

5 Predictive categories in expository text

Angele Tados

From: Coulthard, Malcolm (ed.). *Advances in Written Text Analysis*. London: Routledge, 1994.

The approach to text pragmatics presented here is based on research into the discourse structure of expository text. The initial corpus investigated was drawn from *A Textbook of Economics* (Hanson 1953 [1972]), and a model of discourse analysis was designed (Tados 1981) using the notion of Prediction. The corpus was later expanded to include other areas such as law, stylistics and linguistics.

THEORETICAL ASSUMPTIONS

The model of discourse analysis presented here is based on two basic assumptions. The first is that written text is interactive since two participants are involved: writer and reader, although, of course, 'the exigencies of the medium oblige one of the participants to be only represented at the writing stage, thus complicating the process for both parties' (Sinclair 1980: 255). This means that the writer takes on the roles of both addressor and addressee and incorporates the interaction within the encoding process itself (see Widdowson 1978a: 21).

The second assumption is that the writer is in agreement with the propositions expressed in the text unless s/he specifically signals detachment. So, for instance, if the writer says, 'Every commodity is nothing more than a bundle of services', s/he will be taken to be in agreement with the proposition, but if s/he says 'It has been pointed out by some economists that every commodity is nothing more than a bundle of services', s/he is overtly detaching him/herself from the proposition and attributing it to some other entity. In this latter case, s/he will at a later point be expected to give an evaluation of the proposition expressed.

THE NOTION OF PREDICTION

The term **Prediction** has previously been used in a generalized sense to refer to the activity of guessing or anticipating what will come in the text, an activity based on the reader's common-sense knowledge of the world, of content and formal schemata (Carrell 1983; Swales 1986). As used here, however, the term is much more specific: it refers to an interactional

phenomenon – a commitment made by the writer to the reader, the breaking of which will shake the credibility of the text.

Prediction is thus a prospective rhetorical device which commits the writer at one point in the text to a future discourse act. It is overtly signalled in the text and thus a piece of text which does not have a signal of Prediction cannot unambiguously commit the writer to a certain course of action.

To illustrate the notion of Prediction, let us look at the following example (the sentences have been numbered for convenience and the signals italicized).

- (1) *Two problems* arise in this case. (2) First, there is the universal alibi which exists as long as we have no independent indicator of a change in tastes . . .
- (3) Second, the possibility is admitted in theory that some demand curves might slope upwards.

(Lipsey 1963: 154)

In sentence (1) above there a specific numeral, 'two', followed by a noun of the type I have called Enumerables (see table 5.1, p. 72), whose referents in the first instance are signalled as to follow in the text. The occurrence of such a signal commits the writer to enumeration, which, in this example, comes in sentences (2) and (3).

CATEGORIES OF PREDICTION

Six categories of Prediction were identified in the initial corpus (Hanson 1953 [1972]): Enumeration; Advance Labelling; Reporting; Recapitulation; Hypo-theicality; Question. Each of these categories consists of a *pair*, the first, predictive, member (symbol V), signals the prediction which has to be fulfilled by the second, predicted, member (symbol D). A member may consist of one or more sentences in a member (see Tadros 1981, 1985).

But what do we mean by sentence? Here it is necessary to extend the notion of sentence to include not only what is traditionally conceived of as a sentence boundary, but also other stops not traditionally regarded as terminal signals – the dash and the colon – since these latter can be taken as sentential terminal signals when they separate a V from a D member. The reason for extending the traditional notion is that the dash and the colon are capable of marking major discourse patterns.

In what follows, the categories of Prediction will be discussed. Examples are drawn from the following texts:

- (1) *The Sound Pattern of English*, Chomsky and Halle 1968 (C&H).
- (2) *Salmond on Jurisprudence*, Fitzgerald 1966 (F).
- (3) *A Textbook of Economics*, Hanson 1953 [1972] (H).
- (4) *An Introduction to Positive Economics*, Lipsey 1963 (L).
- (5) *Economics: an Introductory Analysis*, Samuelson 1948 [1964] (S).
- (6) *Towards an Analysis of Discourse*, Sinclair and Coulthard 1975 (S&C).
- (7) *Stylistics and the Teaching of Literature*, Widdowson 1975 (W).

