PREDICTING THE CONTENT OF TEXTS

Paul Nation
Reader in Applied Linguistics
Victoria University
Wellington

Prediction of the content of a listening or reading text or paragraph is often recommended as a teaching technique and comprehension strategy (Nuttall, 1982; Watts, 1986). This prediction is usually drawn from the topic or heading of the text or from the first sentence of a paragraph. The justification for such an activity is that prediction stimulates and helps develop relevant previous knowledge which will make comprehension of the spoken or written text easier and more meaningful. The requirements for prediction are (1) a systematic or generalisable knowledge base to use in predicting and (2) the ability to recognize cues in the topic, heading or first sentence that signal what knowledge base should be used. The purpose of this article is to look at a range of knowledge bases for prediction and the types of clues that could direct access to them.

Prediction is usually a conscious procedure and will not often be used unless there is strong strategic motivation to do so. It is thus wise to see two possible goals for prediction activities. One goal is to see the activity as an end in itself. That is, the goal of the activity is to develop skill at predicting the content of the text from a few clues before the full text is read or listened to. Another goal is to see the prediction activity as a means to gradually build up the knowledge bases that among other things allow successful prediction. This second goal sees prediction skills as one of the outcomes of the prediction activities but also encourages other outcomes such as metacognitive awareness of text structure to use in writing, the development of self-questioning strategies for comprehension, and improved notetaking skills. For teachers in training, this second goal could also include the outcomes of skill in question making and designing information transfer activities. Observing learners predicting or predicting by teachers can be a useful teacher development activity.

Using a sample text, let us now look at five bases for prediction. The text chosen for the examples here is "Limestone caves" by Gillian Crook (New Zealand School Journal 1978, Part 4, No. 2: 14-16) part of which is reproduced in the appendix. It is a non-fiction text aimed at readers from 10 to 12 years old. The five bases are indefinite noun groups, conjunction relationships, topic type, genre, and background knowledge. When looking at each of these, the following points are covered: previous research and theory on the basis for prediction, how to use it, the knowledge needed to make use it, types of classroom activity, and the value of the basis for prediction and other purposes.

INDEFINITE NOUN GROUPS

The first sentence of the first paragraph is "Limestone is just one of the many kinds of rocks found in New Zealand".

The English article system signals a major distinction between specified and non-specified noun groups. The most frequent English word is the accounting for almost seven percent of the words on any page of a text. In the majority of its uses, the signals that the following noun is special in some

way because it has been mentioned before, it is one we all know about from previous experience, or it is specified by an accompanying adjective, preposition group, or some other item like a reduced or full relative clause following it.

Non-specified nouns do not have the in front of them. They may be, for example, indefinite plurals, uncountable nouns with no preceding article, or a singular countable noun preceded by a. It is likely that non-specified nouns in the predicate of a sentence will be described more specifically in the following sentence or sentences. Put another way, a typical movement of information is from generalization to specific cases or details. In our example, the indefinite noun group "just one of the many kinds of rocks found in New Zealand" has the headword one. One here is indefinite and unspecified because it is one of many and we do not know which specific one it is. We should expect that this will he described for us in the following sentence as indeed it is . " It is called a sedimentary rock. " There are only two sentences in the first paragraph.

The first sentence of the second paragraph is " Sedimentary rocks are made up of second-hand material " . Note that the indefinite noun group in the predicate of the preceding sentence " a sedimentary rock" is described in the sentence which follows it which is the first sentence of the next paragraph . The indefinite noun group in the predicate of this sentence is "second-hand material " . We would expect this noun group to be explained in the rest of the paragraph.   Our prediction is correct as the rest of the paragraph explains in four sentences how sedimentary rock is formed from older rock .

