
Around the clause *cohesion and discourse*

9.1 *The concept of cohesion*

In two of the foregoing chapters we have been concerned with the textual organization of the clause. In Chapter 3 we described the thematic structure, based on the functions of Theme and Rheme; and in the last chapter we described the information structure, based on the functions of Given and New.

Both theme and information are realized as configurations of structural functions, though with two important differences between them. Theme is a system of the clause; and it is realized by the sequence in which the elements of the clause are ordered — Theme comes first. Information is not a system of the clause: it has its own domain, the information unit, which typically corresponds to a clause but not necessarily so; and it has its own realization in the form of tonic prominence — which typically comes at the end of the information unit, but again not necessarily so. It is these differences that make it possible for thematic and informational patterns to be combined in so many varying ways. Theoretically, there is no reason why Given and New should not also have been organized as a system of the clause and realized by the sequence of the elements — like Theme and Rheme, only based on final instead of initial position, with the New always coming last. Notice, however, that this arrangement would have greatly curtailed the potential of these two systems in the language, since they would have been combinable in only one way, with the Theme always selected from within the Given, and the New always selected from within the Rheme. As it is, Theme + Rheme and Given + New are **typically** combined in this way, but at the same time they are independent of each other: it is possible for the same element to be both Theme and New, and this is a meaningful choice. In other words, theme and information are related by the ‘good reason’ principle: other things being equal, the information unit is also a clause, hence a thematic unit; the New follows the Given, and thus the **FOCUS** of information, which is the culmination of the New, also forms the culmination of the Rheme.

Theme and information together constitute the internal resources for structuring the clause as a message — for giving it a particular status in relation to the surrounding discourse. But in order that a sequence of clauses, or clause complexes, should constitute a text, it is necessary to do more than give an appropriate internal structure to each. It is necessary also to make explicit the external relationship

between one clause or clause complex and another, and to be able to do so in a way which is not dependent on grammatical structure.

We have described the pattern of structural relationships between clauses (Chapter 7); these are what produce clause complexes. A clause complex corresponds closely to a SENTENCE of written English; in fact it is the existence of the clause complex in the grammar which leads to the evolution of the sentence in the writing system. But the clause complex has certain inbuilt limitations, from the point of view of its contribution to the texture of a discourse. The things that are put together in it have to be clauses; and they have to occur next to one another in the text. These are inherent in the nature of grammatical structure.

As we saw, a very wide range of semantic relationships is encoded through nexuses within the clause complex. But in order to construct discourse we need to be able to establish additional relations within the text that are not subject to these limitations; relations that may involve elements of any extent, both smaller and larger than clauses, from single words to lengthy passages of text; and that may hold across gaps of any extent, both within the clause and beyond it, without regard to the nature of whatever intervenes. This cannot be achieved by grammatical structure; it depends on a resource of a rather different kind. These non-structural resources for discourse are what are referred to by the term COHESION.

There are four ways by which cohesion is created in English: by reference, ellipsis, conjunction, and lexical organization. We can illustrate all of these from the following text.

Little Boy Blue, come blow your horn!
 The sheep's in the meadow, the cow's in the corn.
 Where is the boy that looks after the sheep?
 He's under the haycock, fast asleep.
 Will you go wake him? No, not I!
 For if I do, he'll be sure to cry.

The use of *he . . . him . . . he* to refer back to 'the boy that looks after the sheep' is an instance of reference. The forms *no not I* and *if I do* exemplify ellipsis; they have to be interpreted as *no I (will) not (wake him)* and *if I (wake him)*. The word *for* expresses a conjunctive relationship between 'I will not' and 'if I do he will cry'. The word *sheep* in line three reiterates *sheep* in line two; *cow* relates to *sheep*, *corn* to *meadow*, and *wake* to *asleep*; these are all examples of lexical cohesion. We will first summarize these, and then devote a section to each in turn.

(1) REFERENCE. A participant or circumstantial element introduced at one place in the text can be taken as a reference point for something that follows. In the simplest case this means that the same thing comes in again, like *the boy who looks after the sheep . . . he . . . him . . . he* above. But it may also mean that it serves as a basis for comparison, like *Henry . . . someone else* in *Henry can't play today. We'll have to find someone else*, where *someone else* means 'someone other than Henry'.

(2) ELLIPSIS. A clause, or a part of a clause, or a part (usually including the lexical element) of a verbal or nominal group, may be presupposed at a subsequent place in the text by the device of positive omission — that is, by saying nothing,

where something is required to make up the sense. Either the structure is simply left unfilled, as in *not I for I will not wake him*, which is ellipsis properly so called; or else a placeholding element is inserted to signal the gap, like the *do* in *for if I do*, which is referred to as SUBSTITUTION.

(3) CONJUNCTION. A clause or clause complex, or some longer stretch of text, may be related to what follows it by one or other of a specific set of semantic relations. These relations are basically of the same kind as those which obtain between clauses in an expanded clause complex, as described in Chapter 7 under the headings of elaboration, extension and enhancement. The most general categories are those of apposition and clarification, addition and variation, and spatio-temporal, manner, causal-conditional and matter.

(4) LEXICAL COHESION. Continuity may be established in a text by the choice of words. This may take the form of word repetition; or the choice of a word that is related in some way to a previous one — either semantically, such that the two are in the broadest sense synonymous, or collocationally, such that the two have a more than ordinary tendency to co-occur. Lexical cohesion may be maintained over long passages by the presence of keywords, words having special significance for the meaning of the particular text.

These resources collectively meet the text-forming requirements referred to earlier. They make it possible to link items of any size, whether below or above the clause; and to link items at any distance, whether structurally related or not. Note, however, that they meet these requirements in different ways. Reference is a relationship between things, or facts (phenomena, or metaphenomena); it may be established at varying distances, and although it usually serves to relate single elements that have a function within the clause (processes, participants, circumstances), it can give to any passage of text the status of a fact, and so turn it into a clause participant. For example *that* in the following passage:

'I'm just one hundred and one, five months and a day.'
'I can't believe *that*!' said Alice.

Ellipsis (including substitution) is a relationship involving a particular form of wording, either a clause or some smaller item; it is usually confined to closely contiguous passages, and is particularly characteristic of question + answer or similar 'adjacency pairs' in dialogue. For example, *so* in Alice's reply:

'... if you've seen them so often, of course you know what they're like?'
'I believe so,' Alice replied thoughtfully.

Conjunctive relations typically involve contiguous elements up to the size of paragraphs, or their equivalent in spoken language; conjunction (in this sense) is a way of setting up the logical relations that characterize clause complexes in the absence of the structural relationships by which such complexes are defined. For example *then* in the Gnat's answer:

'Supposing it couldn't find any?' she suggested.
'Then it would die, of course.'

Finally reiteration and collocation are relations between lexical elements: most typically between single lexical items, either words or larger units, e.g. *locomotive* (word), *steam engine* (group), *in steam* (phrase), *steam up*, *get up steam* ('phrases' in the dictionary sense); but also involving wordings having more than one lexical item in them, such as *maintaining an express locomotive at full steam*. Lexical ties are independent of structure and may span long passages of intervening discourse; for example

[the little] voice was drowned by a shrill scream from the engine

where *engine* was separated from the latest previous occurrence of a related lexical item (*railway journey*) by thirty-six intervening clauses.

Many instances of cohesion involve two or three ties of different kinds occurring in combination with one another. For example:

'You don't know much,' said the Duchess; 'and that's a fact.'

Alice did not at all like the tone of this remark, and thought it would be as well to introduce some other subject of conversation.

where the nominal group *this remark* consists of a reference item *this* and a lexical item *remark*, both related cohesively to what precedes. Similarly in *some other subject of conversation*, both *other* and *subject* relate cohesively to the preceding discussion, which was about whether or not cats could grin. Typically any clause complex in connected discourse will have from one up to about half a dozen cohesive ties with what has gone before it, as well as perhaps some purely internal ones like the *that* by which the Duchess refers back to the first part of her own remark.

Cohesion is, of course, a process, because discourse itself is a process. Text is something that happens, in the form of talking or writing, listening or reading. When we analyse it, we analyse the product of this process; and the term 'text' is usually taken as referring to the product — especially the product in its written form, since this is most clearly perceptible as an object (though now that we have tape recorders it has become easier for people to conceive of spoken language also as text). So it is natural to talk about cohesion as a relation between entities, in the same way that we talk about grammatical structure, for example the structure of the clause. In the last resort, of course, a clause (or any other linguistic unit) is also a happening; but since a clause has a tight formal structure we do not seriously misrepresent it when we look at it synoptically as a configuration. The organization of text is semantic rather than formal, and (at least as far as cohesion is concerned; we are not going into questions of register structure in this book) much looser than that of grammatical units. We shall represent cohesive relations simply by additions to the structural notation. But it is important to be able to think of text dynamically, as an ongoing process of meaning; and of textual cohesion as an aspect of this process, whereby the flow of meaning is channelled into a tracable current of discourse instead of spilling out formlessly in every possible direction.