Enumeration

Enumeration is a category of Prediction in which the V member carries a signal that commits the writer to enumerate. There is of necessity more than one D member.

Criteria for V membership of Enumeration

Each criterion is both a sufficient and a necessary condition.

- (1) Where a structure has either
 - (a) a plural subject followed by a verb which demands a complement followed by a colon,
 - or
 - (b) a free clause followed by a clause binder (a word which joins a bound clause to a free clause (Sinclair 1972: 25).
- (2) Where a sentence includes a cataphoric textual place reference item such as *the following* or *as follows* in association with a plural noun.
- (3) Where a sentence includes an Enumerable (see definition below, and table 5.1) in association with a numeral, provided the information is presented as new to the context.

Before proceeding further, let us explain some terms.

'Enumerable' comprises both what we might call 'sub-technical' nouns (e.g. *advantages, reasons, aspects*, etc. as distinct from *men, women and children*) as well as discourse reference nouns (e.g. *examples, definitions, classifications*). See table 5.1 below. The important point to bear in mind is that the referents of such nouns are, in the first instance, textual, that is, other stretches of language.

'Numeral' can be exact, such as *two, three, four*, or inexact, such as *a few, several, a number of*.

'New' is glossed as that which is assumed not to be recoverable from the context. For instance, 'There are three reasons for . . .' is presented as new to the context, whereas 'The three reasons mentioned above . . .' is presented as recoverable, and hence the structure does not predict Enumeration although this may still occur.

Three types of Enumeration have been established, using the criteria above:

Type (a) Enumeration: This is isolated on criterion 1.

The major points are:

(S&C, p. 61)

This is possible under conditions when:

(H, p. 157)

In the first example the Signal of Enumeration is the colon following a structure with a plural subject and a verb that demands a complement. In the second example the colon follows a bound clause binder, 'when'. In either case not only is Enumeration predicted, but also that fulfilment will follow straightaway. Thus a syntactically incomplete sentence terminating with a colon requires syntactic completion which is provided discursively by the D member of Enumeration.

Type (b) Enumeration: This type is isolated on criterion 2. In type (b) the V member is a syntactically complete sentence, although it may have a colon. The signal is the occurrence of the textual place items *the following/as follows* when in association with a plural noun.

The following, for example, are all short story openings:

(W, p. 64)

In the above example Enumeration is predicted and it will follow without delay since the colon allows no interruption.

Type (c) Enumeration: This type is isolated on criterion 3. The V member is a syntactically complete sentence, but unlike type (a), the colon is not crucial, although it may occur. What is crucial, however, is the occurrence of a numeral, exact or inexact, in association with the Enumerable. Enumerables found in the corpus are given in table 5.1.

Table 5.1 Enumerables

adjuncts	classes	effects	motives	sources
advantages	concepts	elements	objections	stages
angles	conditions	examples	periods	suggestions
aspects	consequences	factors	points	terms
attempts	courses	features	policies	things
branches	criticisms	forms	problems	trends
categories	difficulties	influences	propositions	types
causes	disadvantages	kinds	qualities	varieties
circumstances	drawbacks	meanings	reasons	views
			senses	ways

The term 'question of law' is used in *three distinct* though related *senses*. (F, p. 66)

It will be noted that in the example above there is an exact numeral *three*, whereas in the example below the numeral is inexact, a *number of*:

In addition to insurance, there are a *number of ways* by which risks can be reduced.

(H, p. 17)

So far we have been concentrating on the V member. We will now briefly indicate how we recognize a D member of Enumeration.

Cognitively, of course, the D member will have to correspond to the V member – that is, a reason signalled must be a reason given. But to the unwary readers, of whom there are many, this might not be readily available, so, in order to help the reader recognize the enumerated text, the writer will use certain devices such as: special features of layout, numbering, punctuation, sequencing markers (*first, second, etc.*), lexical repetition and grammatical parallelism (identical sentence structures). In the example below the writer uses sequential markers (first, second) in the D members as well as grammatical parallelism and lexical repetition.

V It is useful to divide linguistic universals into *two categories*.

D (i) There are, first of all, certain 'formal universals' that determine the structure of grammars and the form and organization of rules;

D (ii) In addition, there are 'substantive universals' that define the sets of elements . . .