The knowledge required for this kind of prediction is not great.There is no need for a grammatical description of indefinite noun groups as learners can he taught to recognise these intuitively . The quality of the prediction is limited by the fact that it is based on noun groups and is thus of the nature "We will be given some more information about x". This means that the prediction is not likely to contain a lot of detail . In our sample text, "Limestone caves" (note that the title is an indefinite noun group), this approach to prediction is useful for 9 out of the 15 first sentences .   It is misleading in one sentence and not helpful in three others because they do not contain indefinite noun groups in the predicate, for example "Once limestone has been formed and pushed up to make land, it begins, very slowly, to wear away again". Predicting frorn the verbs in these cases would be quite successful.

CONJUNCTION RELATIONSHIPS

Each sentence in a text is usually in some kind of relationship with the sentences near it. These relationships may be signalled by a conjunction, an adverb or even a verb or noun. More often, however, they are not formally signalled. Halliday and Hasan (1967) have called these relationships , conjunction relationships, and they can be of several kinds, such as cause and effect, contrast, time sequence, alternative, amplification, exemplification, or exclusion. Nation (1984) has argued that there are two major types of relationships and thus two major types of paragraphs. In one type, the sentences are of roughly equal importance, often in an inclusion (And, And, And) relationship or a time sequence relationship (Then, Then, Then) . In the other, generally less common, type, the sentences are predominantly in weighted or unequal relationships, such as cause-effect, contrast, amplification, exemplification, or exclusion.   In these kinds of relationships one item in the relationship is given more importance for the message of text than the other. For example, in a contrast

relationship, the second part of the contrast always carries more weight than the first . For example, "You are a good student, but you did not perform well on the test" has a negative message because the negative clause is the second item in the contrast. "You did not perform well on the test, but you are a good student" has a more positive message. In a paragraph dominated by this kind of relationship, one sentence will stand out as being the one given the most emphasis by the writer. Let us now look at how these ideas can be applied to prediction.

One way of predicting what will follow a sentence is to ask, " What conjunction marker could be placed after the sentence?" For example, is the sentence "Sedimentary rocks are made up of second-hand material" likely to be followed by However, Then, As a result, That means or On the other hand? The most reasonable prediction is That means which indicates a very common, often unmarked, relationship, the amplification relationship. In this relationship, all or part of the preceding sentence is looked at in more detail in the following sentence. It can be described as a generalization - detail relationship. The prediction of the relationship also indicates whether the paragraph will be one made up of roughly equal sentences or one containing weighted relationships resulting in the sentence being the main sentence .

The amplification (That means . . .) relationship is a weighted relationship with the most important sentence always coming first. This indicates a strong likelihood that the rest of the paragraph will be involved with providing explanation and detail about the idea of sedimentary rocks being made up of second-hand material.

In order to use conjunction relationships as a basis for prediction, it is necessary to be aware of most of the range of possible relationships and typical signals. Nation (1984) makes use of ten relationships derived from Halliday and Hasan (1967). It is useful, but not necessary, to be aware of the weighting that each relationship involves. When learning the relationships it is not necessary to learn names for the relationships. Instead, signal words can be used, as in "the but relationship", or "the then relationship".

As a classroom activity, prediction using conjunction relationships would involve learners looking at the first sentence of a paragraph arnd deciding what signal or relationship would follow that sentence. "I think the first word of the next sentence would be "However". "The next sentence will be in contrast with this sentence". Learners would be encouraged to justify their decisions and to elaborate them in terms of the message in the first sentence. Knowledge of conjunction relationships is useful when guessing words from context, deciding on the main idea in a paragraph or text, and learning new conjunction markers.

The value of this knowledge basis for prediction is that it can he used in virtually every circumstance because in coherent text each sentence enters into a relationship with the sentences near it. It allows a more detailed prediction and a wider range of predictions than indefinite noun groups do.

TOPIC TYPE

Johns and Davies (1983) developed a set of 12 categories for classifying non-fiction text. The categories include among others physical structure, process (how things develop or are produced),

characteristics (what things are like), theory (as in experimental studies), instruction (how to do something), and state/situation (what happened). It is often possible to predict the topic type from the heading of the text. Here are some examples from New Zealand School Journals.

