

The Representation of Derivable Information in Memory:  
When What Might Have Been Left Unsaid Is Said

Rand J. Spiro, Joseph Esposito, and Richard J. Vondruska  
Center for the Study of Reading  
University of Illinois at Urbana-Champaign

It is now widely accepted that natural language comprehension is a constructive process. Information in discourse interacts with a variety of impinging contextual factors (including, most prominently, the comprehender's pre-existing knowledge) in an active, creative process that results in understandings not derivable by any solely linguistic or logical analysis (c.f., Bransford & McCarrell, 1975; Spiro, 1977, in press). Acceptance of the constructive view of comprehension entails a concomitant delimitation of the range of possible theories of mental representation. Knowledge structures must possess some capability for detecting the pragmatic, as well as logical, implications of the incomplete data contained in discourse (c.f., Charniak, 1974; Minsky, 1975; Rumelhart & Ortony, 1977, Schank & Abelson, 1977). In other words, knowledge structures must contain considerable information about the way the world usually works. This characteristic of representation is useful and efficient because natural and social contexts do produce sufficient constraints on worldly events and ideas as to make them, to a limited extent, orderly and predictable.

However, a point often overlooked is that these same knowledge structures, with their information about the world's orderliness, may allow for more efficient processing and memorial representation of explicit information in discourse, in addition to their role in deriving implicit information. This paper will be concerned with the psychological processing of (imperfectly) predictable or derivable information that is nevertheless explicit in discourse.

#### Predictable Information in Discourse

Despite the fact that most research on inferential processes in comprehension has been concerned with generation of implicit information, much inferentially related information is embodied explicitly in discourse. We are referring here primarily to pragmatic inferences, i.e., implications that are usually but not necessarily true. Language is infrequently characterized by absolute redundancy; semantic content is rarely "repeated," except for special purposes such as emphasis. However, pragmatic inferences are only imperfectly predictable. If you read that a karate champion hit a block, uncertainty is reduced by also reading that the block broke, despite the fact that that outcome

is usually to be expected. Similarly, it would not be considered unusual when relating the events at a birthday party to mention that there was a cake with candles blown out by the celebrant. Many things go in stereotyped ways but require explicit mention because the stereotype does not describe all possible cases. Throughout this paper, "predictable" is used as a shorthand for "imperfectly predictable, or characterized by significantly less than perfect uncertainty."

How is explicit but predictable information processed? As was mentioned above, attention has been primarily devoted to the processing of implicit predictable information, leaving little guidance on the present issue. However, in a variety of theoretical orientations, there is a common implication about how predictable information would be dealt with: simply put, explicit information, whether predictable or not, receives sufficient processing to be encoded in long-term memory. For example, Kintsch (1974) assumes "that subjects process and store [an inference] whether or not it is presented explicitly" (p. 154). It is difficult to imagine discourse representation theorists, who argue for the explicit representation in memory of implicit inferences (e.g., Frederiksen, 1975, Meyer, 1974), arguing that explicit inferences are not represented. In schema theories (e.g., Rumelhart & Ortony, 1977), explicit discourse information is used to bind schema variables, again suggesting that predictable information would receive explicit mental representation. If anything, one would expect existing theories to predict that explicit inferences would receive a stronger memorial representation than unpredictable information, given their greater contextual support. For example, in their associative network model, HAM, Anderson and Bower (1973) argued that the greater the number of interconnections between information, the greater the likelihood that information within the interconnected network would be recalled. This view will be referred to as the "storage of explicit inferences" (SEI) hypothesis.

An alternative hypothesis is that predictable information, however central to a discourse, is taken for granted, processed only superficially and receives an attenuated cognitive representation or no enduring representation

at all. If needed subsequently, it can be derived. This view will be referred to as the "superficial processing of explicit inferences" (SPEI) hypothesis. Processing explicit inferences in such a manner has the advantage of a cognitive economy of representation (besides a likely reduction in processing time). Most information that is acquired will never be used again. It would then seem to be more efficient to devote extra processing effort to the occasions when the information is needed (i.e., by deriving it when remembering) rather than exerting effort toward stable encoding at the time of comprehension.

