seems to be that these verbs denote actions which can serve as signals. If the body part is not explicitly mentioned, then the signalling function of the action is highlighted *(Mary shrugged, Mary pouted)',* if the body part *is* mentioned, the action itself is highlighted, and this may, or may not, be intended also to carry the conventional message (cf. *John shrugged his shoulders to dislodge the parrot* and *?John shrugged to dislodge the parrot).* The impossibility of **Mary smiled her lips* or **Mary frowned her forehead* is presumably due to the fact that these are basically intransitive verbs but the question remains of why this should be so.

12.7 Specifying co-occurrence restrictions

In this section some of the problems of stating the co-occurrence regularities of words will be discussed, without, perhaps, all of them being resolved.

Classically, selectional restrictions were stated in the form of semantic categories to which lexical partners had to belong (recall that most selectees are nouns). Furthermore, these categories were of the classical variety, with sharp boundaries and necessary and sufficient criteria for membership. So, for instance, in the case of *X drank Y* and *X poured Y into Z,* the selectional restrictions of both *drink* and *pour* require that Y denote a liquid. Violation of the restriction leads to anomaly. Hence, the following are normal:

(34) John drank the milk.
(35) John poured the milk into the cup.

(36) Mary drank the beer.
(37) Mary poured the petrol into the can.

while the following are not:

(38) ?John drank the bread.
(39) ?John poured the cabbage into the pan.
(40) ?Mary drank her wedding ring.
(41) ?Mary poured the cup into the milk.

In some cases (but probably not any of the above), the anomaly can be resolved by reinterpreting the sentence as a metaphor:

(42) Mary drank in John's words.

If the patient (i.e. the thing affected) of either of these verbs is not specified, then the feature [LIQUID] will be transferred to them; thus, in each of the following, a normal interpretation would be that the patient is in liquid form:

(43) Mary drank the medicine.
(44) John poured the butter over the meat.
(45) The aliens were drinking a purplish substance.

This is all very well, so far as it goes. However, consider, first, the following:

(46) Mary poured the sugar into the bowl.
(47) The lorry poured the bricks onto the road.

By no stretch of the imagination can the sugar and bricks be considered to belong to the category of liquids, yet these sentences are not as anomalous (are they at all?) as they ought to be. One possibility is that we have misidentified the selectional restriction: perhaps the restriction for *pour,* at least, should require that the patient is capable of flowing. This would seem reasonable for sugar, but is it plausible for bricks? Do they flow? Here we seem to be stretching the meaning of*flow* somewhat.

Second, consider the following:

(48) Mary drank the petrol.
(49) John drank the sulphuric acid.

Are these normal? The patients are certainly liquids. If not, is this a sign that the selectional restrictions as specified are inadequate? If we think of drinking as a purely physical activity—the ingestion of liquids—then these are not odd. If Mary took in some petrol in the way that people normally take in water, then we would have to describe her action as *drinking.* However, there are other aspects to drinking: people usually drink to satisfy a thirst, for nourishment, or for enjoyment. Drinking harmful liquids is definitely eccentric.

A way of accommodating both these types of case is to take a view akin to the prototype view of categories. There are no hard and fast rules for combining words, combinations are not either normal or anomalous, they are more or less normal. We can therefore say, for instance, that *pour* has a preference for liquid patients; that is, the more the patient behaves like a liquid, the more normal the result will be (or, the better an example of the use of *pour* we will have). In a sense, bricks in large enough numbers falling out of a lorry, and from a distance, have some of the characteristics of "flowing", and to that extent resemble a liquid. In the case of drinking, there are prototypical and less prototypical instances of drinking. To characterize prototypical instances, we need to bring in more than just the physical nature of what is drunk. For these reasons, it is better to speak of selectional preferences. And yet there is still a problem here. In some sense, liquidness is more essential to drinking than harmlessness. It is necessary, for drinking (or pouring) to occur, that the patient should be sufficiently liquid-like; it is not necessary that the liquid should be nourishing, therapeutic, or thirst quenching. So we haven't completely got rid of necessity.