9.2 Reference

(1) It seems quite likely that reference first evolved as an 'exophoric' relation: that is, as a means of linking 'outwards' to some person or object in the environment. So, for example, the concept of 'he' probably originated as 'that man over there'.

In other words we may postulate an imaginary stage in the evolution of language when the basic referential category of PERSON WAS DEICTIC in the strict sense, 'to be interpreted by reference to the situation here and now'. Thus *I* was 'the one speaking': *you*, 'the one(s) spoken to'; *he, she, it, they* were the third party, 'the other(s) in the situation'.

The first and second persons *I* and *you* naturally retain this deictic sense; their meaning is defined in the act of speaking. The third person forms *he, she, it, they* can be used exophorically; but more often than not, in all languages as we know them, such items are ANAPHORIC: that is, they point not 'outwards' to the environment but 'backwards' to the preceding text. The following is a typical example:

Peter, Peter, pumpkin eater,
Had a wife and couldn't keep her.
He put her in a pumpkin shell
And there he kept her very well.

Here *he* and *her* are anaphoric, 'pointing' respectively to Peter and to his wife.

An anaphoric relationship of this kind creates what we are calling cohesion. Presented with one of these words, the listener has to look elsewhere for its interpretation; and if he has to look back to something that has been said before, this has the effect of linking the two passages into a coherent unity. They become part of a single text.

The quality of texture depends partly on cohesion and partly on structure. If the pronoun and its referent are within the same clause complex, this is already one text by virtue of the structural relationship between the clauses; the cohesion merely adds a further dimension to the texture. If on the other hand there is no structural relationship, the cohesion becomes the sole linking feature, and hence critical to the creation of text. The cohesive relationship itself is not affected by considerations of structure; *Peter . . . he* form an identical pattern whether they are within the same clause complex or not. But they carry a greater load in the discourse if they are not.

A text is the product of ongoing semantic relationships, construed by a variety of lexicogrammatical resources. If 'Peter' runs through the narrative structure of the discourse, then whether he is mentioned by name or by 'pro-name' or not at all he will provide a source of coherence. Whatever requires the listener or reader to store and retrieve what has gone before has this effect. But the third person forms *he, she* etc. are the main referential resources, since they are both anaphoric and explicit. We can leave Peter out altogether; but this is possible only under certain structural conditions, as in *Peter . . . had a wife and (he) couldn't keep her*. This is anaphoric, but not explicit. Or we can go on calling him *Peter*, which is explicit, but not anaphoric: since it does not require you to retrieve him from elsewhere, if we go on calling him *Peter* every time you will begin to wonder whether we are still talking about the same guy. To keep him in the picture, we need to use PERSONAL reference items (see list in Table 9(1) below).

Table 9(1)

(1) <i>Personals</i>		Head		Deictic
Class \ Function		Determinative	Possessive	
Singular	Masculine	he/him	his	his
	Feminine	she/her	hers	her
	Neuter	it	[its]	its
Plural		they/them	theirs	their
(2) <i>Demonstratives</i>		Head	Deictic	Adjunct
Class \ Function				
Specific	Near	this/these	this/these	here (now)
	Remote	that/those	that/those	there (then)
Non-specific		it	the	
(3) <i>Comparatives</i>		Deictic/ Numerative	Epithet	Adjunct/ Submodifier
Class \ Function				
General	Identity	same, equal, identical &c.		identically, (just) as &c.
	Similarity	similar, additional &c.	such	so, likewise, similarly &c.
	Difference	other, different &c.		otherwise, else, differently &c.
Particular		more, fewer, less, further &c.; so, as &c. + numeral	bigger &c.; so, as, more less &c. + adjective	better &c.; so, as, more, less &c. + adverb

(2) The second type of reference item is the DEMONSTRATIVE, *this/that, these/those* (cf. the brief account given in Chapter 6). Demonstratives may also be either exophoric or anaphoric; in origin they were probably the same as third-person forms, but they retain a stronger deictic flavour than the personals, and have evolved certain distinct anaphoric functions of their own.

The basic sense of 'this' and 'that' is one of proximity; *this* refers to something as being 'near', *that* refers to something as being 'not near'. The 'that' term tends to be more inclusive, though the two are more evenly balanced in English than their equivalents in some other languages. Proximity is typically from the point of view of the speaker, so *this* means 'near me'. In some languages, as pointed out earlier, there is a close correspondence of demonstratives and personals, such that

there are three demonstratives rather than two, and the direction of reference is near me (*this*), near you (*that*) and not near either of us (*yon*). This pattern was once widespread in English and can still be found in some rural varieties of Northern English and Scots. In modern standard English *yon* no longer exists, although we still sometimes find the word *yonder* from the related series *here*, *there* and *yonder*; but another development has taken place in the meantime.

Given just two demonstratives, *this* and *that*, it is usual for *that* to be more inclusive; it tends to become the unmarked member of the pair. This happened in English; and in the process a new demonstrative evolved which took over and extended the 'unmarked' feature of *that* — leaving *this* and *that* once more fairly evenly matched. This is the so-called 'definite article' *the*. The word *the* is still really a demonstrative, although a demonstrative of a rather particular kind.

Consider the following examples:

- (a) The sun was shining on the sea.
- (b) This is the house that Jack built.
- (c) Algy met a bear. The bear was bulgy. The bulge was Algy.

In (a) we know which 'sun' and which 'sea' are being referred to even if we are not standing on the beach with the sun above our heads; there is only one sun, and for practical purposes only one sea. There may be other seas in different parts of the globe, and even other suns in the heavens; but they are irrelevant. In (b) we know which 'house' is being referred to, because we are told — it is the one built by Jack; and notice that the information comes after the occurrence of the *the*. In (c) we know which bear — the one that Algy met; and we know which bulge — the one displayed by the bear; but in this case the information had already been given before the *the* occurred. Only in (c), therefore, is *the* anaphoric.

Like the personals, and the other demonstratives, *the* has a specifying function; it signals 'you know which one(s) I mean'. But there is an important difference. The other items not only signal that the identity is known, or knowable; they state explicitly how the identity is to be established. So

my house = 'you know which: the one belonging to me'
 this house = 'you know which: the one near me'

but

the house = 'you know which — the information is there somewhere if you look for it'

In other words, *the* merely announces that the identity is specific; it does not specify it. The information is available elsewhere. It may be in the preceding text (anaphoric), like (c) above; in the following text (CATAPHORIC), like (b); or in the air, so to speak, like (a). Type (a) are self-specifying; there is only one — or at least only one that makes sense in the context, as in *Have you fed the cat?* (HOMOPHORIC).

Thus *the* is an unmarked demonstrative, while *this* and *that* are both 'marked' terms — neither includes the other. Their basic deictic senses are 'near' and 'remote' from the point of view of the speaker. But they are also used to refer within the text. The 'near' term *this* typically refers either anaphorically, to something that has been mentioned immediately before, or by the speaker, or is in some way or

other being treated as 'near', as in (a) below; or cataphorically, to something that is to come, as in (b):

- (a) "You may look in front of you, and on both sides, if you like," said the Sheep; "but you can't look *all* round you — unless you've got eyes at the back of your head."
But these, as it happens, Alice had *not* got.
- (b) "The great art of riding, as I was saying, is — to keep your balance. Like this, you know —"

He let go the bridle, and stretched out both his arms to show Alice what he meant.

(Example (b) is EXOPHORIC in the immediate context, but cataphoric in the text.) The singular *this* is also used to refer in the same way to extended passages of text, as in (c):

- (c) "Come back!" the Caterpillar called after her. "I've something important to say!"
This sounded promising, certainly: Alice turned and came back again.

The 'remote' term *that* refers anaphorically to something that has been mentioned by the previous speaker, now the listener, as in (d), or is being treated as more remote or from the listener's point of view, as in (e):

- (d) "But he's coming very slowly — and what curious attitudes he goes into!" . . .
"Not at all," said the King. "He's an Anglo-Saxon Messenger — and those are Anglo-Saxon attitudes."
- (e) "I'll put you through into Looking-glass House. How would you like *that*?"

Again, the singular *that* often refers back to an extended passage of text, as in (f):

- (f) "If *that's* all you know about it, you may stand down," continued the King.

where *that* refers to the whole of the preceding interrogation taking up two pages of the story. Note that the reference item *it* is similarly used for text reference, as in (g):

- (g) "So here's a question for you. How old did you say you were?"
Alice made a short calculation, and said "Seven years and six months."
"Wrong!" Humpty Dumpty exclaimed triumphantly. "You never said a word like it."