#### Experiments on the Representation of Explicit Inferences

There are considerable problems in designing an empirical test of the hypothesis that explicit pragmatic inferences in discourse are not represented in long-term memory. If one merely tests memory for the inference, failure to remember could be attributed to not storing the information or to storing and then forgetting it, if the inference is remembered, it could be because it was stored and then retrieved, or it may have been generated at the time of test without having been stored.

Spiro and Esposito (1977) developed a paradigm not subject to the ambiguities of interpretation of the more simple design discussed above. The primary manipulation of interest involved subsequently vitiating the force of an earlier explicit inference. If the inference is not stored, certain predictable errors in recalling it should be made.

In the first experiment, subjects were presented stories which contained information A, B, and C such that B was strongly implied by A except in the presence of C. For example, the A, B, and C elements in one story (about a demonstration by a karate champion) could be paraphrased as follows:

- A: The karate champion hit the block.
- B: The block broke.
- C: He had had a fight with his wife earlier. It was impairing his concentration.

C was either presented prior to A and B (C-Before), after A and B (C-After), or not at all (No-C). When C was not included in the story, if SPEI is correct, the B element should be taken for granted, processed only superficially, and not stably represented. It would be derivable if needed. However, if C is presented after A and B, memory for B should be impaired since B was not stored and C will block its derivation from A at the time of test. On the other hand, if C occurs in the text prior to A and B, then B is not strongly implied by A. B cannot be taken for granted with the assumption that it can be generated later if needed. Here B should be stably represented and memory for B should not be impaired.

However, if SEI is correct, memory for B should not be affected by whether C is before or

after A and B, since B is stored whether it is implied by A (C-After) or not implied by A (C-Before). Two objections to this argument can be made. The information might be stored, but remembering C might lead to a decision that the memory for B must be mistaken (a kind of output interference). However, C is present whether it occurs before or after A and B, so such an explanation would not account for differential effects of C-placement. The other possibility is that B is represented in C-After, but the representation is altered or corrected when the C information is encountered. This possibility was investigated in the second experiment.

In the first experiment, the following predictions of the SPEI hypothesis were tested. More errors in response to questions about the presented predictable information (B) should be made in the C-After than in the C-Before conditions. Errors can be erroneous judgments that nothing about the implied information was presented, called B-Mention errors (e.g., the story did not mention whether the block was broken), or, when the subject believes that something about B was mentioned, remembering incorrectly what was specifically said in the direction of conforming with the C information, called B-Incorrect errors (e.g., it said in the story that the block did not break when he hit it). Confidence in errors of the latter kind were also analyzed. If subjects are as confident about these errors as they are about their accurate responses, it would be even more difficult to maintain the hypothesis that the explicit inferences were represented.

In the No-C condition, B-Mention errors may occur since B would not be represented according to the SPEI hypothesis. The more important prediction regarding the No-C condition is that B-Incorrect errors should not occur more often than in the C-Before condition. Otherwise, the differences between C-Before and C-After might be attributable to heightened accuracy due to greater salience of the implied information in the former condition rather than greater inaccuracy due to a failure to store the implied information in the latter condition.

College subjects read eight target vignettes each containing A and B information, and C information included or not and placed as a function of which of the three conditions subjects were randomly assigned to. C information was always on a separate page from the A and B information, and subjects were instructed to not look back after reading a page. After reading all the vignettes, the subjects were tested for their memory for the vignettes. Of particular interest were the two types of questions, mentioned above, concerning the B information (remember, B was always explicit in the stories).

The results supported the hypothesis that pragmatic inferences presented in text are superficially processed and do not receive a

stable and enduring representation in memory. In the C-After condition, subjects tended either to report that the inference was not presented in the text or that the opposite of the inference was presented. Furthermore, confidence in these errors was as high as confidence in correct memories. It is difficult to retain the notion that inferences are deeply processed and stably encoded when the C-After manipulation can produce errors like remembering the block was not broken when the karate champion hit it. The results cannot be attributed to interference produced by the inference-vitiating C information at output, since the C-Before subjects would also be subject to such interference. Neither can the results be attributed to differential availability of C at output, perhaps due to primacy/recency effects related to the position of C in the text, since the information was almost always recalled. Also, unimportance of the B information is not a viable alternative since B tended to be central to the story (e.g., in a story about a karate champion's performance, information about his success in the demonstration is certainly important).