"The earth breaks forth: volcanoes in Hawaii" (Process and Characteristics)

"Flax: the rise and fall of an industry" (State-situation and Process)

"How to make your own fold-up chair" (Instruction)

"Gavin Thorley gives some advice on middle-distance running" (Instruction and principle)

Although Jonns and Davies list 12 topic types, typically only a few of these occur in a subject area such as Linguistics, Botany, or History.

The heading <u>Limestone Caves</u> suggests that we could meet the following topic types:

Characteristics   What are limestone caves like?
Process   How are limestone caves formed?

There is a vague possibility that other topic types could be met too. The text may tell about an exciting event such as an exploration involving a limestone cave (state/situation), or limestone caves may be used to demonstrate a principle or prove an hypothesis (theory).

Each topic type has its own particular constituents. So, determining the topic type gives rise to more specific predictions. The constituents of the process topic type are:
state or form of object/material location + time or stage + instrument or agent + property or structure + action (Johns and Davies, 1983 p.7).

We can turn these constituents into questions relating to limestone caves.

<u>What are the stages involved in the development of a limestone cave?</u>

<u>What material is involved at each stage?</u>

<u>Where (location) and when does each stage occur?</u>

<u>How long (time) does each stage last?</u>

<u>What (instrument) acts on the cave at each stage to bring about change?</u>

<u>What is the cave like (property or structure) at each stage?</u>

<u>What happens (action) at each stage?</u>

Instead of questions an information transfer grid could be used.

Stage	Material and Structure	Location	Time	Instrument and Action

Here are the questions that focus on the constituents of the Characteristics topic type.
What are the features of the thing described?
What is the proof that some of these features exist?
What general category does this thing fit into?
What other information is there about this thing?

By looking at the first sentence of each paragraph, it is possible to predict what constituents of the topic type are dealt with in the paragraph. For example, paragraph 1 begins "Limestone is just one of the many kinds of rocks found in New Zealand." This is a part of the characteristics topic type and will indicate the classification group that limestone fits into. Paragraph 2 is a little more difficult.

The first sentence, "Sedimentary rocks are made up of second-hand material ", could introduce a continuation of the characteristics topic type which would present features of sedimentary rocks. However, it could also introduce the process topic type which would describe the process by which sedimentary rocks are formed, and in fact, this is what it does

To use topic type as a basis for prediction, learners must be familiar with the few topic types that are relevant to their subject area This may not need to be more than five or six types These can be learned as questions or information transfer grids. The appendix contains some examples.

The strength and also the weakness of a topic type approach to prediction is that it sets up detailed far-reaching content predictions If the topic type of the text has been correctly determined, then these predictions will be very effective and will allow the systematic incorporation of background knowledge into the prediction. If the topic type has not been correctly determined or if the topic type changes without clear signalling, then misleading prediction is likely.

Familiarity with topic types has many benefits besides prediction. Questions or information transfer diagrams based on topic types can be used for reading and listening comprehension activities, for information gathering as a preparation for writing or formal speaking, and for notetaking.

GENRE

The "Limestone Caves" text occurs in a New Zealand School Journal which is a publication provided to all New Zealand schools. The journals contain a balanced mixture of articles on factual topics, and stories and poems. They are almost all specifically written for the School Journal and they are written in language that is engaging and accessible for young native speakers of English. Every article is accompanied by pictures.

This context in which the text appears allows some predictions to be made. First, there are likely to be attempts to relate the information about limestone caves to what the reader already knows, and second, the text is likely to be systematic and careful in its presentation of information. This should make predicting content more straight-forward.

A genre-based approach places importance on the function or purpose of a text because this affects what a reader should get from it and how the ideas in it are expressed. The title "Limestone Caves" indicates that this text is a "report" (Martin, 1989) and the general reference of the first sentence "Limestone is just one of the many kinds of rocks found in New Zealand", confirms this . A report tells us what a class of things is like. Reports make general statements.