The locative demonstratives *here* and *there* are also used as reference items; *here* may be cataphoric, as in (g) above, or anaphoric and 'near' as in (h); *there* is anaphoric but not 'near', as in (j), where it means 'in what you said':

- (h) "I think you ought to tell me who *you* are, first."
"Why?" said the Caterpillar.
Here was another puzzling question; . . .
- (j) "Suppose he never commits the crime?" said Alice.
"That would be all the better, wouldn't it?" the Queen said, . . .
Alice felt there was no denying *that*. "Of course it would be all the better," she said: "but it wouldn't be all the better his being punished."
"You're wrong *there*, at any rate," said the Queen.

The temporal demonstratives *now* and *then* also function as cohesive items, but conjunctively rather than referentially (see Section 9.5 below).

(3) There is a third type of reference that contributes to textual cohesion, i.e. COMPARATIVE reference. Whereas personals and demonstratives, when used anaphorically, set up a relation of co-reference, whereby the same entity is referred to over again, comparatives set up a relation of contrast. In comparative reference, the reference item still signals 'you know which'; not because the same entity is being referred to over again but rather because there is a frame of reference — something by reference to which what I am now talking about is the same or different, like or unlike, equal or unequal, more or less.

Any expression such as *the same, another, similar, different, as big, bigger, less big*, and related adverbs such as *likewise, differently, equally*, presumes some standard of reference in the preceding text. For example, *such, another, more* in (a), (b) and (c):

- (a) "Why did you call him tortoise, if he wasn't one?" Alice asked.
 "We called him Tortoise because he taught us," said the Mock Turtle angrily: "really you are very dull!"
 "You ought to be ashamed of yourself for asking such a simple question," added the Gryphon.
- (b) "At the end of two yards," she said, putting in a peg to mark the distance, "I shall give you your directions — have another biscuit?"
- (c) "I like the Walrus best," said Alice: "because, you see, he was a *little* sorry for the poor oysters."
 "He ate more than the Carpenter, though," said Tweedledee.

Like personals and demonstratives, comparative reference items can also be used cataphorically, within the nominal group; for example *much more smoothly than a live horse*, where the reference point for the *more* lies in what follows.

Table 9(1) summarizes the principal categories of reference item in English.

As has already been made clear, there is no structural relationship between the reference item and its referent. In order to mark the cohesive relationship in the text, we can devise some form of notation such as that shown in Figure 9-1.

9.3 *Ellipsis and substitution*

Reference is a relationship in meaning. When a reference item is used anaphorically, it sets up a semantic relationship with something in the preceding text; and this enables the reference item to be interpreted, as either identical with the referent or in some way contrasting with it.

Another form of anaphoric cohesion in the text is achieved by ellipsis, where we presuppose something by means of what is left out. Like all cohesive agencies, ellipsis contributes to the semantic structure of the discourse. But unlike reference, which is itself a semantic relation, ellipsis sets up a relationship that is not semantic but lexicogrammatical — a relationship in the wording rather than directly in the meaning. For example, in

- Why didn't you lead a spade?
 —I hadn't got any.

the listener has to supply the word *spades* in order to make sense of the answer.

Alice looked on with great interest as the King took an enormous memorandum book
 out of his pocket, and began writing. A sudden thought struck her and she took hold
 of the end of the pencil, which came some way over his shoulder, and began writing
 for him
The poor King looked puzzled and unhappy, and struggled with the pencil for some
 time without saying anything; but Alice was too strong for him, and at last he panted
 out, "My dear! I really *must* get a thinner pencil. I can't manage this one a bit, it writes
 all manner of things that I don't intend —"
 "What manner of things?" said the Queen, looking over the book (in which Alice had
 put "The White Knight is sliding down the poker. He balances very badly") "That's
 not a memorandum of *your* feelings!"

R:C = reference : comparative
 R:D = reference : demonstrative
 R:P = reference : personal

↖ = anaphoric (cohesive)
 ↗ = cataphoric
 ↑ = exophoric/homophoric

Fig. 9-1 Text analysed for reference

Sometimes an explicit indication may be given that something is omitted, by the use of a substitute form; for example *one* in

I've lost my voice.
 —Get a new one.

The substitute serves as a place-holding device, showing where something has been omitted and what its grammatical function would be; thus *one* functions as Head in the nominal group and replaces the Thing (with which the Head is typically conflated). Ellipsis and substitution are variants of the same type of cohesive relation. There are some grammatical environments in which only ellipsis is possible, some in which only substitution is possible, and some, such as *I preferred the other [one]*, which allow for either.

There are three main contexts for ellipsis and substitution in English. These are (1) the clause, (2) the verbal group and (3) the nominal group. We shall consider each of these in turn.

(1) The clause. Ellipsis in the clause is related to mood, and has been illustrated already in Chapter 4. Specifically, it is related to the question-answer process in dialogue; and this determines that there are two kinds: (a) yes/no ellipsis, and (b) WH- ellipsis. Each of these also allows for substitution, though not in all contexts. We will consider the yes/no type first.

(a) yes/no ellipsis: (i) the whole clause. In a yes/no question-answer sequence the answer may involve ellipsis of the whole clause, e.g.

- Can you row?
 — Yes. [I can row]
 Is that all?
 — No. [that is not all]

The first clause in such a pair is not necessarily a question; it may have any speech function, e.g.

- Have another biscuit?
 — No, thank you. [I won't have another biscuit]
 You're growing too.
 — Yes [I'm growing too], but I grow at a reasonable pace.

Corresponding in meaning to *yes* and *no* are the clause substitutes *so* and *not*. (Etymologically the word *yes* contains the substitute *so*; it is a fusion of (earlier forms of) *aye* and *so*.) In certain contexts these substitute forms are used: (i) following *if* — *if so*, *if not*; (ii) as a reported clause — *he said so*, *he said not*; (iii) in the context of modality — *perhaps so*, *perhaps not*. Examples (and cf. Chapter 7, Section 7.5.3 above):

- "Are you to get in at all? That's the first question, you know." It was, no doubt; only Alice did not like to be told so. [that that was the first question].
 Does your watch tell you what year it is?
 — Of course not. [Of course my watch does not tell me . . .]
 I dare say you never even spoke to Time!
 — Perhaps not. [Perhaps I never even spoke to Time]
 If you've seen them so often, of course you know what they're like.
 — I believe so. [I believe I know what they're like]
 If I like being that person, I'll come up; if not [if I don't like being that person], I'll stay down here till I'm somebody else.
 But they should be five times as *cold*, by the same rule —
 — Just so. [They *are* five times as cold]

The general principle is that a substitute is required if the clause is **projected**, as a report; with modality (*perhaps*) and hypothesis (*if*) being interpreted as kinds of projection, along the lines of:

- he said so — I thought so — I think so — it may be so — perhaps so — let us say so — if so

In addition, the substitute *not* is used when the answer is qualified by a negative in some way:

- I shouldn't be hungry for it, you know.
 — Not at first [you wouldn't be hungry for it at first], but . . .

where a positive clause is simply presupposed by ellipsis:

- Would you like to see a little of it?
 — Very much indeed. [I should very much indeed like to see a little of it]

(a) yes/no ellipsis: (ii) part of the clause. As an alternative to the ellipsis of the whole clause, there may be ellipsis of just one part of it, the Residue. For example:

- Must* a name mean something?
 — Of course it must. [mean something]
 I can't believe that.
 — Can't you? [believe that]
 "The horror of that moment," the King went on, "I shall never, *never* forget!"
 "You will [forget the horror of that moment], though," the Queen said, "if you don't make a memorandum of it."
 Take pen and ink and write it down.
 — I will [take pen and ink and write it down], if I can remember it so long.
 Hold your tongue!
 — I won't! [hold my tongue]

With a declarative response, if there is a change of Subject only, we may have substitute *so*, *nor* in initial position (= 'and so', 'and not') followed by the Mood element.

- Of course you know your A B C ?
 — To be sure I do. [know my A B C]
 — So do I. [know my A B C]
 I haven't the slightest idea.
 — Nor have I. [the slightest idea]

The order is Finite ^ Subject (to get the Subject under unmarked focus). If the Subject is unchanged, so that the focus is on the Finite, the order is Subject ^ Finite:

- I want to be a Queen.
 — So you will [be a Queen], when you've crossed the next brook.

The negative has various forms:

- They've never replied.
 — So they haven't/Nor they have/Neither they have [replied]

Not infrequently, the Residue is substituted by the verbal substitute *do*, as in:

- They say an apple a day keeps the doctor away.
 — It should do [keep the doctor away], if you aim it straight.

If the focus is on the Residue (and hence falls on *do*), the substitute form is *do so*:

- Alice very obediently got up, and carried the dish round, and the cake divided itself into three pieces as she did so. [as she carried the dish round]

(b) WH- ellipsis: (i) the whole clause. In a WH- sequence the entire clause is usually omitted except for the WH- element itself, or the item that is the response to the WH- element:

- I think you ought to tell me who you are, first.
 — Why? [ought I to tell you who I am]
 It writes all manner of things I don't intend.
 — What manner of things? [does it write]
 What did they draw?
 — Treacle. [they drew treacle]
 They're at it again.
 — Who? [who are at it again?]
 — The lion and the unicorn, of course. [are at it again]

The substitute *not* may appear in a WH- negative, as in *Don't look now. — Why not?* Substitution is less likely in the positive, except in the expressions *how so?*, *why so?*.

