are generally not jointly produced in the process of interaction. Nevertheless, here, too, there is obviously interpersonal involvement. The writer seeks to engage the reader and does so by making appeal to the conventions that define particular genres which are assumed to be common knowledge. So if I open up my newspaper to read a football report, or look up a menu in a cookery book, or consult a manual of instructions for assembling equipment, I will have some idea of what to expect because I know how football reports, menus, and instructions are typically written. Or, if I venture into writing myself and decide to offer an article to a learned journal, reporting on a piece of research, I know that I will be expected to structure it in conformity to the genre which has been established by the **discourse community** of scholars in this area of learning as appropriate for this kind of written communication. This might require me to begin by locating my own study in the context of current research in the field, to follow a certain format in describing the design of my investigation, to present my findings in a certain way, and so on. Compliance with such requirements can be seen as a condition of membership of this particular discourse community, and of publication in this particular journal.

Conclusion

How strict the compliance has to be will vary, of course. Genre conventions are by no means rigidly fixed and always adhered to: they are naturally subject to variation and change because there will always be some room for individual manoeuvre. People will have schematic knowledge of what is typical of particular genres and this will prime their expectations. But these expectations may need to be subsequently adjusted. The ideational and interpersonal schemata we have been considering in this chapter are relatively stable knowledge structures or states of mind, customized or conventionalized as normal in a particular community. But we also need to consider how they are put to work, made operational in the production of actual text. This is the concern of the next chapter.

5
Co-textual relations

Information structure

As was pointed out earlier (in Chapter 2), one of the things we do when we use language is to formulate a proposition, to make reference to some state of affairs. Let us suppose that we want to express a proposition about a certain event, a demonstration, for example, and the actions of the police in dispersing the crowd. English allows for the possibility of expressing our proposition in different ways, for example:

The police dispersed the demonstrators.
The demonstrators were dispersed by the police.

If we think of these as linguistic forms, we recognize the first as an active and the second as a passive sentence with *the police* being the subject in the first case and *the demonstrators* in the second. But if we think of these expressions as utterances, they are textual variants, different ways of distributing the propositional information. In terms of textual structure, the first piece of information, which here takes the form of a subject noun phrase, is said to be the **theme** and the rest of the utterance the **rheme**. But what if we want to go on and say something else? How would we order the information in the next utterance? We could follow the same pattern and start with the same theme (T):

The police(T) dispersed the demonstrators. Some of the law officers(T) ...

Or we might reverse the order and thematize the previous rheme (R):

The police(T) dispersed the demonstrators(R). The banners they were carrying(R→T) ...

Or we could, of course, introduce a different theme altogether:

The police(T1) dispersed the crowd. The motorcade(T2) ...

It needs to be noticed, however, that whether a theme is different or not is by no means easy to determine. It may be a matter of simply identifying a semantic connection (recognizing *police* and *law officers* as synonymous, for example), but schematic knowledge may also be involved. There is, for example, no semantic link between the words *demonstrator* and *banner*: to recognize the thematic connection here, you need to invoke what you know about demonstrations.

The alternative thematizations we have been considering are brought about by changes in grammatical structure, and the use of active and passive forms is only one way of signalling theme and rheme. There are other possibilities, for example the use of sentences like:

What the police did(T) was to disperse the demonstrators.

But thematization can also be brought about by changing the sequence of constituents in a sentence without affecting its structure, as in:

The police(T) dispersed the demonstrators early in the day.
Early in the day(T), the police dispersed the demonstrators.

English, then, in common with other languages, makes provision within its encoding conventions for expressing propositional meaning in different ways. The question arises as to why the users of the language would want to take advantage of this provision. Theme and rheme can be easily identified as a feature of text, but what is it that motivates the use of one textual variant rather than another? What communicative intention lies behind them: in other words, what might their discourse function be? Here things are not so easy.

An item of information might be given the status of theme because the first-person text producer, **P1**, assumes it to be **given**, that is to say, already known by the second-person receiver, **P2**. In this case, the theme simply confirms common knowledge and sets the scene for the **new** information to be provided in the rheme. Conversely, the theme could signal the main **topic** that P1 wants to talk about, with the rheme representing **comment** on that topic. P1's use of the theme/rheme sequence to signal given/new information is clearly P2-oriented and co-operative: it expresses, we might say, consideration of P2's position (*Let me remind you of what you know already* ...). But the use of theme/rheme to signal topic/comment is quite different. This is P1-oriented, with P1 asserting his/her own position (*This is what I am going to talk about* ...).

