DOCUMENT RESUME

ED 234 361 CS 007 296

AUTHOR	Danks, Joseph H.
TITLE	Text Comprehension Processes in Reading: Appendix.
INSTITUTION	Kent State Univ., Ohio.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Sep 82
GRANT	NIE-G-78-0223
NOTE	388p.; For related document, see CS 007 295. Parts may be marginally legible.
PUB TYPE	Reports - Research/Technical (143) -- Collected Works - General (020)
EDRS PRICE	MF01/PC16 Plus Postage.
DESCRIPTORS	Adults; Beginning Reading; *Cognitive Processes; Comparative Analysis; Learning Theories; Oral Reading; Primary Education; *Reading Comprehension; *Reading Instruction; *Reading Processes; *Reading Research; *Reading Strategies; Research Methodology
IDENTIFIERS	*Reader Text Relationship

ABSTRACT

The readings in this collection were prepared to accompany a report of a series of experiments conducted to determine what information readers, both skilled adults and children beginning to read, use when they read to understand a story. Titles of the readings are are (1) "Oral Reading: Does It Reflect Decoding or Comprehension?" (2) "Models of Language Comprehension," (3) "Experimental Psycholinguistics," (4) "Comprehension in Listening and Reading: Same or Different?" (5) "An Interactive Analysis of Oral Reading," (6) "Comprehension of Prose Texts during Reading," (7) "Comprehension Processes in Oral Reading," (8) "Integration of Sentence Meanings in Stories," (9) "Comprehension of Metaphors: Priming the Ground," (10) "An Information Processing Analysis of the Cognitive Processes Involved in Oral Reading," (11) "Reading Comprehension Processes in Polish and English," (12) "A Comparison of Reading Comprehension Processes in Polish and English," and (13) "Memory and Metamemory Processes: Levels of Processing and Cognitive Effort in the Retention of Prose." (FL)

ED234361

Text Comprehension Processes in Reading: Appendix

Joseph H. Danks

Kent State University

National Institute of Education

NIE-G-78-0223

September 1982

CS 007 296

Contents

Danks, J. H., & Fears, R. Oral reading: Does it reflect decoding or comprehension? In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading, Vol. 3. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979. Pp. 89-108.

Danks, J. H. Models of language comprehension. Polish Psychological Bulletin, 1978, 9, 183-192.

Danks, J. H., & Glucksberg, S. Experimental psycholinguistics. Annual Review of Psychology, 1980, 31, 391-417.

Danks, J. H. Comprehension in listening and reading: Same or different? In J. Danks & K. Pezdek, Reading and understanding. Newark, Del.: International Reading Association, 1980. Pp. 1-39.

Danks, J. H., & Hill, G. O. An interactive analysis of oral reading. In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981. Pp. 131-153.

Danks, J. H. Comprehension of prose texts during reading. XXIInd International Congress of Psychology, Leipzig, 1980. Abstract guide, Vol. 1, p. 162.

Danks, J. H., Bohn, L., & Fears, R. Comprehension processes in oral reading. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), The process of language understanding. Chichester: Wiley, in press.

Danks, J. H., Bohn, L., End, L., & Miller, R. Integration of sentence meanings in stories. Psychonomic Society, St. Louis, 1980. (Abstract.) Bulletin of the Psychonomic Society, 1980, 16, 150.

End, L. J., & Danks, J. H. Comprehension of metaphors: Priming the ground. Midwestern Psychological Association, Minneapolis, 1982.

Danks, J. H. An information-processing analysis of the cognitive processes involved in oral reading. American Educational Research Association, New York, 1982.

Kurkiewicz, D., Kurcz, I., & Danks, J. H. Reading comprehension processes in Polish and English. Polish Psychological Bulletin, 1981, 12, 25-31.

Danks, J. H., & Kurcz, I. A comparison of reading comprehension processes in Polish and English. Submitted for publication, 1982.

Bohn, E. A. Memory and metamemory processes: Levels of processing and cognitive effort in the retention of prose. Submitted for publication, 1982.

In L. B. Resnick & P. A. Weaver (Eds.), Theory and Practice of Early Reading, Vol. 3. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1979.

4 Oral Reading: Does It Reflect Decoding or Comprehension?

Joseph H. Danks
Ramona Fears
Kent State University

How do most elementary school teachers determine whether a child can read? It is likely that the teacher hands the child a book at the child's estimated reading level and asks the child to read. That the reading is to be aloud is usually not stated but understood implicitly by both teacher and child. What reading activity is commonly found in most traditional lower-elementary-grade classrooms?--children reading aloud individually or in unison. Oral reading provides the teacher with a quick evaluation of each child's progress and provides the child with practice on at least some aspects of reading. Although oral reading is a widely used procedure, we lack well-developed theories concerning what specific components of the reading process are assessed and what reading skills are developed by oral reading.

TWO HYPOTHESES ABOUT ORAL READING

A general model of reading that is commonly assumed proposes that print input is first decoded into a phonological code that has most of the characteristics of an oral verbal input. This code is then comprehended by the usual routines of language comprehension that the child has developed for speech. These two stages of reading are usually seen as discrete stages that can be taught independently. This assumption has led to a division of reading instruction into teaching decoding and comprehension. Decoding depends only minimally on comprehension, although some investigators posit "downstream" effects on the basis of top-down models of comprehension. For example, a major aspect of Goodman's (1967) informed-guessing model

is that previously comprehended material facilitates the decoding of print. However, these downstream influences of comprehension on decoding are usually not thought to be essential for successful decoding but only helpful when the contextual information is available. On the other hand, comprehension is necessarily dependent on decoding for the representation on which to operate. Except for providing the input for comprehension, however, the decoding process does not directly affect the comprehension process [Eds. note: See also Frederiksen, and Goodman and Goodman, Volume 1, this series].

Given this rough two-stage model of reading, where does oral production fit? The motor production aspect of oral reading must be "tacked on" to the model of the reading process, because there is no production component explicit in the model. There are two general hypotheses as to when oral production is initiated in the reading process. These hypotheses are illustrated in Fig. 4.1.

The decoding hypothesis is that oral production is initiated immediately following decoding (point A in Fig. 4.1). Oral production is initiated on the basis of the phonological code that is the output of decoding. (Oral reading then serves as practice during the initial decoding stage.) In this case, there might be no comprehension of the text at all; or oral production and comprehension might progress in parallel; or comprehension might occur much later than the oral production, perhaps as the reader hears him- or herself talk. Children sometimes imitate speech without comprehending it, so it would not seem unreasonable that beginning readers might initiate oral production on the basis of the phonological code without comprehending the message first.

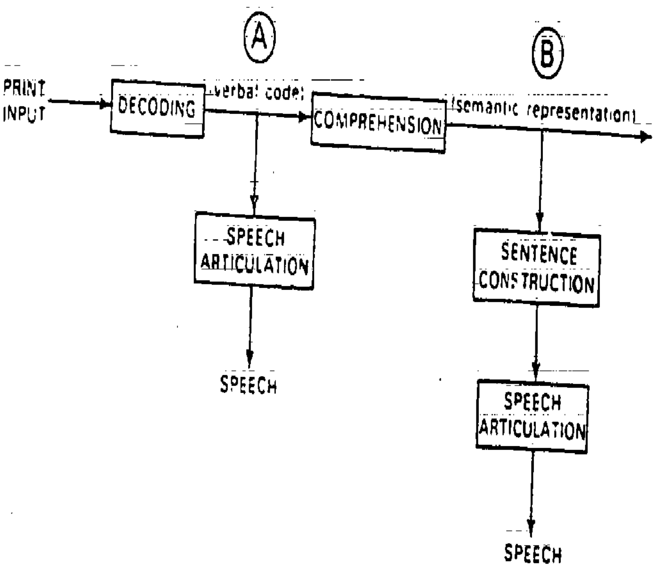


FIG. 4.1. Two hypotheses about oral reading performance.

The comprehension hypothesis is that oral production is initiated only after comprehension processes have constructed a semantic representation of the message (point B in Fig. 4.1). Oral production is initiated from the semantic representation and not from the phonological code that existed at an earlier point in the process. In fact the verbal code may no longer exist at the time at which oral production is initiated. The oral production process beginning with the semantic representation resembles sentence production in its essential components. A speaker has an idea to express, and he or she translates that idea into lingusitic form and then expresses it in speech. In oral reading, the semantic representation of the printed message constitutes the idea that serves as the input to the production process.

These two hypotheses represent classes of hypotheses, with variations depending on the specific conceptions of decoding, comprehension, and speech production. For example, comprehension is described as if it were a single process with a fixed beginning and ending. However, comprehension may be a loose collection of processing strategies rather than a single routine. The possible variations in these two classes of hypotheses will become evident. For the moment, they serve as convenient touchstones for conceptualizing the question of how oral production meshes with the reading process. The two hypotheses can be differentiated further by comparing two types of readers that appear to embody each hypothesis.

Word Callers

The decoding hypothesis appears to be supported by reading disabled children labeled *word callers* (Smith, Goodman, & Meredith, 1976). These are children who can read aloud but who do not understand what they have read. One would assume on the basis of their oral reading performance that they understand. However, when they are tested for comprehension of the message, they have at best a minimal understanding of what they have just read. Word callers' understanding is not improved when they are permitted to read silently. Although it is not clear what they are doing when they are silent reading other than staring at the page, their understanding is not increased. So the problem is not that their normal comprehension process is disrupted by the additional task of having to read aloud. In terms of our general model, then, word callers support the decoding hypothesis in that their oral production must be initiated immediately following decoding but before comprehension occurs.

At one time or another, many adults have had the experience of reading aloud without comprehending. When learning a foreign language, many people pass through a phase when they can read aloud in the second language but not understand what they read. Or when reading some particularly difficult text like a philosophic treatise, one might read it aloud to allow more

time to think about what is being said but still might not understand what is written. It seems possible, then, that at an early point in reading acquisition, some beginning readers might be able to read aloud on the basis of their decoding skills but not comprehend the message.

There is considerable dispute, however, over whether word callers really exist and over what the criteria should be for so labeling a child. Goodman (1973) claimed that "remedial reading classes are filled with youngsters in late elementary and secondary schools who can sound out words but get little meaning from their reading [p. 491]," although he adduced no statistics to support that claim. Other reading specialists claim that the number of true word callers is exceedingly small, because children who are labeled "word callers" by classroom teachers actually have poor decoding skills or poor language comprehension skills.

What criteria should be considered in classifying a child as a word caller? According to the traditional definition, the child must be able to read aloud reasonably well and not understand what was read. What is meant by *reasonably well?* At minimum, the child must read at close to the typical rate of comprehending readers with about the same number of errors and with normal intonation. There is some question of whether word callers can meet these criteria. For example, poor readers tend to read with a list intonation (Clay & Imlach, 1971). Reading with a list intonation is a clear clue to a lack of comprehension, because decoding punctuation and combining that information with the meaning of the passage leads to intonation patterns more typical of speech. What is meant by *not understand?* The key test is whether the child can understand the passage if it is presented aurally. If adults had the foreign language text or the philosophy essay read aloud, they would not understand it any better than when they read it themselves. If the child does not comprehend when listening, the problem may be attributable to a general language or conceptual deficit rather than to a deficiency in reading-specific comprehension skills or in the coordination of decoding and comprehension processes.

Word callers may be related to a class of children labeled "hyperlexics," who are superficially similar to word callers in their reading behavior (Mehegan & Dreifuss, 1972; Silverberg & Silverberg, 1967, 1968–1969). The common distinguishing feature is that "they manifested an unusual and premature talent in reading [aloud] against a background of generalized failure of development, or marked impairment, of other language functions" (Mehegan & Dreifuss, 1972, p. 1106). Their reading is a voracious compulsion that frequently develops in the preschool years. In addition, they are frequently retarded, autistic, or hyperkinetic. Perhaps these children represent an extreme instance of word calling mixed with an intellectual or

emotional disturbance, or they may be a qualitatively different type of reade Given the disagreements about both word callers and hyperlexics, caref investigations and descriptions of both are needed.

Dialect Speakers

Readers whose oral reading appears to support the comprehensio hypothesis are those whose oral language dialect is different from the diale of the primers. The most salient example in the United States are childre who speak a Black dialect and learn to read from primers printed in standa English. When asked to read aloud, they produce numerous "errors" in or production; that is, their speech does not match the speech that one wou expect, based on the print. However, their deviations are not arbitrary wi respect to the meaning of the text. Many of these "errors" do not change t meaning of the text but are a translation of the message into the reade dialect. [Eds. note: See Shuy, Volume 1, this series, and Simons, this volum

Dialectal variation occurs at all linguistic levels, although phonem miscues are the most frequent in oral reading (Burke, 1973). Although Bla children may "mispronounce" a printed word (e.g., /ro/ for *road*), th comprehend the meaning of the printed word (Melmed, 1973). Many ot Black English responses in oral reading are morphological variations, such dropping regular past-tense and third-person singular endings on verbs a plural and possessive markers on nouns (Rosen & Ames, 1972a, 197 Weber, 1973). Lexical substitutions also occur; for example, *bucket* for *p* and *gym shoes* for *sneakers* (Burke, 1973). However, we have found studies that have reported cases where dialectal variations changed mean more than oral reading errors of standard English speakers changed meani

To translate the text into their oral dialect without a change in mean necessitates that Black children first comprehend the printed text. A forti they must decode the standard English text correctly before comprehend it. It is incorrect to claim that these children have deficient decoding skills fact, it is inappropriate to label these children as having a reading probl They know how to read. The variation in oral reading results from a variat in speech that is different from standard English.

At the very least, oral reading is an inappropriate assessment tool w applied to these children unless the "errors" are interpreted in terms of child's own dialect. Hunt (1974–1975) scored Black children's response the Gray Oral Reading Test both according to the manual and correcting dialectal responses. She found an increase of only 0.1 grade level between two scoring systems, although that difference was statistically signific However, the better readers (as defined by the standard scoring of the test

who were still below grade level) were helped more by the dialectal scoring. They gained one-half grade level on the average with some children gaining more than a whole grade level.

Using oral reading as an instructional device in the classroom must be tempered by a teacher who understands that a child who translates into his or her own dialect is reading correctly (Goodman, 1965a; Goodman & Buck, 1973; Labov, 1967). Otherwise, the teacher may underestimate the child's level of reaching achievement and may put undue pressure on the child by constantly "correcting" his or her oral productions. The child also may be confused by not understanding why the teacher is correcting what he or she is reading correctly (Fasold, 1969).

These children who read orally by translating the standard English of the primers into their own dialect provide convincing evidence for the comprehension hypothesis. Accurate comprehension (and therefore accurate decoding) must have occurred before the child initiated oral production.

RESEARCH ON ORAL READING

Given the identification of these two types of readers, one of whom apparently supports each hypothesis, what empirical evidence is available? Three sources of evidence are reviewed. One source is the analysis of oral reading errors. The errors or miscues are evaluated with respect to the reading processes that underlie the performance. A second source of evidence is studies of eye-voice span. If the eye-voice span varies with the semantic, syntactic, or conceptual difficulty of the text, then the reader may be comprehending the meaning before initiating the oral production. The third source of evidence is a task in which the text material is altered—for example, a misspelled word, a wrong part of speech, a semantically anomalous word, or a logical inconsistency. Whether the oral reader is disrupted by a particular type of alteration in the text indicates whether he or she is using that type of information to process the text.

Oral Reading Errors

A major problem with the literature on oral reading errors has been the lack of agreement on a classification system to analyze the errors (Weber, 1968). The classification schemes reflect the investigators' underlying assumptions about the nature of the reading process. Those who view oral reading primarily as a performance skill score as errors hesitations, poor enunciation, and inappropriate intonation and phrasing (Weber, 1968). Those who view oral reading as a reflection of underlying processes have focused on the

graphic-phonetic similarity and syntactic-semantic acceptability as two major determinants of oral reading errors.

