

## GUIDELINES TO CHECK COHERENCE IN EXPLANATORY TEXTS

Graziella Tonfoni - Bertram Bruce (\*)

### RIASSUNTO

Questo articolo intende presentare un modello per il controllo e la verifica dei livelli di coerenza nei testi di carattere esplicativo. Attraverso la identificazione di diversi elementi e fattori quali determinatori della coerenza, tale modello fornisce uno schema per la identificazione dei problemi e le relative strategie di ristabilimento della coerenza.

### ABSTRACT

This paper deals with the problem of coherence in Explanatory Texts. The model which has been presented gives a set of evaluation criteria in order to establish the level of coherence which has been reached by a certain text, to check whether such level is adequate and to increase the level of coherence if such is not the case by using the guidelines proposed by the computational model.

### 1. INTRODUCTION

A substantial part of recent research on reading has focused on the importance of well-structured texts in both comprehension of a text and subsequent recall of the information the text contains (see Anderson, Osborn & Tierney, 1981). This research has produced a variety of methods for analyzing texts of various types: stories, non-fictional narratives, persuasive essays, and explanatory texts. Of all these, texts whose purpose is to explain (or inform) are perhaps the most important for learning in school. Yet, while there is a rich literature describing methods for story analysis (Bruce & Newman, 1978; Mandler & Johnson, 1977; Rumelhart, 1975; Stein-Glenn, 1978; Wilensky, 1978), the methods for analysis of explanatory texts are either partial or overly general. Important contributions to resolving this problem have been

made by (Anderson, Spiro & Andersonb, 1978; Goetz & Armbruster, 1980; Tierney, Mosenthal & Kantor, 1980). Armbruster, 1980; Tierney, Rosenthal & Kantor, 1980).

There are many reasons to want to have better methods of analysis for explanatory texts. Tonfoni (1982), for example, has shown that coherent explanatory texts are easier to comprehend, summarize and remember. Problems that children have in learning from texts may be attributable to the incoherence of the texts they are expected to master. It would help in teaching if one could determine whether and in what ways a particular text lacked coherence. Moreover, knowing just how a text failed to be coherent could be useful in teaching strategies to a reader for coping with the "inconsiderateness" (Anderson, in press) of the text. Publishers, who produce texts, could well use guidelines that would ensure coherence in the context of satisfying the numerous demands placed upon them in producing school texts.

Those who select texts for children to read could likewise use help in determining what makes a text coherent.

This paper draws from research on coherence in explanatory texts (Tonfoni, 1982). It presents a set of principles for examining texts that can be used to determine relative coherence and also to point to areas in which a particular text might be improved. The principles, or guidelines, define a method of analysis that is consistent with, but more precise than, our intuitive notion of coherence. We hope that these guidelines will be improved as the result of application of the guidelines to school texts and further research on text analysis.

### 2. WHAT DO WE MEAN BY EXPLANATORY TEXT?

We can define an Explanatory Text as a text whose main goal is to explain something, to provide information, or to teach something about a specific topic.

(\*) G. Tonfoni: Bolt Beranek and Newman, Cambridge, MA, USA - Università di Bologna, Italia.