Now that we have decided that the text is an example of the report genre, we can make predictions from two related sources which Swales (1990) calls our content schemata (the type of information that we could expect to find in such a text and what we already know about the topic) and our formal schemata (the language and organisation used to present the information).

Our content schemata would lead us to expect that we would find out what limestone caves are like, what they are made of, how they are made, and where we can find them. We would also expect to be told of some famous examples. Our predictions could be more detailed if we already knew something about these subtopics, for example, how they are made. The first sentence of each paragraph would more likely indicate what subtopic or which part of a subtopic was being dealt with in that paragraph. Familiarity with topic types would systematize our knowledge of content.

Our formal schemata would lead us to expect that the text would be in the present tense, but where there was a move from generalisation to specific examples, there would be a shift to the past tense. Pronoun usage may also signal the shift (Fountain and Nation, 1976). The limestone caves text however stays firmly within generalisation all the time with present tense verbs and you and we with general reference. In Biber's (1990) typology, Limestone caves is likely to be an example of "Learned exposition" which is characterised among other things by a predominance of nouns rather than pronouns, relative clauses, nominalisations, and active verbs.

In order to use genre as a basis for prediction, it is necessary to be aware of the questions that a genre approach should seek to answer and the possible range of answers. The following list is a useful beginning.

1   What is the function of the text? A text may have more than one function, for example to inform and entertain, or to inform and to sell.
2   Where does the text occur and who is the audience? For example, the text may be in an encyclopaedia for young readers.

3 . What type of text is it? Martin (1989) and Derewianka (1990) see the following general types as being useful for school children to be aware of.

Recount   (telling what happened - a focus on particular events)

Procedure (telling how something gets done - a general focus)

Description (telling what a particular thing or individual is like)

Reports (telling what classes of things are like)

Explanations (justifying and saying why something happens)

Expositions   (presenting a viewpoint with arguments)

4   What is the "tone" of the text? Is it formal or informal, detached or involved?
5   What types of conjunction relationship occur in the text?
6   What kinds of nouns and verbs are used? Are the nouns people or things? Are they general or particular? Do the verbs mainly describe happenings, perception and feeling, behaviour, description, saying?
7   What are the grammatical features of the text? What tense is used? Is the passive used? Are the sentences complex or simple? Are there complicated noun groups?

The answers to each of these questions from question 2 on need to be followed by another question - How does this reflect the function of the text? For example, if a lot of the verbs in the text are in the passive, how does this passivization support the communicative function of the text?

When predicting what will occur in a text, the 7 groups of questions listed above provide a range of aspects to cover Note how the aspects to cover in a genre approach include the bases of indefinite noun groups, conjunction relationships, and topic type that were covered in earlier sections of this paper.

Swales (1990) shows that it is possible to usefully focus on much more narrowly defined genre, such as the genre of research article introductions and academic course descriptions. This narrowing of focus results in more detailed predictions. For a particular group of learners it is thus useful to see where such a focus is appropriate.

A genre based approach provides a richer view of discourse than topic types because it stresses the relationship between communicative purpose and form (both linguistic and rhetorical form), and content. Topic type on the other hand is concerned only with the information not the form, although formal clues provide a useful indication of topic type.

Because division into different genre categories is based on consideration of a variety of factors including formal clues, rhetorical arrangement, type of information, and communicative purpose, there is a large number of possible genre divisions. These divisions can be as broad as categories such as descriptions, procedures, or recounts, and as detailed as prospectuses or house advertisements.

A genre approach to text has many benefits for reading and writing sensitively with a focus on the purpose of language use.

BACKGROUND KNOWLEDGE

The knowledge that we already have of a topic from our previous experience provides a strong basis for prediction. If we already know a lot about limestone caves, then it will not be difficult for us to predict what may occur in a text about them. Such prediction is not usually systematic, and the texts we need more help with are those where our previous knowledge of the topic is not great. Combined with systematic approaches to prediction such as those described here, background knowledge is an extremely useful basis for prediction. In all predicting activities learners should be encouraged to make the fullest possible use of background knowledge to elaborate and support systematic prediction.