(C&H, p. 4)

Advance Labelling

Advance labelling is a term used here to refer to a category of Prediction in which the writer both labels and commits him/herself to perform a discourse act. Thus, if a writer says 'Let us distinguish between x and y', he is committed to showing us the distinction between the items concerned; if the writer says 'This can be illustrated by the following diagram', a prediction is set up that s/he will produce the promised diagram.

Criteria for V membership of Advance Labelling

Four criteria are given below, all of which must be satisfied to qualify for inclusion:

- (1) The sentence must contain a labelling of an act of discourse.
- (2) The labelling of the act must be prospective.
- (3) The role of the actor is not assigned elsewhere, and, therefore, remains as the writer's.
- (4) The sentence labelling the act must not include its performance. Advance Labelling is realized by (a) linear text, (b) by non-linear text, a 'table', 'diagram', 'graph' or the like, or (c) by non-linear text followed by linear text. We will now exemplify each type.

Type (a) Advance Labelling

V This analysis leads us to make *the important distinction* between real income and money income.

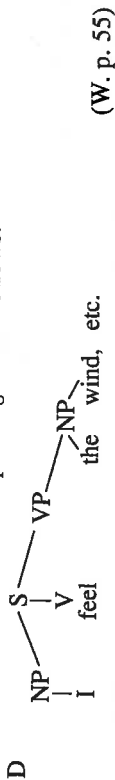
D Money income measures a person's income in terms of some monetary unit, . . .; real income measures a person's income in terms of the command over commodities which the money income confers.

(L, p. 140)

In the above example the act labelled in advance is 'to make the important distinction' and this sets up the prediction that the two terms 'real income' and 'money income' will be distinguished. The prediction is fulfilled in the D member.

Type (b) Advance Labelling

9 V We can show this in a simple diagram as follows:



Type (c) Advance Labelling

V Consider now the following cost schedule of a firm:

TABLE XXVIII
The Cost Schedule of a Firm (2)

Output	Total cost	Average cost	Marginal cost
Units	£	£	£
20	270	13.5	-
30	330	11.0	6
40	400	10.0	7
50	500	10.0	10
60	630	10.5	13
70	840	12.0	21

Db The table shows that if average cost is falling, marginal cost will be less than average cost; if, however, average cost is rising, the marginal cost will be greater than average cost. It also shows ...

(H, p. 231)

Reporting

We mentioned earlier that a basic assumption is that the writer avers the opinions and ideas of the text so long as s/he does not specifically detach him/herself from the embedded propositions expressed. The writer detaches him/herself from propositions by attributing them to others. This detachment predicts involvement, which means that the writer will come again into the text in order to declare his/her state of knowledge as regards what s/he is reporting. I have termed this 'Evaluation' to be taken in the broad sense in which Labov and Fanshel (1977) use the term:

The term *evaluation* here appears as a superordinate term that includes agreement, disagreement and more extended types of evaluation; it comprises both cognitive and evaluative types of response.

(ibid.: 101)

Criteria for V membership of Reporting

All the criteria given below must be satisfied to qualify a sentence as a realization of Reporting:

- (1) The sentence must contain at least one Report Structure. The typical Report Structure is a pair of reporting and reported clauses as in 'Those who support the bargaining Theory of Wages assert that ...' (H, p. 315), but a quoting/quoted pair can also occur as in the example on p. 80 below, and a specialized adjunct as in 'In their view', or 'According to Alfred Marshall'.
- (2) The sentence must contain propositional content which is attributed to others.
- (3) The writer must detach himself from what he is reporting, i.e. if he says 'As x said' or 'x has rightly pointed out' there is no detachment here from the Report Structure. There is no Prediction of Evaluation because the evaluation has already occurred.
- (4) The position of the reporting clause in its sentence and paragraph must be taken into account. Where the report is the only one in the paragraph and it comes at the end it is not predictive but is interpreted as a comment.

Reporting verbs and verb phrases that occurred in the corpus are given in table 5.2 in their base forms. A glance at the list in table 5.2 indicates its heterogeneous nature: while grammatically most of the items can take a *that* complement, quite a number take a nominal-group complement which may be followed by an appositional *that* clause, for example, 'He put forward the view that'; yet others are admitted to the group on condition that they combine with sub-technical or metadiscoursal nouns 'placing factors of production' and 'making points' or 'suggestions'.