B. Bruce: Bolt Beranek and Newman, Cambridge, MA, USA.

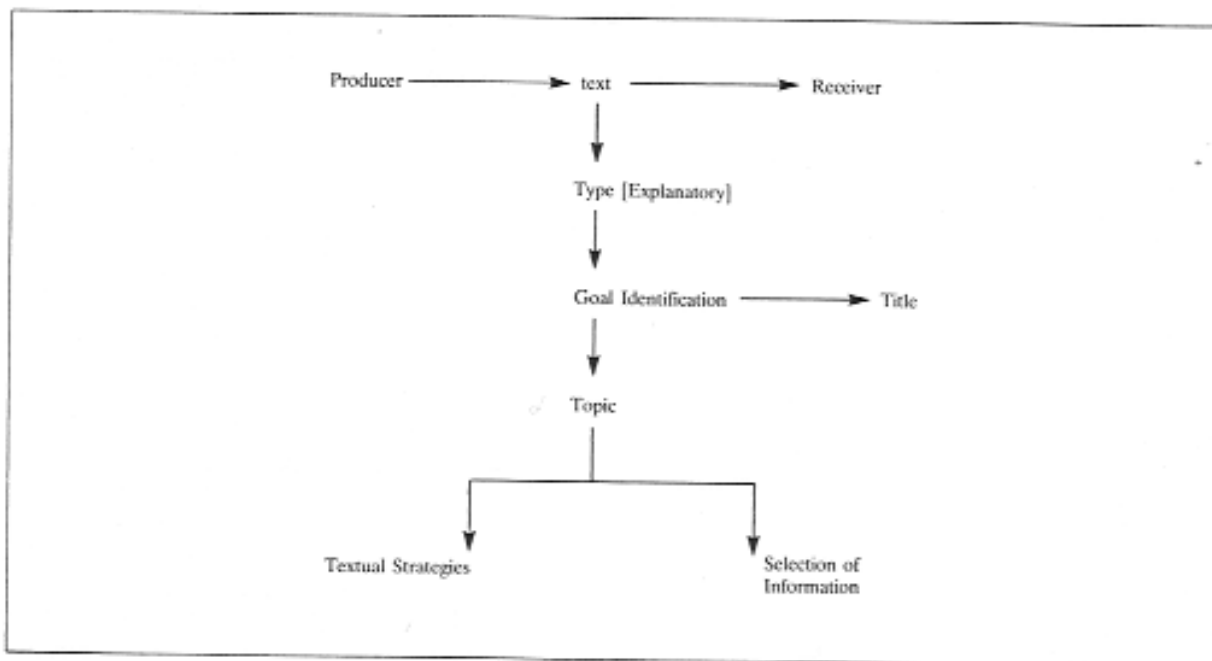


Fig. 1

Writing a text that has explanatory goals implies the use of specific constraints and rules on Text-production relative to those goals. The author must be aware of which information he or she will select to have the reader understand from reading the text.

Explanatory texts can have different levels of specificity and different amounts of technical details; different topics determine a different choice of writing strategies.

An Explanatory Text can also have different levels of redundancy.

Redundancy can help the reader remember important facts given in the text; too much redundancy can be misleading and can create problems in text-understanding.

Writing an Explanatory Text requires a selection of the most relevant facts or pieces of information which must be given about a specific topic. The identification of the relevant information in turn requires the use of what we call Focus-Strategies in order to have the reader understand and select such information as opposed to other redundant or not strictly relevant information which might also be in the text.

Consider figure 1:

In order to produce a satisfactorily explanatory text, which can be understood by a text-receiver, the text-producer must identify one or more goals by choosing a topic. After having chosen a topic he or she

will select the information which is most relevant.

The goal of an Explanatory Text is typically represented by its title. Any selection of information necessarily implies a set of specific textual strategies. Some important strategies in Explanatory Text Producing are those which can improve Text-Understanding. Focalization strategies, as well as topic-maintenance strategies are critical, because they force the Receiver/Learner to receive and understand the Text in a specific way.

### 3. WHAT IS COHERENCE IN AN EXPLANATORY TEXT?

We might say that an Explanatory Text is coherent when it is logically consistent and when its parts fit together into a unified whole. But, this definition is insufficient. A text that satisfies the definition may still be incoherent if the reader cannot determine exactly what the author wanted to say.

Thus we should add that an Explanatory Text is coherent only when the Reader (Text-Receiver) can identify the Author's (Text-Producer) goal which means he or she can identify the topic and understand the text in the manner that the Text Producer desired (see Adams & Bruce, 1982).

But the definition is still incomplete. An Explanatory Text is coherent only when all its parts are causally or temporally related and when there are no gaps

