

## ON THE INFORMATIVENESS REQUIREMENT

Rachel GIORA\*

Grice's (1975) maxims of Quantity and Relation can be viewed as postulating two opposite principles for the regulation of discourse. The requirement to be informative (Quantity) and extend information can be viewed as constrained by a requirement of an opposite nature – the demand to be relevant and to keep to the same (given) point.

To account for the opposing constraints on text formation noted above, I suggest that we consider the dynamics of concept formation. Specifically, I propose the application of the principles governing categorization to text organization. In a previous work of mine (Giora (1985b)) categorization was shown to reflect text organization in relation to its least informative constituent – the Discourse Topic (DT) proposition. Such organization conforms to the Relevance Requirement. Here I intend to focus on the status of the more informative constituents in a linguistic set. I will show that while texts tend to begin with a generalization – the least informative message in a given set, the Informativeness Condition on texts requires that they proceed along an informativeness axis and end with the most informative message in that set.

### 1. Introduction

The main thrust of psychological research into text structuring deals with narratives in schematic terms (Rumelhart (1975), Mandler (1984), Shen (1985) *inter alia*). They suggest that narratives are organized in terms of spacial/temporal or logical coherence. In a previous work of mine (Giora (1985b)) I discuss non-narrative/expository texts. On the basis of Mandler's (1984) distinction between schematic and categorial organization I suggest that we view non-narrative/expository texts as organized around the principle of similarity.<sup>1</sup>

A category is formed when two or more entities share similar features. However, it is not necessary that all category members share a set of common

\* I am deeply indebted to Tanya Reinhart, Josef Grodzinsky, Yeshayahu Shen and Nils Erik Enkvist for their most critical and insightful comments.

Author's address: R. Giora, Department of General Studies, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel 69978.

<sup>1</sup> The attempt to apply categorial organization to narrative texts was exercised by Bransford and Franks (1971) (as reported in Posner (1986)). Bransford and Franks showed that subjects store overall summaries rather than individual examples. This study can be taken to show that only prototypes (overall summaries) are retained.

features. It is enough that they bear family resemblance whereby adjacent members share common features (Wittgenstein (1953), Rosch and Mervis (1975), *inter alia*). As shown by Rosch and Mervis (1975), Smith and Medin (1981) and Tversky (1977), categories can be described as organized along the inclusion principle. Such organization exhibits vertical ordering. Superordinate categories, such as furniture for example, govern (or include) subordinate categories, such as chair, table etc. Those, too, govern other sets of subordinate categories, such as kitchen chair or tea table. As for their horizontal/internal structuring, Rosch and Mervis showed that this is hierarchical too. Highly typical members, called prototypes, are considered best 'exemplars'. They are the members that best represent the category. They share the largest sets of the category common and distinctive features. In other words, they are the members of the set which are similar to all or most of the category members. Such hierarchical grading typifies all sorts of categories, including those which seem to lack grading by definition, e.g., the set of even numbers (Armstrong et al. (1983)). Barsalou (1983) further shows that such internal structuring in relation to prototypes typifies *ad hoc* categories as well. He explains this on a similarity comparison process that imposes graded structure on any category regardless of its type.

Informationally, however, prototypes are redundant. Sharing the highest amount of common and distinctive features, the prototypes represent the redundancy structure of the category. Prototypes are not only the 'good' exemplars – the best representatives of the category. Functionally, they constitute a reference point, relative to which the other members get measured and classified. Items are either members or non-members relative to the similarity they bear to the prototypes. Prototypes can cognitively function as a reference point since they are the most accessible items of the set. When people think of a category member, they, as a rule, think of the most typical members of that category.

Given the reference-point function of the prototype as the element governing the set, it is plausible to assume an internal structuring that exhibits an accessibility hierarchy. That is, an organization in which similarity is the measure suggests an accessibility grading that is prototype-dependent: The top governing entry is the most accessible member and the bottom/marginal constituent is least accessible in the set. Informationally, such grading reflects an informativeness hierarchy ranging from the least to the most informative member in the category.

In sum, prototypes share the greatest amount of common and distinctive features with the other members of the set and thus exhibit the category redundancy structure. Cognitively, they are most accessible and function as the category reference point relative to which the other category members get classified.

In my attempt to show that non-narrative texts get organized along

categorial principles, I demonstrated (Giora (1985b)) that texts, like category concepts, are also hierarchically organized. As the principle of organization is similarity, the internal structuring is determined by a prototypical proposition called Discourse Topic (DT). The DT proposition is a generalization bearing the highest amount of conceptual intersections with the other propositions in the set. Like the prototype concept of natural categories it exhibits the redundancy structure of the semantic unit and thus best represents it. Linguistically it is awarded accessibility marking which indicates that it is a salient text constituent, suitable to function as a reference point. At the text surface level, DTs occupy text initial position which best guarantees their cognitive function as the entry organizing the text. (It was found that where they are not text initial, text processing is slowed down.) Such organization in relation to the DT proposition of the text conforms to the Relevance Requirement. Each incoming message is processed in terms of, or is measured for its similarity to the DT/prototypical proposition

Viewing non-narrative/expository texts as reflecting categorial organization suggests, among other things, that such texts progress from the most accessible/least informative message – the DT proposition – to the least accessible/most informative one. That is, while the top of the hierarchy is the most accessible/abstract constituent, the bottom constituent must be the least accessible/abstract constituent. Within categorial structuring, such ordering follows from organization in terms of similarity. It reflects top-bottom processing. Ordered according to the similarity principle, that is, proceeding from the most to the least accessible/informative constituent, the text accords with processing strategies. As shown by Rosch (1973) and Sanford and Garrod (1981) for example, category inclusion and storage, evidenced by speed of retrieval, reflect ease of processing. Where category members are processed in terms of the category feature sets, the least informative the item the more quickly it is retrieved.

