

Topic Levels

Joseph E. Grimes
Cornell University and
Summer Institute of Linguistics

Now that the sentence is no longer the edge of our world, we see more clearly than ever how totally responsive speech is to the situation that calls it forth and to the people involved in it. Bare content is shaped and packaged to meet many requirements at once.

I have tried to sort out two broad categories of these requirements. The first, cohesion, is hearer oriented. It is the influence on form of the speaker's own assumptions about what the hearer knows at each instant of the communication process. The second, staging, is speaker oriented. It reflects how the speaker calibrates the importance of different parts of what he himself intends to say. I find it helpful to tie down the discussion of both cohesion and staging to differences in linguistic form; there may be other psychological or philosophical overtones to cohesion and staging that have no such direct repercussions, but getting at those overtones is another matter.

One of the areas where we are making progress in the linguistic study of discourse is in seeing how speaker and hearer always seek a common ground of reference. This area, however, is hidden in a terminological thicket. Charles Hockett (1959:201) originally identified it as topic, in which, in his terms, 'the speaker announces a topic and then says something about it.' Gundel (1974) has followed this usage, and I think it the best label for now even though it gets confused with topicalization, which may or may not be part of the same package. Much earlier some of the Prague School theorists (summarized in Danes 1974), and later Halliday (1967), used theme for a similar concept, and now Grosz (1977) has used focus for something not very different. Each of these three terms, topic, theme, and focus, has also been used for at least two other kinds of phenomena by reputable linguists, so eventually we are going to have to put together a road map to all the alternatives; but just because the terms are confused is no reason to conclude that no headway is being made.

The idea that I am going to continue to refer to as topic is this: for communication to succeed, speaker and hearer have to establish common ground. This common ground is usually a presumed agreement about the identity of certain objects in the real world. It may also be agreement about certain events or about certain relations that hold. As far as its linguistic expression goes, I think it significant that its formal makeup appears to revolve around nominals most of the time, treating things that are not necessarily objects as if they were.

One reason the common ground phenomenon seems so important is that without its narrowing effect, the hearer might not be able to manage the numerous semantic alternatives that could be developed from each expression in the text constructed by the speaker. Gundel has pointed out the utility in this regard of a formulation attempted by Searle (1969:126):

For any speaker S, any object X and any predicate P, it is a necessary condition of S's having predicated P of X in the utterance of a sentence containing P, that X should have been successfully referred to in that utterance and all the presuppositions of P should be true of X.

Searle's X is very much like what I am calling the topic, in that unless the hearer can relate to it referentially, he can neither agree nor disagree with whatever else may be said about it.

Gundel illustrates how even an isolated sentence like George ate a plate of shrimp cannot be assimilated as part of a real communication unless the hearer has some way of knowing which of the people named George is being referred to; once he knows that, he can react with yes he did or no he didn't or oh, the sign of a new constellation of information in memory (Winograd 1972).

Grosz has made a useful distinction

between the hearer's memory for concepts, which tends to be long term and global to a text, and the hearer's memory for linguistic form, which tends to be short term and local to a segment of text. Her distinction interlocks with the one Halliday and Hasan (1976) make between reference and substitution. What they call reference identifies concepts, objects outside of language, and even pieces of language itself, as they are mentioned once they have been introduced. What they call substitution includes ellipsis; it refers to the reactivation of stretches of speech from earlier in the text in order to talk about situations that have not already been referred to, but which have enough in common with others that have been referred to that the same linguistic expression can apply to the new situation.

It is the first of these, memory for concept rather than form, from which the speaker appears to take what he hopes is common ground between himself and his hearer. This selection from the field available for global reference is what is behind Gundel's observation that the topic has to be accessible to the hearer. (The most accessible things are characteristically the standard elements of the communication situation: I, you, here, and now.)

In the course of a text, be it monologue or dialogue, the referential common ground that is used as the core of communication may change. This is true of the global topic and apparently of local topics as well. The initial core of reference may be designated very simply, for example by a single noun phrase, with no differentiation of parts or functions at the beginning. Gundel even shows that the topic of some sentences may be implicit, not mentioned in that sentence, but nevertheless to be taken into account if the sentence is to make sense.