Using specially constructed word lists, Shankweiler and Liberman (1972) found that optical confusability, as exemplified by reversals of letter sequence and orientation, was a much less significant factor in producing oral reading errors than were orthographic factors, such as position of the sound segments and phoneme-grapheme correspondences. Initial segments were better read than medial or final ones, and consonants were read better than vowels. Errors on vowels were predicted by the number of possible orthographic representations.

Using word lists as opposed to prose precludes any evaluation of syntactic and semantic determinants. Shankweiler and Liberman (1972) justified their use of word lists by the fact that there were significant correlations (averaging .70) between error scores on oral reading of lists of words and error scores on the Gray Oral Reading Test in each of four groups of children. They concluded that "the problems of the beginning reader appear to have more to do with the synthesis of syllables than with scanning of larger chunks of connected text [p. 298]." However, since the word list data accounted for only about 50% of the variance on the Gray Oral Reading Test, considerable variance remains to be explained by syntactic and semantic components of connected text. Goodman (1965b) reported that many words that were missed when they appeared on a list of isolated words were read correctly when they appeared in a story context. In fact, first graders read 64%, second graders read 75%, and third graders read 82% of the missed words correctly given the syntactic and semantic constraints of the story.

The term *semantic constraints* is usually used to refer to the meaning of the sentence constraining what lexical items might meaningfully complete the sentence. *Semantic constraints* also can be used to refer to access to the meaning of a word in the lexicon, or "mental dictionary." Two experiments with isolated words are particularly relevant. Perfetti and Hogaboam (1975) reported that more skilled comprehenders were more rapid at word recognition (and pronunciation) than were less skilled comprehenders, even when all words were known to both groups. The difference between the groups was larger for infrequent than for frequent words. Golinkoff and Rosinski (1976) used a somewhat different task in which automatic semantic access would interfere with the subject's performance on picture naming. They found that although less skilled comprehenders were weak on decoding their semantic access skills were not impaired. The results of these two studies are inconsistent with respect to whether more and less skilled comprehenders differ in semantic access. However, the fact that variation in semantic access affected naming responses indicates that semantic access occurred before the naming response was initiated.

Other investigators have compared the effects of graphic and syntactic-semantic constraints on errors in connected text. They have uniformly reported that oral reading errors are represented more accurately as alternatives that are syntactically and semantically plausible than as alternatives that match graphic constraints. Out of 7,674 substitution errors committed by first graders, Clay (1968) reported that 72% were syntactically appropriate but that only 41% could be attributed to grapheme-phoneme correspondences. Biemiller (1970) found that first graders' graphically similar substitution errors were less frequent than were contextually appropriate substitutions. A similar finding has been reported by Weber (1970a, 1970b); however, good first-grade readers were more influenced by graphic similarity than were poor readers. Visual graphic form appears to constrain the choice of a response from a set of possible words previously limited by syntactic and semantic constraints. The graphic form of a word does not appear to be a primary factor per se.

That the large majority of substitution errors in oral reading tend to be syntactically and semantically appropriate is well documented. In one case, oral reading errors of college students were not just syntactically appropriate but were predicted by a formal grammar, an augmented transition network (Stevens & Rumelhart, 1975). Studies by Biemiller (1970), Clay (1968), Cohen (1974-1975), Goodman (1965b), and Weber (1970a, 1970b) reported that first graders' oral reading errors tend to be grammatical and meaningful for the context up to the point of the error. Frequently, the error is grammatically and semantically consistent with the remainder of the sentence as well. If not, self-correction usually occurs (Clay, 1969; Goodman, 1965b; Weber, 1970a). Weber (1970a) reported that ungrammatical errors were more graphically similar to the printed word than were grammatical errors, illustrating a trade-off between these two determinants. Biemiller (1970) studied first graders longitudinally through the first grade. He found an initial phase in which the errors were semantically appropriate to the prior context but were minimally related to the graphic form of the stimulus word. During a second phase called the nonresponse phase, the number of errors dropped, and there was an increase in graphically similar substitutions whereas the percentage of contextual (semantically appropriate) substitutions remained constant. Finally, in the third phase, the percentage of contextual substitutions increased whereas the percentage of graphic substitutions remained stable. A comparable sequence has been reported by Clay (1969) for self-corrections by first-grade children (5 years old in New Zealand). Initially, errors were self-corrected only if they did not make sense in the context; then visually dissimilar errors were corrected; and finally, both factors were relevant, so that only a substitution that made sense and was graphically similar was left uncorrected.

There is some evidence that certain types of oral reading errors may be a partial function of the instructional program. The children observed by Biemiller, by Weber, and by Clay were receiving reading instruction in meaning-emphasis programs. Cohen (1974-1975) analyzed the oral reading errors of first graders being taught with a code-emphasis approach. Her results were a bit different from Biemiller's in that she found only a few readers who went through an initial phase of giving contextual responses. Instead, most started out in a brief nonresponse phase. In the next phase, these children produced a significant number of nonsense words. Evidently, the emphasis on sounding out words and attempting to pronounce them induced children to make up words based on the graphic stimulus. Following the phase in which nonsense errors predominated, the children began producing meaningful word substitutions as the context gained in importance.

It appears that the primary conclusion to be derived from studies of oral reading errors is that readers proceed through the comprehension stage before initiating oral production. How else could syntactic and semantic constraints have such a powerful effect on oral reading errors unless such were the case? Not only is this conclusion true for skilled adult readers, but it also holds for beginning readers and for good readers as well as poor readers. Thus, the comprehension hypothesis is substantially correct, and word callers exhibit a clear reading disability that is qualitatively different from typical reading.

Although this logic is appealing, we argue that it is incorrect. From the fact that a child makes a grammatically and semantically appropriate error, one cannot conclude with certainty that he or she has comprehended the intended meaning of the text before initiating oral production. The child may have constructed an interpretation or meaning for the prior text and filled in unknown, missing, or unsampled words on the basis of the constructed meaning. On some occasions, the constructed meaning may be the same as the textual meaning, but this correspondence does not necessarily indicate that the child obtained that meaning by processing the text word through to its semantic representation and then substituting a synonym at production. If a child accurately comprehended the text prior to making an error (the comprehension hypothesis), then not only would the error be acceptable in the sentence up to that point: it would be a close paraphrase as well. A syntactically and semantically acceptable error is not necessarily a close paraphrase. For example, substituting *car* for *cat* in "The girl saw the cat run across the road" does not yield a close paraphrase of the sentence even though *car* is syntactically and semantically acceptable. There would be no need to correct an error that is a close paraphrase, because there would be no inconsistency with the remainder of the sentence. However, substitution

errors are frequently self-corrected because they are ungrammatical or inconsistent with the remainder of the sentence (Clay, 1969; Goodman, 1965b; Weber, 1970a). Thus, the self-correction phenomenon suggests that although the substitution errors are syntactically and semantically acceptable, they are not close paraphrases. The child predicts a meaning, produces an oral response that is appropriate for that meaning, realizes that the substitution is inconsistent with the later text, and then repeats to correct the error.

Our argument is that there are at least two possible mechanisms by which syntactically and semantically appropriate errors are produced in oral reading. One is what we originally suggested as the comprehension hypothesis; that is, the error is an output error in which the message was correctly comprehended but then was translated into the reader's idiolect. This mechanism results in close paraphrases and describes what happens with dialect readers. The second possible mechanism is that the substitution is generated on the basis of the preceding context. In this case, the text word is not actually read, but a response is produced based on the semantic representation of the preceding text. This might occur primarily under speed pressure or when the word is unknown but a response is required. This latter mechanism corresponds in many respects to Biemiller's (1970) first phase, in which the first graders respond with a contextually appropriate response that may not have a close correspondence to what is actually printed. "The actual graphic display takes second place to grammatical acceptability. Reasonable as this might be as a tactic for the young reader, he [or she] must sooner or later read what is actually written rather than what he invents" (Gibson & Levin, 1975, p. 281).

The difference between the two mechanisms may be reflected by comparing the reading errors with the child's own idiolect. If the oral production is an accurate translation of the message of the printed text into the child's idiolect, then one would be justified in concluding that the "error" is an oral production change. If a Black child deleted markers for past tense, third-person singular, plurality, or possession (Rosen & Ames, 1972a, 1972b), then one reasonably could conclude that the "errors" were oral production changes. Taking an example from Weber (1970a), however, if a reader substituted *dimes* for *money*, it is more likely that he or she ignored the graphic stimulus and generated the error from the prior context.

Perhaps the two mechanisms can be differentiated by the level of the error. Pronunciation and morphological errors would be the result of oral production differences, but more complex syntactic and semantic substitutions would be contextually based responses. This interpretation probably is not adequate. Consider another example from Weber (1970a). She called "I will see what *is it*" an ungrammatical response to the printed sentence, "I will see what *it is*." It is possible that the erroneous response was

"grammatical" in the child's idiolect. The failure to invert the subject and predict in embedded relative clauses is characteristic of one stage in children's acquisition of embedded clauses (Menyuk, 1969). Although the failure of inversion more typically is found in nursery-school children than in first graders, any one of several factors—for example, slight oral language delay in this particular child—could explain its presence. Although no information was reported about this child's idiolect, the example does illustrate that we cannot identify any particular level of oral reading error as either an output error or a contextual error without a comparison with the child's oral language.

From the fact that a child produces syntactically and semantically appropriate errors in oral reading, the teacher should not infer automatically that the child is comprehending and therefore reading adequately. The source of the errors must also be determined. Only after a comparison with the child's idiolect, rather than a comparison with an adult's responses to the same graphic stimulus, can one determine whether the error is a translation based on a veridical semantic representation or a guess based on the preceding text.

Eye-Voice Span

When moderately skilled readers read aloud, the eye is fixated on the line of print somewhat in advance of the word being vocalized. This difference (typically measured in words) is called the eye-voice span. The eye-voice span is influenced by a number of factors—age and skill level of the reader (Levin & Turner, 1968), difficulty of perceptual processing (Resnick, 1970), syntactic structure (Levin & Kaplan, 1968, 1970; Rode, 1974-1975; Schlesinger, 1968), difficulty of the material (Buswell, 1920; Fairbanks, 1937), and task demands for the reading (Levin & Cohn, 1968). For the current discussion, the most important conclusions drawn from among the studies are that the eye-voice span is responsive to syntactic structure and tends to terminate at phrase and clause boundaries. The usual interpretation has been that readers read in phrase or clause units. The reader may actively construct a hypothesis about what is being read and then test that hypothesis against the printed text. Thus, if reading is an active sampling, constructing, and testing process, then one would expect hypothesis generation to be defined by syntactic and semantic units. If so, the reader comprehends the material before initiating oral production; in fact, it is the semantic representation that permits correct continuations after the visual stimulus has been removed from view.

This interpretation of eye-voice span is subject to the same objection that we raised to the usual interpretation of oral reading errors. Even if the eye-voice span is influenced by the preceding context and comprehension of the preceding meaning, one does not have certain evidence that the reader

comprehended the printed material before vocalizing it. One may object that our argument rests on the premise that the reader has comprehended the preceding material. Although such may indeed be the case, the reader may have comprehended the preceding material only after he or she had vocalized it, perhaps by having comprehension and oral production run in parallel or perhaps by comprehending the oral output itself. In any case, comprehension prior to initiation of oral production is not required by our premise.

If readers in eye-voice span experiments base their responses in part on informed constructions of preceding material, then errors should be made in the eye-voice span. Rode (1974-1975) reported such errors for third, fourth, and fifth graders. On 15% of the trials, an erroneous word was substituted for a printed word between two (or more) correct words. She found that 62% of these errors were syntactically and semantically appropriate, and even fewer violated the syntax of the sentence. In a recognition test following eye-voice span measurements, Levin and Kaplan (1968) found extremely few false-positive responses (0.1%), indicating that readers were not purely guessing. In both studies, subjects responded at least partially on the basis of the preceding text.

Text Alterations

Although the analysis of spontaneously occurring oral reading errors provides a wealth of information, not infrequently the type of error needed to answer a specific question is not committed. We then are faced with the uncertainty of not knowing whether the reader was not influenced by that particular aspect of reading or whether the text we chose for reading did not give the reader the opportunity to commit such an error. The uncertainty can be alleviated by inserting inconsistencies of the type we wish to study the text. Various types of structure in the text (e.g., graphic, syntactic, semantic) are used by one or more components of the reading process. Altering a specific type of structure will disturb any component that uses that information. Moreover, the types of structure and the corresponding processing components can be ordered from perceptual processing, using graphic information, to more abstract levels, such as components that use syntactic, semantic, and prose structure. If we are interested in whether semantic access occurs, a very infrequent word, one unlikely to be known by the reader, or a pronounceable nonsense word can be inserted in the text. Or if we are interested in whether the reader is integrating sentence meaning, we can insert a word that produces an inconsistency in meaning. If oral reading is disrupted in either instance, then we can reasonably conclude that the text was processed at least by the components using the changed structure.

Siler (1973-1974) attempted to differentiate between syntactic and semantic determinants of oral reading errors. He introduced semantic

disruptions (a word that was the correct part of speech but that was anomalous in the context) and syntactic disruptions (an inversion of one pair of words). Syntactic disruption produced a larger effect on both oral reading time and oral reading errors than did the semantic one. Both types of disruption had an effect on oral reading, but it is difficult to draw a comparative conclusion because there is no common scale on which to compare the relative magnitudes of the syntactic and semantic manipulations (Danks, 1969; Dooling & Danks, 1975).

Lazerson (1974-1975) had college students read Caxton's preface to the *Eneydos* (1490), which was printed in Late Middle English with variable spellings and an archaic syntactic-semantic system. In some conditions, Lazerson corrected only the syntax-semantics to conform with Modern English; in some conditions, both the spelling and the syntax-semantics were corrected. Archaic syntax-semantics increased oral reading time, and the addition of variable spelling increased it even more, but there were no differences in comprehension. The variable spelling and the archaic syntax-semantics probably affected the performance system but not reading comprehension per se.

In both Siler's and Lazerson's experiments, the disruptions in oral reading were measured in terms of total time and total errors in reading a passage. These overall measures demonstrate that processing involved syntactic and semantic components, but they are too gross to determine whether processing occurred before or after oral production had been initiated. If oral reading "disfluencies" (i.e., disruptions) are measured relative to a specific alteration in the text, then the point of initiation of oral production can be specified more precisely. Where the oral reading disfluency occurs relative to the change in the text provides a means of deciding between the decoding and comprehension hypotheses. If the disfluency occurs before the reader has uttered the altered part of the text, then the text must have been processed to that level prior to initiating production. However, if the disfluency occurs after the altered section has been uttered, then the processing at that level occurred much later, perhaps even in response to the oral output itself.

Three experiments have used a more precise procedure of introducing specific alterations and measuring oral reading disfluencies in the immediately surrounding text. These three experiments introduced alterations corresponding to three levels of processing—lexical access, syntactic and semantic integration, and intersentence integration.

To determine the effects of disrupting lexical access, Miller (1975) introduced four types of modifications into paragraphs—infrequent words, pronounceable nonsense words with and without syntactic markers, and phonologically impossible sequences. He measured substitution errors in the original text surrounding the inserted word; hesitations, incorrect intonation, and other performance variables were not measured. In second graders' oral

reading, there were increased errors on the two words immediately preceding and following the inserted word, but there was no effect due to the type of text alteration. Since none of the inserted words were available in the "mental dictionary" of the child (including the infrequent words), the reader was unable to locate the word before pronouncing it. Thus, attempts at semantic access occurred prior to oral output. The type of alteration did not make a difference, because the same process, lexical access, was disrupted by all alterations.