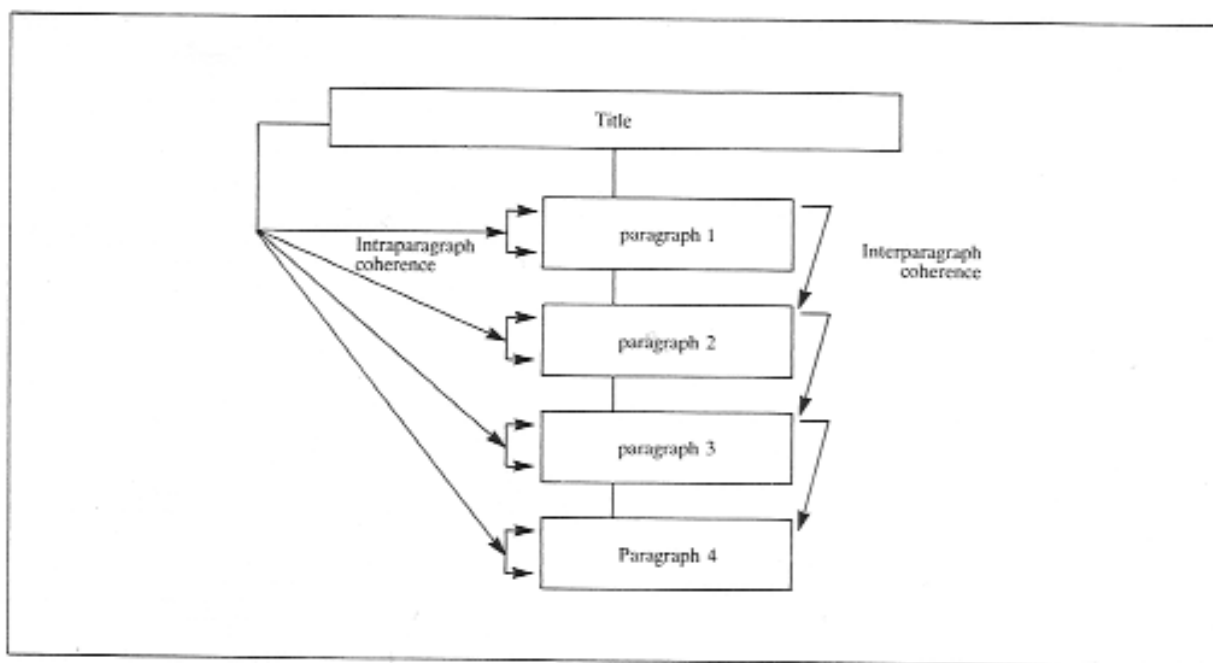


Fig. 2

which cannot be filled by the use of information given previously in the text or likely to be known by the "ideal reader" (Fillmore, 1980). Furthermore, the information selection must be compatible with the Text Receiver's presuppositions and expectations. Thus, coherence is the complex result of a set of variables which determine and affect any text production.

Coherence in explanatory texts can be checked at various levels. For example, we can distinguish between local coherence (the coherence of a sentence or of a short set of sentences) and global coherence (the coherence of a paragraph or a whole text with respect to the topic given by the title). The coherence of an explanatory text can be better defined depending on the nature of its topic; for example, a history text will base its coherence on logical-chronological relations, a descriptive text will base its coherence on logically ordered sequences.

By distinguishing different levels and degrees of coherence within the same text, we can also consider different levels of understanding of a text. Which are the most relevant features of an explanatory text? An explanatory text has to make evident the topic and the information which has been selected and given. Whenever a condition on coherence is violated, confusion is likely to result.

An explanatory text might be incoherent at some level, although it is coherent at some others. For example, an explanatory text could be coherent with re-

spect to its topic, but syntactically disconnected, or it could be syntactically connected but not coherent with respect to the given topic, or its parts might not be logically - chronologically ordered, or textual coherence could be compromised by a set of abrupt topic-shifts (1). All those aspects create troubles in understanding and can all be considered as a result of an incorrect use of textual strategies.

To summarize, we can say that any attempt to define textual coherence has to be specified by identifying different degrees of coherence in the different levels of a text.

#### 4. DIFFERENT LEVELS OF A TEXT

A text can be defined as a set of sentences which are syntactically and semantically related (Halliday-Hasan 1976).

Syntactic relations include such things as anaphora (e.g., pronominalization) or the use of connectors. Connectors are linguistic elements that explicitly link sentences together.

They can represent logical or temporal relations (e.g., "since", "in spite of", "after that"). Semantic relations include such things as topic maintenance or

(1) Topic-shift is the change of the topic from one paragraph into another or from a sentence into another.

topic domain maintenance. This means that, once a topic has been given by the title of the text, the text must keep it as the focus of attention. If the topic shifts, the new topic has to be part of the main topic-domain and has to be relevant to the main topic. By referring back to the syntactic coherence, we can easily identify different levels within an explanatory text:

1. Sentence level (correctness of a sentence);
2. Paragraph level (coherence in relating and chaining a set of sentences by a correct use of connectors and connecting strategies);
3. Textual level (coherence in relating and chaining a set of paragraphs by a correct use of connecting strategies).