(b) WH- ellipsis: (ii) part of the clause. Sometimes in a WH- clause, or its response, the Mood element is left in and only the Residue is ellipsed. For example, with WH- Subject:

- They're at it again.
 — Who are? [at it again]
 Who can untie this knot?
 — I can. [untie that knot]

Similarly if the WH- element is part of the Residue:

- Don't look now.
 — Why shouldn't I? [look now]

Thus clausal ellipsis and substitution occurs typically in a dialogue sequence where in a response turn everything is omitted except the information-bearing element. Examples of such responses would be:

- (a) in a yes/no type environment:
 (i) polarity only: *yes no so not* (in *I think so/not* etc.)
 (ii) mood: *will you? I will* etc.
 (iii) mood + polarity: *so do I nor do I so he was* etc.
- (b) in a WH- type environment:
 (i) WH- only: *who? where? John over there* etc.
 (ii) WH- + polarity: *why not? not me* etc.
 (iii) WH- + mood: *why didn't they? I could tomorrow* etc.

A clause consisting of Mood only, such as *I will*, could equally occur in either environment; typically, in a yes/no environment, the focus would be on *will*, which bears the polarity ('Will you . . . ?' — *I will.*), whereas in a WH- environment, the focus would be on *I*, which carries the information ('Who will . . . ?' — *I will.*).

The elliptical or substitute clause requires the listener to 'supply the missing words'; and since they are to be supplied from what has gone before, the effect

is cohesive. It is always possible to 'reconstitute' the ellipsed item so that it becomes fully explicit. Since ellipsis is a lexicogrammatical resource, what is taken over is the exact wording, subject only to the reversal of speaker-listener deixis (*I* for *you* and so on), and change of mood where appropriate.

(2) The verbal group. Since the verbal group consists of Finite plus Predicator, it follows automatically that any clausal ellipsis in which the Mood element is present but the Residue omitted will involve ellipsis within the verbal group. There is no need to repeat the discussion of this phenomenon.

Substitution in the verbal group is by means of the verb *do*, which can substitute for any verb provided it is active not passive, except *be* or, in some contexts, *have*. The verb *do* will appear in the appropriate non-finite form (*do*, *doing*, *done*). Examples:

- Does it hurt?
 — Not any more. It was doing last night.
 Have the children gone to sleep?
 — I think they must have done.

As we have seen, this *do* typically substitutes for the whole of the Residue (or, what amounts to the same thing, when the verb is substituted by *do*, the rest of the Residue is ellipsed).

Since there are no demonstrative verbs — we cannot say *he thatted*, *he whatted*? — this need is met by combining the verb substitute *do* with demonstratives *that*, *what*. For example:

- A shower of little pebbles came in at the window, and some of them hit her in the face.
 "You'd better not do that again!"
 The next thing is, to get into that beautiful garden — how is that to be done, I wonder?
 I shall sit here, on and off, for days and days.
 — But what am *I* to do?

The form *do not* functions as a single reference item. (For the difference between reference and ellipsis-substitution, see the note at the end of the present section.)

(3) The nominal group. Ellipsis within the nominal group was referred to in Chapter 6, where it was shown that an element other than the Thing could function as Head; for example *any* in

- Have some wine.
 — I don't see any wine.
 — There isn't any.

There is a nominal substitute *one*, plural *ones*, which functions as Head; it can substitute for any count noun (that is, any noun that is selecting for number, singular or plural); for example,

- That's a joke. I wish you had made it.
 — Why do you wish I had made it? It's a very bad one. [a very bad joke]
 This here ought to have been a *red* rose-tree, and we put a white one [a white rose-tree] in by mistake.

Like *do* in the verbal group, the nominal substitute *one* is derived by extension from an item in the structure of the full, non-elliptical group — in this case the indefinite numeral *one*, via its function as Head in a group which is elliptical as in

- I vote the young lady tells us a story.
— I'm afraid I don't know one.

The parallel development of the two substitutes, verbal *do* and nominal *one*, is as shown in Table 9(2):

Table 9(2)

	As Modifier	As elliptical Head	As substitute Head
Verbal <i>do</i>	he does know does he know he doesn't know	perhaps he does surely he doesn't	he may do he never has done
Nominal <i>one</i>	one green bottle a green bottle	there was one there wasn't one	a green one ten green ones

In some instances the nominal substitute fuses with a Modifier, as in *mine*, *none* in the following:

- Take off your hat.
—It isn't mine. [my hat]
—Stolen!
—I keep them to sell. I've none [no hats] of my own.

These can be analysed as elliptical, the elements *my*, *your*, *no* etc. having a special form when functioning as Head.

We remarked earlier that ellipsis-substitution is a relationship at the lexicogrammatical level: the meaning is 'go back and retrieve the missing words'. Hence the missing words must be grammatically appropriate; and they can be inserted in place. This is not the case with reference, where, since the relationship is a semantic one, there is no grammatical constraint (the class of the reference item need not match that of what it presupposes), and one cannot normally insert the presupposed element. Reference, for the same reason, can reach back a long way in the text and extend over a long passage, whereas ellipsis-substitution is largely limited to the immediately preceding clause.

But the most important distinction, which again follows from the different nature of the two types of relationship, is that in ellipsis-substitution the typical meaning is not one of co-reference. There is always some significant difference between the second instance and the first (between presupposing item and presupposed). If we want to refer to the same thing, we use reference; if we want to refer to a different thing, we use ellipsis-substitution: *Where's your hat? — I can't find it. — Take this (one)*. Each can take on the other meaning, but only by making it explicit: *another hat* (reference, but different), *the same one* (substitution, but not different). Thus reference signals 'the same member' (unless marked as different by the use of comparison); ellipsis-substitution signals 'another member of the same class' (unless marked as identical by *same*, etc.). The difference is most clear-cut in the nominal group, since nouns, especially count nouns, tend to have

"Being so many different sizes in a day is very confusing."

"It isn't ^{E.c} " [very confusing]

"Well, perhaps you haven't found it ^{S.c} (so) yet, but when you have [to be very confusing]

to turn into a chrysalis – you will ^{S.v} (do) some day, you know – and [turn into a chrysalis]

then after that ^{E.c} into a butterfly, I should think you'll feel it a little [you have to turn]

queer, won't you?"

"Not a bit ^{E.c} " [I shall (not) feel it
(a bit) queer]

"Well, it would feel very queer to me "

"You! Who are you?"

"I hardly know ^{E.c} , sir, just at present " [who I am]

"So you think you're changed, do you?"

"I'm afraid I am ^{E.v} , sir " [changed]

E.c = ellipsis clausal
 E.v = ellipsis verbal
 E.n = ellipsis nominal

S.c = substitution clausal
 S.v = substitution verbal
 S.n = substitution nominal

Fig. 9-2 Text analysed for ellipsis and substitution

clearly defined referents; it is much less clear-cut in the verbal group or the clause.

Within the nominal group, 'another member' means a new modification of the Thing; Deictic (*this one, another one, mine*), Numerative (*three, the first (one)*), or Epithet (*the biggest (one), a big one*). In the verbal group, it means a new specification of polarity, tense or modality through the Finite element (*did, might (do), hasn't (done)*); and there is a slight tendency for ellipsis to be associated with change of polarity and substitution with change of modality. This tendency is more clearly marked with the clause, where ellipsis adds certainty (yes or no, or a missing identity), whereas substitution adds uncertainty (if, maybe, or someone said so); this is why, in a clause where everything is ellipsed except the modality, it is quite usual to use a substitute (*possibly so, perhaps so*) unless the modality is one of certainty — here we say *certainly* (elliptical), rather than *certainly so*.

Figure 9-2 is a short text marked for ellipsis and substitution. For the sake of the exposition, the ellipsed items have been shown at the side, although this is not a necessary part of the analysis.

9.4 Conjunction

We saw in Chapter 7 that the fundamental logical-semantic relations of expansion and projection take many different forms in combination with other features. An example is given, in Appendix 3, of the causal relation expressed in a variety of grammatical guises. Most of the encodings presented there are structural: the causal

relation is realized in the structure of a clause, or of a hypotactic clause nexus. Examples are also given, however, of non-structural relations, where cause and effect are in different sentences but the relationship is still made explicit; for example

She didn't know the rules. Consequently she died.
She died. For she didn't know the rules.

Here the relationship of cause constitutes a cohesive bond between the two clauses; and it is expressed by the words *consequently* and *for*.

