THEORETICAL NOTES

On Throwing Out the Baby
with the Bathwater:
A Reply to Black and Wilensky's
Evaluation of Story Grammars*

JEAN M. MANDLER

University of California, San Diego

NANCY S. JOHNSON

State University of New York at Buffalo

A number of criticisms of a recent paper by Black and Wilensky (1979) are made. (1) In attempting to assess the observational adequacy of story grammars, they state that a context-free grammar cannot handle discontinuous elements; however, they do *not* show that such elements occur in the domain to which the grammars apply. Further, they do not present adequate evidence for their claim that there are acceptable stories not accounted for by existing grammars and that the grammars will accept nonstories such as procedures. (2) They state that it has been proven that under natural conditions children cannot learn transformational grammars, which is a misrepresentation of the learnability proofs which have been offered. (3) Most important, they take an unduly narrow approach to story understanding by claiming that people only understand story content and do not have knowledge of story structure which is useful in comprehension or memory. Counterevidence from the literature is cited which indicates that such knowledge is both useful and used, and a number of methods for assessing the *psychological* adequacy of structural models are discussed.

Recently Black and Wilensky (1979) have criticized in some detail the notion of a story grammar as an approach to understanding how people comprehend and remember stories; they also criticize the approach on abstract formal

grounds. Although we have some reservations about focusing too exclusively on linguistic arguments, even within such a framework their arguments are open to question. The present paper identifies some inadequacies in their analyses and goes on to discuss the issue of what one should expect a structurally oriented approach to contribute to our understanding of stories and how people process them.

I. THE FORMAL AND EMPIRICAL ADEQUACY OF
OF STORY GRAMMARS

Black and Wilensky (hereafter referred to as B&W) characterize the formal properties of story grammars in terms of known constraints on classes of grammars in order to evaluate the adequacy with which such models could, in principle, account for various types of stories. They begin by correctly noting that a finite state grammar is observationally inadequate as a characterization of story structure because such a system does not allow self-embedding, a point which others have made as well (e.g., Johnson & Mandler, 1980). However, their remaining analyses of the relations between types of grammar and the structure of traditional stories are misleading. In particular, they have ignored a crucial factor in their assessment of the observational adequacy of these theories, namely, the domain over which the theories range. For example, B&W claim that a story grammar based on context-free phrase structure rules would be inadequate because it could not "correctly represent" discontinuous elements.[1] Although this observation about context-free rules is technically correct, its relevance obviously depends on whether or not discontinuous elements occur in the domain to which the grammars were intended to apply.[2] B&W cite four examples of stories to represent cases in which an ongoing episode is temporarily interrupted by another unrelated episode. However, none of their examples was drawn from the oral tradition; instead, one was invented by B&W to illustrate their point, and the remaining three ("The Arabian Nights," "The Decameron," and "The Canterbury Tales") are all *collections* of stories, placed within an encompassing framework by a sophisticated author to interest or amuse a reader.

For reasons discussed at length elsewhere (Johnson & Mandler, 1980; Mandler & Johnson, 1977), we have explicitly limited our grammar to stories from the oral tradition.[3] This is indeed a major restriction on the domain, but one that is both sensible and psychologically motivated. We have suggested that

---

[1]Black and Bower (1980) make the same argument, and our reply applies to that article as well.

[2]B&W imply that a context-sensitive grammar could handle discontinuous embeddings, but see Postal (1964) who argues that even context-sensitive grammars do not provide a desirable account of such sequences.

[3]Rumelhart (1977) restricts his story schema to the class of "problem-solving" stories.

traditional stories are limited to a relatively small set of underlying constituent structures, a limitation which may have arisen through constraints imposed by oral transmission. Whatever their origin, however, the limited number of forms for traditional stories should enable people to engage in more top-down, expectation-driven processing than is possible with many other types of connected discourse, and this feature of traditional stories makes them a particularly interesting and useful domain for psychologists to study. The structure of a folk tale, myth, or fable is not the same as the structure of a novel or the modern art form of the ''short story,'' and is certainly not the same as the structure of a collection of stories *taken as a whole;* hence, we believe that B&W's examples are irrelevant. To make their argument legitimately, B&W would need to examine folk tales from the oral tradition (such as those gathered in anthropological collections, books of fairytales, and so forth) and find instances in which one episode is arbitrarily interrupted by another unrelated episode. However, B&W did not carry out such an enterprise, and we doubt that it would be successful.[4]

Concerning the empirical adequacy of story grammars, B&W raise two issues: Are there acceptable stories not accounted for by story grammars; and do the grammars ever generate nonacceptable stories? In considering the first question, B&W state that the deletion and movement rules in Johnson and Mandler (1980) are incomplete, since one can produce comprehensible deletions and movements that are not allowed by the grammar. Furthermore, they argue that this deficiency could not meaningfully be alleviated by adding additional rules because, they claim, any type of story unit may be deleted or moved as long as its original content or position remains inferable.