Referring to semantic coherence we can again distinguish different levels in explanatory text:

1. Sentence level
2. Paragraph level
3. Textual level
4. Main Topic-Title level

We can represent a text as in figure 2. According to this schema we can distinguish different levels of coherence:

1. Intraparagraph coherence: how sentences of a paragraph are related by the use of connectors and connective strategies.
2. Interparagraph coherence: how different paragraphs of a text are related by the use of connective strategies.
3. Textual coherence: how the whole text is related to the topic given in the title.

Texts can be ranked with respect to their level of coherence, for example, a text coherent at level (1), but not at levels (2) or (3) would be less coherent than one which was coherent at all three levels. We can represent different levels of coherence in Figure 3.

	T <sub>1</sub>	T <sub>2</sub>	T <sub>3</sub>	T <sub>4</sub>	T <sub>5</sub>	T <sub>6</sub>	T <sub>7</sub>	T <sub>8</sub>
Intraparagraph Coherent	X	X	X	X				
Interparagraph Coherent	X	X			X		X	
Textual Coherent	X			X	X			X

Fig. 3

The highest coherent text is T<sub>1</sub> the lowest is T<sub>8</sub>. We can have intermediate levels of coherence (2 degrees - 1 degree).

## 5. PRINCIPLES AND RULES FOR COHERENCE IN EXPLANATORY TEXTS

In an explanatory text "what is difficult to understand" or the new information which has to be learned by the Text-Receiver can be defined as the topic of the text. The topic of an explanatory text is relevant in the following rules:

(a) **Topic identification:** the topic has to be given in the title and has to be identifiable within a paragraph, and within the texts.

(b) **Topic maintenance:** the topic given by the title has to be kept in the paragraphs and in the text.

(c) **Topic-maintenance:** if the identified topic of a text is not kept within the whole text, those paragraphs which don't keep it have to have another topic which is nevertheless part of the main topic-domain and which is strictly related (logically-temporally) to it.

(d) **Topic-shifting:** the main topic can be changed from a paragraph to another; the new topic is not part of the main topic-domain only if it respects: (e) Adherence principle, (f) Reestablishment principle.

(e) **Topic-adherence:** if the main topic is changed between paragraphs and the new topic is not part of the main-topic-domain, it has to be related to (logically-temporally) and adherent to the main one.

(f) **Topic-reestablishment:** if the main-topic is changed between paragraphs and the new topic is not part of the main topic-domain we have a set of informations, which we can define as "digression". At the end of the digression, the main topic has to be reestablished.

We can represent topicalness and the set of principles in the diagram of figure 4.

In order to maintain coherence, an ideal explanatory text must use correct connectors and connective strategies by following a logical or temporal ordering.

(g) **Logical ordering:** sequences and paragraphs must follow by respecting causal logical relations (especially expository texts).

(h) **Temporal ordering:** sequences and paragraphs must follow by respecting temporal sequencing (specifically expository narrative texts).

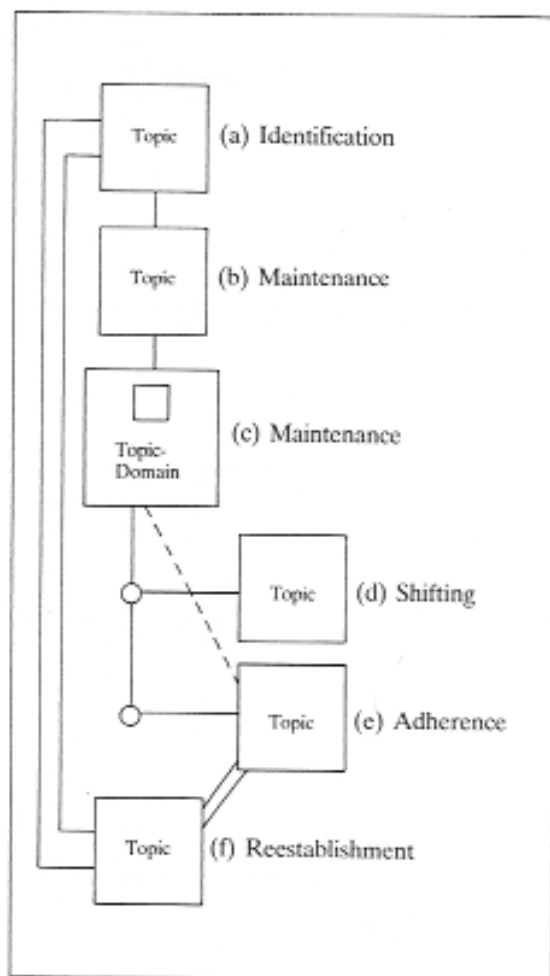


Fig. 4

The ideal coherent text would satisfy all of the following principles:

(i) **Specificity:** the information given within each paragraph and more generally within the entire text has to be specifically related to the topic or topic-domain, which means to the title.