Before checking the hypothesis that texts proceed from the least to the most informative message it is appropriate that the notion of informativeness be more explicitly clarified.

## **2. On defining informativeness**

The extent to which the requirements of Relevance and Informativeness regulate text construction can be illustrated by the informational aspects of visual perception. Attneave (1954) shows that when we begin to consider perception as information handling process, much of the information received by an intelligent organism is redundant. He shows that predictability, i.e., lawfulness, is a redundancy structure and that information lies where unpredictable change occurs. Thus, in visual perception, homogeneity (of color, of

direction) which conceptually will be termed similarity, reflects redundancy. Information, however, is concentrated along contours (i.e., regions where colors change abruptly) and further concentrated at those points on the contour at which its direction changes most rapidly (i.e., at angles or peaks of curvature). It is further noted that around the object's boundaries those informative loci are indispensable to an objective and appropriate representation of the visual object. "It appears likely", Attneave contends, "that a major function of the perceptual machinery is to strip away some of the redundancy of stimulation, to describe or encode incoming information in a form more economical than that which it impinges on the receptors" (1954: 189). Visual perception, then, relies heavily upon both redundancy and informativeness and serves to illustrate by analogy the extent to which both redundancy and informativeness are necessary where verbal information processing is concerned.<sup>2</sup> While attempting to define the notion of informativeness of verbal material I will allude to the two notions of redundancy and informativeness as handled by classical theories of information.

### *2.1. Theories of informativeness*

According to theories of categorization, informativeness is quantifiable. To say that a category member is informative in a given set is to say that it has more features (information) than necessary for category inclusion. As mentioned above, the psychological theories of categorization (Rosch (1973), Rosch and Mervis (1975), Tversky (1977), Smith and Medin (1981), for example) view the category set of common and distinctive features as reflecting the redundancy structure of the category. That is, the information that various members of the category have in common is considered redundant in that category. By contrast, that information which members do not share and which is additional to that set of common/similar features, is informative in that category. A categorial organization implies that that information which category members do not share and which determines their distinctiveness within the given category, constitutes informativeness in that category.

<sup>2</sup> It seems relevant to consider studies of visual perception research concerning informativeness. In terms of spontaneous attention, for example, Berlyne (1960, 1966) distinguishes two classes of stimuli that attract spontaneous attention: physical properties, such as the presence of many contours; and collative properties, such as novelty, complexity, or significance. Berlyne (1958) and Day (1965) show that novel, complex, and incongruous objects are always fixated in preference to others, and that subjects spend more time looking at such stimuli when given control of a device which presents pictures successively (Berlyne (1957), Berlyne and Lawrence (1964)). It is significant to note that linguistic material, described here as informative, has been traditionally treated in terms of novelty and complexity (or heaviness) and significance (or 'importance') (Firbas (1975), Van Dijk (1979), Bolinger (1975)) and in terms of focus of attention (Erteschik-Shir and Lappin (1979) *inter alia*).

Linguistic research, too, deals with informativeness in quantitative terms.<sup>3</sup> Sperber and Wilson (1984) for example, measure informativeness (which they term 'relevance') against the number of contextual implications a message has in a context. The greater the amount of contextual implications a message yields in a context (weighed against the amount of processing effort needed), the more informative it is. Consider their example whereby (a) is more 'relevant' i.e. informative than (b) given the context of 1–5, in which (a) has four contextual implications (7–10) while (b) has one (6). It should be noted that both (a) and (b) require the same amount of processing, since their logical structures are identical:

- (1) People may buy more than one ticket.
  - (2) A ticket costs \$1.
  - (3) The person who bought the green ticket wins the prize.
  - (4) The prize is \$100,000.
  - (5) Anyone who wins \$100,000 can fulfil the dreams of a lifetime.
- 
- (a) James bought the green ticket
  - (b) John bought the blue ticket.
- 
- (6) The blue ticket cost John \$1.
  - (7) The green ticket cost James \$1.
  - (8) James wins the prize.
  - (9) James wins \$100,000.
  - (10) James can fulfil the dreams of a lifetime.

According to classical information theories (Shannon (1951), Attneave (1959) *inter alia*) a message is informative relative to the number of uncertainties it either reduces or eliminates relative to a question. Given that a question has a number of possible equally probable answers, the amount of uncertainty involved in that question corresponds to the number of the alternative answers. Particularly, the measurement of informativeness is computed in terms of a unit called bit (the abbreviation of binary digit). Within this approach, for a message to be informative it must reduce the amount of alternative answers by half, namely, by one bit. To illustrate this, consider the checkerboard example (Attneave (1959)). The goal of the game is to discover which of the 64 possible squares is on the mind of the questioner. It is easy to show that exactly six questions are always necessary and sufficient to locate the square. What is essential is that every question should reduce the alternatives by half:

<sup>3</sup> Perry's (1979), too, is a quantitative proposal. For him the most informative message is the one with the maximum of clues needed for an understanding of what is unresolved in the text. It is the one that links the highest number of disparate items. Both Perry's and Sperber and Wilson's approaches can be reformulated in terms of classical theories of information.