As before, B&W's question is one of observational adequacy and again raises the issue of how to test it. The approach taken by B&W has been to generate fragments of event sequences involving goal-directed behavior and to argue that, because such fragments are ''comprehensible,'' they are well formed (or presumably could be used in composing a complete story that would be well formed). We question the appropriateness of this approach for two reasons. First, ''comprehensibility'' alone is not a sufficient criterion for deciding that a given piece of text has the structural characteristics of a story. Without a specific limitation on the domain from which examples and counterexamples are to be drawn, one could declare that any bit of comprehensible prose is a ''story'' and cite counterexamples to a particular grammar or other structural characterization.

---

[4] It can be argued that even the literary examples provided by B&W do not represent arbitrary interruptions of one episode by another. In ''The Decameron,'' for example, tales are generally told as an ''outcome'' of an ''attempt'' to get one character to provide amusement for the others, and are followed by an ''ending'' in which listeners react to the story itself. That is, although the content of the embedded episodes is arbitrary the embeddings themselves are not. In contrast, the truly discontinuous type of structure described by B&W would require that the embedded stories be told at arbitrary moments and that the embedding episode then proceed as if the embedded episode had not occurred.

Second, one cannot, in principle, assess the adequacy of a story grammar by producing a *fragment* of text in isolation and showing that it is comprehensible. This is not a question of whether argument by counterexample is appropriate but a restriction on the nature of such argumentation that is essentially no different than that used in linguistics. Judgments of well-formedness based on sentence fragments cannot guarantee that those fragments when joined will result in a well-formed sentence. For similar reasons, tests of well-formedness involving stories are more appropriately based on complete stories than on story fragments.

For example, B&W exemplify a category of "attempt deletion" with the following fragment: "John needed a book from the library and it was soon in his possession." They state that the attempt itself (John's getting the book) is deletable because it is inferable. It may indeed be inferable (although it is no longer clear that John himself executed an attempt). However, the issue of observational adequacy concerns whether there are, in fact, traditional stories that omit the attempt category. The answer to this question is found by examining the appropriate domain of stories, which B&W have not done. Our examination of the domain of traditional stories, which contributed to our characterization of transformational rules, showed that goals are often deleted, whereas attempts are not; similarly goals and attempts are occasionally inverted in the sequence of statements in such stories, but this is not the case for attempts and outcomes.[5]

In any case, since we are engaged in a psychological, rather than a linguistic enterprise, we would suggest that more psychological methods of assessing observational and descriptive adequacy are in order. Although one might collect judgments of well-formedness for stories that have been generated with and without the categories in question, it might be more productive to ask whether people have more trouble comprehending or remembering a story from which such material has been deleted than one which includes it. Again, B&W have not carried out such tests, nor have they cited the available evidence. Current evidence, although far from complete, supports the hypothesis that both deletion and movement of basic story units can disrupt reading time (Haberlandt, 1980) and retrieval, even when the stories remain comprehensible and therefore the missing information presumably is inferable (Mandler, 1978; Mandler & De-Forest, 1979; Stein & Nezworski, 1978; Thorndyke, 1977). In addition, these studies have shown that certain kinds of units (notably simple reactions and goals) are much more likely to be omitted in recall than others (consistent with Johnson and Mandler's deletion rules), and that recall tends to follow the canonical order specified by the grammars even when the input order has moved

---

[5]Children's self-generated stories show similar characteristics. Stein and Glenn (1980) found no instances of attempt–outcome inversion in a large corpus of stories generated by 5- to 13-year-old children and in only 6 percent of the stories which were well enough formed to have episodic structure were attempts omitted. In these cases a higher-order goal path was expressed, which summarized an attempt and outcome in a single statement.

constituents from their normal place. Such findings suggest that there *are* restrictions on deletion and movement of story constituents.