(l) **Pertinence:** any kind of information given within the paragraph of the text which is not specifically related to the topic must be pertinent and be related at least to the topic-domain. Any digression, however, to be justified and cannot be inserted randomly.

(m) **Focus Evidence:** the topic of the paragraph or of the text must be the focus and has to be recognizable.

(n) **Style Continuity:** the style used in the different paragraphs and in the whole text has to be homogeneous.

(o) **Explicit Coreference:** any anaphoric pronoun or noun must have an explicit coreferential antecedent and must be totally unambiguous.

(p) **No Gapping:** paragraphs as well as sentences cannot violate (g) and (h) which means that information has to be given explicitly and by respecting logical-chronological sequencing.

We can represent an explanatory text with respect to the whole set of principles which have been stated about coherence in the schema of figure 5.

Any Expository text can be evaluated by considering this whole set of principles if we want to state the degree of coherence according to the schema 9 fig. 4.

To summarize, principles (a), (b), (c), (d), (e), (f), determine intraparagraph coherence, (g), (h), (l), (m), (n), (o), (p) determine interparagraph coherence, (i) determines textual coherence.

## 6. COHERENCE AND UNDERSTANDING, COHERENCE AND LEARNING

The coherence of an explanatory text is strictly related to the way information about the topic has been selected and ordered. In other words, since the main goal of an Explanatory Text is to give a set of facts and to teach something about a specific topic, the text has to be produced in the best possible way in order to facilitate the reader's understanding.

According to the distinction made between intraparagraph, interparagraph and textual coherence we can identify "local coherence", "extended local coherence" and "global coherence", which corresponds to the previous categorization.

Local coherence determines a good understanding of a part of a text but does not imply a global understanding of a text.

Extended coherence can help in memorizing and storing more facts, but still is not sufficient to reach a global understanding of a text.

Global Text-Understanding can be reached only by texts which have a 3 degree coherence. Studies (Tonfoni, 1982) have shown that only 3 degree coherent texts can be really understood and remembered. This hypothesis has been supported by a set of experiments done with high school students in Italy (Bologna) by giving them texts having different levels of coherence and by checking their understanding of the texts through summarization and question asking. The most coherent texts turned out to be the best understood, summarized, and remembered.