Miller and Isakson (1976) assessed intrasentence integration by substituting verbs in sentences. Semantic integration was disrupted by replacing a transitive verb with another transitive verb that was semantically unacceptable. For example, *paid* replaced the verb *planted* in the sentence: "The old farmer planted the bean seeds in the rich, brown soil." Both syntactic and semantic integration were disrupted by substituting a semantically anomalous intransitive verb—for example, *went*—in the foregoing sentence. Subjects were groups of fourth graders who had been divided into good and poor comprehenders (more than or less than one-half year above or below grade-level placement on the reading comprehension subtest of the Iowa Test of Basic Skills, respectively; all readers were within one-half year of grade placement on the vocabulary subtest). The only effect of the verb substitutions was increased oral reading errors at the verb position by good comprehenders. Only the readers who performed at a relatively high level of comprehension skill processed the semantic and syntactic information. The results suggest that integration occurred prior to the initiation of the oral response, because the production of the altered word itself was disrupted. Perhaps with even more skilled readers, the disruption would occur one or more words prior to uttering the altered word.

In a recently completed study, we[1] assessed the disruptive effect of a conceptual inconsistency or contradiction. Within a paragraph, a sentence was altered such that it was inconsistent with a single critical word in the next sentence but not inconsistent with the rest of the paragraph either preceding or following. For example, the first sentence was replaced in the following pair of sentences (taken from Patton's memoirs of World War II): "I then told him that, in spite of my most diligent efforts, there would be some raping. I should like to have the details as early as possible so that we can hang these men. He said that this was . . . ." The replacement was: "I told him there would unquestionably be some helping by the soldiers," which produced an inconsistency with the critical word *hang* in the second sentence. Note that the second sentence containing the critical word was not altered in any way. There were 10 experimental paragraphs, and the location of the text alteration varied in the paragraphs. Two groups of 10 college students read these passages aloud, and the readings were tape-recorded. Half the passages for each group were altered to produce the inconsistency, and half were left unchanged. Complementary sets of changed and unchanged passages were presented to the two groups of readers. The intervals between saying each of the 5 words preceding and following the critical word in the second sentence, as well as the length of time to say the critical word, were measured by playing back the tape at one-fourth the recording speed. Because the time distributions were skewed, the data were transformed logarithmically.

Subtracting the control group means from the experimental group means, the curve depicting differences in interword time intervals across position in the sentence showed a significant disruption 2 and 3 words following the critical word. The time to say the critical word itself was longer for the experimental group than for the control group, a difference that was significant across both readers and passages.

The results suggest that the reader had comprehended the material prior to initiating oral production of the critical word itself. Detection of the inconsistency required detailed and integrated comprehension. It depended not just on access to the lexical item in semantic memory and not just on comprehension of the sentence currently being uttered but on integration with the semantic representation of the preceding sentence as well. This integration with prior context requires additional time to accomplish (Dooling, 1972). Even if one assumes a constructive or top-down comprehension process, the match between the expected meaning of the second sentence and the actual meaning must have occurred at an abstract level of representation, because the inconsistency could not have been detected on a perceptual level or by comparing individual words or phrases. Thus, comprehension must have occurred well ahead of oral production.

THE TWO HYPOTHESES REVISITED

Our conclusion is that neither the decoding nor the comprehension hypothesis holds all the time but that the particular processes involved in oral reading are reader and task specific. A given reader with specific materials and a definite purpose for reading processes the text to the extent that he or she is capable and to an extent consistent with the implicit or explicit purposes. The reader then initiates the oral production process at that point. Word callers are unable to progress beyond the decoding stage before initiating their oral production, but dialectal speakers are able to comprehend the text before initiating production. Reader limitations, textual variation, and purpose affect processing in somewhat different fashions. The reading level of the reader, whether limited by level of acquisition or by skill, is a

[1]This experiment was designed in collaboration with Karen Fischer. We thank Mark Germano for his assistance in conducting the experiment.

limitation of the system. It sets an upper limit on the level of processing that the reader is able to attain. The other two factors, difficulty of the text and the purpose of reading, result in variation in the level of processing.

Reading Level. The level of reading skill sets an upper bound on the processing that a reader can accomplish prior to the initiation of oral production. A limitation resulting from level of acquisition is most often indexed by the age or the grade level of the child. Good and poor readers typically are defined by whether or not they exceed or fall below the grade level corresponding to their chronological age.

The level of processing is not independent of rate of processing and memory limitations. There is evidence that children do not differ from adults in the absolute size of their memory capacity but that children use their memory stores less efficiently (Chi, 1976). This impoverished ability to select and store relevant information is particularly critical in reading where integration of information must occur over a span of input. The reader must develop strategies for the effcient intake and storage of printed information. The rate must be fast enough that the requisite information exists simultaneously in memory so that it can be integrated. However, if the initial stages of the reading process are not sufficiently fluent or are not sufficiently automatic to proceed with minimal or no attention, then the rate of input will not be sufficiently rapid to overcome memory limitations. The beginning phases of reading acquisition are occupied with the practice of decoding skills. When they become sufficiently automatic, attention can be directed to the later comprehension stages (LaBerge & Samuels, 1974). [Eds. note: See Perfetti and Lesgold, Volume 1, this series.]

With increasing age, readers become better able to control the strategies necessary for processing efficiently so that they can be directed toward the particular task set for them. Just as they can better control the strategies involved in the efficient storage and retrieval of material, they can control the reading process to particular ends.

Purpose or Task. The reading task influences the level of processing accomplished. Keeping in mind that the level of reading acquisition sets an upper limit and that the general developmental level of the reader may determine his or her ability to control the reading strategies, a reader can focus on decoding, on comprehension, or on oral production. If the reader expects to be tested on his or her knowledge of the content of the passage, then the reader will attempt to comprehend the passage more thoroughly. However, if the reader expects to be evaluated solely on oral production, then he or she may focus attention on decoding and pronunciation, thereby ignoring comprehension.

The schoolchild who must read for the teacher with the class listening probably pays particularly attention to decoding and oral production so as not to make a mistake. Typically, the teacher and the other children follow the text, so they know immediately if a mistake is made. Pehrsson (1974) tested fifth graders under such conditions. When the teacher focused attention on correct decoding and oral production, reading rate and comprehension decreased, as one might expect. But unexpectedly, oral reading errors increased. Conversely, Pehrsson found that if the children were permitted to read aloud without interruption and had to retell what they had read, then comprehension increased.

Text Difficulty. The text can vary in difficulty at several different levels. The type font may make it difficult to discriminate the letters, the vocabulary may be difficult, the syntactic structure may be complex, and the ideas and conceptual organization may be abstract or obscure. These levels of difficulty interact with the level of reading of which the reader is capable. If the reader is concentrating attention on decoding, either because of ability limitations or task orientation, then the complexity of the syntax and the difficulty of the conceptual structure will not have an effect on oral production. If there is an inconsistency at a higher level than the reader is capable of processing, then there will be no disfluency in oral production.

Comprehension processes in oral reading cannot be evaluated by using a list of words. Lexical access can be assessed, but accessing meaning of isolated words is but a small part of the processes involved in the comprehension of paragraphs. In prose comprehension, words must be amalgamated for sentence meaning, and sentences must be integrated for textual meaning.

In summary, the decoding hypothesis of oral reading holds in certain contexts with particular materials and for certain types of readers, and the comprehension hypothesis holds in others. All three factors interact to determine the specific level of processing of the text. Whether oral production is initiated prior to or after comprehension is determined by these factors.

POSTSCRIPT

In his comments on our chapter, Trabasso discusses the lack of clear definitions of decoding and comprehension. Decoding typically refers to the translation of print input into an appropriate phonological code. Comprehension refers to the process of extracting meaning from the phonological code. Neither of these definitions is precise enough to know what operations to use to investigate each. Trabasso correctly asserts that procedural definitions of decoding and comprehension are needed.

The basic reationale for our discussion of the research in oral reading is that oral reading reflects processing at a variety of levels. Oral reading errors were classified as being related to the graphic properties of the stimulus or as related to syntax or semantics. Several levels of processing in reading were identified in terms of alterations in text, the effects of which were measured by disfluencies in oral reading. This processing hierarchy effectively has eliminated the need for a distinction between decoding and comprehension. Which processing levels are involved in decoding, and which are involved in comprehension? The distinction is no longer formally necessary, because the levels of processing in reading have been defined by the particular operations used to disrupt each processing level.

The original question as to whether oral production is initiated after decoding or after comprehension has been divided. One question is whether or not a particular level of processing (as defined by a manipulation of the text) is involved in oral reading. Any disruption in oral reading provides evidence of processing at that level. The second question is whether oral production is initiated before or after processing at a particular level has occurred. This question is answered by the point of disfluency in oral reading relative to the point of alteration in the text. Disfluencies prior to when the altered text is uttered indicate that processing at that level has been completed; disfluencies after the altered text has been uttered indicate that oral production was initiated prior to processing at that level. Using this rationale, one's model of reading is implicit in the selection of what processing levels are interesting to manipulate.

ACKNOWLEDGMENT

The research was supported in part by Grant No. MH 21230-02 from the National Institute of Mental Health.

REFERENCES

Biemiller, A. The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, 1970, *6*, 75–96.

Burke, C. Dialect and the reading process. In J. L. Laffey & R. Shuy (Eds.), *Language differences: do they interfere?* Newark, Del.: International Reading Association, 1973.

Buswell, G. T. An experimental study of the eye-voice span in reading. *Supplementary Educational Monographs*, 1920, (Whole No. 17).

Chi, M. T. H. Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition*, 1976, *4*, 559–572.

Clay, M. M. A syntactic analysis of reading errors. *Journal of Verbal Learning and Verbal Behavior*, 1968, *7*, 434–438.

Clay, M. M. Reading errors and self-correction behaviour. *British Journal of Educational Psychology*, 1969, *39*, 47–56.

Clay, M. M., & Imlach, R. H. Juncture, pitch, and stress as reading behavior variables. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 133–139.

Cohen, A. S. Oral reading errors of first grade children taught by a code emphasis approach. *Reading Research Quarterly*, 1974–1975, *10*, 616–650.

Danks, J. H. Grammaticalness and meaningfulness in the comprehension of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 687–696.

Dooling, D. J. Some context effects in the speeded comprehension of sentences. *Journal of Experimental Psychology*, 1972, *93*, 56–62.

Dooling, D. J., & Danks, J. H. Going beyond tests of significance: Is psychology ready? *Bulletin of the Psychonomic Society*, 1975, *5*, 15–17.

Fairbanks, G. The relation between eye-movements and voice in the oral reading of good and poor silent readers. *Psychological Monographs*, 1937, *48*(3, Whole No. 215), 78–107.

Fasold, R. W. Orthography in reading materials for Black English speaking children. In J. C. Baratz & R. W. Shuy (Eds.), *Teaching Black children to read*. Washington, D.C.: Center for Applied Linguistics, 1969.

Gibson, E. J., & Levin, H. *The psychology of reading*. Cambridge, Mass: MIT Press, 1975.

Golinkoff, R. M., & Rosinski, R. R. Decoding, semantic processing, and reading comprehension skill. *Child Development*, 1976, *47*, 252–258.

Goodman, K. S. Dialect barriers to reading comprehension. *Elementary English*, 1965, *42*, 853–860. (a)

Goodman, K. S. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, *42*, 639–643. (b)

Goodman, K. S. Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 1967, *6*, 126–135.

Goodman, K. S. The 13th easy way to make learning to read difficult: A reaction to Gleitman and Rozin. *Reading Research Quarterly*, 1973, *8*, 484–493.

Goodman, K. S., & Buck, C. Dialect barriers to reading comprehension revisted. *Reading Teacher*, 1973, *27*, 6–12.

Hunt, B. C. Black dialect and third and fourth graders' performance on the Gray Oral Reading Test. *Reading Research Quarterly*, 1974–1975, *10*, 103–123.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, *6*, 293–323.

Labov, W. Some sources of reading problems for Negro speakers of nonstandard English. In A. Frazier (Ed.), *New directions in elementary English*. Urbana, Ill.: National Council of Teachers of English, 1967.

Lazerson, B. H. The influence of highly variable spelling upon the reading performance of skilled readers of Modern English. *Reading Research Quarterly*, 1974–1975, *10*, 583–615.

Levin, H., & Cohn, J. A. Studies of oral reading: XII. Effects of instructions on the eye-voice span. In H. Levin, E. J. Gibson, & J. J. Gibson (Eds.), *The analysis of reading skill*. Washington, D.C.: U.S. Government Printing Office, 1968.

Levin, H., & Kaplan, E. L. Eye-voice span (EVS) within active and passive sentences. *Language and Speech*, 1968, *11*, 251–258.

Levin, H., & Kaplan, E. L. Grammatical structure and reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.

Levin, H., & Turner, A. Sentence structure and the eye-voice span. In H. Levin, E. J. Gibson, & J. J. Gibson (Eds.), *The analysis of reading skill*. Washington, D.C.: U.S. Government Printing Office, 1968.

Mehegan, C. C., & Dreifuss, F. E. Hyperlexia: Exceptional reading ability in brain-damaged children. *Neurology*, 1972, *22*, 1105–1111.

Melmed, P. J. Black English phonology: The question of reading interference. In J. L. Laffey & R. W. Shuy (Eds.), *Language differences: Do they interfere?* Newark, Del.: International Reading Association, 1973.
Menyuk, P. *Sentences children use.* Cambridge, Mass.: MIT Press, 1969.
Miller, J. W. Disruptive effect: A phenomenon in oral reading. *Reading Horizons,* 1975, *15,* 198-207.
Miller, J. W., & Isakson, R. L. *The effect of syntactic and semantic violation on high and low reading comprehenders.* San Francisco: American Educational Research Association, 1976.
Pehrsson, R. S. V. The effects of teacher interference during the process of reading or how much of a helper is Mr. Gelper? *Journal of Reading,* 1974, *17,* 617-621.
Perfetti, C. A., & Hogaboam, T. Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology,* 1975, *67,* 461-469.
Resnick, L. B. Relations between perceptual and syntactic control in oral reading. *Journal of Educational Psychology,* 1970, *61,* 382-385.
Rode, S. S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly,* 1974-1975, *10,* 124-142.
Rosen, C. L., & Ames, W. S. An exploration of the influence of dialect on the oral reading of sentences by Black children, grades three through six. *Reading World,* 1972, *11,* 201-209. (a)
Rosen, C. L., & Ames, W. S. An investigation of the influence of nonstandard dialect on the oral reading behavior of fourth grade Black children under two stimuli conditions. In J. A. Figural (Ed.), *Better reading in urban schools.* Newark, Del.: International Reading Association, 1972. (b)
Schlesinger, I. M. *Sentence structure and the reading process.* The Hague, Netherlands: Mouton, 1968.
Shankweiler, D., & Liberman, I. Misreading: A search for causes. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Mass.: MIT Press, 1972.
Siler, E. R. The effects of syntactic and semantic constraints on the oral reading performance of second and fourth graders. *Reading Research Quarterly,* 1973-1974, *9,* 583-602.
Silverberg, N. E., & Silverberg, M. C. Hyperlexia—specific word recognition skills in young children. *Exceptional Children,* 1967, *34,* 41-42.
Silverberg, N. E., & Silverberg, M. C. Case histories in hyperlexia. *Journal of School Psychology,* 1968-1969, *7,* 3-7.
Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in school* (2nd ed). New York: Holt, Rinehart & Winston, 1976.
Stevens, A. L., & Rumelhart, D. E. Errors in reading: An analysis using an augmented transition network model of grammar. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition.* San Francisco: Freeman, 1975.
Weber, R. M. The study of oral reading errors: A survey of the literature. *Reading Research Quarterly,* 1968, *4,* 96-119.
Weber, R. M. First-graders' use of grammatical context in reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading.* New York: Basic Books, 1970. (a)
Weber, R. M. A linguistic analysis of first-grade reading errors. *Reading Research Quarterly,* 1970, *5,* 427-451. (b)
Weber, R. M. Dialect differences in oral reading: An analysis of errors. In J. L. Laffey & R. Shuy (Eds.), *language differences: Do they interfere?* Newark, Del.: International Reading Association, 1973.