The other example of empirical inadequacy which B&W offer is based on their claim that current story grammars would incorrectly accept as a story a procedural passage which they "adapted" from the procedural descriptions studied by Graesser (1978). It is interesting to note that in order to demonstrate that a story grammar will accept a procedural description, B&W turned one of Graesser's procedures into a story! They added exactly those structural components that traditional stories have and that procedural expositions lack, such as a setting and beginning event which introduce a protagonist, who then reacts to this event with a simple reaction and a goal, and so on. Graesser's procedures consist only of lists of directions, and could not be parsed by the existing story grammars. Such lists are the general form for procedures. Indeed, we will all be surprised if the next time we open a package of new fishing gear, we find the instructions for use couched in the following form: "It is fishing season in Connecticut and your friend has asked you to go fishing. You don't know how to fish, but if you want to impress your friend, here is what you do . . . Congratulations! You have caught a fish!" We do not find it surprising that a procedure written in story form sounds more or less like a story; however, a procedure written in *procedural form* does not include a protagonist, tie its content to a specific locale or time, or provide a resolution.[6]

Finally, B&W state, as have we, that a formally adequate grammar will require the addition of a transformational component to whatever base is specified. However, they argue that this is an unhappy state of affairs because "Wexler and Hamburger have shown that transformational grammars are unlearnable, except under a very artificial set of conditions" (Culicover, 1976), and thus "children would not be able to learn a grammar of that form" (Black & Wilensky, 1979, pp. 219–220). This is a serious misrepresentation, both of Wexler and Hamburger's work and Culicover's presentation of that work (cf. Culicover, 1976; Hamburger & Wexler, 1973; Wexler & Hamburger, 1973), and is particularly puzzling in light of Culicover's statement that "It is quite clear that transformational grammars are learned by children" (Culicover, 1976, p. 291). The thrust of the arguments presented by these authors is that certain *types* of transformational grammars could not be learned by a child with *a particular set* of processing constraints, given *certain* limited kinds of input (e.g., without semantic information), and *a particular criterion* for learning—*not* that transformational grammars are simply unlearnable.

In concluding this section, it should be noted that we have argued that B&W's statements concerning the formal and empirical adequacy of the rules of

[6]Similar comments apply to Black and Bower (1980), in which the same argument—that story grammars will accept procedures—appears.

story grammars are unduly narrow in conception and in some cases wrong. A case can also be made, and is cogently argued by Rumelhart in his reply (1980), that their arguments are largely irrelevant to the function and use of a story grammar. As Rumelhart points out, the most interesting aspect of any grammar is that it assigns a viable constituent structure to productions from a domain. We obviously agree (cf. Johnson & Mandler, 1980) and would add that whatever formalism one uses to describe stories, whether that be rewrite rules, an ATN, or yet another method, *some* characterization of their constituent structure will be a requirement. We turn to this issue in the next section.

II. THE ROLE OF STRUCTURE
IN STORY PROCESSING

B&W conclude from their analyses that the story grammar approach does not address important issues in story understanding and that it is more important to investigate "the knowledge that people use to understand stories." They claim further that people could not use a story schema to help them comprehend a story because they would have had to understand the story before being able to assign its sentences to their correct constituent parts. This argument is tantamount to claiming that stories do not have *any* kind of constituent structure (or if they do, it is not appreciated or used by the listener or reader). Such an extreme approach is not supported by the existing literature on story comprehension, nor does it take account of many facts about retrieval, some of which have been presented by Black himself. Black and Bower (1979) demonstrated that the "episode" constituent determines some aspects of recall of stories. The tendency of people to reorder their recall to conform to an ideal schema, described earlier, is another example of structural knowledge governing processing, as is the finding that when people forget part of a story their substitutions tend to preserve the underlying structure even though the substituted content differs from the original (Mandler & Johnson, 1977). In general, one must assume that an interactive process is at work during comprehension in which expectations based on past experience combine with the actual details of the material to guide the understanding process. An interactive approach also suggests that incoming data are held in working memory long enough to be classified and stored, and that later information is interpreted in light of the classifications that have already been made. Rumelhart (1977) and Rumelhart and Ortony (1977) have provided sophisticated discussions of how an interactive model of comprehension might work, a model which B&W have ignored, substituting instead a painfully inadequate model that no cognitive psychologist would espouse (see Rumelhart, 1980).