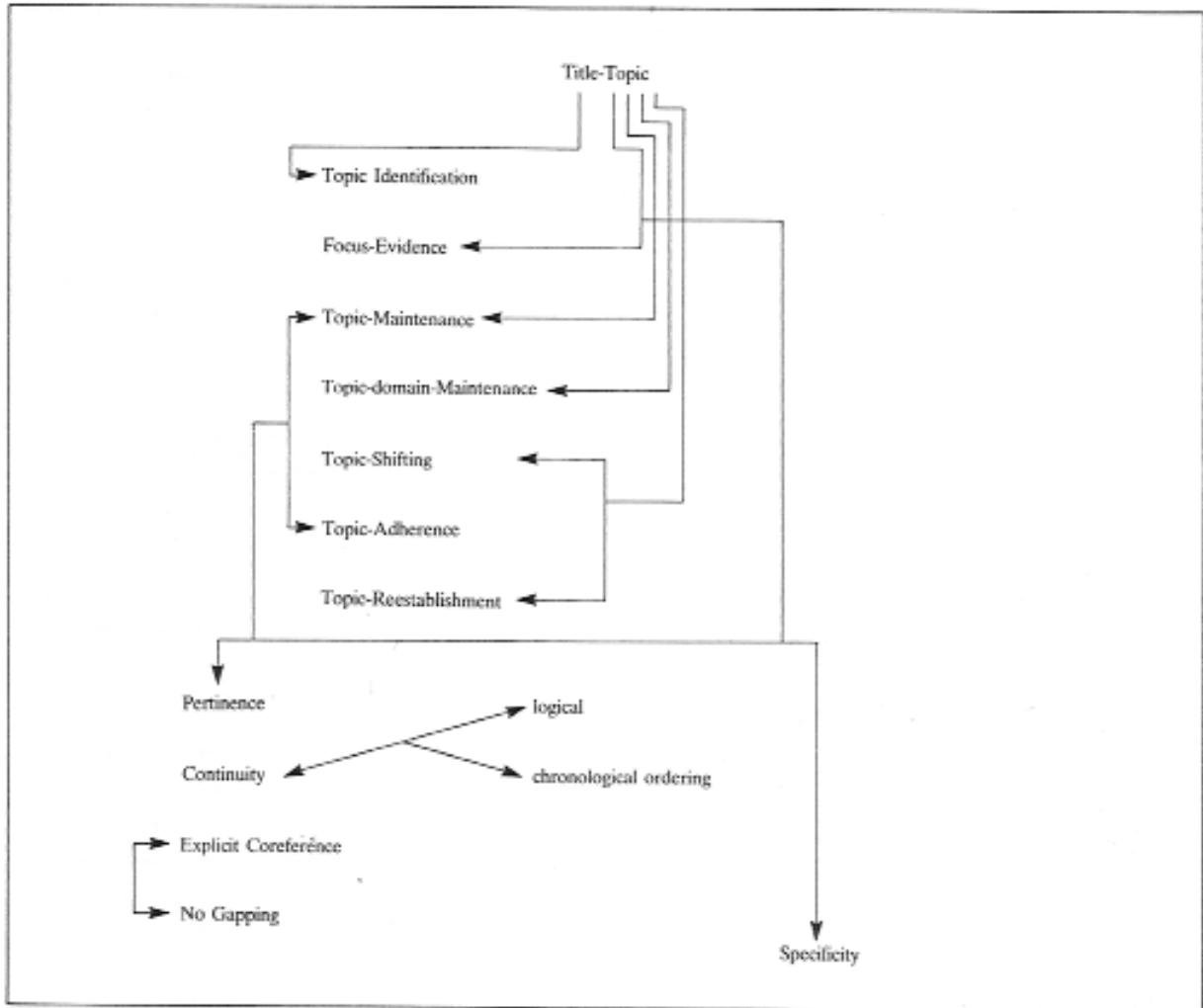


Fig. 5

By understanding, we mean being able to identify the topic and to reorganize the information by selecting and keeping the most relevant elements according to the choice previously made by the Author (Text-Producer).

Text-Understanding can be checked by asking the Text-Receiver (Learner) to summarize it. Summarization shows the level of understanding which has been reached by the Learner.

Text Understanding is a precondition for learning from a text.

Figure 6 summarize what has been stated:

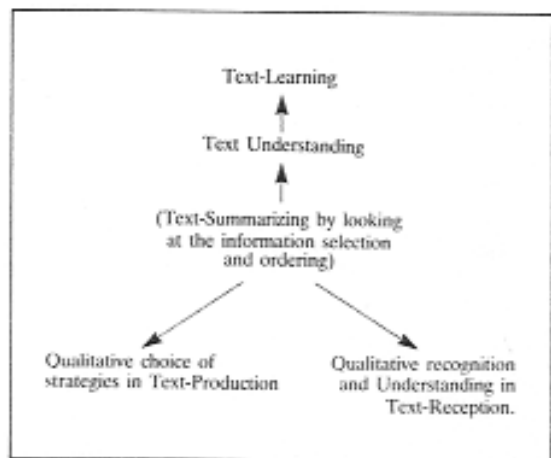


Fig. 6

## 7. HOW TO EVALUATE A TEXT BY LOOKING AT ITS COHERENCE

A text can be evaluated by analyzing its coherence degree.

Referring back to the schema given in figure 3 it is possible to identify 8 possibilities just by looking at the way texts have been produced.