Equally heterogeneous is their semantic behaviour. The list contains both factives (*show, realise, prove, know*) and non-factives (*claim, suggest, think, state*).

The distinction between factives and non-factives is significant for

Table 5.2 Reporting verbs and verb phrases

adopt	discuss	note	regard
agree	develop	notice	reiterate
argue	emphasize	observe	relate
assert	enunciate	oppose	say
assume	expound	place	show
base	formulate	point out	state
believe	imagine	propose	stress
boast	insist	prove	suggest
claim	know	put forward	support
consider	level	realize	think
contend	look (on, upon)	recognize	treat
declare	make	recommend	seek (to relate)
define	modify	refute	wish (to show)
disagree			

prediction. Factives, whether negative or affirmative, presuppose the truth of the proposition embedded in their complement clause, whereas with non-factives nothing is presupposed about the embedded propositions and hence the writer is not committed to their truth. (For more details see Tados, 1981, 1985.) The following example illustrates Reporting:

V *Halliday's* (1970) *discussion* of language structure and function *is pitched* at a different level. *He is concerned* . . . *His approach* . . . Halliday *insists that* . . . *He finds* . . .

(S&C, p. 12)

In the V member the writers are not presenting their own propositions, but rather they are attributing the proposition to Halliday by means of a series of detaching signals. The predicted member is the writers' evaluation, indicating their return to averral.

Recapitulation

The term 'Recapitulation' is used to refer to a member which predicts by recalling information from earlier in the text: 'It was mentioned/stated/pointed out above/in the preceding section'; or by the inferential 'then'. Recapitulation predicts that there will be new information, but not what it will be – the predicted information may take the form of contrastive particulars, further elaboration or explanation. In the V member there is a verb or a nominalization that refers to a discourse act, and generally a textual time or place item such as *already*, *in chapter* . . . , *in the section above*, *so far*, etc.

Criteria for V membership of Recapitulation

- (1) The sentence must contain either (a) a labelling of an act of discourse or (b) the inferential 'then'.
- (2) If (a), the following further criteria apply:
 - (i) the labelling must have a past-tense morpheme in the clause predicator;
 - (ii) the role of the actor must not be assigned elsewhere, but remain the writer's.
- (3) Whether (a) or (b) the sentence must not be paragraph-final, for in that case its function will be that of comment (i.e. reminder of relevance).

Table 5.3 lists verbs and verb phrases occurring in Recall signals.

In the example below, the Recall signal in the V member is 'We have said'. In the D member we find contrastive particulars, explicitly signalled by means of 'however'.

V *We have said that* the underlying representations, lexical as well as phonological, are abstract as compared with phonetic features

D There is, however, one very obvious sense in which the underlying representations are more abstract than the phonetic representations
(C&H, p. 11)

Table 5.3 Verbs and verb phrases in Recall signals

assume	examine	mention
consider	find out	note
deal with	give	notice
define	indicate	point out
discuss	make (+ a nominalization, e.g. reference)	see
emphasize		

Hypotheticality

Like Reporting, Hypotheticality is based on the notion of authorial detachment, but here the writer detaches him/herself from the world of actuality through the creation of a hypothetical world. Hypotheticality presupposes that the writer is aware of the gap between his/her conceptual world and that of the reader, and by means of this device the writer is able to set up a world where there are only two countries, two linguistic theories, in order to confine him/herself to those aspects of a situation that will enable him/her to derive a generalization.