This article has looked at a variety of bases for predicting the content of texts. Those bases are not mutually exclusive and could be seen as the larger components of a syllabus to develop awareness of the systems that lie behind the content and presentation of texts.

The order of presentation in this article has followed the complexity of the information involved in the various bases, with genre being the most complex.

It may be more effective however for some learners to start with topic type where the focus is on information and to go from there to more formal aspects The goal of such study should be to develop an awareness of how the multiplicity of features that make up a text are interdependent. This awareness will result in the reader bringing more to texts and thus getting more from them .

REFERENCES

Biber, D . (1990) A typology of English texts . Linguistics 27: 3 - 4 3 .

Derewianka, B . (1990) Exploring How Texts Work Primary English Teaching Association, Rozelle, NSW, Australia.

Fountain, R . L . and I . S . P . Nation (1976) Formal signals of discourse types. Te Reo 19: 3-23.

Halliday, M.A.K. and R. Hasan (1967) Cohesion in English Longman, London .

Halliday, M. A .K. (1985) An Introduction to Functional Grammar. Edward Arnold, London.

Johns, T. and F. Davies (1983) Text as a Vehicle for information: the classroom use of written texts in teaching reading in a foreign language. Reading in a Foreign Language 1, 1: 1-19.

Martin, J.R. (1989) Factual Writing: Exploring and Challenging Social Reality (Oxford University Press, Oxford.

Nation, I . S . P . (1984) Understanding paragraphs . Language Learning and Communication 3, 1: 61-68.

Nuttall, C . (1982) Teaching Reading Skills in a Foreign Language. Heinemann, London .

Swales, J.M. (1990) Genre Analysis Cambridge University Press, Cambridge .

Watts, N.R. (1986) Developing aural anticipation and prediction strategies. English Teaching Forum 24, 1: 21-23, 29.

APPENDIX 1

LIMESTONE CAVES
Gillian Crook

Limestone is just one of the many kinds of rocks found in New Zealand. It is called sedimentary rock.

Sedimentary rocks are made up of second-hand material. Old rocks are gradually worn down by the weather, heat, rain, wind and frost. The rock cracks and breaks up into smaller pieces. These may get washed into rivers, and here water grinds them down even smaller. These small pieces of rock are carried along by the water and will eventually find their way into lakes or the sea They settle on the bottom as sediments.

Limestone is made of second-hand material too; but it's not always old rock. Limestone is made up of lots and lots of what were once sea creatures.

If you look along the sea shore you'll see animals like mussels, pipis and periwinkles attached to the rocks. The hard part of these animals - the shell - is made mainly from a chemical called calcium carbonate The animals are able to take this chemical out of sea water and use it to make their shells.

When the sea creatures die, most of their soft parts are eaten and the rest decays. The hard parts remain. These hard parts settle to the sea floor as sediments. They may form large shell banks in shallow seas, coral reefs, or they may break up into little pieces and form muds in deeper water. In some places, lots of these shells are washed ashore and form shell beaches. There are quite a few shell beaches around the New Zealand coast.

APPENDIX 2

Some common topic types

Physical structure
  What are the parts?
  Where are the parts located?
  What are they like?
  What do they do?

Instruction
  What are the steps involved?
  What materials and equipment are needed?
  What do we need to be careful about at some steps?
  What is the result of the steps?
  What does this result show?

State / Situation
  What are the people etc involved ?
  What time and place are involved?
  What is the background leading up to the happening?
  What happened?
  What are the effects of this happening?

Principle
  What is the law or principle involved?
  Under what conditions does the principle apply?
  What are some examples of the principle in action?
  How can we check to see that the principle is in action?
  How can we apply the principle?

Theory
  What is the hypothesis?
  What lead to this hypothesis?
  How is it tested?
  What are the results of testing?
  What is the significance of the results?