It is clear that thematization plays a crucial role in organizing information in a text. What its significance is for discourse interpretation in particular instances, however, is not so clear, as we shall see when we return to it later, in Chapter 7.

Text linkage

Theme/rheme assignment is a general way of organizing information and carrying reference over from one proposition to the next. But the linking of theme and rheme across parts of text depends on the identification of other more specific and small-scale connections to establish text continuity. As communication takes place, in speech or writing, what is said at a particular point naturally makes reference to what has been said before and a context is created in the mind and signalled in the text in the process of its production. We saw this earlier in the passage about Tony Blair and Kosovo. The writer assumes that what is mentioned at the beginning will be kept in mind and serve as a context for what follows:

Tony Blair was on his way to Bucharest .... *The Prime Minister* astonished his advisers ...

Here the definite article signals shared knowledge and prompts the reader to refer back to a previous mention that can be connected with it. Further linkage is provided later by the use of a pronoun:

... by suddenly announcing that *he* was going to drum up local support ...

The pronoun, as its name suggests, acts as a **pro-form**, that is to say, it stands in for the fuller expression that precedes it. This kind of text-internal, or **co-textual** connection is known as **anaphora**.

Anaphora and pro-forms

It should be noted that this anaphoric connection is not directly signalled by the linguistic pro-forms themselves. It has to be inferred by identifying and interpreting the cross-reference. Take the pronoun *he*. This encodes the semantic features of singular and masculine and it works anaphorically only if these features can be traced back to realize a reference in common with a previous mention. In the present case, this is easy to do because there is only one possible candidate: *he* is identified as having a referential link to *The Prime Minister* and *Tony Blair*. But there might be, and indeed quite commonly is, more than one possible candidate for referential connection. Consider another pronoun, *its*, as it occurs in the second part of our text:

> The Prime Minister astonished his advisers by suddenly announcing on the aeroplane that he was going to promise Romania early membership of the European Union in return for *its* continued backing.

Its encodes the semantic features of singular and non-human and these features are to be found in *Romania, the European Union*, and even in *membership of the European Union*, so it could be linked linguistically to any of these three. The appropriate anaphoric connection is a matter of inferring which makes most sense pragmatically, which corresponds most closely with the reader's contextual knowledge of the world—in other words, it is a matter of discourse interpretation. The writer here assumes that such contextual knowledge will indeed be activated, but if in doubt, he could, of course, have made this correspondence more textually explicit by simply repeating the previous mention:

> ... in return for *Romania's* continued backing.

Or copying more **semantic features** from the previous mention in a more explicit pro-form as in:

> ... in return for *that country's* continued backing.

We return here to the points made in Chapter 2 about feasibility. For in making intended meaning more evident by a more explicit pro-form, the writer makes it easier to process.

There are occasions, of course, when writers would have been better advised to avoid minimal pro-forms and provide more explicit textualization of their discourse intentions to increase feasibility and avoid misunderstanding, or unintended ambiguity. Consider the following text:

> Unfortunately in the weeks to come autumn leaves will create a dangerous hazard, especially to the elderly when they fall and become a soggy mess on the pavement.

Here the pronoun *they* copies only the semantic feature of plurality (unlike the singular pronouns, *he, she, it*, the plural is not marked for gender in English). There are two plural noun phrases in the text which are linguistic candidates for co-textual linkage: *autumn leaves* and *the elderly*. The intended anaphoric connection (one must suppose) is between *they* and *autumn leaves*. But an alternative interpretation is possible, of course by calling up a fanciful context which sets up an anaphoric connection with the pronoun and *the elderly*, so that it is they who fall and become a soggy mess on the pavement. Hence the comic effect.

Cohesion

The identification of connections that are linguistically signalled, like those between a pronoun and a previous noun phrase, enables us to recognize the **cohesion** of a text. But the personal pronouns we have been considering so far are only one kind of cohesive device. There are many others. Features of a preceding noun phrase can also be copied by pro-forms that consist of nouns of more general or inclusive meaning as when, in the example given earlier, there is a cohesive link between *Romania* and *that country*. We might note that the expression *that country* does not actually refer here to the country itself as a geographical location, but is shorthand for *the government of that country*. That being so, an expression like *its government* or *the government* would also serve as a pro-form to make the cohesive link in this passage. If, however, the word *Romania* were to be

used in reference to the country itself, then a different expression would be called for. Consider the following interaction:

A We went to Romania last summer. Beautiful *country*. We loved *the place*.