One alternative interpretation that remains is that subjects do deeply process and stably encode the presented inference, but "correct" their representation when the inference-vitiating information is presented. If subjects are storing B and then changing or correcting it at the time C is presented, errors on B should occur in the C-After condition no matter how soon the test is administered after reading. However, if the SPEI hypothesis is correct, when delay intervals are brief enough some surface memory for the superficially processed B information may remain, reducing the number of B errors. Accordingly, in the second experiment subjects were tested either immediately after reading each story (Interspersed Questions condition) or, as in the first experiment, after the entire set of stories had been read (Questions-After condition). Again, the C-Before and C-After manipulations were employed.

The results of the second experiment replicated those of the first one in the Questions-After condition. Furthermore, the C-after effect was largely absent in the Interspersed Questions condition, demonstrating that the effect is not due to storing and then changing the representation of the B information (the explicit inference).

#### Related Issues

The discussion of implications of the superficial processing effect will at times be limited to reading rather than listening. Most of the following is of a speculative nature.

#### Representation and Underlying Mechanisms

Assuming some compatible representation system, what characterizes the processes that produce the superficial processing effect? At this time, only speculations about alternative possibilities can be offered. There are three potentially beneficial aspects of superficial processing of explicit predictable information:

cognitive economy (the information need not be specifically stored in long-term memory), speed of processing (you can process and understand such information rapidly), and automaticity of processing (less conscious effort and working memory space are required).

Two simple, preliminary accounts of the first factor, cognitive economy, can be offered. The superficial processing phenomenon appears most compatible with a schema-theoretic mode of representation. Perhaps variable bindings that are default (or at least high probability) values are not explicitly instantiated when they are explicit in discourse (but see the discussion of Determinants of Performance Variability below). However, one should not be overly persuaded by the simplicity of such an account. Other types of representation systems could also account for the phenomenon. For example, a spreading activation model (e.g., Collins & Loftus, 1975) might predict that explicit information is not tagged in memory when it has been recently activated with some greater than criterion strength. This issue will receive further discussion in the next section.

Regarding speed of processing, several possibilities may be offered: the information is actually predicted, perhaps followed by a selective scanning for partial clues of confirmation (e.g., the word "broke" in the karate champion example; perhaps such checks could be made in the visual periphery and, when positive, result in saccades that skip the predicted information), or the expectation may be formed after beginning to read the predictable information followed by skipping ahead to the next linguistic unit ("Oh. They're talking about this now. Well there's no doubt how it will turn out. I can pass this by."); or temporary binding of a schema variable (essentially a verification of fit) may be more rapid than more durable instantiation, or less metacognitive activity (pondering, studying, rehearsing, etc.) may be devoted to predictable information, given its derivability (this also relates to automaticity, obviously). Regarding automaticity, it seems likely that the amount of conscious processing required would be negatively correlated with the goodness of fit to prior knowledge. Thus conscious attempts to make sense of predictable information would be expected less often. Also, related to the suggestions above regarding expectations and rapidity of processing, the operation of some preattentive process (in the sense of Neisser, 1967) is a possibility. Naturally, it may be the case that all of these factors are contributing. However, some of the factors may be mutually exclusive. For example, if default values are processed automatically, an expectation and confirmation process may be redundant.

#### Determinants of Performance Variability

Occurrence of superficial processing and failure to store information probably depends on more than predictability or derivability considered in isolation. For one thing, the

derivability of other information in the discourse will have an effect. The greater the proportion of fit to one's schemata for the discourse as a whole, the more likely it is that conforming information will be left to be derived. If a story takes place in a restaurant, and all the restaurant-related information is typical, then that aspect of the story can be stored with the abstract schema node "typical restaurant activities." However, when the proportion of fit is poor, i.e., some atypical events occur, even typical, predictable events may have to be stored.