In some cases it is difficult to pin down exactly what the co-occurrence constraints are. Take the case of the adjective *avid.* Dictionaries typically mention interest, enthusiasm:

Someone who is avid has an extreme interest in something so that they do it with enthusiasm. *(Collins Cobuild Dictionary.)*
strongly interested, enthusiastic. *(Longman Dictionary of the English Language.)*

These definitions seem to capture the sense of *avid* in, for example: *an avid reader, an avid television viewer, an avid stamp-collector.* But this sense does not rule out the following less normal collocations, which seem to fall under the definitions given: *?an avid footballer, an avid gambler, an avid musician* (although *an avid concert-goer* is OK), *an avid botanist.* Some sense of consumption or acquisition seems to be necessary: compare *?an avid computer hacker,* which has no orientation towards reception, and *an avid net-surfer* which has. Even this is not quite right, because *an avid womanizer* and *an avid drinker* do not feel good either (although the latter case might be explained by the necessity for "interest"): it seems that satisfaction of the basic appetites does not count. It is not clear what sort of account of selectional preferences is called for in such cases. It may be that we could build up a picture of a prototypical avid person in terms of which an account of preference grading could be framed. (The picture is complicated by the slightly different, but none the less related, requirements of avid *for.* This is satisfactory in combination with: *praise, affection, knowledge, recognition',* but less so with: *sex(f), food, exercise, music, money.)*

12.8 Co-occurrence patterns between words

It is a commonplace observation that words prefer some partners to others. And some dictionaries take it upon themselves to impart what they call 'collocational information' to their readers. In this section we shall look at the different factors (not excluding the semantic factors discussed above) which govern the relative frequency of association of two (or more) words, and in the process we shall hope to provide a rationale for a useful lexicographic practice.

The question we shall attempt to answer will be formulated comparatively: Why does A have a greater affinity for X than for Y? This will be helpful in isolating the different factors. The notion of (collocational) affinity refers to the ratio between the actual co-occurrence of two words, and their predicted co-occurrence on the basis of their individual frequencies in the language. The first distinction to be made is between those cases where the reason for A's preference for X over Y is due to a semantic clash between A and Y, and the absence of such a clash between A and X, and those cases where there is no such clash between A and Y, and yet A has a greater affinity for X. We shall begin with the latter type of case.

12.8a Extralinguistic factors

Some of the possible reasons for the greater affinity of A for X rather than Y are not located in the language at all, but in the extralinguistic world. For instance, one reason why *Jane fried the egg* is more frequent than *Jane fried the lettuce* is simply that people in the world are more likely to fry eggs than lettuce. It is not that there is anything about lettuce that prevents it being fried: on the contrary, fried lettuce is delicious. Similarly, the reason oZdhas a greater affinity for *clothes* than for *newspapers* is simply that people tend to throw newspapers away when their day is past, but hang on to clothes a bit longer, so that there are more old examples around. However, frequency in the extralinguistic world is not the only consideration, since something may be very frequent, but not often noticed or realized, and is therefore not often talked about. So, for instance, there are probably more old pebbles in the world than old men, but first, old pebbles do not enter our consciousness very often, that is to say, they have low salience, and second, it is much less easy to gauge the age of a pebble than that of a man, that is to say, this is knowledge that we are less likely to have. A further governing factor is significance: to what extent does it matter whether something is old or not? It may be presumed that the more significance something has, the more it gets talked about. Again, there are probably as many old trees as old men, but it matters little, generally speaking, whether a tree is old or young. But it makes a great deal of difference (generally) whether a man is old or young. This is the probable explanation of why the most frequent collocation of *old* in present-day English is *man.*

12.8.2 Stereotypic combinations

A factor leading to collocational affinity which lies on the border between the linguistic and the non-linguistic is the existence of **stereotypic combinations,** such as the co-occurrence of *beautiful* with *flower(s),* or *dear* with *friend.* This is to be distinguished from what will be called **cliches** below: there, it is a matter of there being a standardized way of saying something (although there are alternatives); here, it is a matter of there being a standardized thing to say, or perhaps more revealingly, a standardized thing to think. This seems more a matter of the culture than of the language as such.