Polish Psychological Bulletin
1978, Vol. 9(4), 183—192

JOSEPH H. DANKS
Institute of Psychology
University of Warsaw

MODELS OF LANGUAGE COMPREHENSION

Models of language comprehension are described in terms of how processing components are organized. Four groups of processing components — speech perception (orthographic decoding of print), lexical access of word meaning, integration of syntactic-semantic information in phrases and clauses, and integration of sentences into larger units — are based on different types of information in the linguistic structure of the text. The organization of these components is described by the properties of temporal sequence and direction of processing, interaction among processing components, and flexibility in adapting to task demands. Two classes of models incorporating opposite extremes of these properties are the interpretive and the constructive models. Research results tend to support a constructive model whose processing components interact flexibly and in parallel.

The basic notion of comprehension is that linguistic information is converted by a listener or reader into a meaning that represents what he understands an utterance or passage to mean. What are the cognitive processes by which the listener or reader affects this conversion? This paper discusses these processes and how they might be integrated into a coherent system for linguistic comprehension.[1]

There are two general orientations for characterizing the comprehension process. In an *interpretive* approach, the meaning resides in the speech signal (or printed word), having been placed there by a speaker (or writer), and is represented as linguistic regularities, e.g., phonological, syntactic, and semantic structure. This information is extracted by a listener (or reader) as the meaning of the utterance. The listener is an active searcher for the meaning that has been encoded by the speaker.

In contrast, in a *constructive* approach, the listener (or reader) is attempting to make sense of his experiences. Speech is one source of data on which he can base his understanding. As he comprehends an utterance, the listener constructs a representation of the utterance by integrating it with preceding utterances, social and physical contexts, and his general knowledge about the world. All of these information sources are used to construct a schema representing his experiences, in this case, his linguistic inputs.

In both approaches, comprehension processes are based in part on linguistic information that can be derived from the text. Components of the comprehen-

This paper was written while the author was a visiting professor at the University of Warsaw under an exchange program between that university and Kent State University. Requests for reprints should be sent to the author at the Department of Psychology, Kent State University, Kent, Ohio 44242, U.S.A.

[1] It is assumed that listening and reading comprehension processes are the same except for the obvious differences of input channel. Although this assumption can be questioned, the discussion of comprehension models here is not vitiated by ignoring possible differences.

sion process differ in their use of these information sources. The basic differences between the interpretive and constructive approaches lie in how these component processes are organized. After describing the processing components, the interpretive and constructive models are examined in terms of four properties that characterize the organization of components.

PROCESSING COMPONENTS

Within a segment of discourse or text, levels of linguistic structure can be identified and ordered in a hierarchy based on the size and abstractness of the units. At the lowest level, phonological structure, acoustic, auditory, and articulatory properties of the speech signal, and visual properties of the letters and words can be specified. At a higher level, words can be identified in terms of their phonological, orthographic, syntactic, and semantic properties. At a still higher level, sentences are analyzed into phrases and clauses, each with an internal structure that specifies the structural relations among the words of the sentence. Finally, the interrelations among the propositions that underlie the constituent sentences represent variously the paragraph structure of printed text, the narrative structure of stories, or the social turn-taking structure of conversation.

Models of comprehension generally assume that components of the comprehension process are related to these levels of structure. There is a general correspondence between the linguistic structure at any given level and the psychological processing that uses that information. Corresponding to the first level is *perceptual processing*, either auditory or visual, in which word units are identified from the speech signal and from print. This level includes all subprocesses from distinctive feature analysis through phoneme and letter identification. The second level is *lexical access*. The identified word is accessed in the mental dictionary so that the correspondence between its phonological or orthographic form and its syntactic and semantic representations is established. The third level is *sentence integration* in which the syntactic properties of words are used to integrate their semantic representations into the meanings of phrases, clauses, and sentences. Subprocesses here are first segmentation of the phrases, then identification of their structure, and finally integration into a new structure. The fourth processing level is *textual integration* of clausal and sentential propositions into larger mental structures or schemata. Text includes both conversation and discourse (speech) and prose (written). There is no theoretical limit to the scope of these schemata, but there is a practical limit defined primarily by memory limitations.

The existence of the various processing components has been established by many investigators: perceptual (Miller & Isard, 1963), lexical access (Fears & Danks, 1977; Marslen-Wilson, 1975; Marslen-Wilson & Tyler, 1975), sentence integration (Danks, 1969; Fears & Danks, 1977; Marks & Miller, 1964; Miller, 1975; Mistler-Lachman, 1972, 1974, 1975; Siler, 1973–74; Tyler & Marslen-Wilson, 1975), and textual integration (Rosenberg & Lambert, 1974).

How are the processing components organized into a comprehension process? Generalizations about organization are difficult because of conflicting conclusions from experimental studies. Part of the conflicting results may be due to differences in experimental task. Different tasks engage different processing components in varying degrees (Perfetti, 1976). The question of organization of processing components is the question of what model best describes the comprehension process. Interpretive and constructive models are described in terms of properties that reflect how the processing components are organized.

PROPERTIES OF MODELS

Models of comprehension differ along four dimensions: directionality of processing, temporal organization of processing components, independence among processing components, and flexibility of organization. These properties are not completely independent, especially as they are organized into a model. Decisions on each dimension tend to be correlated.

Directionality of processing

Most models of language comprehension have assumed a *bottom-up* (from surface features to abstract representation) direction of processing rather than a *top-down* direction. In bottom-up models the speech signal is processed in terms of its acoustic properties, perhaps by relating the extracted features to articulatory mechanisms. Lexical identification follows next, and then syntactic and semantic interpretation and integration. The process ends with integration of the sentence meaning with prior text and the listener's knowledge of the world. For reading the model is little different. It begins with decoding of print into an auditory code. Then the comprehension process proceeds via the same system that comprehends speech. A bottom-up model seems natural because it corresponds to a peripheral-to-central direction of neural processing.

This model has been challenged recently, however, by theorists who hold that comprehension processes proceed in a top-down direction (Goodman, 1967). To reverse the direction of processing to top-down, the listener generates predictions about the input from his knowledge and expectations about topic and speaker. These predictions are tested against the input and, if correct, represent the meaning of the utterance. Thus, information flows from abstract meaning representations to specific surface strings. At the level of speech perception, this direction of processing is represented by analysis-by-synthesis models (Halle & Stevens, 1964; Stevens, 1972).

Temporal organization

Temporal organization of processing components refers to whether the components are ordered serially or whether they operate in parallel. If the processing components operate in sequence, either bottom-up or top-down, then the temporal organization is serial. Serial processing has been the typical assumption for most models of comprehension. The alternative to serial organization is parallel processing. All processing components, or at least several of them, operate at the same time. For example, a word may be processed simultaneously in terms of its phonetic, lexical, syntactic, and semantic properties.

Dependency among components

Dependency relations among processing components refer to whether or not each component proceeds to analyze the input without being influenced by any other component, or whether components exert influence on other components, say by providing critical information. With *parallel* models it is possible to assume complete independence among processing stages. A final decision process must integrate the information derived from each component, but the component processes themselves may remain independent. With *serial* models there must be a limited form of dependency. At minimum, each stage must receive input from the immediately preceding component.

Both serial and parallel models can have extensive dependency relations among the components, however. Parallel models may be wholly interactive if information from each component is immediately available to all other components. Serial models can be interactive in that a later processing component can reject the analysis of an earlier stage and reinitiate processing in the

preceding component. This is the case in 'garden-path' sentences, such as "A carpenter took a plane to London, but forgot the rest of his tools", in which words and phrases in the first part of the sentence must be reanalyzed on encountering an inconsistensy in the last half. A more common type of dependency among stages in a serial model is the addition of feedback loops in which more abstract processing of a word or phrase influences the processing of succeeding words or phrases. In an independent system, one would wait until all words and phrases had been analyzed to integrate their meanings into a single sentence meaning. But with the addition of feedback loops, processing of later words in a sentence can be facilitated by the completed processing of earlier words. If, however, the number of feedback loops added to a serial model results in every component being connected with every other component, the serial model is functionally equivalent to a parallel model.

Flexibility of organization

Can the sequence of processing components be reorganized, or different components highlighted, or strategies invoked in order to respond efficiently and effectively to a specific task? Most comprehension models have been constructed such that there is minimal flexibility in the ordering and structure of the processing components. Usually, there is one fixed order (for serial models) or a single set of processes that invariably apply (for parallel models). With the realization that listeners can extract and encode information at various levels of abstractness, some theorists have included multiple exit points in their serial models (Mistler-Lachman, 1972, 1975). If a task demand can be satisfied by specific information, the listener can halt the comprehension process as soon as sufficient information has been extracted (Perfetti, 1976). But this flexibility is relatively limited because the comprehension process itself does not change.

INTERPRETIVE MODELS

An extreme version of the interpretive model is bottom-up, serial, noninteractive, and inflexible. The stages of processing are ordered strictly with the first stage accepting external input, performing its analysis, and passing along the results to the second stage. This sequence of analyzing information continues until the meaning has been extracted from the input. The model is 'input-driven' in that the reception of an external input provides the impetus for processing. The stages are ordered serially in a sequence that is the same for each input. The only connection between the stages is that the results of each stage serve as input for the next. The model is inflexible since the sequence of stages is predetermined and is not contingent on the input, on the context or purpose, or on the results of the ongoing process. Whatever apparent flexibility a listener or reader has is gained by supplementary processing of the sentence meaning after the comprehension process per se is completed. An example of an extreme interpretive model of reading comprehension is offered by Gough (1972).

There are many experiments, however, that demonstrate top-down effects and that require interaction among components of the comprehension process. Beginning with Pillsbury (1897) through more recent studies of the word superiority effect (Reicher, 1969; Wheeler, 1970), investigators have found the perception of a single letter to be facilitated by being embedded in a word. At a higher level, perception of words depends in part on the syntactic and semantic context in which they appear (Miller & Isard, 1963; Stevens & Rumelhart, 1975; Weber, 1970). Finally, the interpretation assigned to larger units of text — phrases, clauses, sentences, and paragraphs — depends on the context, situation, setting, or theme

(Dooling & Lachman, 1971; see also Bransford & Johnson, 1973; Bransford & McCarrell, 1974).

These context effects frequently are handled in interpretive models by the use of feedback loops. As processing proceeds, the results from more abstract processing levels are returned to earlier stages where the later information can be used to facilitate lower-level processing of later inputs. With the addition of these feedback loops, the interpretive model becomes interactive. A negative consequence is that one needs to connect nearly every stage with every other stage in order to account for the experimental results, so that the model is functionally equivalent to parallel-interactive models.

Other experiments have investigated the nature of top-down and bottom-up effects using the introduction of linguistic violations to disrupt processing. The rationale is to violate the linguistic structure at one or more levels and then to measure differences in how subjects process the altered texts. The basic assumption is that when the linguistic structure is violated at a given level, then cognitive processes that depend on that level of linguistic information will be disrupted, which in turn will lead to differences in performance. Top-down results are an impairment in lower-level decision tasks with violations at higher linguistic levels. Bottom-up results are those that find increased difficulty of comprehension and memory with violations at lower linguistic levels.

The results of these experiments have been mixed. Marslen-Wilson and Tyler (1975) found an effect of violations of sentence structure and meaning on monitoring for individual words (lexical access). Marslen-Wilson (1975) demonstrated that sentence context facilitated restoration of altered syllables (lexical access) in a shadowing task. Also using a shadowing task, Rosenberg and Lambert (1974) found that violating textual structure by deleting sentences, inserting extraneous material, and reordering sentences disrupted the lower-level task of shadowing. In contrast, Mistler-Lachman (1975) required readers to decide whether sentences with different levels of syntactic violations were meaningful or not, and found that the decision times were ordered by level of syntactic violation. Tyler and Marslen-Wilson (1975), using an interrupted sentence recall task (Jarvella, 1970, 1971), obtained effects of intra-sentence syntactic and semantic violations across clausal boundaries.

Fears and Danks (1977) investigated three processing components with the reading counterpart of a shadowing task, oral reading. The level of lexical access was tapped by inserting a pronounceable nonword. The level of syntactic integration was assessed by inserting a word that was both syntactically and semantically inappropriate for the context. The level of semantic integration was manipulated by inserting a word that was semantically anomalous but which was the correct part of speech. Oral production times of the words surrounding the altered words indicated that the peaks of disruption were ordered from nonword to syntactic to semantic violation, supporting a bottom-up order of processing. However, there was substantial overlap in the patterns of disruption, indicating parallel, interactive processing.

The resolution of these varying effects is not possible within a rigid interpretive model, but is possible with a constructive approach as discussed in the next section.

A second aspect of the pure interpretive model that conflicts with experimental data is its lack of flexibility. Finding differences in performance in different experimental situations, some investigators interpreted their findings in terms of depth-of-processing (Craik & Lockhart, 1972). For example, Mistler-Lachman (1972, 1974) used three tasks that were presumed to induce different levels of processing — judging whether a sentence was meaningful, judging whether a sentence followed from a preceding sentence, and pro-

ducing a following sentence. She found that these three tasks did induce different levels of comprehension (1972) and memory (1974).

The model implicitly assumed by most studies using a depth-of-processing rationale is an interpretive model with the addition of multiple exit points following various stages. When sufficient information has been extracted from the sentence input to satisfy the experimental task, the comprehension process terminates without extracting the most abstract level of meaning. The experimental tasks are assumed to be ordered so that successive exit points are activated by each task.

The primary problem with adding multiple exit points to handle task demands is that they permit only 'quantitative' differences in a fixed sequence of processing stages. They do not permit 'qualitative' differences resulting from the strategic reorganization of processing components. Given the wide variety of everyday tasks which listeners and readers face, a more flexible processing system is needed. Do we read a play in the same way that we read a novel? And do we read either in the same way that we read a poem? It is not just a matter of reading each genre to varying depths, but we attack the task of comprehending differently (cf. Gibson & Levin, 1975, pp. 454–465). Is the difference between listening to two conversations simultaneously at a party and listening to a lecture simply a matter of depth of processing? Apparently it is not. We seem to be capable of restructuring the set of processing components by which we comprehend speech and print.

There are a number of experimental studies that demonstrate that the kind of processing and the resulting meaning representation vary qualitatively with task demands. One example of these task related effects is in Danks (1969). Intra-sentence syntactic and semantic violations had differential effects depending on whether the readers were to comprehend the sentence, correct its grammar, or correct its meaning. For another example, Garrod and Trabasso (1973) and Glucksberg, Trabasso, and Wald (1973) concluded that it was impossible to construct a single verification model that would account for all the picture-sentence verification data for the ubiquitous active-passive comparison. Subjects varied their encodings as well as their comparison strategies as a function of the particular verification task they were engaged in. Kolers (1974, 1975) obtained extended memory for the perceptual operation subjects used in reading text that had been transformed in several spatial dimensions. Aaronson and Scarborough (1976) reported that if subjects read for later recall, the reading-time patterns followed variation in syntactic structure, but if the subjects read for comprehension, the reading-time patterns reflected semantic information. Finally, Frederiksen (1975a, 1975b) has found that subjects reading material only for memory as opposed to subjects studying a text to solve a problem for which the text provided information, did not differ in the total amount recalled, but did differ in the types of information and in the amount of derived and inferred information remembered. How the subject processes the text is influenced by task demands. It is not simply a difference in the amount of processing, nor is it just a difference in what is included in the representation. The comprehension process itself is modified to accomodate the specific demands of the immediate task.

Flexibility of processing also is demonstrated by the findings that sentence meanings are appropriate for the particular situation. Unambiguous words are represented differently depending on the sentence context (Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974). For example, *piano* is represented differently in "The man tuned (or lifted) the piano". Glucksberg, Hartman, and Stack (1977) put this idea to a more severe test. They found that the possibility of interpreting a literally false sentence,

such as "Some roads are snakes", as a plausible metaphor hindered subjects' rejection of the sentences as false. Subjects could not avoid constructing the metaphorical meaning simultaneously with comprehending the literal meaning. In fact, they may have even computed the metaphorical meaning first. Thus, a sentence is not given a single interpretation which then is modified to fit the context, but the original comprehension process constructs an appropriate representation.