Occurrence of superficial processing is also likely to be affected by the extent to which the system is taxed. When the system is overloaded, as when there is a large amount of information to be acquired or the time to acquire the information is limited, more superficial processing and leaving of information to be derived probably goes on. Perhaps the system has flexible criteria for derivability, reducing criteria under overload conditions and increasing them when processing load is light (and when demands for recall accuracy are high or when subsequent availability of the information is limited). Briefly digressing, there may be a temptation to confuse superficial processing of derivable information with skimming. However, skimming is a selective seeking and then deep processing of situationally important information (see FRUMP, in Schank & Abelson, 1975) whereas superficial processing involves selectively not processing deeply information perceived as derivable, however important it might be. In other words, the same information that might receive more attention while skimming may receive less attention in normal situations if the information is derivable. This will happen to the extent that skimming results in shallow processing of earlier information that is the basis for the derivability of the later information.

Besides context-based variability in derivability criteria, research in the psychology of prediction indicates the potential operation of a general bias in determining the criterion for derivability and superficial processing. For example, Fischhoff (1975, 1977) has found that when people are told that some event has occurred, they increase their subjective probability estimate of the likelihood that the event was going to occur. Similarly, estimation of how much was known before being given a correct answer increases when the answer is provided. In the case of superficial processing of information in discourse, it is possible that the derivability of information is overestimated after it is explicitly encountered. It seems to be a fairly common experience, for example, to not write down an idea that you are sure will be derivable again later, only to find subsequent derivation impossible. What is being suggested here is a source of forgetting not usually discussed in memory theories: superficial processing of information whose derivability has been overestimated.

### The Form of Expression of Derivable Information

Semantic content, prior knowledge, and task contexts are not the only determinants of perceived derivability. The linguistic form in which information is expressed will sometimes provide signals of what information is already known or can be taken for granted, as when information is expressed near the beginning of a sentence (c.f., Clark & Haviland, 1977, on the given-new strategy). Taking an example from Morgan and Green (in press), compare sentences (1) and (2).

- (1) The government has not yet acknowledged that distilled water causes cancer.
- (2) That distilled water causes cancer has not yet been acknowledged by the government.

In (2) there is a stronger implied presumption of the truth of the proposition regarding distilled water and cancer than there is in (1).

In general, it seems that placing information in a sentence-initial subordinate clause lowers the superficial processing criterion. Consider continuations (3) and (4) of "The karate champion hit the block."

- (3) The block broke, and then he bowed.
- (4) After the block broke, he bowed.

The block's breaking would appear to be more taken for granted in (4) than in (3).

Linguistic signals of predictability or derivability need not be implicit. Consider continuations (5), (6), and (7) of the same sentence as above.

- (5) Obviously, the block broke.
- (6) As you would expect, the block broke.
- (7) Naturally, the block broke.

Words like "clearly" and phrases like "of course" are explicit linguistic signals that information to follow is predictable and can be superficially processed. However, one would expect that such signals could have their effect only for information within an acceptable range of plausibility. That is, a plausible but not predictable continuation may be more likely to be taken (erroneously) as predictable when preceded by a linguistic signal. However, if the information contains salient implausible aspects or something clearly irrelevant, a signalling phrase such as "as you would expect" might result in more attention being devoted to the continuation information.

### Implications for the Nature of Discourse Memory

To the extent that discourse is superficially processed, memory must be reconstructive rather than reproductive. Rather than retrieving traces or instantiations of past experience, the past must be inferred or derived. Just as a paleontologist reconstructs a dinosaur from bone fragments, the past must be reconstructed from the incomplete data explicitly stored. Evidence for such reconstructive

processes has been provided by Spiro (1977), who found a pervasive tendency for subjects to produce predictable meaning-changing distortions and importations in text recall under certain conditions. In general, when subsequently encountered information contradicted continuation expectations derived from a target story, the story frequently was reconstructed in such a way as to reconcile or cohere with the continuation information. This process of inferring the past based on the present was termed accommodative reconstruction. After a long retention interval, subjects tended to be more confident that their accommodative recall errors had actually been included in the story than they were confident about the accurate aspects of their recall. Why should such gross errors occur and then be assigned such high confidence? Part of the answer surely involves their function in producing coherence. Still, it is somewhat surprising that subjects should be so sure they read information that bore not even a distant inferential relationship to what they actually did read.