Let's consider the following text:

*Which People Came from Asia? (2)*

*Many Chinese came to America in the 1840's when gold was discovered in California. Some Chinese looked for gold. Others found that they could make more money working as cooks and in other service jobs. The Chinese were not always welcome in the new land. Many people thought they were different. Some white workers were afraid the Chinese would take away their jobs.*

*Chicago has a Chinatown. It is a community near the Stevenson Expressway, a short distance from downtown.*

*The main business street, Wentworth Avenue, is lined with restaurants and shops which tourists like to visit.*

*Many young Chinese are moving from Chinatown to North Side neighborhoods. Young Chinese women and men work in business and professional jobs. But everyone likes to return to Chinatown on the Chinese New Year. The dragon dances and fireworks make this an exciting event.*

The text which has been presented as an example shows a very low level of coherence:

- |  |     |
|--|-----|
| (a) Topic-identification (People coming from Asia) |     |
| (b) Topic-maintenance                              | NO  |
| (c) Topic-domain-maintenance                       | YES |
| (d) Topic-shifting                                 | YES |
| (e) Topic-adherence                                | NO  |
| (f) Topic-reestablishment                          | NO  |
| (g) Logical ordering                               | NO  |
| (h) Temporal ordering                              | NO  |
| (i) Pertinence                                     | YES |
| (m) Focus-Evidence                                 | NO  |
| (n) Styl-Continuity                                | YES |
| (o) Explicit Coference                             | NO  |
| (p) No Gapping                                     | NO  |
| (i) Specificity                                    | NO  |

(2) This text appears in "Chicago - The City and Its People".  
M. Stated. Benedict Press, Chicago, 1981 p. 51.

The text doesn't actually give information about the topic (Asian People) but only about some of them (Chinese People). The topic is shifted several times without being reestablished.

There is no topic adherence. Paragraphs are neither logically nor temporally ordered. We do have pertinence about the shifted topic (Chinese People) but there is no Focus Evidence. We have Style Continuity but no Explicit Coference due to the continuous shifting. We do have Gapping in information because of the lack of the main topic. Finally, it follows that the text doesn't have any specificity.

It is possible to revise the text so that it conforms to all of the coherence principles. The basic information can be reorganized in order to produce a more appropriate and coherent text as follows:

*Which People Came From Asia?*

*Many people came from Asia to the States: Chinese Koreans, and Japanese. In particular, Chinese people came to America in the 1840's when gold was discovered in California. Some Chinese looked for gold, others found other jobs. The Chinese were not always welcome in the new land because many of the other people already living in the States thought that Chinese were different and some white workers were also afraid the Chinese would take away their jobs.*

*Chinese people used to live in a part of the city called Chinatown. Chicago has a Chinatown which is a community near the Stevenson Expressway, a short distance from downtown. The main business street, Wentworth Avenue, is lined with restaurants and shops that tourists like to visit.*

*Many young Chinese are moving from Chinatown to Northside neighborhoods. Young Chinese women and men work in business and professional jobs but everyone likes to return to Chinatown for the Chinese New Year when the dragon dances and the fireworks made this an exciting event.*

According to our ranking scale we can produce different text having various degrees of adherence to the set of identified principles.

## 8. CONSISTENCY OF THE GUIDELINES WITH INTUITIVE JUDGMENTS

The guidelines given above are far more specific than our intuitive notion of "coherence" and give a finer grained measure of relative coherence of texts. The question arises: Is the resulting measure consistent with our intuitive notion?

To answer this question, we asked ten (non-researcher) adults to rank eight texts in order of coherence.

The texts were all short selections from the Chicago text. Each text represented one of the eight categories shown in Figure 3. The results of the subjects' rankings are shown in Table 1.

Examination of the table shows that the text rated most coherent by the subjects was coherent at all three levels; the text rated least coherent was not coherent at any level. In general, textual and intrapara-graph coherence seem more significant than interpara-graph coherence.

## 9. CONCLUSION

Any Text-Producer must select the information according to the stated principles in the following order:

Identify Topic (Topic-Identification) Specificity;

Organize Information which is pertinent (Logical-Ordering, No Gapping, Pertinence);

Identify Priorities (Focus Evidence) in Information Giving and the Relative Hierarchical Structure;

Choose Correct Style (Continuity Strategies for In/Explicit Coreference Formation; Topic Maintenance, Topic Domain/Topic Shifting, Topic Adherence).

By respecting this set of principles Explanatory Text will be well formed and ready to be read and understood by the Text Receiver/Learner. Learning from a text implies being able to select that information which is the most relevant and retain it in memory. There is no a priori way of recognizing the most relevant information; it has to be made evident by a correct Text-Production. This is why we wanted to underline the importance of a correct Text-Production which will improve teaching methods and learning techniques.