- (1) Is it one of the 32 on the left half of the board? (Yes)
- (2) Is it one of the 16 in the upper half of the 32 remaining? (No).
- (3) Is it one of the 8 in the left half of the 16 remaining? (No).
- (4) Is it one of the 4 in the upper half of the 8 remaining? (No).
- (5) Is it one of the 2 in the left half of the 4 remaining? (Yes).
- (6) Is it the upper one of the 2 remaining? (Yes).

For further explication consider an elaboration on Sperber and Wilson's example (1) below. Given the context of (1a-d), it is obvious that (1f) is more informative than (1e) relative to the question: Who won the prize?

- (1a) People could buy only one ticket.
- (1b) There were 39 blue tickets and 1 green ticket.
- (1c) Forty people bought tickets.
- (1d) The person who bought the green ticket won the prize.
- (1e) Someone bought the green ticket.
- (1f) James bought the green ticket.

Clearly, (1e) does not reduce the number of uncertainties to such an extent as does (1f).

In terms of probability, (1f) is also much less probable than (1e) (by 1/40 to 40/40). The chances that someone bought the green ticket (1e) are high (1). The chances that James bought the green ticket (1f) are low (1/40). In terms of probability, then, an informative message in a set is the least probable message in that set (1f in the 1e-f set in our example).

It should be noted, however, that the notion of probability is not simply a matter of statistics only. There is also the question of psychological probability. The chances that either 'James won the prize' or 'Susan won the prize' might be equal (1/40), provided that both James and Susan were among the forty who bought tickets. However, in a different context, probability should be measured differently. In a math class, for instance, the chance that 'Susan flunked the test' as opposed to 'James flunked the test' when 'Someone flunked the test' is at stake, is much lower, given that Susan is a much brighter student than James. Or, in terms of natural categories, the chance that 'I saw a bird' means 'I saw a chicken' is much lower than 'I saw a robin' (Rosch (1973), Clark and Clark (1977), Sanford and Garrod (1981), Sanford (1985)). Being a prototype 'robin' is a much more accessible exemplar than 'chicken' in terms of bird-category. To formulate it in terms of number of possibilities reduced, the reference to 'chicken' reduces a much greater amount of possible candidates for 'bird' than 'robin'. True, a 'bird' could be either 'robin' or 'pigeon' or 'eagle' or 'chicken'. But as 'robin' is a more accessible (i.e., probable exemplar than the rest, the reference to 'robin' reduces only one alternative, while the reference to 'chicken', being the least prototypical/



accessible among the given examples, reduces all the remaining exemplars. The most informative member in a set, i.e., the member bearing the greatest amount of extra information compared to the prototypical member in that set, is thus the least probable in both psychological terms and in terms of number of uncertainties reduced. As it is least accessible it is least probable, which explains the greater number of uncertainties it reduces relative to a more prototypical member.

No wonder, then, that informativeness is often assigned surprise value. Defined as the least probable message in a set relative to a given question, the informative message is least predictable or most surprising in that set.

Having defined informativeness in terms of class inclusion and number of uncertainties reduced, and having presented an informative message as unpredictable, least probable and surprising, we can proceed now to consider the correlation between amount of information and text ordering. In what follows, a message (x) is more informative than (y) in case it reduces more options relative to a given question/DT.

### 3. Paragraph ordering

Having shown that a text segment begins with the most redundant message in that given segment (Giora (1985b)), my claim here with respect to text ordering is that the text continuum is informatively structured. A coherent text progresses from the least to the most informative message. Where text linear ordering reflects informativeness structure it conforms to the Graded Informativeness Requirement.

However, before checking this claim against empirical findings, consider, first, the ordering of category items by various researchers. The lists in (2) below are graded according to an accessibility/prototypicality scale. Rosch (1973), Armstrong et al. (1983) for instance, present their findings, starting with the most prototypical exemplar (constituting the redundancy structure of the category) and ending with the borderline case exemplar – the most informative i.e., dissimilar member. Note that the least prototypical members mark here the boundaries of the categories indicating their possible inclusion within two neighboring sets (of the same order).

(2) <i>Fruit</i>	<i>Vegetable</i>	<i>Bird</i>
apple	carrot	robin
plum	asparagus	eagle
pineapple	celery	wren
strawberry	onion	chicken
fig	parsley	ostrich
olive	pickle	bat

The point about these lists is that they reflect storage in memory under the most accessible member which represents the category set of common and distinctive features. Specifically, the hierarchy reflects the cognitive distance that obtains between the various members and the prototype. This distance accounts for speed of retrieval, and similarly, for ease of processing in terms of the number of features searched when a decision on category inclusion is to be made.

This principle of similarity which regulates concept organization in memory is viewed here as a more general principle that accounts not only for concept formation procedures but also for the linear structuring of non-schematic/expository texts. Texts, like categorial representation of concepts, are organized in terms of similarity. My claim with respect to the text surface structure is that given a coherent informative text, the semantic hierarchy is correlated with the text linear ordering.