12.8.3 Default patterns (cliches)

A number of factors leading to collocational affinity are, of course, part of the language. We shall make a distinction between patterns of co-occurrence, divergence from which leads to anomaly of some kind, and those where there is not necessarily any anomaly, merely a degree of markedness or heightened salience. An example of the latter type is *barefaced lie,* where *shameless, brazen, unabashed, insolent,* or *blatant* would be semantically compatible, but the choice of one of these would be less 'automatic'. Another example is:

(50) **X** was last night under intense pressure to resign.

Here, the meaning of *intense* would be equally well conveyed by *strong* or *extreme,* but is significantly more likely. Similarly, *fresh allegations* (cf. *new allegations), gross negligence* (cf. *great negligence),* etc.

12.8.4 'Arbitrary' collocational restrictions

It is obvious enough that the meanings of words have an effect on their collocational affinity. A foreigner who knew the meanings of the words would not need to be told that *The farmer killed the rabbit* is more likely to occur in English than *The farmer killed the gate.* It is not that occasions of gate killing are rare in English-speaking countries (but a national pastime elsewhere); it is rather that they are inconceivable anywhere. This is because things have to be alive before they can be killed, and gates are just not living things. Here it is a matter of the satisfaction, or otherwise, of inherent selectional preferences. A person who consistently got this sort of thing wrong would be suspected of either a deficient knowledge of the meanings of the words, or insanity. However, as we saw above, there are also selectional preferences which are arbitrary in the sense of not being predictable from general knowledge. For instance, we say *a high wind* but *heavy rain.* In each case the adjective indicates the degree to which the relevant phenomenon is manifesting itself, and the degree is the same in both cases. But we cannot say *a heavy wind* or *high rain.* There is no inherent semantic incompatibility between "high degree" and "wind": the incompatibility is between the word *heavy* and the word *wind.*

This is information that even a sane foreign learner cannot be expected to have, and should be presented in any dictionary that aims at comprehensiveness. (Notice that it is not entirely clear that the collocational affinities proposed between *kill* and *rabbit,* and between *high* and *wind* will show up as enhanced collocational frequency. In the case of *kill* and *rabbit,* the effect might be masked by the infrequency with which people kill rabbits, or the lack of newsworthiness of such events. In the case of *high* and *wind,* it could well be that the frequency of *high wind* is less than what would be predicted from the separate frequencies of *high* and *wind.* The problem here is what should be counted. If we count word forms, then it is not clear that affinity will be reflected in frequency. If, on the other hand, we look at occurrences of the notion "high wind", then we would expect the form *high wind* to be the most frequent. Or perhaps we should be more specific still, and ask ourselves, given that we wish to express the notion "high wind", and given that we wish to use the word wW, what would be its most likely partner?)

12.8.5 Non-compositional affinities

A special type of affinity holds between lexical items which occur in a non-compositional (e.g. idiomatic) combination such as *pull someone's leg.* Expressions of this sort were discussed in Chapter 4.

Discussion questions and exercises

1. None of the following sentences is ambiguous, although each one contains at least one ambiguous word. Explain carefully how the selection of appropriate senses operates:

(i) A: Are you going to the club tonight?
B: I'll have to go to the bank first.
(ii) Have you booked the right turn?
(iii) She had gained several pounds since she had worn this ensemble last.

2. Identify the degree of clash in the following (i.e. inappropriateness, paradox, incongruity):

(i) She's more than just a pretty countenance.
(ii) The president is said to be unconvinced by the locomotion.
(iii) Mum, it's so nice to be back in my place of domicile again!
(iv) The whole thing was over in an age.
(v) I don't know if he acted from motives of despair or crockery.

3. Consider the selectional restrictions governing the X-position in the following (give a prototype account where appropriate):

a record X X is sad a leisurely X Can you lend me an X? (consider why *tree* is odd in this position)

Suggestions for further reading

This chapter is mostly a development of ideas which first appeared in Cruse (1986), especially chs. 4.12 and 12.2. Cruse (forthcoming *a)* takes a prototype-theoretical approach to syntagmatic sense relations. The notion of 'semantic head' presented here is quite closely parallelled by Langacker's 'profile determinant' (see Langacker 19916). For a structuralist account of selectional restrictions, see Kastovsky (1980). Katz and Fodor (1963) give the first generative version. Jackendoff's 'preference rules' (see, for instance, Jackendoff 1983) yield a prototype-like account of co-occurrence restrictions/preferences.