Rank	Text Coherent	Inter	Intra
1 (best)	YES	YES	YES
2	YES	NO	YES
3	YES	NO	YES
4	NO	YES	YES
5	YES	NO	NO
6	NO	NO	YES
7	NO	YES	NO
8 (worst)	NO	NO	NO

Table 1: Comparison of Rankings in Terms of Coherence with Prior Text Analysis

## 10. REFERENCES

- [1] Adams, M.J., Bruce, B.C. Background knowledge and reading comprehension for J. Langer and M.T. Smith-Burke (eds); Reader Meets Author: Bridging the Gap: A psycholinguistic and sociolinguistic perspective, Dover Del. Internat'l Reading Assoc., 1982. Also as Reading Education Report n. 13, Urbana: Univ. of Illinois, Center for the Study of Reading, January, 1980.
- [2] Anderson, R.C., Spiro, R., Anderson, M.C., SCHEMATA AS A SCAFFOLDING FOR THE REPRESENTATION OF INFORMATION IN CONNECTED DISCOURSE, American Educational Research Journal, 1978, 15, p. 433-440.
- [3] Anderson, R.C., Osborn J., Tierney R.J., LEARNING TO READ IN AMERICAN SCHOOLS: BASAL READERS AND CONTENT TEXTS, University of Illinois, The Center for the Study of Reading, 1981.
- [4] Anderson, T.H., Armbruster, B.B., Kantor, R.N., HOW CLEARLY WRITTEN ARE CHILDREN'S TEXTBOOKS? OF BLADDERWORTS AND ALFA, Center for the Study of Reading, University of Illinois at Urbana-Champaign, 1980.
- [5] Fillmore, C.J., IDEAL READERS AND REAL READERS, in "Georgetown Univ. Round table on Languages and Linguistics", Tannen B, Georgetown Univ. Press, 1981, pp. 248-270.
- [6] Goetz, E.T., Armbruster, B.B., PSYCHOLOGICAL CORRELATES OF TEXTSTRUCTURE, in "Theoretical Issues in Reading Comprehension", 1980, Lawrence, Erlbaum, Ass, Hillsdale, N.J. p. 201-216.
- [7] Bruce, B.C., PLANS AND SOCIAL ACTIONS; Spiro, R.J., Bruce, B.C., Brewer, W.F. THEORETICAL ISSUES IN READING COMPREHENSION, 1980, Lawrence Erlbaum Ass., Hillsdale, N.J. p. 367-382.
- [8] Halliday, M. AK., Hasan, R., COHESION IN ENGLISH; Longman, London 1976.
- [9] Mandler, J.M. Johnson, N.S., REMEMBRANCE OF THINGS PASSED: STORY STRUCTURE AND RECALL, Cognitive Psychology, 9, 111-115, 1977.
- [10] Rumelhart, D.E.; NOTES ON A SCHEMA FOR STORIES, in D.G. Brown and A. Collins (eds) "Representation and Understanding Studies in Cognitive Science", Academic Press, New York 1975.
- [11] Stein, N.L., Glenn, C.G., AN ANALYSIS OF STORY COMPREHENSION IN ELEMENTARY



SCHOOL CHILDREN, in R. Freedle ed. "Discourse Processing" Multi-disciplinary perspective", Hillsdale, N.J. Ablex, 1978.

[12] Tierney, R.J., Rosenthal, J., Kantor, R.N., SOME CLASSROOM APPLICATIONS OF TEXT ANALYSIS TOWARD IMPROVING TEXT SELECTION AND USE, Center for the Study of Reading, University of Illinois at Urbana-Champaign, 1980.

[13] Tonfoni, G., PER UN'IPOTESI DIDATTICA TESTUALE, Bo, Modena, 1982.

[14] Tonfoni, G. DALLA LINGUISTICA DEL TESTO ALLA TEORIA TESTUALE, Milano, Unicopli, 1982.

[15] Wilensky, R. UNDERSTANDING GOAL-BASED STORIES, Ph. D. thesis, Yale University, 1978.

[16] Rumelhart, D.E.: NOTES ON A SCHEMA FOR STORIES, in D.G. Brown and A. Collins (eds) "REPRESENTATION AND UNDERSTANDING STUDIES IN COGNITIVE SCIENCE", Academic Press, New York 1975.

**Note:**

This paper is the result of a research activity done by Graziella Tonfoni and Bertran Bruce at Bolt Beranek and Newman, Cambridge Mass. U.S.A.

This paper circulates as BBN Technical Reports - Bolt Beranek and Newman, Cambridge Mass.

Permission for reproduction must be requested to the authors.