The linear organization of texts is accounted for in terms of ease of processing or cognitive distance, i.e., in terms of number of features assessed when a decision on category inclusion is to be made. Thus, the more similar a textual message is to the DT proposition, the easier and faster the decision on its inclusion within the context set of that text. And likewise, the least similar a message is to the DT, the more difficult it is to decide on its inclusion, given the number of uncommon features to be measured. The most informative/least similar text constituent thus signals the text semantic boundary.

The decision on category inclusion is an assessment of the number of common/similar features against the number of dissimilar features. When an item is too informative, that is, when the number of dissimilar features exceeds the number of common features to the extent it can already be stored under a different entry to which it is more similar, then this item can be either a borderline case member or excluded within the category under discussion. In terms of text well-formedness, the objective of this analogy, the Relevance Condition, which requires that the various propositions in the text be similar to its prototypical/DT proposition – thus determines the amount of informativeness allowed within a text segment. When a text constituent is too informative in terms of number of features searched for similarity to the DT proposition, it might be considered irrelevant in that text. Consider, for example the last sentence in (3) below which seems to start a new segment, and the last sentence in (4) below which sounds irrelevant too. The over informativeness in the end of (4) ridicules the old woman, suggesting that the most important event was her release. Note that the degree of informativeness exhibited in (3) and (4) below will not be treated as irrelevant by Sperber and Wilson:

- (3) It often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter, unexpectedly



came upon another which was far more important than the one he was looking into. The penicillin is a result of such a discovery. The penicillin was accidentally discovered by Fleming in 1928.

- (4) Two people knocked at Klara Rozenberg's in Raanana wishing, very strangely, to buy sheep... But before Klara managed to close the door they attacked her, shut her mouth with a cellotape and tied her to a chair. They beat her ruthlessly. They then wore gloves and...robbed her jewelry, her husband's pistol and 600 IS and ran away. Klara managed to release herself and called the police. 25 Arabs were arrested and released. Klara's granddaughter said later: "Grandma managed to get released ... because she watches Ninja's movies".

Let us now consider empirical findings. First, I checked an accidental collection of 55 passages (Alexander (1976)) for their final text-constituent. The examination revealed that 83.6% of them ended with the most informative message in that text. Those that did not, repeated the initial DT proposition. As is well known, DT repetition functions as a boundary marker (Longacre (1979)).

Next, I conducted experiments. To test native readers' intuitions as to the preferred ordering of text constituents, two different passages were checked, a narrative sequence (Experiment 3 in the appendix, reported also in Giora (1983b)) and a non-narrative expository text (Experiment 1 and 2 below). In view of the categorial organization hypothesis, I predicted that subjects would prefer texts ending with the most informative, i.e., least similar message in the given text (Experiment 1) and those that evolve informatively (Experiment 2).

### *Experiment 1*

*Aim:* to show that native readers prefer a paragraph in which the relatively most informative message occupies paragraph final position.

*Subjects:* 73 (48+25) graduate students majoring in Education and Journalism.

*Materials:* six expository texts (5-10 below) identical in every respect apart from the ordering of the final constituents. In one (5, 7, 9 below) the paragraph ended with the most informative message and in the other (6, 8, 10 below), the most informative message was embedded in pre-final position.

- (5) It often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter, unexpectedly came upon another which was far more important than the one he was looking into. The penicillin is a result of such a discovery.

- (6) It often occurred in the history of science that an important discovery was come upon by chance. The penicillin is a result of such a discovery. A scientist looking into one matter, unexpectedly came upon another which was far more important than the one he was looking into.
- (7) Until a century ago scientists believed that only human beings make use of tools. But a surveillance of a couple of decades of the ways of life of animals has changed their mind. They reached the conclusion that some animals do use tools. Examples are many: there are elephants which scratch their backs with a stick they hold by their trunk; there are drones which cover their nest with earth or pebbles, and many other examples.
- (8) Until a century ago scientists believed that only human beings make use of tools. But a surveillance of a couple of decades of the ways of life of animals has changed their mind. Examples are many: there are elephants which scratch their backs with a stick they hold by their trunk; there are drones which cover their nest with earth or pebbles, and many other examples. They reached the conclusion that some animals do use tools.
- (9) Let's deal with what seems a simple topic, but which is rather complicated. The topic is names. Yes, as simple as that, our name. The first and the second name.
- (10) Let's deal with what seems a simple topic, but which is rather complicated. The topic is names. Yes, as simple as that, the first and the second name. Our name.

*Method:* Subjects were presented with the two versions of the texts in a randomized order and asked about their preferences as to the more natural text ordering. They were told to rely on their intuitions as native readers. Having responded to this request they were asked which order of presentation was easier to recall, which was easier to follow and which presented the information better. Order of the three questions was randomized as well.

*Results:* In accordance with prediction, 79% preferred the informative-final version. Specifically, of the 48 subjects who read (5) and (6), about 70% (68.75%) of the subjects preferred (5), ( $z = 2.59$ ,  $p < 0.05$ ). Of the 25 subjects who read (7–10) 96% preferred (7) ( $z = 4.6$ ,  $p > 0.05$ ) and 72% preferred (9) ( $z = 2.2$ ,  $p > 0.05$ ). In addition, the answers to the three questions concerning recall, comprehension and well-formedness, correlated with their first response.