Once the topic of a text is put into play, that topic may be developed in at least three different ways that have been described so far in the literature: it may be expanded, shifted, or split. Expansion adds things to the core of reference. Shift adds new referents to the core and leaves others off, so that what is taken as common ground at one point in the text differs from what was taken as common ground earlier. Splitting the core results in local topics being brought into play in relation to global topics; or rather, higher level topics are split into a higher level part and a lower level part, a process which if repeated may yield more than two levels in the same text.

Topic expansion is illustrated in a story from Time magazine (June 21, 1976, p. 56). The date of the issue needs to be

taken as the initial topic; there is no other common ground to begin with between the writer and the reader. This is normal in news stories: consider how impossible it is to agree or disagree with anything like The Giants beat the Dodgers until one knows the occasion. The title of this piece is Teton: Eyewitness to Disaster. For a reader who knows his geography, Teton identifies a place, while for one who does not, there is at least a good chance that it is a place name or the name of a person. The idea of referential common ground gives us for starters a reasonable guess at an event that happened the week before the appearance of the magazine, and possibly a place, as a limiting field within which to place the interpretation of the rest of the message.

The text begins with a paragraph set off in italic type and quotation marks, in which the speaker is not identified: "This wet spot on the side of the dam started spurting a little water ..." The noun phrase that begins the sentence is definite, as is the dam contained within it. The definiteness here suggests that the writer expects the reader to be able to find the reference because it is accessible within the limits already set. If he follows that suggestion and takes Teton as the name of the dam and accepts this as identifying something new within that field, his reference succeeds. (The side is legitimately definite once we identify the dam by what Halliday and Hasan call lexical cohesion; dams have sides.) As far as pinning down a core of reference is concerned, the text so far has its topic built up as clearly as if the article had begun much more fully, as for example Last week at a place called Teton where there is a dam, a wet spot appeared on its side.

The text goes on "... and I asked my mother, 'Do you think we should notify the authorities?' She said: 'I don't think ...'" Here the person who is making the report is mentioned explicitly for the first time, as is the mother to whom the question is addressed. This complex of the observers, the wet spot on the dam, the location, and the time persists as the referential core or topic through the course of over a column. It is built up by small references to give an expanded topic.

Topic shift differs from expansion in that some referential elements appear to be dropped from the topic as the text progresses; some things that were treated as part of the common ground in the earlier part of the text are not so treated later. Schank (1977) focuses on the intersection of the referential field of one sentence and that of its successor, and tries to define some (but not all) ways in which that intersection relates to the referential field of the next

sentence. His initial definition of topic as 'any object, person, location, action, state, or time that is mentioned in the sentence to be responded to' is probably too inclusive, because it does not take into account Searle's factor of successful reference; but once a text is begun, Searle's boundary condition is no longer needed, because the reference has been established by the text itself.

For Schank a new topic is 'derived from the original input but is not identical to it', in that reference may shift from a specific element mentioned in the earlier sentence to the class of which it is a part in the later sentence, or vice versa. The element in the later sentence may also be a different conceptualization that is like the first in kind; Schank calls this supertopic. It may also be 'a comment that can be inferred from the interaction of two conceptualizations', or metatopic. Schank suggests more specific rules that he hopes will characterize the way topics shift in the course of a text.

The key concept, however, is his observation that a sentence out of context cannot be said to have a topic, because for him the topic arises only out of the interaction of adjacent sentences by the process of intersection. If he is right, or close to right, it is reasonable that some of the things that are treated as topic earlier in the text be given different treatment later in the same text, because they are no longer taken as part of the referential common ground between the speaker and hearer.

The idea of splitting up the referential field embraced by the topic into higher and lower level topics has been treated in two different ways, each of which may be valid in its place. Grosz recognizes the phenomenon, and gives an intriguing example (1977.23) of a pronoun it which refers back to a global topic last mentioned half an hour earlier, even though a whole series of local topics has come between the pronoun and its antecedent. In her discussion of global and local topics, however, she tends to equate the first with memory for concepts and the second with memory for forms, chiefly because she finds the domain of ellipsis to be restricted to a local segment of text, and to always involve memory for forms.