This type of cohesion is known as conjunction. A range of possible meanings within the domains of elaboration, extension and enhancement is expressed by the choice of a conjunctive Adjunct (an adverbial group or prepositional phrase), or of one of a small set of conjunctions *and, or, nor, but, yet, so, then*, typically (and in the case of the conjunctions obligatorily) in thematic position at the beginning of the clause.

(1) Elaboration. There are two categories of elaborating relation, (a) apposition and (b) clarification. We will consider the appositive type first.

(a) apposition. In this type of elaboration some element is re-presented, or restated, either (i) by exposition, the 'i.e.' relation, or (ii) by example, the 'e.g.' relation. Typical conjunctive expressions of these two kinds are as follows:

- (i) expository: in other words, that is (to say), I mean (to say), to put it another way
- (ii) exemplifying: for example, for instance, thus, to illustrate

(b) clarification. Here the elaborated element is not simply restated but reinstated, summarized, made more precise or in some other way clarified for purposes of the discourse:

- (i) corrective: or rather, at least, to be more precise
- (ii) distractive: by the way, incidentally
- (iii) dismissive: in any case, anyway, leaving that aside
- (iv) particularizing: in particular, more especially
- (v) resumptive: as I was saying, to resume, to get back to the point
- (vi) summative: in short, to sum up, in conclusion, briefly
- (vii) verifactive: actually, as a matter of fact, in fact

(2) Extension. Extension involves either addition or variation. Addition is either positive *and*, negative *nor* or adversative *but*; but since the adversative relation plays a particularly important part in discourse it is best taken as a separate heading on its own. Variation includes replacive *instead*, subtractive *except* and alternative *or* types.

(a) addition

- (i) positive: and, also, moreover, in addition
- (ii) negative: nor

(b) adversative: but, yet, on the other hand, however

(c) variation

- (i) replacive: on the contrary, instead
- (ii) subtractive: apart from that, except for that
- (iii) alternative: alternatively

(3) Enhancement. The various types of enhancement that create cohesion are (a) spatio-temporal, (b) manner, (c) causal-conditional and (d) matter. Each of these will be briefly discussed and exemplified.

(a) spatio-temporal. Place reference may be used conjunctively within a text, with *here* and *there*, spatial adverbs such as *behind* and *nearby*, and expressions containing a place noun or adverb plus reference item, e.g. *in the same place*, *anywhere else*. Here spatial relations are being used as text-creating cohesive devices.

Note however that most apparently spatial cohesion is in terms of metaphorical space; for example *there* in *there you're wrong*; cf. expressions like *on those grounds*, *on that point*. These are actually expressions of Matter. Many conjunctive expressions of the expanding kind are also in origin spatial metaphors; e.g. *in the first place*, *on the other hand* (*hand* involves a double metaphor: 'part of the body' — 'side' [*on my right hand*] — 'side of an argument').

Temporal conjunction covers a very great variety of different relations, the most general categories being as follows:

(i) simple

- [a] following: then, next, afterwards [including correlatives first . . . then]
- [b] simultaneous: just then, at the same time
- [c] preceding: before that, hitherto, previously
- [d] conclusive: in the end, finally

(ii) complex

- [e] immediate: at once, thereupon, straightaway
- [f] interrupted: soon, after a while
- [g] repetitive: next time, on another occasion
- [h] specific: next day, an hour later, that morning
- [j] durative: meanwhile, all that time
- [k] terminal: until then, up to that point
- [l] punctiliar: at this moment

Those that are called 'complex' are the simple ones with some other semantic feature or features present at the same time.

Many temporal conjunctives have an 'internal' as well as an 'external' interpretation; that is, the time they refer to is the temporal unfolding of the discourse itself, not the temporal sequence of the processes referred to. In terms of the functional components of semantics, it is interpersonal not experiential time. Parallel to the 'simple' categories above we can recognize:

(iii) simple internal

- [m] following: next, secondly ('my next point is') [incl. correlatives first . . . next]
- [n] simultaneous: at this point, here, now

- [o] preceding: hitherto, up to now
- [p] conclusive: lastly, last of all, finally

These shade into temporal metaphors of an expanding kind such as *meanwhile*, *at the same time* (*meanwhile let us not forget that . . .*, *at the same time it must be admitted that . . .*).

(b) manner. Manner conjunctives create cohesion (i) by comparison, (ii) by reference to means. Comparison may be (a) positive ('is like'), or (b) negative ('is unlike'):

- (i) comparison
 - [a] positive: likewise, similarly
 - [b] negative: in a different way

(ii) means: thus, thereby, by such means

Expressions of means are however not often conjunctive; those that are are usually also comparative, e.g. *in the same manner*, *otherwise*.

(c) causal-conditional. In many types of discourse the relation of cause figures very prominently as a cohesive agent. Some cause expressions are general, others relate more specifically to result, reason or purpose:

- (i) general: so, then, therefore, consequently, hence, because of that; for
- (ii) specific
 - [a] result: in consequence, as a result
 - [b] reason: on account of this, for that reason
 - [c] purpose: for that purpose, with this in view

Conditionals subdivide into (i) positive, (ii) negative and (iii) concessive.

- (i) positive: then, in that case, in that event, under the circumstances
- (ii) negative: otherwise, if not
- (iii) concessive: yet, still, though, despite this, however, even so, all the same, nevertheless

(d) matter. Here cohesion is established by reference to the 'matter' that has gone before. As noted earlier, many expressions of matter are spatial metaphors, involving words like *point*, *ground*, *field*; and these become conjunctive when coupled with reference items. Typical expressions are:

- (i) positive: here, there, as to that, in that respect
- (ii) negative: in other respects, elsewhere

It is clear that a number of these different types of conjunctive relation overlap with one another. The conjunctive relation of 'matter' is very close to some of those of the elaborating kind, and the concessive ('despite X, nevertheless Y') overlaps with the adversative ('X and, conversely, Y'). Such pairs are characterized by differences of emphasis, and some instances can be assigned to one member or the other; but others cannot, and may be interpreted either way. The categories given here are those which have been found most useful in the interpretation of texts, and

their schematization is such as to relate to other parts of the system of the language. Table 9(3) sets out the conjunctive relations so as to show how they match up with expansion generally.

Secondly the whole phenomenon of conjunction shades into that of reference. Many conjunctives have reference items embedded in them, typically *that* or *this*: *in that case*, *despite this*, *from there on*, etc. In such cases the conjunctive relation can be taken as the predominant one, because it embodies more meaning — more semantic features; any instance which can be assigned to a conjunctive category can be interpreted as such and the reference item ignored.

One question that arises in the interpretation of a text is what to do about conjunction that is implicit. It often happens, especially with temporal and causal sequences, that the semantic relationship is clearly felt to be present but is unexpressed; for example

George Stephenson died on 12 August 1848 . . . He was buried at Holy Trinity, Chesterfield.

where there is obviously a temporal relationship between the two parts; cf. the following where the relation is one of cause:

Hudson decided next to establish himself in London. He bought what was then considered to be the largest private house in London, Albert House,

It is clear that texture is achieved through conjunctive relations of this kind and there is no reason not to take account of it. On the other hand, the attempt to include it in the analysis leads to a great deal of indeterminacy, both as regards whether a conjunctive relation is present or not and as regards which particular kind of relationship it is. Consider the extract:

Around 1823, certain normally staid and sensible firms in the city of London got themselves very worked up about the possibilities of great fortunes to be made in South America. The idea was admittedly very exciting. Everybody knew the old stories, even if many of them were legendary, about the Inca gold mines, about the Spanish conquistadores and the undreamt of mineral wealth which they had found. These mines had been worked by hand, without machines, and long since left abandoned. Think what can now be done, suggested some bright speculator, using all our new and marvellous steam engines!

This is a highly cohesive passage; but it is difficult to say what implicit conjunctive relationship would hold between pairs of adjacent sentences, or between each sentence and anything that precedes it.

It is perhaps as well, therefore, to be cautious in assigning implicit conjunction in the interpretation of a text. It is likely that there will always be other forms of cohesion present, and that these are the main source of our intuition that there is a pattern of conjunctive relationships as well. Moreover the presence or absence of explicit conjunction is one of the principal variables in English discourse, both as between registers and as between texts in the same register; this variation is obscured if we assume conjunction where it is not expressed. It is important therefore to note those instances where conjunction is being recognized that is implicit; and to characterize the text also without it, to see how much we still feel is being left unaccounted for.