*Discussion:* With respect to the first DT proposition of (5) which represents the redundancy structure of the text ('It often occurred in the history of

science that an important discovery was come upon by chance'), the proposition concerning the discovery of the penicillin ('The penicillin is the result of such a discovery') is more informative in terms of number of alternatives reduced. Chance discovery in the history of science – the purported DT of the present text, formulated in the first proposition – alludes to a number of possibilities. The mention of the discovery of the penicillin, which is a specific, special case of the category 'scientific chance discovery', eliminates the other alternatives falling under this category. In the same manner, this proposition is more informative than the middle proposition ('A scientist looking into one matter, unexpectedly came upon another, far more important than the one he was looking into'), as this one is simply a paraphrase, or, rather a repetition of the DT proposition. Though formulated in a slightly more specific manner, exemplifying the general matter by a generic, non-specific agent, it only slightly restricts the number of possibilities. In terms of cognitive distance, the last message requires more processing as it is less similar than the intermediate message. The above experiment thus attests that, given the choice, readers prefer a text ordering that reflects a categorial organization whereby the more informative message follows the less informative one, in accordance with the inclusion principle.

Likewise, with respect to the first and second sentences which form the DT of (7) (scientists changed their mind about the exclusiveness of humans' use of tools) the following proposition ('They reached the conclusion that some animals do use tools') reduces the uncertainty by half – some, rather than all animals use tools. The last proposition is the most informative message as it consists of specific examples of the set of animals. However, in (8), in which the specific examples precede the more general information, this information is both missing before the specifications and is redundant in terms of number of uncertainties reduced where it follows them.

The same holds for the structure of (9) in which the most informative message in terms of number of possibilities reduced is the last one ('the first name and the second name'). 'Our name' which precedes it could denote either our first name or our second name or both. The last message thus reduces the possibilities by two thirds. Where 'our name' appears last (10) it is no longer informative, that is, it neither reduces uncertainties nor does it have additional contextual implications. It is therefore redundant in that position. Its felt absence before the more specific information is explainable in terms of cognitive distance.

## *Experiment 2*

*Aim:* to show that native readers prefer a paragraph in which the intermediate (that is, non-final) messages are ordered along the informative axis, in which the less informative message precedes the more informative one.

*Subjects:* 40 graduate students majoring in Education and Journalism.

*Materials:* three pairs of expository texts (11–16 below). The pairs are each identical in every respect apart from the order of the intermediate constituents. In (11), (13), (15) the more informative message succeeds the less informative one and in the respective matches (12), (14), (16) the reverse order obtains.

- (11) Until a century ago scientists believed that only human beings make use of tools. But a surveillance of a couple of decades of the ways of life of animals has changed their mind. They reached the conclusion that some animals do use tools. Examples are many: there are elephants which scratch their backs with a stick they hold by their trunk, there are drones which cover their nest with earth or pebbles, and many other examples.
- (12) Until a century ago scientists believed that only human beings make use of tools. But they reached the conclusion that some animals do use tools. A surveillance of a couple of decades of the ways of life of animals has changed their mind. Examples are many: there are elephants which scratch their backs with a stick they hold by their trunk, there are drones which cover their nest with earth or pebbles, and many other examples.
- (13) Let's deal with what seems a simple topic, but which is rather complicated. The topic is names. Yes, as simple as that, our name. The first and the second name.
- (14) Let's deal with what seems a simple topic, but which is rather complicated. The topic is our name. Yes, as simple as that, names. The first and the second name.
- (15) In 1952 something happened on Adamelo Mount which was reported by the press and on the radio and shocked world public opinion. At the bottom of one of the glaciers of the mountain a few human bodies were discovered. Investigators decided that those were the bodies of five Italian soldiers who fell into one of the ice pits during the battles which took place there during First World War, in 1916.
- (16) In 1952 something happened on Adamelo Mount which was reported by the press and on the radio and shocked world public opinion. At the bottom of one of the glaciers of the mountain five bodies of Italian soldiers were discovered. Investigators decided that those were the bodies of a few human beings who fell into one of the ice pits during the battles which took place there during First World War, in 1916.

*Method:* Subjects were presented with two versions of the texts (the pairs above) in a randomized order and asked about their preferences as to the more natural text ordering. They were told to rely on their intuitions as native readers.

*Results:* In accordance with prediction, the differences were significant: 81.66% of the subjects preferred the more informative last version. Specifically, 92.5% preferred passage (11) ( $z = 5.37$ ,  $p > 0.05$ ), 77.5% preferred passage (13) ( $z = 3.48$ ,  $p > 0.05$ ) and 75% preferred passage (15) ( $z = 3.16$ ,  $p > 0.05$ ).

*Discussion:* The passages that were preferred exhibit a text continuum that correlates with the Requirement for Graded Informativeness as delineated here (section 1 above). As already stated, the Graded Informativeness Condition on texts requires that text constituents be presented along the informativeness axis, starting with the least informative/most accessible message and ending with the most informative/least similar one. The purpose of these experiments is to show that the Informativeness Requirement is a condition on the text continuum as a whole. In the first experiment here (Experiment 1 above) it is shown that the Informativeness Requirement holds for the last text constituent which constitutes the most informative message in a given text segment. In the second experiment (Experiment 2 above) this is confirmed with respect to intermediate, non-final text constituents.