B Really? We were *there* a year or two ago—didn't think much of *it*.

As we can see here, cohesive pro-forms vary in how much meaning is copied from the previous mention: the very general noun *place* copies less than the more particular noun *country*, but copies more than the pronoun *it*.

But cohesion does not only involve the kind of noun phrase or nominal replacement we have been considering so far. In this exchange, for example, the adverb *there* is clearly a pro-form that links up with the preceding *Romania, country, place*. But what it copies is the concept of location as would be expressed in prepositional phrases like *to Romania* or *in Romania*. And verb phrases can be copied into pro-forms and connected in a similar way. Consider the following:

A We went to Romania for our holidays last summer.
B We *did* too.
A We thought it was a wonderful place.
B Well, we *didn't*.

Here the verb form *did,* simply copies the features of action and past time to make the required cohesive link with the preceding verb phrase. Again, alternatives are available to make the link more explicit if necessary, as in

We were *there then* too.
So did we.
We *did the same.*
Well, we *didn't think so.*

These **cohesive devices,** then, serve to link parts of a text together. It is important to note, however, that they (i.e. these cohesive devices) do so (i.e. link parts of the texts together) so that new content is understood in relation to the context that has been established in the reader's mind by what has been said before.

That is to say, the text design has a discourse function—it is designed to key into context so as to express the message the producer has in mind.

So far we have been considering cohesive pro-forms that have an **anaphoric** function, that is to say, that work retrospectively in that they copy features from preceding expressions in a text. But we should note that pro-forms can also work prospectively as **cataphoric** devices, that is to say they can precede a more explicit mention. Thus, to take a simple example, the pronoun *he* functions anaphorically in:

When the Prime Minister was on his way to Romania, ←*he* astonished his advisers.

But it functions as a cataphoric pro-form in:

When *he*→ was on his way to Romania, the Prime Minister astonished his advisers.

Cohesion and the least effort principle

It needs to be stressed that pro-forms, whether functioning anaphorically or cataphorically, are tokens of meaning which only make sense when their relationship with what has preceded or what follows, is both identified and interpreted. This, as we have seen with cases of ambiguity, is not always successfully achieved. This might lead one to wonder why language users do not always make use of semantically fuller expressions to ensure a more explicit cohesive connection. The difficulty is that this can result in a reduction of feasibility by getting readers to process information they do not need. The general point here is that communication generally operates on a least effort principle and we only use as much language as we need to make the required contextual connection. The problem is always to know how to regulate the degree of co-textual explicitness by the judicial choice of pro-form. How much do writers need to spell out cohesive links, how far can they count on readers making sense of the text without them? Consider, for example, the following newspaper text.

The Muslim Council of Britain has set up an investigation into mosques, women's organizations and Islamic youth centres across the country to root out extremism.

Sir Iqbal Sacranie, secretary general of the council, told The Independent that the council, which has more than 400 affliliates and is the most powerful Muslim body in the country of Britain, has set up the focus groups to locate and combat the terrorist threat. Its early findings will be revealed in a national conference in September, he said.

The move comes amid allegations that the council is failing mainstream Muslims and has its roots in extremist politics.
(*Independent*, online edition, 15 August 2005)

The use of the definite article in *the country* in the first and second paragraphs here, and *the council* in the second signals that we are meant to connect these expressions anaphorically to what has preceded—to *Britain* in the first case, and to *The Muslim Council of Britain* in the second. Similarly, the pronoun *he* links anaphorically to *Sir Iqbal Sacranie*. These connections are easy enough to make and it would seem perverse to spell them out. To do so (i.e. to spell such connections, i.e. the connections between *country* and *Britain* and between *the council* and *the Muslim Council of Britain* and between *he* and *Sir Iqbal Sacranie* ...) would, as is plain to see, make the text unnecessarily cumbersome:

> Sir Iqbal Sacranie, secretary general of the Muslim Council of Britain, told The Independent that the Muslim Council of Britain, which has more than 400 affliliates and is the most powerful Muslim body in the country of Britain ...

However, the definite reference *the focus groups* is rather less straightforward. There is no preceding noun *groups,* nor any noun which can be related semantically to it. In this case, what we have to do is assume that the co-textual verb phrase *has set up* signals that *the focus groups* and *an investigation* are intended to be synonymous so that what is being referred to is the investigation mentioned in the first paragraph. Here we could argue that a more explicit co-textual link would be desirable. A similar point might be made about the anaphoric pronoun *its*.