Criteria for V membership of Hypotheticality

Each of the characterizations given below is both a necessary and a sufficient condition for V membership of Hypotheticality:

- (1) Where a sentence contains a verb like *assume*, *suppose*, *consider* and is subject to the following conditions:
 - (i) the verb is either used in the imperative or is preceded by *let us*;
 - (ii) in the case of *consider* the verb is followed by a nominalization which has no embedded propositional content (for a detailed treatment, see Tados 1980).
 - (2) Where a sentence contains the structure common in mathematics of the setting up of variables: *let* + NP + *be* + NP.
 - (3) Where a sentence contains a fictitious proper name.
 - (4) Where a sentence contains 'if + NP + VP (past verb) + NP + VP (past modal)'.
 - (5) Where a sentence contains 'if + NP + VP (present verb) + NP + VP (present or past modal)', provided that:
 - (i) the noun in the first NP does not make reference to an entity which is actual;
 - (ii) 'if' is not paraphraseable by 'whenever' in that context.
- V *Suppose* the legislator could draft rules that were absolutely clear in application: even so he could not foresee every possible situation that might arise,
- D As it is, legal uncertainty is counterbalanced by judicial flexibility.
(F, p. 40)

This example satisfies criterion 1. The signal of Hypotheticality is the imperative *suppose*.

V In order to simplify discussion of the advantages . . .

Let the two countries be Atlantis and Erewhon, and let the two commodities be cloth (typifying manufactured goods) and wheat (typifying agricultural products).

(H, p. 463)

This example illustrates criteria 2 and 3, since it has the structure *let + NP + be + NP*, together with the two fictitious names *Atlantis* and *Erewhon*.

. . . but the patterns he creates express also the very elusiveness of what he perceives. *If it were not elusive, if it could be brought within the compass of what is conventionally communicable, then . . .*

(W, p. 70)

This example illustrates 4 above. The use of the counterfactual conditional signals from the start an unreal world which is clearly at variance with the real world. This unreal world is demolished on the basis that it does not accurately mirror the real world.

In the example below, both the V and D members will be given. The V member meets criterion 5 above and the D member is a Generalization.

V *If Spenlow has an account with the Eastern Bank, Northampton, and draws a cheque for £25 in favour of Drood, who pays it into his account at the Western Bank, Exeter, this cheque will be cleared through the London Clearing House.*

D All cheques originating from banks in towns other than that where they have been paid in are sent each day to the Head Office of the payee's bank after which they go the London Clearing House.

(H, p. 405)

The signals of Hypotheticality in the V member are both the *if* clause and the fictitious entities: 'Spenlow', 'Drood', 'Eastern Bank' and 'Western Bank'. It is interesting to note that the fictitious banks are located in real cities. The moral of this tale comes in the D member.

In the D member, specific items in V are repeated in less specific terms. Now 'a cheque' or 'this cheque' becomes 'all cheques', 'the Eastern Bank' becomes 'banks' and so on. The function of the D member of Hypotheticality is, thus, to generalize from the Hypothetical statements.

Question

Question is a category of Prediction based on the underlying assumption of writer detachment. The writer detaches himself from the resolution of the disjunction of the proposition posed by the question he asks, and this detachment predicts that he will be involved at some later point to declare his state of knowledge as regards the question.

Criteria for V membership of Question

The following criteria are necessary for a V member of Question:

- (1) The sentence must have interrogative syntax.
- (2) It must occur at section level, not under the heading 'Questions'.
- (3) There must not be more than two interrogative sentences in succession, otherwise there is the implication 'not now, but later'.

It will be observed that in some texts some questions are typographically detached by occupying the position of heading or sub-heading; others do not have the heading status. In the former case the predicted member does not come immediately after the question, there is always some intervening material to prepare the way for the writer's declaration of his state of knowledge. In the latter case, the question is similar to elicitation in that there is a tendency, though this is not always the case, for the D member to follow straightaway. These two types of question are illustrated by the two examples below.

V *Is college worthwhile?*

D Education is one of society's most profitable investments. Human capital yields a return as great or greater than capital in the form of tools and buildings . . .

(S, pp. 119-20)

V *Can this statement be reconciled with a theory of scarcity?*

D Indeed, it can since . . .

(H, p. 7)

The question in the first example occupies heading status, which predicts a delayed D member. A question of this type foreshadows the existence of problems in communication. The writer eliminates the problems by trying to reduce the number of 'D-events' (Labov and Fanshel 1977), using the Socratic question technique. In other words, he tries to ensure that no terms or concepts required for the D member are unfamiliar to the reader. In the D member we find an answer to the question.

The question in the second example is different. It occurs at section level, and is not typographically detached from the rest of the text. The V member is followed by the D member straightaway, that is, there is no intervening material separating the V from the D.