Thus, in (11) above 'changed their mind' (which together with the first sentence constitutes the DT of the given passage) is less informative than the specific conclusion reached in the following proposition where it is made clear that the options are reduced by two; some animals rather than all the animals, for example, use tools too. The reverse order (12 above), in which the specific instance, which already presupposes 'changing their mind', precedes this presupposition, was judged as inappropriate. An order of presentation where a less informative message following the specific one is already redundant. In terms of processing it is already generated (as an inference) for purpose of similarity measuring, to enable the inclusion of the more informative, less accessible or more distant message within the more general category stated in the beginning. It could save the effort of gap-filling were it positioned before the more informative constituent, which explains the preference of the last informative order of presentation of (11).

Likewise in (13), 'names' (which is more informative than the 'simple/complex topic' of the first proposition in terms of number of options reduced) is a more general, less informative category than 'our name', which reduces the number of possible candidates for instantiations of 'name'. And in (15) too, (whose first proposition is the most general/least informative message in the whole set) 'the bodies of five Italian soldiers' is more informative than 'a few human bodies' in terms of the number of possible alternatives reduced.



Being more specific their preferable linear position follows the more general proposition. In terms of processing, the more specific information, that which requires extra processing on account of more dissimilar features, or that which is more distant from the DT proposition in terms of number of features searched, is placed last as it presupposes all that precedes it. Speaking in terms of linear processing, where more general information is missing in the reading process, as in (12), (14), and (16), some mental effort is required to generate it. Where later it follows the more specific information it is already redundant in the sense that it does not reduce uncertainties and is a waste of time.

The above experiments thus show that where texts contain messages of different informative load, their linear ordering conforms with the Graded Informativeness Requirement.

#### **4. Text progression**

The claim that this work makes in relation to text progression is that texts tend to proceed from the most redundant to the most informative message in the text. As for progression from segment to segment, my previous works (Giora (1983a, b, 1985a)) demonstrate that writers reserve the most informative text constituent for text final position to both signal text segmentation and, at the same time, enable text progression. The informative message positioned in the end of a text segment does not only mark the end of the segment but is immediately available for the writer to pick up for further discussion. In my previous works such information was termed Dominant. It was shown to be an appropriate candidate for the function of the next DT. Picked up for further discussion, the previously most informative message is now the generalized, most redundant message in the given text segment. Constituting the DT of that segment, it will be followed by more informative messages that will reduce uncertainties. Such text progression from the final section of one paragraph to the topical, initial section of the following one was shown to be a common text progression.

Consider, for instance, the continuation of (5) above in (17) below. The second paragraph of (17) is a detailed report of the accidental discovery of the penicillin. The second paragraph in (17) takes up as a DT the most informative proposition of the previous paragraph on which it elaborates in detail. The text in (17) thus illustrates the way texts progress along the informative axis even at the paragraph progression level:

- (17) It often occurred in the history of science that an important discovery was come upon by chance. A scientist looking into one matter, unexpectedly came upon another which was far more important than the one he was looking into. The penicillin is a result of such a discovery.



In 1928 the British bacteriologist, Fleming, tried to cultivate a species of germs called staphylococci. One day he discovered that one of the cultures was accidentally contaminated by a fungus called penicillium notatum ...

However, to validate a norm let us consider a case in which it is violated. (18) below is a passage from a daily newspaper where the interviewer is mocking the interviewee by responding to the less informative message in the text segment. The interviewee, herself a journalist, is giving an example of her ability to control herself despite traumatic experiences:

- (18) A-B: I went to interview D.H., the Time reporter. I knew he had a murderous dog ... named Dolar. As a rule the dog is locked away when guests are expected. So, at the villa gate, when I rang twice and nobody answered I opened the gate and walked towards the door. But before I reached the door Dolar attacked me with terrible anger. I knew that a dog like that is trained to murder so I was trying to reach for my throat. I dropped down on the pavement and sacrificed my leg with *my precious boot on it*. Dolar stuck his teeth in my boot. I was sure he would reach my throat in the end. I said to myself, that's your end. That's it.

G: *Was it a leather boot?*

A-B: A snake leather boot. Suddenly, the door opened and D.H. appeared. He was shocked. But as he was a high ranking officer he didn't lose his temper. He jumped over me so that Dolar should let go of my leg. The trouble was that ... it took the dog's jaw some time to release the boot. *I lost a precious boot* but a few minutes later I was already inside and interviewing D.H. as if nothing had happened, and as if all was well.

G: *How's the boot?*

The most informative message relative to the DT (her self-control in view of a dangerous, traumatic experience) is certainly not her loss of her boot. Text progression that deviates from the informative axis seems to violate the Relevance Requirement. The implicature here is of ridicule.

## 5. Summary

My notion of text informativeness is indebted to previous linguistic research concerning information structuring at the sentence level. Historically speaking the notion of informativeness termed Rheme, was first developed by the Praguians and their American followers to characterize that constituent in the

sentence which adds information about the Topic constituent. That constituent by means of which the speaker predicates something about the Topic (Mathesius (1939)) and which constitutes the new information the sentence is to impart (Halliday (1967)), was assigned high Communicative Dynamism (Firbas (1975)). It was further shown (Daneš (1974), Erteschik-Shir and Lappin (1979)) to constitute a possible Topic of the following sentence.