Table 9(3) Synoptic summary of expansion

type of expansion		functional relationship with which expansion is combined	COHESION between clause complexes (non-structural)	INTERDEPENDENCY between clauses in a clause complex			
				paratactic	hypotactic		
⊕ ELABORATION	apposition	expository exemplificatory	In other words For example	that is	which, who	[non-finite clause]	
	clarification	{ various types	Or rather, Anyway, Actually &c.	at least	NON-DEFINING RELATIVE CLAUSE		
⊕ EXTENSION	addition	positive negative adversative	Also Neither However	and nor but	while whereas	besides without	
	variation	replacive subtractive alternative	On the contrary Otherwise Alternatively	only or	except that if not . . . then	besides other than	
⊗ ENHANCEMENT	spatio-temporal	place	extent point(s)	There	there	as far as where(ver)	
		time	extent point(s) prior subsequent { various complex types	Throughout Simultaneously Previously Next Finally, At once, Meanwhile &c.	now then	while when(ever) before, until after, since as soon as &c.	while, in when, on before, until after, since
	manner	means quality comparison	Thus Likewise	so	as, as if	by like, as if	
	causal-conditional	cause	reason result purpose insurance	Therefore Consequently To that end	so, for thus	because in order that in case	with, by as a result of (so as) to, for in case of
		condition	positive negative concessive	In that case Otherwise Nevertheless	then otherwise though	if, as long as unless although	if, in event of without despite
matter	respectiva	In this respect					
class of item { that is being related: by which relationship is realized			clause(complex): prepositional phrase or adverb	independent clause: conjunction	finite or non-finite dependent clause: conjunction, preposition, or relative (noun)		

EMBEDDING of clause as Modifier in nominal group	CIRCUMSTAN- TIATION in clause (as process)	PHASE, CONATION &c. in verbal group complex (TENSE, VOICE in verbal group)		ATTRIBUTION or IDENTIFICATION as relational process in clause
which, who; [non-finite that DEFINING RELATIVE CLAUSE	as ROLE	PASSIVE VOICE is (v ⁿ)	PHASE (a) TIME start, keep (b) REALITY seem, turn out	INTENSIVE 'is' ⊖
whose, of which DEFINING RELATIVE CLAUSE (POSSESSIVE)	with, including without ACCOMPANIMENT instead of except (for)	PAST TENSE has (v ⁿ) OBLIGA- TION has to (v ^o)	CONATION & POTEN- TIALITY try; succeed; can, learn	POSSESSIVE 'has' ⊕
DEFINING RELATIVE CLAUSE (CIRCUMSTANTIAL) (a) CIRCUMSTANCE AS HEAD place (where/that) time (when/that) reason (why/that) &c. (b) CIRCUMSTANCE AS MODIFIER [HEAD] where/at which [noun] when/on which for which about which &c.	for at, in PLACE for at, on before after during &c. TIME by, with [adverb] like MANNER because of CAUSE for in case of in the event of in default of despite CONDITION about MATTER	PRESENT TENSE is (at) (v ⁿ) EXPECT- ATION is to (v ^o)	MODULA- TION (a) TIME begin by (b) MANNER venture hesitate (c) CAUSE happen, remember	CIRCUMSTANTIAL 'is at' (a) CIRCUMSTANCE AS PROCESS occupies, follows, causes concerns &c. (b) CIRCUMSTANCE AS PARTICIPANT is at, in, on, before, like, because of, about &c. ⊗
finite or non-finite rankshifted clause: relative (noun, adverb, or prepositional phrase)	propositional phrase: preposition	verbal group: auxiliary	verbal group complex: verb	nominal group: verb or preposition

"Heat is only the motion of the atoms I told you about "

"**Then** what is cold?"
C:cond

"Cold is only absence of heat "

"**Then** if anything is cold it means that its atoms are not moving "
C:cond

"Only in the most extreme case. There are different degrees of cold" \emptyset A piece of ice is cold
C:ap

compared with warm water. **But** the atoms of a piece of ice are moving – they are moving
C:conc

quite fast. **as a matter of fact** **But** they are not moving as fast as the atoms of warm water
C:ap C:conc

So that compared with the water, the ice is cold **But** even the water would seem cold, if
C:caus C:ad

compared with a red-hot poker **Now** I'll tell you an experiment you ought to try one day "
C:temp

C:ad = additive

C:caus = causal

C:cond = conditional

C:ap = appositive

C:conc = concessive

C:temp = temporal

\emptyset = implicit conjunction

Fig. 9-3 Text analysed for conjunction

Figure 9-3 gives an example of a text showing conjunctive relations. The headings that may be found useful for most purposes of analysis are the general ones of appositive, clarificative; additive, adversative, variative; temporal, comparative, causal, conditional, concessive, matter.

9.5 *Lexical cohesion*

The remaining type of pattern by which a speaker or writer creates cohesion in discourse is the choice of lexical items.

Lexical cohesion comes about through the selection of items that are related in some way to those that have gone before.

(1) Repetition. The most direct form of lexical cohesion is the repetition of a lexical item; e.g. *bear* in

Algy met a bear. The bear was bulgy.

Here the second occurrence of *bear* harks back to the first.

In this instance, there is also the reference item *the*, signalling that the listener knows which bear is intended; and since there is nothing else to satisfy the *the*, we conclude that it is the same bear. But this referential link is not necessary to lexical cohesion; if we had *Algy met a bear. Bears are bulgy*, where *bears* means 'all bears', there would still be lexical cohesion of *bears* with *bear*. In this case, however, there would be only one tie; whereas in the example cited first there are two, one referential (*the*) and one lexical (*bear*).

As the last example shows, in order for a lexical item to be recognized as repeated

it need not be in the same morphological shape. For example, *dine*, *dining*, *diner*, *dinner* are all the same item, and an occurrence of any one constitutes a repetition of any of the others. Inflexional variants always belong together as one item; derivational variants usually do, when they are based on a living derivational process, although these are less predictable. (For example, *rational* and *rationalize* are probably still the same lexical item, though the relationship between them has become rather tenuous; but neither now goes with *ration* — *rational* is closer to *reason*, though not close enough to be considered the same item.)

In Landor's line

I strove with none, for none was worth my strife

there is a strongly felt cohesion between *strife* and *strove*, suggesting that *strive*, *strove* and *strife* are one and the same lexical item.

(2) **Synonymy.** In the second place, lexical cohesion results from the choice of a lexical item that is in some sense synonymous with a preceding one; for example *sound* with *noise*, *cavalry* with *horses* in

He was just wondering which road to take when he was startled by a noise from behind him. It was the noise of trotting horses. . . . He dismounted and led his horse as quickly as he could along the right-hand road. The sound of the cavalry grew rapidly nearer . . .

Here again the cohesion need not depend on identity of reference. But once we depart from straightforward repetition, and take account of cohesion between related items, it is useful to distinguish whether the reference is identical or not, because slightly different patterns appear.

(a) with identity of reference. Here the range of potentially cohesive items includes synonyms of the same or some higher level of generality: synonyms in the narrower sense, and SUPERORDINATES. For example, in

Four-&-twenty blackbirds, baked in a pie.
When the pie was opened, the birds began to sing.

we have one instance of repetition (*pie* . . . *pie*) and one of synonyms (*blackbirds* . . . *birds*). *birds*, however, is at a higher level of generality than *blackbirds*; it is a superordinate term. In fact we might have (disregarding the scansion, of course) any of the following sequences:

four-&-twenty blackbirds . . .	the blackbirds began to sing
"	the birds began to sing
"	the creatures began to sing
"	they began to sing

the reference item *they* being simply the most general of all. Compare *python* . . . *snake* in the verse quoted in Appendix 3 below (. . . *who bought a Python from a man . . . the Snake is living yet*); and *pig* . . . *creature* in the following passage from *Alice*:

This time there could be no mistake about it; it was neither more nor less than a pig, and she felt that it would be quite absurd for her to carry it any further.

So she set the little creature down, and . . .

Such instances are typically accompanied by the reference item *the*. This interaction between lexical cohesion and reference (*the pig . . . the creature . . . it*) is the principal means for tracking a participant through the discourse.

Related to these are examples such as the following, where there is still identity of reference, although not to a participant, and the synonym may not be in the same word class (*cheered . . . applause; cried . . . tears*):

Everyone cheered. The leader acknowledged the applause.

I wish I hadn't cried so much! I shall be punished for it, I suppose, by being drowned in my own tears!

(b) without necessary identity of reference. The occurrence of a synonym even where there is no particular referential relation is still cohesive; for example

There was a man of Thessaly
 And he was wondrous wise.
 He jumped into a hawthorn bush
 And scratched out both his eyes.
 And when he saw his eyes were out
 With all his might and main
 He jumped into a quickset hedge
 And scratched them in again.

where the quickset hedge is not the same entity as the hawthorn bush but there is still cohesion between the synonyms *hedge* and *bush*.

In this type of cohesion we find other semantic relationships, particular variants of synonymy: hyponymy (specific-general) and meronymy (part-whole). Given a lexical set consisting of either hyponyms, where *x*, *y* and *z* are all 'kinds of' *a*, or meronyms, where *p*, *q* and *r* are all 'parts of' *b*, as in Figure 9-4: the occurrence of any pair of items within the set will be cohesive; for example

Elfrida had a beautiful little glass scent-bottle. She had used up all the scent long ago; but she often used to take the little stopper out . . .