There is substantial evidence against interpretive models even when they are modified by feedback loops and multiple exit points. Let us turn to a model that would be suggested by the constructive approach.

CONSTRUCTIVE MODELS

Constructive models are parallel, interactive, and flexible. The contrast between bottom-up and top-down is not meaningful in the context of a model that is both parallel and interactive. As input is perceived by the ears (or eyes), it is stored in a work buffer (or message center, Rumelhart, 1976). This work buffer holds the input as well as the results from the components as they process the input. Memory structures interconnect the information that the individual has available at any given point in time. Episodic information about past experiences, semantic information, encyclopedic knowledge about the world, as well as the current contents of consciousness stored in the work buffer are all part of the memory structures.

The comprehension processes, encompassing phonetic (or graphic) identification, lexical access, syntactic and semantic integration, and textual integration, are not ordered temporally or structurally in terms of importance. All components have direct access to the information contained in the memory structures. The processing components are constantly alert for information in the work buffer on which to operate. Whenever such information is available, one or more of the components acts on that information and returns the results to the work buffer. The information returned by a component is a transformation of a data structure into a different form, either recording the information in a different format or integrating two or more information units into a new one. Thus, the memory structures are changing constantly. The altered memory structures are susceptible to action by other components.

There is no meaning output from the completed comprehension process other than the continual updating of memory structures. The comprehension process is completely integrated with memory and thinking. One does not comprehend an input and only then do something with it, such as store it in memory or use it to solve a problem. Rather the information in the work buffer is continually available for these other cognitive processes. The only 'meaning' that is the output from comprehension is a new memory structure.

How is flexibility gained in a constructive model? When there is sufficient information in the work buffer to satisfy the demands of the situation, a response is produced. Without sufficient information, processing continues until such is secured, or until a time limit is crossed. In either case, one responds on the basis of whatever information is available. If a decision is required on the meaning of a word, or on the meaningfulness of a sentence, or on whether a sentence integrates with the text, a response is produced whenever that information has been produced by the relevant processing component. Differential responses as a function of task result from differential amounts of processing necessary to produce a response. Certain types of processing take longer because the transformation, recoding, and integration of memory structures is more complex. A second way that flexibility may be gained is by differen-

tial attention to one or more of the components. For example, if the task is to indicate for each sentence whether it is meaningful or not, then more attention may be devoted to the component concerned with intra-sentence semantic integration, and less to other components.

It is possible to characterize constructive models as a mix of bottom-up and top-down directions of processing. A top-down, or predictive, direction results from integrating the input with the current contents of memory. The memory structures define the range of possible interpretations that may be assigned to a new input. In everyday situations, one attempts to fit new information into existing memory structures, so that new inputs are comprehended as congruent with them. It is possible to break that connection, particularly in laboratory experiments. To the extent that the input is either inconsistent with existing information, predictive constraints perforce cannot operate, and bottom-up, or interpretive, processing dominates. The degree to which the comprehender has prior knowledge that is related to the information in the input determines the extent to which top-down forces can operate.

A second factor that affects the mix of interpretive and predictive orientations is the motivation for the subject to integrate the input with his memory structures. Given insufficient motivation, the subject may keep the input isolated (Spiro, 1977). For complete comprehension, the subject must be motivated by the task and the instructions to integrate the input with his memory structures. Otherwise, there is incomplete comprehension that deviates from that typically used.

A third factor affecting the interpretive-predictive mix is the availability of and fluency with the components. A child learning to read does not have complete control over some of the components, such as visual processing of graphic information. An adult learning a second language has to learn new processing strategies to meet the demands of the new language. A child during first language acquisition can be characterized in similar terms. As the individual uses and practices the processing components, they become more and more automatic (LaBerge & Samuels, 1974). Until the organization of processing components becomes automatic, attentional demands may force a serial, bottom-up organization.

The experimental results described earlier are more consistent with a constructive model. Firstly, both bottom-up and top-down effects are predicted as a function of the specific experimental task because all processing components are operating in parallel on the information in the work buffer. Sometimes low-level analysis is required to produce data structures for more abstract components. Other times the existing conceptual structures are sufficiently congruent with the new input so as to specify its meaning without low-level analysis, for example, when skimming an article on a familiar topic. Secondly, a constructive model is able to adapt flexibly to the specific situation. Responses are based on the current contents of the work buffer. The decision-making process monitors the work buffer for the information needed. The criterion set for a response, then, can be adapted to the task. Additionally, those processing components whose output is necessary for a response can be accentuated. The variety of experimental results in the research literature represents this flexibility. It is possible to specify models for specific experimental tasks, for example, for picture-sentence verification tasks (Carpenter & Just, 1975), but not for comprehension generally. There is no single comprehension process. Rather there are many processing components that are adapted strategically for particular comprehension situations.

REFERENCES

(The *Journal of Verbal Learning and Verbal Behavior* is abbreviated *JVLVB*.)

Aaronson D. & Scarborough H. S.,

Performance theories for sentence coding: Some quantitative evidence, *J. exp. Psychol.: Human Perc. and Perf.*, 1976, **2**, 56–70.

Anderson R. C. & Ortony A., On putting apples into bottles: A problem of polysemy, *Cogn. Psychol.*, 1975, **7**, 167–180.

Barclay J. R., Bransford J. D., Franks J. J., McCarrell N. S. & Nitsch K. E., Comprehension and semantic flexibility, *JVLVB*, 1974, **13**, 471–481.

Bransford J. D. & Johnson M. K., Considerations of some problems of comprehension, in: W. G. Chase (ed.), *Visual information processing*, New York, Academic Press, 1973.

Bransford J. D. & McCarrell N. S., A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend, in: W. B. Weimer & D. S. Palermo (eds.), *Cognition and the symbolic processes*, Hillsdale, N. J., Lawrence Erlbaum, 1974.

Carpenter O. & Just M., Sentence comprehension: A psycholinguistic model of verification, *Psychol. Rev.*, 1975, **82**, 45–73.

Craik F. I. M. & Lockhart R. S., Levels of processing: A framework for memory research, *JVLVB*, 1972, **11**, 671–684.

Danks J. H., Grammaticalness and meaningfulness in the comprehension of sentences, *JVLVB*, 1969, **8**, 687–696.

Dooling D. J. & Lachman R., Effects of comprehension on retention of prose, *J. exp. Psychol.*, 1971, **88**, 216–222.

Fears R. & Danks J. H., *Levels of processing in oral reading*, Midwestern Psychol. Assoc., Chicago, 1977.

Frederiksen C. H., Acquisition of semantic information from discourse: Effects of repeated exposures, *JVLVB*, 1975a, **14**, 158–169.

Frederiksen C. H., Effects of context-induced processing operations on semantic information acquired from discourse, *Cogn. Psychol.*, 1975b, **7**, 139–166.

Garrod S. & Trabasso T., A dual-memory information processing interpretation of sentence comprehension, *JVLVB*, 1973, **12**, 155–167.

Gibson E. J. & Levin H., *The psychology of reading*, Cambridge, Mass., MIT Press, 1975.

Glucksberg S., Hartman D. E. & Stack R., *Metaphor comprehension is an automatic and parallel process*, Psychonomic Soc., Washington, 1977.

Glucksberg S., Trabasso T. & Wald J., Linguistic structures and mental operations, *Cogn. Psychol.*, 1973, **5**, 338–370.

Goodman K. S., Reading: A psycholinguistic guessing game, *J. Reading Specialist*, 1967, **6**, 126–135.

Gough P. B., One second of reading, in: J. F. Kavanagh & I. G. Mattingly (eds.), *Language by ear and by eye*, Cambridge, Mass., MIT Press, 1972.

Halle M. & Stevens K. N., Speech recognition: A model and a program for research, in: A. Fodor & J. J. Katz (eds.), *The structure of language*, Englewood Cliffs, N. J., Prentice-Hall 1964.

Jarvella R. J., Effects of syntax on running memory span for connected discourse, *Psychon. Sci.*, 1970, **19**, 235–236.

Jarvella R. J., Syntactic processing of connected speech, *JVLVB*, 1971, **10**, 409–416.

Kolers P. A., Two kinds of recognition, *Canad. J. Psychol.*, 1974, **28**, 51–61.

Kolers P. A., Specificity of operations in sentence recognition, *Cogn. Psychol.*, 1975, **7**, 289–306.

LaBerge D. & Samuels S. J., Toward a theory of automatic information processing in reading, *Cogn. Psychol.*, 1974, **6**, 293–323.

Marks L. E. & Miller G. A., The role of semantic and syntactic constraints in the memorization of English sentences, *JVLVB*, 1964, **3**, 1–5.

Marslen-Wilson W. D., Sentence perception as an interactive parallel process, *Science*, 1975, **189**, 226–228.

Marslen-Wilson W. D. & Tyler L. K., Processing structure of sentence perception, *Nature*, 1975, **257**, 784–786.

Miller G. A. & Isard S., Some perceptual consequences of linguistic rules, *JVLVB*, 1963, **2**, 217–228.

Miller J. W., Disruptive effect: A phenomenon in oral reading, *Reading Horizons*, 1975, **15**, 198–207.

Mistler-Lachman J. L., Levels of comprehension in processing of normal and ambiguous sentences, *JVLVB*, 1972, **11**, 615–623.

Mistler-Lachman J. L., Depth of comprehension and sentence memory, *JVLVB*, 1974, **13**, 98–106.

Mistler-Lachman J. L., Queer sentences, ambiguity, and levels of processing, *Mem. & Cogn.*, 1975, **3**, 395–400.

Perfetti C. A., Niveaux de compréhension des phrases, in: S. Ehrlich & E. Tulving (eds.), *La mémoire sémantique*, *Bull. de Psychologie*, Numéro spécial annuel 1976, 346–355.

Pillsbury W. B., A study in apperception, *Am. J. Psychol.*, 1897, **8**, 315–393.

Reicher G. M., Perceptual recognition as a function of meaningfulness and stimulus material, *J. exp. Psychol.*, 1969, **81**, 275–310.

Rosenberg S. & Lambert W. E., Contextual constraints and the perception of speech, *J. exp. Psychol.*, 1974, 102, 178-180.

Rumelhart D., Toward an interactive model of reading, *Tech. Rep.* 56, CHIP, Univ. of Calif., San Diego, 1976.

Siler E. R., The effects of syntactic and semantic constraints on the oral reading performance of second and fourth graders, *Reading Res. Quart.*, 1973-74, 9, 583-602.

Spiro R. J., Remembering information from text: The 'state of schema' approach, in: R. C. Anderson, R. J. Spiro & W. E. Montague (eds.), *Schooling and the acquisition of knowledge*, Hillsdale, N. J., Lawrence Erlbaum, 1977.

Stevens A. L. & Rumelhart D. E., Errors in reading: An analysis using an augmented transition network model of grammar, in: D. A. Norman & D. E. Rumelhart (eds.), *Explorations in cognition*, San Francisco, Freeman, 1975.

Stevens K. N., Segments, features, and analysis by synthesis, in: J. F. Kavanagh & I. G. Mattingly (eds.), *Language by ear and by eye*, Cambridge, Mass., MIT Press, 1972.

Tyler L. K. & Marslen-Wilson S., *Memory and levels of processing in psycholinguistic context*, Midwestern Psychol. Assoc., Chicago, 1975.

Weber R. M., First-graders' use of grammatical context in reading, in: H. Levin & J. P. Williams (eds.), *Basic studies on reading*, New York, Basic Books, 1970.

Wheeler D., Processing in word recognition, *Cogn. Psychol.*, 1970, 1, 59-85.

(*Received May 1978*)

Ann. Rev. Psychol. 1980. 31:391–417

EXPERIMENTAL PSYCHOLINGUISTICS

◆33

Joseph H. Danks[1]

Department of Psychology, Kent State University, Kent, Ohio 44242

Sam Glucksberg[1]

Department of Psychology, Princeton University, Princeton, New Jersey 08540

CONTENTS

[1]Order of authorship was decided by a flip of a coin. We thank Nancy Cantor, Robert Chabot, Thomas K. Landauer, George A. Miller, Linda Ohlweiler, Robert Schwalm, and Lee Sengbush for helpful comments on various drafts. Our work was supported in part by Grant No. NIE-G-78-0223 from the National Institute of Education to Kent State University, and by Public Health Service Grant No. MH23401 to Princeton University.

0066-4308/80/0201-0391$01.00

INTRODUCTION

The main trends in current psycholinguistics began when George Miller (1962) introduced Noam Chomsky's work to psychologists. One immediate effect was a shift to mentalistic concepts and theorizing. That cognitive approach still prevails, despite occasional calls for a neobehavioristic revival (Staats 1976). A second profound effect was to shift attention from words, and the associations among them, to syntax as the central object of inquiry and a source for organizing principles.

The transition from a word-based associative psycholinguistics to a sentence-based cognitive enterprise was reflected in the contrast between Rubenstein & Aborn's (1960) review and Johnson-Laird's (1974). The central theoretical concept in 1960 was the response hierarchy, and the bulk of the data base concerned the statistical and probabilistic properties of selected language corpora. Despite such a simplified psycholinguistics, or perhaps because of it, optimism ran high and mechanical speech recognition was thought to be just around the corner. By 1974 the complexities of language comprehension had been recognized, and attention was focused on the sentence as the basic unit of analysis. The notion of sentence-as-unit was so firm that Johnson-Laird could assert confidently that "the fundamental problem in psycholinguistics is simple to formulate: what happens when we understand sentences?" (1974, p. 135).

The trend in recent years has been away from the sentence in isolation. There are at least two reasons for this shift. First, just as word comprehension cannot be fully understood outside of sentences, sentence comprehension cannot be fully understood outside of larger discourse contexts. In order to build an adequate model of comprehension we need to incorporate processes beyond the level of the sentence. Second, mechanisms of word and sentence perception are not fully adequate for understanding how larger units, such as conversations and stories, are processed. For a number of reasons, many investigators have claimed that we cannot hope to understand language comprehension without incorporating problem-solving strategies, world knowledge, and rules of social discourse. Syntax and semantics alone are not sufficient (Searle 1975, Shaw & Bransford 1977). One result of this trend has been far less reliance on theoretical linguistics for insights and working hypotheses (cf Halle et al 1978) and more attention to cognitive and social psychology, sociolinguistics, and artificial intelligence (see Winograd 1976, Levelt & Flores d'Arcais 1978). We eventually may have, as the title of a recent paper suggests, "psycholinguistics without linguistics" (Johnson-Laird 1977b).

The general commitment to understanding natural language comprehension raises fundamental questions for an experimental psychology of language. What are the useful units of analysis? Can we learn anything important from laboratory experimentation with artifically constrained language materials, or must we learn to cope with the complexity and richness of naturally occurring language? Will tachistoscopes and reaction-time clocks be replaced completely by videotape recorders? The field is divided on this issue. At one extreme are microanalyses of naturally occurring discourse (e.g. Labov & Fanshel 1977). At the other is formal modeling of highly constrained laboratory tasks (e.g. Carpenter & Just 1975). J. R. Anderson (1976) has gone so far as to argue that the domain of inquiry is so complex that the most we can hope for is plausibility in our theories and some measure of practical application. Wexler (1978) provides a critical analysis of this issue.