Spiro suggested that the basis for such an effect may be in the way information is treated at the time of comprehension; namely, it is superficially processed and not stored in long-term memory. Then, when remembering, individuals should know (at least tacitly) that considerable amounts of predictable or derivable information they have encountered will not be available in memory. In that case, recall would typically involve deriving a lot of missing information. Accordingly, it would not be surprising that subjects faced with memories that lack coherence would assume that missing reconciling information was presented but only superficially processed at comprehension. The information could then be derived at recall with high confidence. Hence the capacity for restructuring the past based on the present.

Individual Differences

A final caveat should be offered regarding the superficial processing effect, but also applicable to all research on schema-based processes in comprehension and memory. The assumption is usually made that there are no qualitative differences between individuals in the manner in which discourse is processed. However, Spiro and his colleagues have recently found that reliable style differences can be predicted in children (Spiro & Smith, 1978) and in college students (Spiro & Tirre, in preparation). Some individuals appear to be more discourse bound, tending toward over-reliance on bottom-up processes. Others are more prior knowledge bound, tending toward over-reliance on top-down processes. For the adult bottom-up readers, prior knowledge obviously must be used to a certain extent in comprehension. However, where use of prior knowledge is more optional, e.g., in providing a scaffolding for remembering information (Anderson, Spiro, & Anderson, 1978), the bottom-up readers capitalize less. Whether the latter type of individual will evince less knowledge-based superficial processing (again an optional use of prior knowledge) is a question currently under investigation.

References

Anderson, J. R., & Bower, G. H. Human associative memory. New York: Wiley, 1973.

Anderson, R. C., Spiro, R. J., & Anderson, M. C. Schemata as scaffolding for information in text. American Educational Research Journal, 1978, in press.

Bransford, J. D., & McCarrell, N. S. A sketch of a cognitive approach to comprehension. In W. B. Weimer and D. S. Palermo (Eds.), Cognition and the symbolic processes. Hillsdale, N.J.: Erlbaum, 1975.

Charniak, E. Organization and inference in a frame-like system of common sense knowledge. In proceedings of Theoretical issues in natural language processing. Cambridge, Mass. Bolt Beranek & Newman Inc., 1975.

Clark, H. H., & Haviland, S. E. Comprehension and the given-new contract. In R. Freedle (Ed.), Discourse processing. Hillsdale, N.J.: Erlbaum, 1978.

Collins, A. M., & Loftus, E. F. A spreading activation theory of semantic processing. Psychological Review, 1975, 82, 407-428.

Fischhoff, B. Hindsight ≠ Foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 288-299.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974.

Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Morgan, J. L., & Green, G. M. 'Pragmatics and reading comprehension. In R. J. Spiro, B. C. Bruce, and W. F. Brewer (Eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, N.J.: Erlbaum, in press.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

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

Schank, R. C., & Abelson, R. P. Scripts, plans, goals, and understanding. Hillsdale, N.J.: Erlbaum, 1977.

Spiro, R. J. Remembering information from text: The "State of Schema" approach. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1977.

Spiro, R. J. Constructive processes in text comprehension and recall. In R. J. Spiro, B. C. Bruce, and W. F. Brewer (Eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, N.J.: Erlbaum, in press.

Spiro, R. J., & Esposito, J. Superficial processing of explicit inferences in text (Tech. Rep. No. 60). Urbana, Ill.; Center for the Study of Reading, University of Illinois, 1977.

Spiro, R. J., & Smith, D. Distinguishing subtypes of poor comprehenders: Patterns of over-reliance on conceptual- vs. data-driven processes (Tech. Rep. No. 61). Urbana, Ill.. Center for the Study of Reading, University of Illinois, 1978.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 1975, 7, 371-458.

Meyer, B. J. F. The organization of prose and its effects on memory. Amsterdam: North Holland, 1975.

#### Footnote

This research was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116.