Meyer (1974) and Clements (1976), however, find the global-local phenomenon operating independently of ellipsis or other substitute-like memory for form. They construct a topical hierarchy consisting of a global topic, whatever local topics are talked about as part of the discussion of the global topic, and whatever lower level topics are talked about as part of the discussion of those

local topics, in what apparently gives a recursively definable topic tree of unrestricted depth. In this model, psychological tests of recall show that subordinate position in the topic tree regularly gives worse recall than superordinate position.

The definition of topic used by Meyer and Clements is earlier than Gundel's, so that one could expect the variance in their results to be reduced by attention to her principles for recognizing topics. Topic for Clements is more like Halliday's theme, ordinarily the first thing in the sentence. His topic hierarchy comes from three rules:

- (i) Topic rule:
Identify the topic of each clause and simple sentence.
- (ii) Old/new rule:
Decide whether the topic is new (never previously mentioned) or old (mentioned in an earlier topic or comment). If new -- assign it one level below the previous topic. If old -- assign it the same level as its first mention.
- (iii) Coordination rule:
If a topic is coordinated with an earlier topic or comment, assign it the same level as that earlier topic or comment.

The work of Clements and Meyer lends credence to the idea that there may be a hierarchy of topics in a text, all referential in Halliday and Hasan's sense, rather than dependent only on short term memory for form. Some observations from Koine Greek, the vernacular Greek of the first century before Christ to the third century after, appear to bear this out.

Word order is used much less in Koine than it is in English to specify grammatical relations, because the case system of nouns carries that load. One of the functions which word order expresses seems to be that of identifying shifts in topic (Grimes 1975). Noun phrases in the nominative case that precede the main verb of a clause regularly make that nominative the topic; that is, they signal the reader to take it as part of the referential core that is to be the common ground between him and the writer.

The conjunctions of Koine Greek also play a part in the topical structure. There are three kinds of conjunctions: coordinating, subordinating, and resumptive. The coordinating conjunctions include most of the ones translatable as 'and' or 'but'. The subordinating conjunctions include the 'because' and 'if' varieties. The resumptive ara, dio, and occasionally idou mean something like

'now, back to the main point'. They reset the topic level from wherever it was to the global topic.

Working along with the conjunction system is the system of grammatical subordination. It uses relativizers, complementizers, participles, and verbal nouns to signal that certain propositions are peripheral in the author's perspective on what he has to say. All these subordinating grammatical mechanisms are used constantly in the ancient Koine documents, even though it would be perfectly possible to express what they express in strings of independent clauses with no subordination.

The result is texts with a richly elaborated grammatical structure, with some clauses at four or even five levels of subordination. The general distinction between global and local applies, but recursively: at any level of subordination, it appears possible to have yet another level of subordination attached.

When we turn, however, to languages that make distinctions of topic level explicit in their pronominal systems, we find no greater elaboration than a distinction between global and local topics. Bacairi of Brazil (Wheatley 1973) is such a language. The terms Wheatley uses in his description are 'thematic - athematic' and 'focal - nonfocal'; but it seems clear from perusal of his paper that we would now want to call the thematic category the topic, and the focal category global. Someone who has been identified as part of the common ground between speaker and hearer for a text as a whole is referred to by the pronoun maca 'thematic focal animate'; someone who is topic for a local segment but is not the global topic at the same time is referred to by auaca 'thematic nonfocal animate'; someone who is global topic, but is referred to within a stretch that has a local topic active, is mauanca or maunca 'athematic focal animate', and those who are topic neither at the global level nor at the local level are referred to by uanca 'athematic nonfocal animate'. There are inanimate counterparts for all four pronouns.