Being singular there is no cohesion here with the plural noun phrase *the focus groups* so it is presumably meant to make a connection with *an investigation into mosques, women's organizations* ... (and not to *The Muslim Council of Britain*). We could make the text more cohesive, therefore, by rewriting it as follows:

> [The council] has set up the focus groups of the investigation to locate and combat the terrorist threat. The early findings of this investigation will be revealed in a national conference in September, Sir Iqbal Sacranie said.

Or, if we wanted to relate the reference to focus groups more explicitly to the three areas of investigation mentioned in the first paragraph (and which are presumably to be focused on), we might rewrite the text along the following lines:

> [The council] has set up groups to focus investigation on mosques, women's organizations, and Islamic youth centres in order to locate and combat the terrorist threat.

An alternative way of doing this would be to provide a lexical connection in the first paragraph by using the word *focus*. For example:

> The Muslim Council of Britain has set up an investigation which will focus attention on mosques, women's organizations, and Islamic youth centres across the country to root out extremism.

Cohesion and coherence

As we have seen, then, how far texts are made cohesive depends on the judgement of the producers of the texts (in speech or writing) about what meanings they can assume the text receivers can work out for themselves by invoking what the they know about the world. Cohesive devices are only aids to understanding and can only be effective to the extent that they enable readers (or listeners) to construct meaning that makes contextual sense to them, in other words to the extent that the cohesion in the text enables them to derive a **coherent** discourse from it. It follows

from this that it is possible for a text to be cohesive but incoherent. For example:

The process may seem complicated but actually it is not really, so long as you prepare things in advance and know what has to be done in what order. Some of the things you need you may already have, but others, of course, you may need to get. They are not always readily available and when they are they can be quite expensive. But the final result will make all the effort and cost worthwhile.

Here we have a text that is co-textually well connected with cohesive devices: *it* relates anaphorically to *the process*, *others* and *they* to *things*, *cost* links semantically to *expensive*, and so on. The trouble is that the reader cannot key the text into a context so as to make sense of it (*process*—what process? *things*—what things?). We cannot tell what this text is supposed to be about because we have no schematic frame of reference to refer these terms to. We could provide one, by giving the text a title: for example *Cooking Chicken Biryani*. Once a frame of reference is provided, the contextual connection can be made, and the meaning then falls coherently into place.

How far you can make coherent sense of a text depends, then, on how far you can relate it to a frame of reference. This is obviously true of this rather curious example (specially invented to make the point). But it is also true of any other text. Take the one we were looking at earlier about the Muslim Council. Here too an external frame of reference is presupposed. Consider the phrase *the terrorist threat*. The definite article signals shared knowledge, but there has been no mention of such a threat in the text itself. It is assumed that readers will already know about it and will relate what is said in the text to the situation in Britain at the time, the bombings in London in the summer of 2005, the danger of Islamic extremists and so on.

A frame of reference is what was referred to in the preceding chapter as an ideational schema. A passage can also be cohesive as text but lack coherence as discourse because it does not key the reader into any familiar schema of an interpersonal kind. Consider the following:

We spent our holidays in Romania. This is a country where grapes are grown. They are a kind of fruit. So are bananas. Fruit contains vitamins, and these are essential for a healthy life. So is regular exercise. Jogging is good for you. We do it every day ...

There is no shortage here of cohesive links across the different parts of the text. *They* is a pro-form that links to *grapes*, *so* is a pro-form that serves as an anaphoric token of both *a kind of fruit* and *essential for a healthy life*, *jogging* makes a semantic connection with the preceding *exercise*, and so on. But cohesive though it is, it is a very odd kind of text, and it is difficult to make any coherent sense of it. Unlike the previous passage, it is not that we cannot key the text into a familiar frame of reference, but rather that the frame of reference keeps shifting: first it seems that the passage is to be about holidays in Romania, and then the topic shifts to grapes, and then fruit in general, and then vitamins and then ... The problem here is not so much that we cannot understand what is being said but that we cannot see the point of saying it. In other words, we do not recognize its purpose because it does not correspond with the communicative conventions of any genre we are familiar with. And so we find it incoherent as discourse, cohesive though it is as text.

Conclusion

The general point to be made is that no matter how cohesive a text may be in terms of internal co-textual links that can be identified, the extent to which it is interpreted as coherent discourse will always depend on how far it can be related externally to contextual realities, to the ideational and interpersonal schemata that readers are familiar with in the particular socio-cultural world they live in.