At the text level what I define as informative was termed 'important' (Van Dijk (1979, 1985)) and 'relevant' (Van Dijk (1979), Sperber and Wilson (1986)). While 'important' was not formally defined, 'relevant', in the sense assigned to it by Sperber and Wilson was (see section 2 above). My definition of informativeness of verbal material relies on notions developed within classical theories of information which I integrate into concept formation theories. Thus, given the probabilistic view of categorization (Rosch (1973), Rosch and Mervis (1975), Rosch et al. (1976), Smith and Medin (1981), Tversky (1977), for example), the amount of informativeness of a category member is measured relative to the extra amount of information it bears compared to the prototypical – most redundant – member and relative to the number of other category members reduced which could be possible alternatives in a certain context. To put it in textual terms, a verbal message (x) is more informative than (y) relative to a given question/DT, in case it has more information than (y) compared to the DT/most redundant message in the text, and in case it reduces more uncertainties than (y) relative to that DT.

I have tried to show here that the Gricean requirement of Relevance constrains the requirement of Informativeness. A speaker can be informative only to the extent that her/his message can be included within the scope of the DT s/he has taken up. Specifically, I have treated text organization in categorial terms and thus defined informativeness in terms of categorial membership whereby a highly informative member is treated as such relative to the most prototypical, i.e., redundant member of the set.

Assuming that expository texts reflect categorial organization (Giora (1985b)), my claim with respect to their linear ordering is that informative messages tend to follow less informative ones. In previous work (Giora (1983a, b, 1985a)) I accounted for the final positioning of the most informative message in terms of recency effect. I showed that informative messages are placed in the end of textual segments for communicative purposes so as to be immediately available for further discussion. Here I further show that this ordering holds for the text continuum as a whole and that it follows from the principle of inclusion within a category.

## Appendix

### Experiment 3

*Aim:* To show that native readers prefer a paragraph in which the relatively most informative message occupies final position.

*Subjects:* 78 graduate and undergraduate students majoring in Maths, Literature, and Journalism.

*Materials:* Two narrative texts, identical in every respect except for the ordering of the most informative constituent. In one (1 below) it appeared in the end of the passage, while in the other (2 below) it was embedded in pre-final position, in accordance with the temporal order of events.

- (1) There are dark days in London when the city is swept by stale darkness and nobody dares leave his place. And on just such a day a man was urgently called to attend his sick son in a faraway hospital at the end of the city. The man opened the door and stepped outside into the dark to look for help. But there was nobody there. Suddenly, a strange hand was laid upon his shoulder, and a man whose face was hidden by the dark night said to him: I shall lead you. And indeed, the stranger led the man from one end of the city to the other, unhesitatingly. Only here and there he confidently remarked: here to the left, mind the steps, watch the slope. When finally they arrived at the hospital, the father said: How can you manage in the dark? – The darkness and the fog don't bother me, said the stranger, because I am blind.
- (2) There are dark days in London when the city is swept by stale darkness and nobody dares leave his place. And on just such a day a man was urgently called to attend his sick son in a faraway hospital at the end of the city. The man opened the door and stepped outside into the dark to look for help. But there was nobody there. Suddenly, a strange hand was laid upon his shoulder, and a man whose face was hidden by the dark night said to him: I shall lead you. The father asked him: How can you manage in the dark? – The darkness and the fog don't bother me, said the stranger, because I am blind. And indeed, the stranger led the man from one end of the city to the other, unhesitatingly. only here and there he confidently remarked: here to the left, mind the steps, watch the slope, until finally they arrived at the hospital.

*Method:* Subjects were presented with the two versions of the text in a randomized order and asked about their preferences as to the more natural text ordering. They were told to rely on their intuitions as native readers.

*Results:* In accordance with prediction, 78% of the subjects preferred the informative-final version ( $z = 4.8$ ,  $p < 0.05$ ).

*Discussion:* With respect to the question (constituting the purported DT) as to how the father would be able to attend his sick son, or rather, what superpower could be summoned to help the father, the preferred version (1 above) provides the most

informative answer at the very end of the text. It should be noted that the text unfolds information gradually, preserving the most informative message which reduces all other uncertain alternatives, to the end. The first informative message concerning the question posed in the beginning ('The man stepped outside ... to look for help') is not informative enough. So is the second ('Suddenly a strange hand ... I shall lead you') which still does not resolve the mystery. The final message ('I am blind') is specific enough to disperse optional hypotheses as to the identity of the strange man whose 'power' or 'capability' helped resolve the initial problem.

This final information is also most surprising as it is least probable given the set of superior if not superhuman powers that could be included as members in the set of possible alternatives. 'The blind man' is least expected, as blindness is least prototypical in that set of 'capabilities'.

In the inappropriate passage (2 above), the answer to the question as to how the father would reach his sick son precedes the information, somehow presupposed, that he would indeed reach him. The information concerning his arrival in the hospital is highly predictable, given the knowledge about the blind man.