On the thematic side -- that is, in the pronouns used for topics -- there is a situational or deictic use that confuses the picture. The pronoun inara 'thematic nondeictic animate' is only textually defined, with an anaphoric antecedent that is taken always from the preceding sentence, not from the topical structure of the text as a whole. On the other hand there is a pronoun mira 'thematic deictic animate' which applies only to animate things that can be seen and are near the speaker. In between the nondeictic and the deictic are the situational uses of maca and auaca. The first denotes someone

far away but in sight, and the second denotes someone nearer to the speaker but not as near as mira.

There are situations in which these two pronouns appear to flip their reference; what they actually do is change the basis of the reference from the situation to the text. Thus I can begin a text by identifying a boy over there as maca because he is relatively far away, and a woman standing closer as auaca because she is not so far away. But if what I have to say revolves around the woman rather than the boy, I will switch after a few sentences to the textual definition of the pronouns and use maca for the woman because she is more central to what I have to say, and auaca to the boy when I treat him as a local topic. However, we have no information about the use of these pronouns to topicalize at more than two levels at a time.

Longuda of Nigeria has a less elaborate pronoun system with respect to topics than that of Bacairi (Newman, in press), but nevertheless it distinguishes topic from nontopic. There is actually a series of pronoun-aspect particles that appear in sequence in texts to identify actions of the central character and distinguish them from the actions of others. In some parts of a text the character singled out by the pronouns is the one the text is about; he is treated as the global topic. Where local topics are introduced, however, the topic pronoun set switches over to the local topic, and the referent of the global topic is referred to by the same nontopic pronoun as all the other characters.

For example, in a story about a rabbit, most of the story uses the topic pronoun series for the rabbit and a nontopic pronoun for everybody else, including an elephant who interacts with the rabbit. One section, however, is about what the elephant did. In that section it is the elephant who gets the topic pronoun series, and the rabbit gets the nontopic series, even if it is the rabbit whom the story is about globally. When the section about the elephant ends, the topic pronoun series reverts to the global topic, the rabbit.

Evidence from languages like these proves the linguistic realism of a distinction between one kind of topic and another. It also agrees in the main with the kinds of analysis we have been making for English; apparently we are fairly close to the right track, in terms of those languages that put this kind of thing right out on the surface. What is not yet clear is the number of levels of topic we may deal with in a text: Clements's analysis of English and my analysis of Greek point toward more than two simultaneous levels of topic, while

the languages that distinguish levels of topic in their pronoun systems seem to support two levels.

Bibliography

- Clements, Paul. 1976. The effects of staging on recall from prose. Cornell University Ph.D. thesis.
- Danes, Frantisek, ed. 1974. Papers on functional sentence perspective (Janua Linguarum series minor no. 147). The Hague: Mouton.
- Grimes, Joseph E. 1975. 'Signals of discourse structure in Koine'. In George MacRae, ed. Society for Biblical Literature 1975 Seminar Papers, Vol. 1. Missoula, Montana: Scholars Press, 151-164.
- Grosz, Barbara. 1977. The representation and use of focus in dialogue understanding (Technical Note 151). Menlo Park: Stanford Research Institute.
- Gundel, Jeannette. 1974. The role of topic and comment in linguistic theory. University of Texas Ph.D. thesis.
- Halliday, Michael A. K. 1967. 'Notes on transitivity and theme in English, Part 2'. Journal of Linguistics 3:199-244.
- and Ruqaiya Hasan. 1976. Cohesion in English. London: Longman Group Limited.
- Hockett, Charles F. 1959. A course in modern linguistics. New York: The Macmillan Company.
- Meyer, Bonnie J. F. 1974. The organization of prose and its effects on recall. Cornell University Ph.D. thesis.
- Newman, John F. in press. 'Participant orientation in Longuda folktales'. In Joseph E. Grimes, ed. Papers on Discourse. Dallas: Summer Institute of Linguistics.
- Schank, Roger. 1977. 'Rules and topics in conversation'. Cognitive Science 1:4.421-442.
- Searle, John. 1969. Speech acts. Cambridge: Cambridge University Press.
- Wheatley, James. 1973. 'Pronouns and nominal elements in Bacairi discourse'. Linguistics 104:105-115.
- Winograd, Terry. 1972. Understanding natural language. New York: Academic Press.