It does not seem constructive to take an "either–or" approach in studying the kinds of knowledge people have about stories (or other forms of discourse) and how they use that knowledge to facilitate comprehension and memory. No

one disputes that using real world knowledge, understanding goals and plans, identifying whether or not causal connections are reasonable, and so forth, are crucial for understanding a story. As B&W note, story grammars provide only the bare structural outlines for story plots, with no content specifications that would ensure that an attempt is consistent with a goal, or a goal consistent with a beginning. It is here, of course, that knowledge of typical goals and how to achieve them, and ways in which attempts can produce or fail to produce particular types of outcomes, is of major importance. What people working within the story grammar framework claim, however, is that *in addition to* content specifications, stories have specifiable structures and that people have knowledge about such structures which they use in the course of comprehension and retrieval. A story grammar is one way of capturing those regularities. Although the literature cited earlier documents the importance of structural factors, none of us is under the illusion that a story grammar will be the ultimate description of story structure. Nevertheless, to explore and understand just what that structure is requires a theoretical framework within which to probe and test. To deny structure altogether is inappropriate. What *is* appropriate and interesting is to frame a theory of story structure and systematically to test its implications.

For example, we can assess the "psychological reality" of the units which various theories posit. Some work has already been accomplished at the level of the "episode" category (Black & Bower, 1979; Glenn, 1978; Haberlandt, 1980; Mandler, 1978). However, a detailed examination of the lower-level units and their posited interrelationships is also needed. There are several ways in which this examination might be accomplished, and these are currently being carried out. One is to show that people can make structural judgments about the units from which stories are composed (Pollard-Gott, McCloskey, & Todres, 1979). Another is to show that the posited units influence processing time by demonstrating, for example, that comprehension or retrieval of sentences within units is faster than for sentences that cross unit boundaries. To test the utility of deletion or movement rules, one might want to show that judgments of well-formedness vary as a function of the position of story units, or that the posited units are better comprehended when they are moved as a unit than when they are moved separately. All these issues are subject to empirical test. It would seem that then and only then can one legitimately evaluate the adequacy of a particular theory of story structure and the extent to which people make use of such knowledge in the course of comprehending and remembering.

REFERENCES

Black, J. B., & Bower, G. H. Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior*, 1979, *18*, 309–318.

Black, J. B., & Bower, G. H. Story understanding as problem-solving. *Poetics*, 1980, *9*, 223–250.

Black, J. B., & Wilensky, R. An evaluation of story grammars. *Cognitive Science*, 1979, *3*, 213–230.

Bower, G. H., Black J. B., & Turner, T. J. Scripts in memory for text. *Cognitive Psychology,* 1979, *11,* 177–220.

Culicover, P. W. *Syntax.* New York: Academic Press, 1976.

Glenn, C. G. The role of episodic structure and of story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior,* 1978, *17,* 229–247.

Graesser, A. C. How to catch a fish: The memory and representation of common procedures. *Discourse Processes,* 1978, *1,* 72–89.

Haberlandt, K. Story grammars and reading time of story constituents. *Poetics,* 1980, *9,* 99–118.

Hamburger, H., & Wexler, K. N. Identifiability of a class of transformational grammars. In K. J. J. Hintikka, J. M. E. Moravcsik, &. P. Suppes (Eds.), *Approaches to natural language.* Dordrecht, Holland: D. Reidel, 1973.

Johnson, N. S., & Mandler, J. M. A tale of two structures: Underlying and surface forms in stories. *Poetics,* 1980, *9,* 51–86.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes,* 1978, *1,* 14–35.

Mandler, J. M., & DeForest, M. Is there more than one way to recall a story? *Child Development,* 1979, *50,* 886–889.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology,* 1977, *9,* 111–151.

Pollard-Gott, L., McCloskey, M., & Todres, A. K. Subjective story structure. *Discourse Processes,* 1979, *2,* 251–281.

Postal, P. Constituent structure: A study of contemporary models of syntactic description. *International Journal of American Linguistics.* Publication 30 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. Bloomington, IN: Indiana University, 1964.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Rumelhart, D. E. On evaluating story grammars. *Cognitive Science,* 1980 *4,* 313–316.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Stein, N. S., & Glenn, C. G. Syntactic structures and real-world themes in children's generated stories. Unpublished manuscript, 1980.

Stein, N. S., & Nezworski, T. The effects of organization and instructional set on story memory. *Discourse Processes,* 1978, *1,* 177–193.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology,* 1977, *9,* 77–110.

Wexler, K. N. & Hamburger, H. On the insufficiency of surface data for the learning of transformational languages. In K. J. J. Hintikka, J. M. E. Moravcsik, & P. Suppes (Eds.), *Approaches to natural language.* Dordrecht, Holland: D. Reidel, 1973.