## References

- Alexander, L.G., 1976. *Fluency in English*. London: Longman.
- Armstrong, Sharon Lee, Lila R. Gleitman and Henry Gleitman, 1983. On what some concepts might not be. *Cognition* 13: 263-308.
- Attneave, Fred, 1954. Some informational aspects of visual perception. *Psychology Review* 61: 183-193.
- Attneave, Fred, 1959. *Applications of information theory to psychology*. New York: Holt, Rinehart and Winston.
- Barsalou, W. Lawrence, 1983. Ad hoc categories. *Memory and Cognition* 11: 211-227.
- Berlyne, D.E., 1957. Conflict and information-theory variables as determinants of human perceptual curiosity. *Journal of Experimental Psychology* 53(6): 399-404.
- Berlyne, D.E., 1958. The influence of complexity and novelty in visual figures on orienting responses. *Journal of Experimental Psychology* 53(6): 399-404.
- Berlyne, D.E., 1960. *Conflict, arousal and curiosity*. New York: McGraw-Hill.
- Berlyne, D.E., 1966. Curiosity and exploration *Science* 153: 25-33.
- Berlyne, D.E. and G.H. Lawrence, 1964. Effects of complexity and incongruity variables on GSR, investigatory behavior and verbally expressed preference. *Journal of General Psychology* 71: 21-45.
- Bolinger, Dwight L., 1975. *Aspects of language* (2nd ed.). New York: Harcourt Brace Jovanovich.
- Bransford, J. and J.J. Franks, 1971. Abstraction of linguistic ideas. *Cognitive Psychology Review* 89: 727-761.
- Clark H. Herbert and Eve V. Clark, 1977. *Psychology and language*. New York: Harcourt Brace Jovanovich.
- Craig, Colette, ed., 1986. *Noun classes and categorization*. Amsterdam: Benjamins.
- Daneš, F., 1974. 'Functional sentence perspective and the organization of texts'. In: F. Daneš, ed., *Papers on functional sentence perspective*. The Hague: Mouton.
- Day, M.E., 1965. *Exploratory behavior as a function of individual differences and level of arousal*. University of Toronto. (Ph.D. dissertation.)
- Erteschik-Shir, Nomi and Sahlom Lappin, 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6: 41-86.

- Firbas, J., 1975. 'On defining the thematic and non-thematic section of the sentence'. In: H. Ringbom et al., eds., *Style and Text*. Stockholm: Skriptor. pp. 314-334.
- Giora, Rachel, 1983a. Segment cohesion. *Text* 2: 155-182.
- Giora, Rachel, 1983b. 'Functional paragraph perspective'. In: Janos S. Petöfi and Emel Sözer, eds., *Micro and macro connexity of texts*. Hamburg: Buske. pp. 153-182.
- Giora, Rachel, 1985a. Informational function of the linear ordering of texts. Tel Aviv University. (Ph.D. dissertation.)
- Giora, Rachel, 1985b. A text-based analysis of non-narrative texts. *Theoretical Linguistics* 12: 115-135.
- Grice, H. Paul, 1975. 'Logic and conversation'. In: Peter Cole and J. Morgan, eds., *Syntax and semantics* 3. New York: Academic Press. pp. 41-58.
- Halliday, M.A.K., 1967. Notes on transitivity and theme in English: II. *Journal of Linguistics* 3: 199-244.
- Jackendoff, Ray, 1985. *Semantics and cognition*. Cambridge, MA: MIT Press.
- Kahneman, Daniel, 1973. *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Longacre, R.E., 1979. 'The paragraph as a grammatical unit'. In: Talmy Givón, ed., *Syntax and semantics* 12. New York: Academic Press.
- Mandler, Jean, 1984. *Stories, scripts and scenes: Aspects of schema theory*. Hillsdale, NJ: Erlbaum.
- Mathesius, U., 1939. On the so-called functional sentence perspective. *Slovo a Slovesnost* 5: 171-174. (In Czech.)
- Perry, Menakhem, 1979. Literary dynamics: How the order of the text creates its meaning. *Poetics Today* 1: 35-64, 311-361.
- Posner, Michael, 1986. 'Empirical studies of prototypes'. In: C. Craig ed., 1986.
- Rosch, Elinor, 1973. 'On the internal structure of perceptual and semantic categories'. In: T.E. Moore, ed., *Cognitive development and the acquisition of language*. New York: Academic Press. pp. 111-144.
- Rosch, Elinor and B. Carolyn Mervis, 1975. Family resemblance in internal structure of categories. *Cognitive Psychology* 7: 573-605.
- Rosch, Elinor, W.D. Gray, D.M. Johnson and P. Boyes-Braem, 1976. Basic objects in natural categories. *Cognitive Psychology* 8: 382-439.
- Rumelhart, H. David, 1975. 'Notes on a schema for stories'. In: D.G. Bobrow and A.M. Collins, eds., *Representation and understanding*. New York: Academic Press. pp. 211-236.
- Sanford, J. Anthony, 1985. *Cognition and cognitive psychology*. London: Weidenfeld and Nicolson.
- Sanford, J. Anthony and Simon C. Garrod, 1981. *Understanding written language*. Chichester: Wiley.
- Shannon, C.E., 1951. Prediction and entropy of printed English. *Bell System Technical Journal* 30: 50-64.
- Shen, Yeshayahu, 1985. The structure of action in short narrative stories. Tel Aviv University. (Ph.D. dissertation.)
- Smith, Edward and Douglas Medin, 1981. *Categories and concepts*. Cambridge: Harvard University Press.
- Sperber, Dan and Deirdre Wilson, 1984. 'On defining relevance.' (Unpublished ms.)
- Sperber, Dan and Deirdre Wilson, 1986. *Relevance*. Oxford: Blackwell.
- Tversky, Amos, 1977. Features of similarity. *Psychological Review* 84: 327-352.
- Van Dijk, Teun A., 1979. Relevance assignment in discourse comprehension. *Discourse Processes* 2: 113-126.
- Van Dijk, Teun A., ed., 1985. *Discourse and communication*. Berlin: Walter de Gruyter.
- Wittgenstein, Ludwig, 1953. *Philosophical Investigations*. New York: Macmillan.