Complex patterning

The six categories of Prediction discussed above should not leave the reader with the impression that texts are neatly structured into V and D members. The fact is that these members, through various combinations, are capable of yielding an interesting variety of complex patterning, when applied to long stretches of text. (For details see Tadros 1981, 1985.) Suffice it to give a few illustrations of the way Predictive categories are interrelated.

(i) Recapitulation preceding Advance Labelling:

V1 *We have examined* the economic forces operating to determine the level of national income – the balance of saving and investment.

V2 *We now turn* to the problems of how the level of national income has fluctuated, and how economists try to forecast the future.

(S, p. 250)

V1 *The previous section presented* a downward view showing how units at each rank had structures realised by units at the rank below.

V2 *This section begins* at the lowest rank and discusses the realisation and recognition of acts; . . .

(S&C, p. 62)

V1 *We have considered* aspects of literary use of language which depend on a combination of what is kept distinct in the code.

V2 *Let us now briefly review the converse:* aspects of literary discourse which depend on dividing what is normally compounded.

(W, p. 62)

(ii) Recapitulation preceding Question

V1 *So far we have been pointing out* certain linguistic peculiarities of this poem as a text.

V2 *What relevance do they have* for an understanding of the poem as discourse, as an act of communication?

(W, p. 57)

V1 *It has already been seen* that a change in demand can bring about a change in supply, and that a change in supply, may cause a change in demand.

V2 *Can it be* that the supply and demand curves are even more intimately related and, indeed are responsive to the same influences.

(H, p. 129)

(iii) Advance Labelling preceding Question

V1 *Now we must clarify* the term 'command'.

V2 *How do* commands differ from requests, wishes and so on?

(F, p. 26)

(iv) Reporting preceding Question

V1 *It is frequently asserted* today that we are living in an age of plenty, because larger quantities than ever before of all kinds of goods and services are being produced.

V2 *Can this statement* be reconciled with a theory of scarcity?

(H, p. 7)

(v) Advance Labelling preceding Hypotheticality

V1 *Let us be* somewhat more precise about Convention 2.

V2 *Suppose* that a formative belongs to the syntactic categories animate, nonhuman, exception to rule n.

(C&H, p. 174)

CONCLUSIONS

The approach to text pragmatics presented here has practical pedagogical applications, since it emphasizes the interactional relationship between writer and reader in discourse. The interaction is manifested through the use of the Predictive Categories examined. Since these Predictive Categories are common across a range of disciplines, they can be fruitfully exploited in the teaching of reading and writing to students of various disciplines.

In the area of reading, it is very important to make students aware of signals of Prediction in order to enhance their reading efficiency. They must be trained, for instance, to recognize signals of Advance Labelling so that they look for the fulfilment of the act labelled; or to exploit signals of Enumeration in order to get at the enumerated items. Through Recapitulation they are forewarned that they should link up bits of information or ideas so that they would not lose the thread of the argument; and by means of signals of Hypotheticality they are alerted to the Generalization at the end of the hypothetical excursion.

The idea of authorial detachment from propositions through Reporting is of particular significance in expository text. Students should thus be trained to distinguish between what a writer 'thinks', 'believes', 'claims', and what he says others 'think' 'believe' or 'claim'. Inability to recognize such signals leads the student to produce a statement like 'Television has made American life better' as a paraphrase of 'When television was first introduced into American society, writers and social scientists thought that this new invention would better American life'. From the above it is clear that the detaching signal was missed by the student.

Predictive Categories are also pertinent to the teaching of writing. Students must be trained to fulfil their commitments to the reader. For instance, if a student signals that s/he is going to compare X and Y, s/he should not simply produce separate descriptions of the items concerned, leaving it to the reader to arrive at the comparison him/herself; or if a student signals that there are three reasons for X, he should be committed to that number. And, when using visual material, students often throw in tables without warning or explanation. They rarely announce or interpret a table or graph: these things just occur in their writing without any reference.

The use of signals of writer detachment from propositions, or text reporting, is particularly important for students in the writing of theses or research papers. How often do students switch from text reporting to text averral without signalling! And, of course, it is an easy step from removing detaching signals, when reviewing your literature, to plagiarism. Thus it is crucial to train students to signal Reporting as well as to provide the predicted Evaluation.