Units of Analysis and Processing Interactions

The search for *the* basic unit of language has largely been abandoned, and with good reason. Many different segments or components of the speech signal, from distinctive features to clauses, sentences, and paragraphs, can be treated as a unit given an appropriate context and task (McNeill & Lindig 1973). More interesting and productive questions can be asked about the interactions among units, or more precisely, among the various levels and sources of information, than are available in any given language processing task. Such interactions typically are referred to as task or top-down effects (Rumelhart 1977a, Danks 1978). A distinction commonly made in artificial intelligence between top-down and bottom-up information processing systems continues to be applied fruitfully in psycholinguistics (Marslen-Wilson & Welsh 1978). A top-down, or knowledge-driven system uses higher-level knowledge to facilitate processing of incoming data. A bottom-up, or data-driven system relies primarily on the information carried by the input signal itself with little or no constraint from prior knowledge or context. Bottom-up systems can function quite adequately, at least at initial stages of processing when the input signals are clear and unambiguous. When the input is noisy or ambiguous, then information from wider contexts can be used to compensate for poor signal quality. Since normal speech inputs are frequently noisy and ambiguous, interactive bottom-up and top-down processing is likely the norm.

This general distinction appears in various forms in the contemporary literature. Some of these forms, in addition to bottom-up vs top-down (Bobrow & Brown 1975), are data-limited vs resource-limited (Bobrow & Norman 1975), interpretive vs constructive (Frederiksen 1975a), literal meaning vs intended or conveyed meaning (Clark & Lucy 1975), interpretive semantics vs generative semantics (Maclay 1971), context-free vs context-dependent (Bobrow & Norman 1975), and code emphasis vs meaning

emphasis (Cohen 1974–1975). Surely there is something in the air that gave rise to so many versions of the same basic distinction. That something might well be the idea that word, sentence, and discourse processes interact continuously as we attempt to interpret speech, print, and other perceptual events. If so, then a fundamental problem in psycholinguistics is to specify just what kinds of information are used during the comprehension of continuous speech and print and precisely how these different information sources interact. To clarify this issue, we review relevant research at the word, sentence, and discourse levels of processing.

WORDS: REPRESENTATIONS AND PROCESSING MECHANISMS

How words are represented in memory and how these representations function in language comprehension tasks have been the foci of three relatively independent lines of research: word recognition, semantic memory, and psychological semantics.

Word Recognition

If one were to judge from the contents of recent reviews and texts, studies of word recognition processes would seem to be outside the mainstream of experimental psycholinguistics (e.g. Clark & Clark 1977; Levelt & Flores d'Arcais 1978). One reason is the emphasis in most such studies on reading rather than continuous speech recognition (e. g. Coltheart 1978). Nevertheless, word recognition research does address a central issue, namely, the extent and mechanisms of interactions between prior knowledge and the stimulus input.

Two distinct types of models take opposing views on this issue. One type, exemplified by Morton's logogen model, permits incoming sensory information to interact continuously with available contextual information and prior knowledge during word recognition (Morton 1969; Morton & Long 1976). When sufficient information from any one or combination of these sources accumulates, a word is consciously recognized. The relative influence of stimulus and contextual information can vary with circumstances. For example, contextual information has a greater effect when stimulus input is degraded than when it is clear and unambiguous (Massaro et al 1978). In contrast, a sequential search model proposed by Forster (1976) typifies a strictly bottom-up view of word recognition. Words are accessed phonologically (or graphemically) via a search through frequency-sorted classes. Syntactic, semantic, and other contextual information can have no

influence until a word-candidate is found by the search process. Therefore, such information can act only to confirm or disconfirm a word choice.

Both of these types of models treat words as indivisible units, albeit in different ways. Because of this, they each have difficulties in accounting for the ways in which people deal with mispronunciations. When people are asked to detect mispronunciations or to shadow speech, two interesting phenomena appear. First, the greater the distortion from real words, the better people are at detecting and accurately repeating those distortions (Cole 1973). The logogen model has no mechanism which would be sensitive to degrees of distortion because logogens consist only of whole, real words and only an output word is available to consciousness. The second phenomenon is the other side of the coin. In continuous speech shadowing, people sometimes fail to detect mispronunciations. They pronounce words correctly even when the stimuli are distorted, especially when the correction is semantically and syntactically congruent. For example, "compsiny" might be shadowed as "company." These fluent phonemic restorations occur at very short shadowing latencies, often before a whole word has been completely heard (Marslen-Wilson 1975). A sequential search model that treats words as whole units to be accessed before contextual information is consulted is inconsistent with these findings.

Marslen-Wilson & Welsh (1978) propose an interactive model of word recognition during continuous speech which can accommodate the available data. This model posits the simultaneous and continuous use of two sources of information—the incoming stimulus and all other available knowledge, such as syntactic, semantic, and contextual information. As the sequence of phonemes is processed in real time, a pool of word candidates that matches the initial phonemes is activated. Simultaneously, all word candidates which fail to satisfy the available contextual constraints are eliminated. Word candidates continue to be eliminated as additional phonological and contextual information becomes available until a best alternative remains. Thus it is possible for words to be recognized before they have been completely heard. This in turn can produce (*a*) failures to detect mispronunciations and (*b*) fluent restorations of distorted phonemes, especially when these occur relatively late in a word. This kind of model is somewhat different from strictly serial processing models, including analysis-by-synthesis models, that invoke contextual knowledge only after some preliminary analysis of input (e.g. Massaro 1975). Interactive models are of particular interest for psycholinguistics because they can be extended quite naturally to sentence and discourse processing and because they can provide a useful framework for incorporating knowledge from various levels into a unified theory of language comprehension.

Semantic Memory

Although the term semantic memory is quite general, the research on this topic has been rather narrowly conceived. The primary concern has been with simple category relations among a restricted set of concrete noun concepts, such as *bird, robin,* and *chicken.* The prevalent assumption is that we can discover how individual word meanings are represented in memory independent of how they may be used in discourse contexts. This is an essentially reductionist, bottom-up approach to semantic memory, and is well represented by Smith (1978). Kintsch (1979) questions the utility of this approach, arguing that semantic memory is our knowledge of the world and is thus propositional rather than word-based. Holding this issue in abeyance, what have we learned about the representation of word meanings?

One persistent issue has been whether word meanings and the relationships among them are stored explicitly in memory, or whether they must be computed afresh whenever needed. Two classes of semantic memory models take contrasting positions on this issue. Network models (Glass & Holyoak 1974-1975) assume that class-inclusion relations are explicitly stored in memory. People decide that statements like "A robin is a bird" are true by retrieving that information from memory (Meyer & Schvaneveldt 1976). Set-theoretic models (Smith et al 1974) assume that such relations among concepts are not stored explicitly and so are not available for direct access. Instead, one decides that a robin is a bird by comparing the semantic features of the relevent concepts, i.e. do *bird* and *robin* share criterial features?

The formal distinction between network and set-theoretic models may be vacuous because any set-theoretic model can be reformulated as a network model and vice versa (Hollan 1975; see also Woods 1975). Nevertheless, at least two interesting issues remain. The choice between retrieval and comparison processes in verifying class-inclusion statements is not resolved by such structural reformulations (Rips et al 1975). Secondly, whether the ordinary categories named by concrete nouns are well defined or fuzzy is also not resolved. Both set-theoretic and network models posit that common noun concepts are well defined, although this is not a necessary assumption of either class of models. The former do so in terms of defining features (Smith et al 1974); the latter in terms of associative links between exemplars and category nodes (Collins & Loftus 1975, J. R. Anderson 1976).

The data from semantic memory experiments which might speak to these two questions consist almost entirely of reaction times to respond "true" or "false" to simple statements of the form "Some/all S are P." Basically, comparisons between two kinds of true and two kinds of false statements provide the fundamental findings. True statements about typical category members, such as "A robin is a bird," are verified faster than statements about atypical category members, such as "A chicken is a bird." False statements which involve related concepts, such as "A butterfly is a bird," take longer to disconfirm than statements with unrelated concepts, such as "A table is a bird." These effects of typicality and semantic relatedness are extremely robust, appearing in a variety of experimental paradigms including production as well as comprehension tasks (see Kintsch 1979 for a comprehensive review).

Robustness and generality may well be these results' strongest claims to fame. With respect to the retrieval vs comparison question, the data do not discriminate between the alternatives. After all, there are only four data points to be explained. If one opts for simplicity and parsimony, then simple decision models of the sort proposed by Homa & Omohundro (1977) and by McCloskey & Glucksberg (1979) are sufficient. They account for the data with minimal representation and processing assumptions. If intuitive plausibility and generality are weighed heavily, then Collins & Loftus's (1975) spreading activation model, which treats assumptions as if they grow on trees, can account for virtually all findings. However, the patterns of results provided by semantic memory experiments seem to reflect both general and task-specific decision processes more than they do the context-free structure of semantic representations (Wickelgren 1977, Corbett & Wickelgren 1978).

The second question, whether natural categories are well-defined or fuzzy, also is not resolved adequately within the standard semantic memory paradigms. The obtained pattern of typicality effects can be accounted for by models which assume well-defined categories (Smith et al 1974) and by those which assume fuzzy category boundaries (McCloskey & Glucksberg 1979). To answer this question we must turn to research in wider contexts.

Psychological Semantics

Several developments in the last 6 years have raised questions about the definiteness of word meanings and category relations. In linguistics, Labov (1973), among others, argued that word meanings were inherently vague and context-dependent, much as Wittgenstein (1953) had claimed 20 years earlier. In cognitive psychology, Rosch and her colleagues have concluded that natural categories do not have well-defined boundaries. Categories are defined in terms of family resemblances rather than by criterial features (Rosch & Mervis 1975, Rosch 1978). Experimentally, demonstrations of the context sensitivity of word meanings have become commonplace (Barclay et al 1974, Anderson & Ortony 1975, Caramazza & Grober 1976). The

important theoretical generalization is that vagueness and flexibility are inherent characteristics of word meanings and are not restricted to a small class of ambiguous words. Polysemy is not an exception; it is the norm.

How can we deal with vagueness and polysemy? In linguistics the problem can be solved by ignoring it. The scope of investigation is restricted to lexical semantics and problems of reference; discourse and pragmatics are ignored (e.g. Lyons 1977; and see a critical review by G. Miller & K. Miller 1979). In philosophy, the problem can be partially handled by adopting a nominal view of naming (Kripke 1972). Put simplistically, the reason that entity *X* is named *N* is because it has been so named. Where linguistics deals only with lexical meanings, philosophy seems to deal only with the arbitrary nature of name assignments. Neither approach seems particularly helpful for a psychology of natural language use. The central questions for psychology remain: (*a*) How are word meanings represented in the mental dictionary, or more broadly, the mental encyclopedia? (*b*) How does one know when to apply, or use, a particular word?

The subset of the lexicon that has received the most attention is the 56 category names and their exemplars gathered by Battig & Montague (1969), e.g. *vegetable: beans, carrots, peas, peanuts,* etc. There is general agreement that such natural categories should not be represented by criterial features, but there is little agreement on just how they should be represented (Medin & Schaffer 1978, Martin & Caramazza 1979). Whatever the final solution to the representation problem, the available data indicate that concrete noun meanings are more or less vague. Category membership statements of the form "An olive is a vegetable" seem to have varying degrees of truth. People can reliably rate the relative truthfulness of such statements (Oden 1977). Further, people have difficulty deciding whether or not atypical category members belong in ordinary categories, e.g. "A peanut is a vegetable." People disagree with one another and are inconsistent with themselves when they make such decisions (McCloskey & Glucksberg 1978). These findings in considered judgment tasks parallel the effects of typicality on reaction time in speeded verification tasks.

Similarity relations among concepts also seem to be less rigid and definite than had been supposed (Rosch 1975). It had seemed intuitively obvious that the similarity of *A* to *B* would be judged the same as the similarity of *B* to *A*. However, there are clear cases where this assumption fails. For example, people judged that North Korea is more similar to Red China than the reverse (Tversky 1977). Tversky argues that geometric models of similarity relations, and hence geometric-spatial models of word meanings (e.g. Hutchinson & Lockhead 1977), are seriously undermined by such asymmetrical similarity relations. One implication is that semantic-feature representations are preferable to geometric-spatial representations because

they allow freedom for representing asymmetries. Krumhansl (1978) shows that one can do so in spatial models as well if assumptions are made about varying densities within semantic spaces. A more general implication of similarity asymmetries is that pragmatic and contextual factors are inherently involved in the comprehension of similarity statements, just as they seem to be in the comprehension of class-inclusion statements.

Perhaps the most interesting programmatic proposal for studying word meanings in general was offered by Miller & Johnson-Laird in their book, *Language and Perception* (1976). Words are assumed to represent preexisting concepts. The meaning of a word is neither a set of semantic feat' s nor a node in a semantic network. Instead, a word's meaning consists of a set of decision procedures that govern its use and application and a set of relationships between its meanings and the meanings of other words. These decision procedures are not limited to linguistic operations. Perceptual and functional properties of objects, actions, and events are integral to the decision procedure. For example, part of the meaning of the verb *meet* consists of a test which must be satisfied before "X meets Y" can be asserted, namely, *does y touch x?* This approach is called procedural semantics and is in principle compatible with artificial intelligence models of language processing in that it consists of instructions to perform specified operations (Johnson-Laird 1977a,b). If these procedures can vary as a function of context, then such an approach is also compatible with the notion of ill-defined word meanings or a fuzzy semantics. More generally, a procedural semantics rejects the utility of an autonomous semantics and explicitly incorporates world knowledge into the mental lexicon. The distinction between lexical and practical knowledge is ignored (Miller 1978). This kind of model is fully compatible with an interactive top-down model of language comprehension because it provides for the continuous availability of both linguistic and extralinguistic information.

In an analogous argument, E. V. Clark & H. H. Clark (1979) propose an intentional view of verb comprehension that integrates linguistic information with world knowledge and social context. Although they retain a linguistic core for word meanings, they nevertheless blur the distinction between linguistic and nonlinguistic sources of information. Both kinds of information must be used when people cope with the interpretations of words used in novel ways, as when nouns are used innovatively as verbs. How do we understand statements such as "He *porched* the newspaper" or "He *Houdinied* his way out of the box"? Clark & Clark argue that a speaker uses such words when there is good reason to believe that the listener can compute his intended interpretation from both linguistic and nonlinguistic knowledge. A meaning must be constructed; because it is novel, it cannot be retrieved from a semantic memory. Verbs in particular seem to require

this sort of integration between lexical and contextual information, but so do other parts of speech, e.g. modals (Johnson-Laird 1978), conjunctions (Fillenbaum 1974a), and quantifiers (Borges & Sawyers 1974, Hersh & Caramazza 1976).

From concrete nouns to verbs to the vaguest of adjectives, the representation of word meanings seems vague and fuzzy. Integration of linguistic and pragmatic-contextual information would seem to be a necessary component of word comprehension. Precisely how different types and sources of information are represented, accessed, and integrated remains to be specified.

SENTENCES: REPRESENTATIONS AND PROCESSES

How is the meaning of a sentence represented, and how are such representations derived? These two related questions have been addressed in two general ways: (*a*) by studies of sentence memory, and (*b*) by studies of sentence perception and comprehension.

Products of Comprehension: Sentence Memory

Two contrasting views of sentence comprehension dominate the literature. They are analogous, in important respects, to the bottom-up vs interactive views of word recognition. The bottom-up, interpretivist view of sentence comprehension claims that sentences are processed in sequential stages. First, a literal or canonical representation of a clause or a sentence is derived from the speech signal. Second, this representation is interpreted as a function of the social and discourse context. This is analogous to word recognition models which posit that words are first accessed via a search process and then interpreted as a function of context (e.g. Forster 1976). Sentence processing models of this type rely heavily on formal linguistic representations, such as those derived from transformational generative grammars (Fodor et al 1974, Schlesinger 1977). These linguistic models provide the format for the initial, literal representations required by an interpretivist view.

In contrast, a constructivist view provides a sentence-processing analog of a fully interactive model of word recognition, where social and discourse contexts influence recognition and comprehension decisions at the outset (Shaw & Bransford 1977). Instead of involving two discrete stages—the first context-free, the second interpretive—a constructive comprehender uses sentence and context information interactively and flexibly to arrive at an understanding of a speaker's intended meaning. On this view, linguistic processing per se has neither temporal nor informational priority.