She knelt down and looked along the passage into the loveliest garden you ever saw. How she longed to get out of that dark hall, and wander about among those beds of bright flowers and those cool fountains, . . .

where *stopper* is a meronym of *bottle*, and *flowers* and *fountains* are co-meronyms of *garden*. Examples of hyponymy:

Then they began to meet vegetation — prickly cactus-like plants and coarse grass . . .

The chessmen were walking about, two and two!

"Here are the Red King and the Red Queen," Alice said . . .

where *plants* and *grass* are co-hyponyms of *vegetation*, and *Red King* and *Red Queen* are co-hyponyms of *chessmen*. There is no very clear line between meronymy and hyponymy, especially with abstract terms; and a given set of items may be co-hyponyms of one term but co-meronyms of another — for example *chair*, *table*, *bed* are 'kinds' (hyponyms) of *furniture*, but 'parts' (meronyms) of *furnishings*; *forward*, *half-back*, *back* are 'kinds' of *players* but 'parts' of a *team*, and so on. But since either relationship is a source of lexical cohesion it is not necessary to insist on deciding between them.

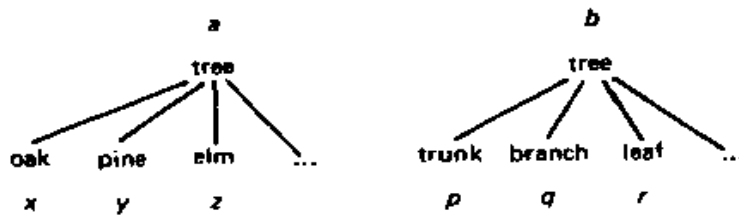


Fig. 9-4 Hyponymy and meronymy

Finally a special case of synonymy is its opposite, antonymy. Lexical items which are opposite in meaning, namely antonyms, also function with cohesive effect in a text. For example, *woke* and *asleep* in

He fell asleep. What woke him was a loud crash.

(3) Collocation. At the same time there are other instances of lexical cohesion which do not depend on any general semantic relationship of the types just discussed, but rather on a particular association between the items in question — a tendency to co-occur. This ‘co-occurrence tendency’ is known as COLLOCATION. For example,

A little fat man of Bombay
 Was smoking one very hot day.
 But a bird called a snipe
 Flew away with his pipe,
 Which vexed the fat man of Bombay.

There is a strong collocational bond between *smoke* and *pipe*, which makes the occurrence of *pipe* in line 4 cohesive.

Clearly there is a semantic basis to a collocation of this kind; a pipe is something you smoke, and the words *pipe* and *smoke* are typically related as Range to Process in a behavioural process clause. Hence *pipe* here will be interpreted as ‘the pipe that he was smoking at the time’. But the relationship is at the same time a direct association between the words; if *pipe* is in the text then *smoke* may well be somewhere around, at least with considerably greater probability than if we just pulled words out of a hat on the basis of their overall frequency in the language. We get ready for it, so to speak; and hence if it does occur it is strongly cohesive.

As a matter of fact, even where there is a relation of synonymy between lexical items, their cohesive effect tends to depend more on collocation, a simple tendency to co-occur. Of course if both relationships are present they reinforce each other; but if a pair of synonyms are not regular collocates their cohesive effect is fairly weak, whereas words which are closely associated but without any systematic semantic relationship are nevertheless likely to have a noticeably cohesive effect. This is because collocation is one of the factors on which we build our expectations of what is to come next.

So for example there is a strong collocational bond between *cold* and *ice*, but not nearly so strong between *cold* and *snow*, though it would make just as good sense; *snow* is more likely to conjure up *white*. We collocate *friends* and *relations*, and also *friends* and *neighbours*; but not very often *relations* and *neighbours*, although

family and *neighbourhood* seem to be associated. The extreme cases of such collocational patterns are to be found in fixed phrases and clichés, like *flesh and blood*, *stretch of the imagination*; but these actually contribute little to cohesion, since they are so closely bound together that they behave almost like single lexical items.

Notice finally that collocations are often fairly specifically associated with one or another particular register, or functional variety of the language. This is true, of course, of individual lexical items, many of which we regard as 'technical' because they appear exclusively, or almost exclusively, in one kind of text. But it is also noteworthy that perfectly ordinary lexical items often appear in different collocations according to the text variety. For example *hunting*, in a story of the English aristocracy, will call up *quarry* and *hounds* (or, at another level, *shooting* and *fishing*); in an anthropological text, words like *gathering*, *agricultural* and *pastoral*; as well as, in other contexts, *bargain*, *souvenir*, *fortune* and suchlike.

Figure 9-5 is an example of a text marked for lexical cohesion, using the categories of repetition, synonymy and collocation.

9.6 *The creation of texture*

We have identified the following features as those which combine to make up the 'textual' component in the grammar of English:

(A) structural

- 1 thematic structure: Theme and Rheme (Chapter 3)
- 2 information structure and focus: Given and New (Chapter 8)

(B) cohesive (Chapter 9)

- 1 reference
- 2 ellipsis and substitution
- 3 conjunction
- 4 lexical cohesion

These are the resources that give 'texture' to a piece of discourse, without which it would not be discourse. In order to do this, these resources are deployed in certain ways; ways which vary considerably according to the register of the text, but about which it is possible to make some general observations as well.

We do not ordinarily meet with language that is not textured. What we call 'nonsense' is something we disagree with; but it is perfectly adequate as discourse — otherwise there would be nothing with which to disagree. (We have the notion of 'incoherent', but this usually refers to the slurred speech of the temporarily deranged.) People go to great lengths to interpret as text anything that is said or written, and are ready to assume any kind of displacement — some error in production, or in their own understanding — rather than admit that they are being faced with 'non-text'. Like everything else we have been investigating, this is an unconscious process; we are not aware of making such adjustments when we listen or read. But it is sometimes brought to consciousness by marginal instances which one has to work hard at decoding: strange children, foreign learners, faulty translations and the like.

One way to see how these resources work is to deconstruct a text, destroying its

Peter rushed straight up to the monster and aimed a slash of his sword at its side. That **stroke** never reached the Wolf. Quick as **lightning** it turned round, its eyes flaming, and its mouth wide open in a howl of anger. If it had not been so **angry** that it simply had to **howl** it would have got him by the **throat** at once. As it was – though all this happened too **quickly** for Peter to think at all – he had just **time** to duck down and plunge his **sword**, as hard as he could, between the **brute's** forelegs into its heart. Then came a horrible, confused **moment** like something in a **nightmare**. He was tugging and **pulling** and the Wolf seemed neither alive nor **dead**, and its bared **teeth** knocked against his forehead, and everything was blood and heat and hair. A **moment** later he found that the **monster** lay **dead**.

Coll = collocation
Rep = repetition
S: = synonymy

S id = identity (of reference)
S hyp = hyponymy
S mer = meronymy

Fig. 9-5 Text analysed for lexical cohesion

textual patterns one by one. Here is the North Star text (see Chapter 8, Section 8.5) after surgery:

The magnet is at the North Star. The earth attracts the North Star. The earth does not attract the stars which are not the North Star. The stars which are not the North Star move around.

In this case we simply removed the cohesion and selected the unmarked options in the various textual systems. If instead we were to select an option at random (as distinct from the unmarked option), we might end up with something like the following version of the 'silver' text (see Appendix 1):

With silver we, Anne, are dealing in this job. What needs to have love is silver. Silver is loved by the people that buy silver. It is silver that silver has a lovely gleam about. The people who love beautiful things are usually people if people come in.

This is, of course, an artificial exercise, set up for purposes of highlighting the textual component of meaning. In real life the different 'metafunctions' are so closely interwoven into the fabric of discourse that it is difficult to conceive of one being disturbed while the others remain unaffected – although certain aphasic conditions may approximate to such a pattern.

In the remainder of this section we attempt a brief summary of the part played by the features listed above in the creation of text. We will group them under four headings: (1) theme and focus; (2) lexical cohesion and reference; (3) ellipsis and substitution; (4) conjunction; with a note on text structure at the end.

(1) Theme and focus. These are the manifestations in English of what the Prague linguists of the 1930s, who were the first to explore this area of grammar systematically, called 'functional sentence perspective (FSP)'.

(a) The choice of Theme. The choice of Theme, clause by clause, is what carries forward the development of the text as a whole. This point was made in Chapter 3 and is illustrated by the texts examined there and in Appendix 1.

The patterning of clause Themes throughout a text tends to differ from one register to another. In narrative and expository texts it is quite likely for the same participant (whose 'sameness' is expressed lexicoreferentially; see (2) below) to remain as topical Theme for a certain stretch of discourse: either a protagonist in the tale, if it is narrative, or that which is being expounded, in an expository context. In texts with a more stepwise structure, involving sequences of instructions or logical argument, one is more likely to find the Theme of one clause selected from within the Rheme of the clause preceding; and there are likely to be conjunctive Themes. In dialogue, there may be alternation of Themes, especially between *I* and *you* representing speaker and listener; and Finite and WH- Themes, in interrogative clauses.