These two views of sentence comprehension stress different aspects of sentence memory. The two-stage interpretivist view seeks evidence for memorial residues of an initial, literal interpretation. The constructivist interactive view seeks evidence for an elaborated representation, with little if any residues of raw, unprocessed verbatim information. The arguments are that if we remember verbatim or linguistic information, then sentence processing must involve an initial linguistic representation which is later available for interpretation. However, if we remember only gist, together with elaborated constructions such as inferences, then sentence processing need not involve an initial linguistic representation. The basic issue, however, is not whether sentence understanding requires extralinguistic information. Rather, the two opposing views differ on just when, in a comprehension sequence, extralinguistic contextual knowledge is used. When the issue is posed in this way, then the sentence memory literature seems irrelevant to the controversy. Either class of models could allow for virtually any form of final memorial representation of fully processed sentences. The predilections for sparse versus rich memorial representations seem more a reflection of preference than of substance. In any case, the available data are consistent with virtually any class of models for sentence representation, including propositions (Kintsch 1974, Ratcliff & McKoon 1978), associative networks (Anderson & Bower 1973, Anderson 1976), words (Hayes-Roth & Hayes-Roth 1977), and images (Danks & Sorce 1973, Marschark & Paivio 1977, Potter et al 1977).

The processing issue aside, what do people normally remember of sentences? The evidence strongly suggests that people can forget much verbatim, specific-item information and still retain gist (e.g. Barclay & Reid 1974, Pezdek & Royer 1974), just as Sachs had argued in 1967. Memory for verbatim information seems particularly fragile with long delays (Anderson & Paulson 1977), and recognition memory is poor when distractor and target items are highly confusable (James & Hillinger 1977). Memory for verbatim information is robust when that information is important. Memorized prose and verse can be retained verbatim for years (Rubin 1977). After all, paraphrases such as "my love is like a red red carnation" just won't do. Similarly, jokes and insults often depend upon specific wording for their import, and they too should be remembered verbatim. Keenan et al (1977) transcribed seminar discussions and then tested participants' recognition memory one day later. As expected, memory for meaning was quite good. Verbatim memory was also good, especially for statements with personal significance, such as sarcasm, personal criticism, and witty remarks. It is, of course, just these kinds of speech acts which cannot survive paraphrasing, and so it is not surprising that people remember the original wording. If they did not, then they could not retain the original meaning.

People may also remember specific wordings even when they are not particularly important. Hayes-Roth & Hayes-Roth (1977) replicated

Sachs's (1967, 1974) findings of recognition confusions between verbatim and paraphrased test items. In addition, they obtained higher confidence ratings and faster verification latencies for verbatim items than for paraphrases. It would seem that some specific lexical information is normally preserved after sentences have been understood. Other normal products of sentence comprehension include specific instantiations of general terms (Anderson & Ortony 1975), integrations of simple sentences into complex wholes (Hupet & LeBouedec 1977), inferred information (Fillenbaum 1974b), and presuppositions and implications (Harris & Monaco 1978). In short, people can remember many kinds of information after listening to or reading sentences (cf Craik 1979). The central issue for psycholinguistics is not what can be remembered, but how those memorial representations are generated in the first place.

Sentence Comprehension Processes

The most fully elaborated form of a two-stage interpretivist sentence processing model is the clausal processing hypothesis (Fodor et al 1974, Hurtig 1978). The structural component of this hypothesis is that sentences are segmented perceptually into clauses. The processing component is that integration and interpretation of word and phrase meanings are postponed until the ends of each clause (Marslen-Wilson et al 1978). Extralinguistic processing, such as the application of contextual information, is performed primarily at clause boundaries, when a completed linguistic representation is available for interpretation.

The primary evidence for perceptual segmentation of sentences into clauses comes from click-location experiments. People judge the location of nonlinguistic sounds, such as clicks, within spoken sentences. Perceived displacements of clicks into deep-structure clause boundaries have been interpreted as evidence for perceptual segmentation of sentences into deep-structure linguistic units (Bever & Hurtig 1975). The interpretive problems of click-location data have been discussed thoroughly elsewhere (Johnson-Laird 1974, Levelt 1978). The major problem is that clause boundaries are normally confounded with other variables, such as serial position of words (Reber 1973) and intonation patterns (Geers 1978). Suffice it to say that sentences may be perceptually segmented into clauses without being comprehended and interpreted on a clause-by-clause basis. Certainly, perceptual clause segmentation does not necessarily imply that interpretive work is restricted to, or concentrated in, clause boundaries (Townsend & Bever 1978).

A more direct implication of the clausal processing hypothesis is that lexical ambiguities are not resolved until clause endings (Bever et al 1973, Olson & MacKay 1974). If this is true, then both meanings of an ambiguous word are always accessed regardless of prior context, with a choice made only at the end of a clause. The alternative hypothesis is that prior context can restrict initial lexical access to the contextually appropriate sense of a word. This question has been extensively investigated with a phoneme-monitoring paradigm. People listen to sentences and respond as quickly as possible to predesignated target phonemes. Phoneme detection latencies are taken to reflect the relative processing difficulty of the word immediately preceding the target phoneme. Lexically ambiguous words appearing just before a target phoneme usually increase detection latencies (Cairns & Kamerman 1975). Prior context apparently did not eliminate this effect (Foss & Jenkins 1973), suggesting that ambiguity resolution did indeed await clause boundaries. More recent evidence, however, suggests that initial lexical access is affected by prior context. Mehler et al (1978) and Newman & Dell (1978) correctly pointed out that many phoneme-monitoring studies had failed to control for potential artifacts like the length, frequency, and phonemic composition of words immediately prior to target phonemes. These factors may have adventitiously increased detection latencies following ambiguous words. Blank & Foss (1978) explicitly controlled for the artifacts discussed by Mehler et al and by Newman & Dell, and found that prior context could facilitate immediate semantic processing of unambiguous words. These findings, together with those of Swinney & Hakes (1976), who did find effects of prior context on ambiguous word processing, argue against a strong form of the clausal processing hypothesis.

Further evidence against a strong form of the clausal processing hypothesis was reported by Marslen-Wilson et al (1978), using a word-detection paradigm. A target word could appear either immediately before or after a clause boundary. Irrespective of whether the target words were monitored for sound (rhyming) or for meaning (category membership), there were no effects of word location. At least insofar as word recognition and interpretation are concerned, position within clauses seems to be irrelevant (Rips et al 1978). These results, along with others (e.g. Rosenberg & Lambert 1974, Lindig 1976, Danks et al 1978), are inconsistent with interpretive models that posit literal representations as a necessary first step in a comprehension sequence before contextual information is brought into play.

One critical assumption of such interpretivist models is that there are nonarbitrary "literal" interpretations that are relatively context-free. However, even in impoverished laboratory situations, sentence-encodings are highly context-sensitive. For example, how sentences are coded in a sentence-verification task will vary as a function of the form of linguistic input (Banks et al 1975), the serial position of test sentences (Garrod & Trabasso

1973), whether pictures precede sentences or vice versa (Glucksberg et al 1973), and individual differences among subjects (MacLeod et al 1978). A particularly clear demonstration of task-specific processing strategies was reported by Aaronson (1976; Aaronson & Scarborough 1976, 1977). Readers were allowed to pace presentation of sentences, one word at a time. Readers who had been instructed to understand the sentences used a different pacing strategy than those who had been instructed to recall the sentences verbatim. Clearly, different laboratory tasks impose different sentence-processing demands, and people seem capable of varying their sentence encoding strategies to suit those demands.

Perhaps the clearest test cases for a literal-first interpretivist model are those sentences whose "literal" meanings do not coincide with their appropriate or intended meanings. Indirect requests are one such class of sentences. The statement "Can you pass the salt?" has at least two interpretations. It can be a question about someone's ability to perform an action, or it can be a request for that action to be performed. An interpretivist comprehension sequence for such sentences would be (*a*) derive a literal meaning, (*b*) check that meaning against the context, (*c*) if it fits plausibly, stop; (*d*) if not, seek an alternative nonliteral meaning that does fit. Clark & Lucy (1975) used such sentences in a sentence verification task and found that people required more time to verify indirect requests than direct ones. However, they tested each sentence in isolation. Gibbs (1979) found similar results with sentences in isolation. With the same sentences embedded in appropriate story contexts, appropriate indirect interpretations were understood more quickly than direct but inappropriate ones. These data are inconsistent with a literal-first, intended-optional comprehension sequence.

Metaphors provide another interesting set of test cases. The standard interpretivist view of metaphor comprehension is an exact parallel of Clark & Lucy's (1975) indirect request model. A metaphorical interpretation of a sentence is optional because it will occur only after a literal interpretation has failed to make sense. Therefore, metaphors are not only optionally interpreted, but they must also take more time and effort than ordinary, literal sentences. There are serious principled problems with such a view (cf Pollio et al 1977). For example, what are the nonarbitrary literal interpretations of such statements as "a woman without a man is like a fish without a bicycle" [attributed to Gloria Steinem, in G. Miller (1979)]? There are also severe empirical problems. Some types of metaphors seem to be processed nonoptionally. Glucksberg et al (1977) presented sentences of the form "some X are Y" in a sentence verification task. Embedded within a long list of simple sentences were literally false but metaphorically true sentences such as "some roads are snakes." People took significantly longer to respond "false" to such sentences. The availability of a "true" metaphorical meaning seemed to produce Stroop-like interference, presumably because people could not inhibit their understanding of the metaphor.

Within discourse contexts, metaphors may be understood as quickly as literal statements. Ortony et al (1978) compared the time required to understand how a target sentence either literally or metaphorically fit with (*a*) an impoverished prior context, or (*b*) a fully adequate prior context. With impoverished contexts, literal and metaphorical usages took 4 and 5 sec, respectively, to understand. With adequate contexts, literal and metaphorical usages were understood equally quickly. It would seem that literal sentence meanings have neither temporal nor informational priority over alternative nonliteral meanings. The alternative readings of a sentence may be equally accessible, just as the alternative readings of word senses may be equally accessible, provided that there is sufficient contextual information to guide comprehension. It should be noted that comprehension in the Ortony et al study of metaphor was operationally defined in terms of integration with a larger unit of discourse, i.e. a prior context. Within such larger units, it seems unlikely indeed that canonical, literal sentence representations would always be generated, only to be discarded when contexts so dictate. The alternative—a constructivist interactive model—would seem more efficient and plausible. Unfortunately, however, this plausibility may have been bought at the price of vagueness.

DISCOURSE COMPREHENSION

Conversations and prose narratives are coherent to us because we use various cohesive devices to link constituents together, and because conversations and texts usually display some global organization. The relevance of these text properties for studies of human memory has been described in Craik's recent review (1979). We consider representative research on how cohesion and organizational structure function in discourse comprehension.

Comprehension Strategies and Discourse Cohesion

During conversation, listeners use inferences about a speaker's intentions to help arrive at coherent understandings. These inferences are based on knowledge of conversations in general as well as on the speaker's utterances. Grice (1975) proposed that successful interpersonal communication depends on an implicit cooperative principle. This principle subsumes four conversational postulates, or maxims: be informative, truthful, relevant, and perspicuous. These maxims, if obeyed by speakers, and if believed to

be obeyed by listeners, provide a basis for making inferences about a speaker's intended meaning.

One instantiation of the relevancy maxim is the given-new contract. This presumes that utterances provide new information, and the contract requires speakers to signal which parts of an utterance are given (old) information, and which are new (Clark & Haviland 1977). This signaling can be accomplished in alternative ways, using linguistic as well as paralinguistic cues. A linguistic cue is illustrated in the sentence sequence "Horace got some beer out of the car. The beer was warm." The use of the definite article for the second mention of beer signals that it is the same beer that Horace had taken out of the car. The existence of the beer is given; that it was warm is the new information. When such cues are not used appropriately, then listeners must engage in additional inferential activity, as in "Horace got some picnic supplies out of the car. The beer was warm." Here, both the existence of the beer and its temperature are new, and listeners take longer to comprehend that sentence (Haviland & Clark 1974).

Other linguistic devices that can signal given-new information are ellipsis, pronominalization, word order, and stress (Halliday & Hasan 1976). MacWhinney & Bates (1978) examined how several such devices are used in English, Hungarian, and Italian. As we might have expected, the three languages differed markedly in the relative availability and strategic use of these devices. However, all three had ways to mark given-new information, and these were used to implement the given-new contract.

Detailed studies of how people establish reference correspondences provide further evidence that discourse comprehension involves considerable inferential and problem-solving work. When people converse there is an implicit understanding that they are talking about the same things (Clark & Marshall 1978). How are such understandings, or reference correspondences, established? At one extreme, people use word substitutions and pronouns, e.g. "Look at all that white fluffy stuff. It's pretty, but I wish the snow would melt." The words "stuff," "snow," and "it" all have the same referent, and listeners have no trouble in making this identification (Rochester & Martin 1977). Other referential identifications require more complex inferences, as in the beer bridge mentioned above.

Coherence emerges only in part from a listener's efforts to infer a speaker's intended meaning. It also derives partly from the organization of a conversation (Schenkein 1978). Topic organization, topic shifts (Goodenough & Weiner 1978), and the regulation of turn taking (Sacks et al 1974, Duncan & Fiske 1977) are signaled by a variety of linguistic and nonlinguistic cues, such as idea completeness, facial expressions, gestures, and intonation. These cues can be used to discriminate between a speaker pausing (*a*) to think, (*b*) to breathe, or (*c*) to allow a listener a turn to talk.

Interactive devices like these are not available to readers of stories and texts, and so readers must rely more heavily on linguistic conventions and knowledge schemata. Many textual linguistic devices are the same as those used in conversation, including definite reference, pronominalization, and anaphora. Readers also assume, with good reason, that successive sentences are related to one another unless otherwise marked (Haberlandt & Bingham 1978).

When a text does not provide explicit bridges between successive sentences, listeners and readers make whatever inferences are required to provide coherence (Crothers 1978, 1979; Warren et al 1979). Consider these alternative versions of a sentence sequence: "The millionaire was murdered (died). The killer escaped." If "died" is used, then an inference is required to bridge between "the killer" and the millionaire's death. This inference is not required if the word "murdered" is used. Using an eye-monitoring technique, Carpenter & Just (1977) found that readers spend about 0.5 sec longer inspecting the phrase "the killer" when "died" is used instead of "murdered." Readers also regress to the preceding sentence to confirm their inference. Similarly, distortions of topical and inferred relations cause disruptions of fluent shadowing (Lindig 1976) and of oral reading performance (Danks et al 1978).

Do readers draw inferences whether or not an inference is required? Clark (1977, 1978) suggests that only those that actually contribute to coherence are drawn, and Hildyard & Olson (1978) report that this ability may develop early. Fourth-grade children not only draw those inferences that are required for story coherence, they also discriminate between those that are useful for story understanding and those that are not.

Most of the cohesion-establishing devices we have discussed are used within relatively small units of discourse, e.g. referential correspondences between pairs of adjacent sentences. More global organizational aids may be found in larger text structures.

Text Coherence: Schemata and Grammars

The potential importance of schemata for lending coherence to narratives is nicely illustrated by vague texts. The sentence "The haystack was important because the cloth ripped" is syntactically well formed, contains no anomalous semantic relations, yet is quite difficult to interpret. If we infer, or are told, that the sentence is about a parachuting accident, then the meaning becomes clear (Bransford & McCarrell 1974). Both comprehension and recall performance are dramatically improved when such thematic information is provided, especially if it is provided either just before or immediately after the text. If recall is delayed and the thematic information given just before recall, it is far less effective (Cofer et al 1976, Dooling &

Christiaansen 1977a,b). In general, the effects of alternative themes and contexts on paragraph and story memory parallel their effects on sentence memory. For example, Anderson & Pichert (1978) used ambiguous stories with alternative thematic perspectives. A tour of a house could be described from the perspective of either a prospective buyer or a burglar. Recall protocols revealed bias in the expected directions, whether the perspective was established before or just after the story. This suggests that thematic information can affect both selective encoding and selective retrieval (Hasher & Griffin 1978).