(b) The choice of focus. The choice of information focus, by contrast, expresses the main point of the information unit, what it is that the speaker is presenting as news; the pattern of focus throughout the text likewise expresses the main point of the discourse. In speech, the focus is realized by tonic prominence; it typically falls on the final lexical element, in the clause or in whatever unit is matched with the information unit, although it can be 'marked' and put anywhere. In writing, the principle is that (i) the information unit is a clause, unless some other unit is clearly designated by the punctuation; and (ii) the focus falls at the end of the unit, unless some positive signal to the contrary is given, either by lexical cohesion (no focus on repeated word) or by grammatical structure (predication: *it is . . . that . . .*).

(c) The combination of Theme and focus. Since the unmarked place of focus is at the end of the information unit, and since the unmarked information structure is 'one information unit one clause', this gives a kind of diminuendo-crescendo movement to the typical clause of English: the downward movement from initial, thematic prominence being caught up in the upward movement towards final, informational prominence, as shown in Figure 9-6. Note how this gives to each message (Chapter 3) the character of a move in an exchange (Chapter 4).

The two kinds of prominence are complementary. The Theme, as pointed out in Chapter 8, is speaker-oriented prominence; it is 'what I am starting from'. The New, which culminates in the focus, is listener-oriented prominence: it is 'what I am asking you to attend to'. As the clause moves away from the first peak, it moves towards the second; and this imparts a small-scale periodic or wave-like movement to the discourse. Larger-scale periodicity may then be superimposed on this, for example by a similar overall pattern in the paragraph.

(2) Lexical cohesion and reference. An important characteristic of many varieties

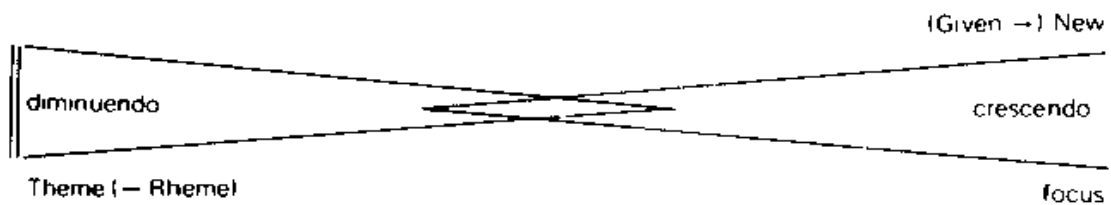


Fig. 9-6 From speaker to listener: the wave-like effect of thematic and focal prominence

of text is the referential chain, produced by a combination of lexical cohesion (repetition and synonymy) and reference. A typical chain from a narrative might be:

A little boy called John . . . John . . . he . . . the lad . . . him . . .

These are sometimes called 'participant chains'; but they are not restricted to participants in the sense of persons — they may be objects, institutions, abstractions, passages of text: anything that can have a participant role in a transitivity structure. Similar chains, though less frequent and less extensive, can be formed with circumstantial elements, and even with the process itself, e.g. *run away . . . do that . . . do it . . . get away . . . escape altogether*.

What gives the text its coherence, however, is not simply the presence of such chains but their interaction one with another. If the tokens (individual occurrences) in one chain relate to the tokens in another chain by some grammatically definable relationship (most typically, perhaps, a relationship in transitivity, because that is where the most highly structured configurations are found), this is strongly cohesive; for example Process *drown* + Medium *fish . . . deadly stonefish . . . it* in text 1 in Chapter 4, Section 4.8 above. Typically such interlocking chains overlap, one taking over from another, like *drown* + *mermaid* → *drown* + *fish* → *fish* + *eat* in the same text; and this is one of the sources of the dynamic flow of discourse.

Like other text-forming patterns, these referential chains and their interlocking chain complexes vary in kind and extent from one register to another. They have been most studied in narrative, but they feature in other types of text besides.

(3) Ellipsis and substitution. If reference, and referential chains, are more typical of narrative, ellipsis and substitution are more characteristically found in dialogue, where the typical sequence is based on pairs, or triads, or longer structures, that are related not so much by ideational as by interpersonal meaning: request → assent, question → answer → acknowledgment, statement → challenge → justification → qualified acceptance, and so on. In sequences of this kind the dynamic comes from the constant shifting in the role relationships among the interactants; and this means that, rather than (or, at least, in addition to) the persistence of identical referents, there is likely to be the sort of 'same but different' semantic relation that is typically maintained by ellipsis or substitution: the same process but different polarity or modality, the same class of entity but different member, different deixis or so on.

Typically this kind of cohesion is also accompanied by cohesion among lexical items; this may perhaps depend, relatively, more on collocation and less on structural semantic relations like synonymy, the cohesive force of collocation being much more localized. In the same way the textual 'reach' of ellipsis and substitution is considerably shorter than that of reference. On the whole, types of cohesion with

a more local effect, ellipsis/substitution and collocation, tend to be associated with dialogue; those with a more global effect, reference and synonymy, with monologue; although these are no more than very general tendencies.

(4) Conjunction. The difference between conjunction and the other text-forming resources is that conjunctive relations are essentially relations between messages or between larger complexes that are themselves constructed out of messages. As has been shown, the logical-semantic relationships that are coded in the form of conjunction are also manifested in many other ways (see also Appendix 3 below).

As a cohesive resource conjunction works in two ways, once again corresponding to the distinction between the ideational and the interpersonal metafunctions.

(a) External (ideational) conjunction. This sets up a relationship between processes. A simple pattern of this kind is that of a sequence of events shown as following one another in time, e.g. *first* [this happened], *next* [that happened], *finally* [the other happened]. All the conjunctive relations set out in Section 9.4 above may function in this way.

(b) Internal (interpersonal) conjunction. This sets up a relationship between propositions or proposals; for example *first* [I say this], *next* [I say that], *finally* [I say the other]. Here the semantic relations are between the steps in an argument, not between phenomena of experience.

Not all conjunctive categories have an 'internal' interpretation; and in some cases, particularly elaboration and certain types of extension, it is often hard to tell the internal and the external apart. Despite these indeterminate instances, the distinction is a valid one, and important to the creation of texture. Different registers vary both in their overall use of conjunction and in their orientation to that of an internal or external kind.

The line between conjunction and paratactic expansion is a fuzzy one; many instances could be interpreted as either. This is reflected in the fact that, in writing, it is often possible to write either So . . . (which we should interpret as conjunction) or . . . , so . . . (which we should interpret as parataxis), with little difference in meaning. But this kind of overdetermination is found throughout the linguistic system, and particularly in the grammar of very general and fundamental semantic relations such as those of expansion and projection.

(5) Text structure. With the clause complex, described in Chapter 7, we reached the upper limits of grammatical structure. The sentence, evolving as a unit of written language, embodies the unconscious awareness of that upper limit.

This does not mean that there are no lexicogrammatical relations obtaining over larger domains; as we have been seeing in this chapter, the semantic relations of coreference, synonymy, expansion and so on are manifested in lexicogrammatical items and patterns just as systematically as the semantics of processes or speech functions. But whereas the latter are realized through grammatical structures, the former are not, or not necessarily. There is no structural relationship between, say, two occurrences of a lexical item, or between *John* and *he* — the members of such pairs are not linked in any constructional pattern. It is this non-structural relationship to which we give the name of cohesion.

Is there then no structure above the clause complex? There is; but not

grammatical structure. A text has structure, but it is semantic structure, not grammatical. Just as a syllable has a phonological structure, and a clause has a grammatical structure, a text has a semantic structure; but while the concept of structure is the same, the level at which it is 'coded' is different. So a text does not consist of clause complexes. It consists of elements of its own, which vary from one register to another: narrative, transactional, expository and so on. Each has its own elements and configurations — which are (or whose own smaller constituents are) realized as clauses or clause complexes in the same way that, say, morphemes, which are the smallest constituents in the grammar, are realized as syllables or syllable complexes.

For a text to be coherent, it must be cohesive; but it must be more besides. It must deploy the resources of cohesion in ways that are motivated by the register of which it is an instance; it must be semantically appropriate, with lexicogrammatical realizations to match (i.e. it must make sense); and it must have structure. But to say this is not in any way to imply that it must be homogeneous, univocal or 'flat'. Discourse is a multidimensional process; 'a text', which is the product of that process, embodies not only the same kind of polyphonic structuring as is found in the grammar (for example in the structure of the clause, as message, exchange and representation), but also, since it is functioning at a higher level of the code, as the realization of semiotic orders 'above' the language, all the inconsistencies, contradictions and conflicts that can exist within and between such higher-order semiotic systems. Because it has this potential, a text is not a mere reflection of what lies beyond; it is an active partner in the reality-making and reality-changing processes.