The effects of thematic information on story recall have usually been interpreted in terms of Bartlett's (1932) concepts of schemata and reconstructive memory. What kinds of schemata do people use to interpret stories and to guide their reconstructions? A dominant set of schemata are our understandings of interpersonal and social interactions, and of people's psychological states and motivations in everyday situations (Bower 1978). An explicit formalization of this kind of information has been attempted by Schank & Abelson (1977) in the form of scripts. A script lists, in hierarchical form, the expected and appropriate sequences of actions and events in specified contexts. For example, a restaurant script would prescribe an event sequence which includes entering, being seated, ordering, eating, paying the bill, and leaving. If we read or hear a narrative with one or more script-prescribed events omitted, such as leaving a tip, we tend to insert it in subsequent recall. Bower et al (1979) found that people tended to agree on what are sensible and plausible scripts and on how to partition scripts into scenes. People also tended to remember script events in canonical order, to fill in routine script events if they are left unstated, and especially to remember salient, unexpected additions. In other words, people share a great deal of knowledge about what happens in familiar social situations and use that knowledge when they read and remember stories.

Scripts and other scenario-like schemata derive from the contents of narratives. They are analogous to the semantics and pragmatics of sentences. In contrast, structural descriptions that are analogous to idealized syntactic descriptions have been proposed in the form of content-free story grammars. Some of these adopt formal rule structures like those of generative grammars, particularly those that have been applied to relatively standardized story formats such as folk tales and children's stories (Rumelhart 1975, Mandler & Johnson 1977, Thorndyke 1977, Stein & Glenn 1979). More general structural descriptions that are applicable to any prose texts have also been proposed. These attempt to characterize how people impose hierarchical structure on propositions that may be derived from texts (Kintsch 1974, Frederiksen 1975b, Grimes 1975, Meyer 1975, van Dijk 1977).

Despite principled differences among these proposals, they share the assumptions (*a*) that prose material is encoded in hierarchically organized format, and (*b*) that the basic unit of encoding is propositional. Four basic empirical findings provide support for such text grammars in general: (*a*) Information at higher, more general levels in a hierarchical representation is remembered better than lower-level details (Meyer & McConkie 1973, Kintsch et al 1975, Gentner 1976). (*b*) When people are asked to summarize a given text, the summaries resemble delayed-recall protocols in that only higher-level information appears consistently, and it does so in structured order (Kintsch 1977, Rumelhart 1977b, Kintsch & van Dijk 1978, Glenn 1978). (*c*) Prose that is presented in grammatical or canonical order is easier to understand and to remember than prose that is presented out of order (Meyers & Boldrik 1975, Kintsch et al 1977, Mandler 1978, Stein & Nezworski 1978). (*d*) Theoretical prose structures predict the comprehension and recall performance of people who already have the appropriate schemata better than that of people who do not, such as children (Poulsen et al 1979), people with alien cultural backgrounds (Kintsch & Green 1978), and people who lack important technical knowledge (Krulee et al 1979).

Unfortunately, these data are not sufficient to discriminate among the various competing alternatives (cf Rubin 1978). Indeed, many of these results were originally reported in 1894 by Binet & Henri (cf Thieman & Brewer 1978). The difficulties associated with formal syntactic analyses of sentences have not been resolved by applying analogous formal structures to discourse. As with sentences, people tend to pay attention to and remember socially and personally relevant information, with little regard for abstract context-independent structures. Anyone who has taught lecture classes knows only too well that students often remember the "wrong" information, such as jokes and extraneous remarks, better than the main ideas. This intuition has been ruefully confirmed by Kintsch & Bates (1977).

Should formal text analyses be abandoned? We think not. Formal representation of text structures can be useful when it is coupled with explicit processing models. Kintsch & van Dijk (1978) provide an encouraging example in their attempt to account for the detailed characteristics of summary and recall protocols. They report some success in modeling both summarization and recall performances by using selected constraints that are derived from a general information processing theory, together with their specific propositional description of the texts. Their model has been extended to permit assessment of the relative readability of texts with some success as well (Kintsch & Vipond 1979). Perhaps the most useful outcome of current approaches to text comprehension will be more practical than theoretical. The kinds of questions asked, and the kinds of answers provided, should find applications in designing and assessing instructional

texts, and may provide guidelines for designing human-machine interaction systems.

TRENDS AND FUTURE DIRECTIONS

Our review has been highly selective and perhaps idiosyncratic. To the extent that it accurately reflects the field, two general trends can be discerned. One trend is exemplified by a concern with language processing within larger psychological contexts, and by an emerging consensus that continuous interactions among sources of information is a central characteristic of such language processing. In effect, this implies rejection of the standard, linguistically based model of language processing in which comprehension was accomplished in sequential stages, from lower-level units to integration with context. From one point of view, this represents a withdrawal of psycholinguistics from interdisciplinary cooperation. Formal linguistic theory no longer provides candidates for idealized, literal, context-free sentence representations.

A second trend, which has been implicit in our review, is a movement toward more interdisciplinary interaction in the larger enterprise of cognitive science. There are signs, including two new journals—*Cognitive Science* and *Discourse Processes*—that the separate disciplines of psychology, linguistics, computer science, anthropology, sociology, and philosophy might begin to recognize shared interests in and complementary contributions to the study of human mental life. Whether this promise will become reality is unclear. Johnson-Laird strikes an appropriate note of caution in his argument for the utility of artificial intelligence as an approach to the study of language: "Psychological processes take place in time, and so, too, do the operations of computers. Perhaps the metaphor can be pushed no further than that, but there does not seem to be any other equally viable alternative" (1977a, p. 213).

Social interactions in everyday situations also take place in time, and these may provide the models for more detailed and specific analyses of how linguistic, conceptual, pragmatic, and interpersonal mechanisms interact in natural language use. The next 5 years should see progress beyond demonstrations of context effects and arguments based upon plausibility. In the meantime, we have learned more about the complexity of normal language processes and about the general characteristics of potentially adequate language-processing theories. There is an emerging consensus that such theories will have to provide mechanisms for interactions of information from the most specific and sensory to the most general and conceptual. Such a theory will not be specific to psycholinguistics, but to human conceptual processing in general.

RECENT REVIEWS AND TEXTS

For aspects of psycholinguistics not covered here, see the following: neurolinguistics (Whitaker 1976, Miller & Lenneberg 1978, Walker 1978); speech perception (Darwin 1976, Kavanagh & Strange 1978, Pisoni 1978), speech production (Garrett 1975, Rosenberg 1977, Deese 1979); referential communication (Asher 1979); language acquisition (Abrahamsen 1977, de Villiers & de Villiers 1978); sign language (Schlesinger & Namir 1978, Klima & Bellugi 1979).

General introductions to the field are provided by several recent texts, including Glucksberg & Danks (1975), H. S. Cairns & C. E. Cairns (1976), H. H. Clark & E. V. Clark (1977), Foss & Hakes (1978), Palermo (1978), and a second edition of Slobin (1979).

Literature Cited

Aaronson, D. 1976. Performance theories for sentence coding: Some qualitative observations. *J. Exp. Psychol. Hum. Percept. Perform.* 2:42–55

Aaronson, D., Scarborough, H. S. 1976. Performance theories for sentence coding: Some quantitative evidence. *J. Exp. Psychol. Hum. Percept. Perform.* 2:56–70

Aaronson, D., Scarborough, H. S. 1977. Performance theories for sentence coding: Some quantitative models. *J. Verb. Learn. Verb. Behav.* 16:277–303

Abrahamson, A. A. 1977. *Child Language: An Interdisciplinary Guide to Theory and Research.* Baltimore: Univ. Park Press. 381 pp.

Anderson, J. R. 1976. *Language, Memory, and Thought.* Hillsdale, NJ: Erlbaum. 546 pp.

Anderson, J. R., Bower, G. H. 1973. *Human Associative Memory.* Washington, DC: Winston. 524 pp.

Anderson, J. R., Paulson, R. 1977. Representation and retention of verbatim information. *J. Verb. Learn. Verb. Behav.* 16:439–51

Anderson, R. C., Ortony, A. 1975. On putting apples into bottles—A problem of polysemy. *Cognit. Psychol.* 7:167–80

Anderson, R. C., Pichert, J. W. 1978. Recall of previously unrecallable information following a shift in perspective. *J. Verb. Learn. Verb. Behav.* 17:1–12

Asher, S. R. 1979. Referential communication. In *The Functions of Language and Cognition,* ed. G. J. Whitehurst, B. J. Zimmerman, 175–97. New York: Academic. 313 pp.

Banks, W. P., Clark, H. H., Lucy, P. 1975. The locus of the semantic congruity effect in comparative judgments. *J. Exp. Psychol. Hum. Percept. Perform.* 1:35–47

Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., Nitsch, K. 1974. Comprehension and semantic flexibility. *J. Verb. Learn. Verb. Behav.* 13:471–81

Barclay, J. R., Reid, M. 1974. Characteristics of memory representations of sentence sets describing linear arrays. *J. Verb. Learn. Verb. Behav.* 13:133–37

Bartlett, F. C. 1932. *Remembering: A Study in Experimental and Social Psychology.* Cambridge: Cambridge Univ. Press. 317 pp.

Battig, W. F., Montague, W. E. 1969. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *J. Exp. Psychol. Monogr.* 80:(3, Pt. 2) 1–46

Bever, T. G., Hurtig, R. R. 1975. Detection of a nonlinguistic stimulus is poorest at the end of a clause. *J. Psycholing. Res.* 4:1–7

Bever, T. G., Garrett, M. F., Hurtig, R. 1973. The interaction of perceptual processes and ambiguous sentences. *Mem. Cognit.* 1:277–86

Binet, A., Henri, V. 1894. La mémoire des phrases (mémoire des idées). *Année Psychol.* 1:24–59

Blank, M. A., Foss, D. J. 1978. Semantic facilitation and lexical access during sentence processing. *Mem. Cognit.* 6:644–52

Bobrow, D. G., Norman, D. A. 1975. Some principles of memory schemata. In *Representation and Understanding: Studies*

in *Cognitive Science*, ed. D. G. Bobrow, A. Collins, pp. 131–49. New York: Academic. 427 pp.

Bobrow, R. J., Brown, J. S. 1975. Systematic understanding: Synthesis, analysis, and contingent knowledge in specialized understanding systems. See Bobrow & Norman 1975, pp. 103–29

Borges, M. A., Sawyers, B. K. 1974. Common verbal quantifiers: Usage and interpretation. *J. Exp. Psychol.* 102:335–38

Bower, G. H. 1978. Experiments on story comprehension and recall. *Discourse Processes* 1:211–31

Bower, G. H., Black, J. B., Turner, T. J. 1979. Scripts in memory for text. *Cognit. Psychol.* 11:177–220

Bransford, J. D., McCarrell, N. S. 1974. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In *Cognition and the Symbolic Processes*, ed. W. B. Weimer, D. S. Palermo, pp. 189–229. Hillsdale, NJ: Erlbaum. 450 pp.

Cairns, H. S., Cairns, C. E. 1976. *Psycholinguistics: A Cognitive View of Language*. New York: Holt, Rinehart & Winston. 252 pp.

Cairns, H. S., Kamerman, J. 1975. Lexical information processing during sentence comprehension. *J. Verb. Learn. Verb. Behav.* 14:170–79

Caramazza, A., Grober, E. 1976. Polysemy and structure of the subjective lexicon. In *Georgetown University Round Table on Languages and Linguistics*, ed. C. Rameh, pp. 181–206. Washington DC: Georgetown Univ. Press. 279 pp.

Carpenter, P. A., Just, M. A. 1975. Sentence comprehension: A psycholinguistic processing model of verification. *Psychol. Rev.* 82:45–73

Carpenter, P. A., Just, M. A. 1977. Reading comprehension as eyes see it. In *Cognitive Processes in Comprehension*, ed. M. A. Just, P. A. Carpenter, pp. 109–39. Hillsdale, NJ: Erlbaum. 320 pp.

Clark, E. V., Clark, H. H. 1979. When nouns surface as verbs. *Language*. In press

Clark, H. H. 1977. Inferences in comprehension. In *Basic Processes in Reading: Perception and Comprehension*, ed. D. LaBerge, S. J. Samuels, pp. 243–63. Hillsdale, NJ: Erlbaum. 370 pp.

Clark, H. H. 1978. Inferring what is meant. See Levelt & Flores d'Arcais, pp. 295–322

Clark, H. H., Clark, E. V. 1977. *Psychology and Language*. New York: Harcourt Brace Jovanovich. 608 pp.

Clark, H. H., Haviland, S. E. 1977. Comprehension and the given-new contract. In *Discourse Production and Comprehension*, ed. R. O. Freedle, pp. 1–40. Norwood, NJ: Ablex. 345 pp.

Clark, H. H., Lucy, P. 1975. Understanding what is meant from what is said: A study in conversationally conveyed requests. *J. Verb. Learn. Verb. Behav.* 14:56–72

Clark, H. H., Marshall, C. 1978. Reference diaries. In *Theoretical Issues in Natural Language Processing-2*, ed. D. L. Waltz, pp. 57–63. New York: Assoc. Comput. Mach. 253 pp.

Cofer, C. N., Chmielewski, D. L., Brockway, J. P. 1976. Constructive processes and the structure of human memory. In *The Structure of Human Memory*, ed. C. N. Cofer, pp. 190–203. San Francisco: Freeman. 213 pp.

Cohen, A. S. 1974–1975. Oral reading errors of first grade children taught by a code emphasis approach. *Reading Res. Q.* 10:616–50

Cole, R. A. 1973. Listening for mispronunciations: A measure of what we hear during speech. *Percept. Psychophys.* 13:153–56

Collins, A. M., Loftus, E. F. 1975. A spreading-activation theory of semantic processing. *Psychol. Rev.* 82:407–28

Coltheart, M. 1978. Lexical access in simple reading tasks. In *Strategies of Information Processing*, ed. G. Underwood, pp. 151–216. London: Academic. 455 pp.

Corbett, A. T., Wickelgren, W. A. 1978. Semantic memory retrieval: Analysis by speed accuracy tradeoff functions. *Q. J. Exp. Psychol.* 30:1–15

Craik, F. I. M. 1979. Human memory. *Ann. Rev. Psychol.* 30:63–102

Crothers, E. J. 1978. Inference and coherence. *Discourse Processes* 1:51–71

Crothers, E. J. 1979. *Paragraph Structure Inference*. Norwood, NJ: Ablex. 124 pp.

Danks, J. H. 1978. Models of language comprehension. *Pol. Psychol. Bull.* 9:183–92

Danks, J. H., Fears, R., Bohn, L., Hill, G. O. 1978. *Comprehension processes in oral reading*. Presented at Ann. Meet. Psychon. Soc., San Antonio, Texas

Danks, J. H., Sorce, P. A. 1973. Imagery and deep structure in the prompted recall of passive sentences. *J. Verb. Learn. Verb. Behav.* 12:114–17

Darwin, C. J. 1976. The perception of speech. In *Handbook of Perception: Language and Speech*, ed. E. C. Carterette, M. P. Friedman, 7:175–226. New York: Academic. 501 pp.

Deese, J. 1979. Pauses, prosody, and the demands of production in language. In *Temporal Variables in Speech: Studies in Honor of Frieda Goldman-Eisler*, ed. H. W. Dechert, M. Raupach. The Hague: Mouton. In press

de Villiers, J. G., de Villiers, P. A. 1978. *Language Acquisition*. Cambridge: Harvard Univ. Press. 312 pp.

Dooling, D. J., Christiaansen, R. E. 1977a. Episodic and semantic aspects of memory for prose. *J. Exp. Psychol. Hum. Learn. Mem.* 3:428–36

Dooling, D. J., Christiaansen, R. E. 1977b. Levels of encoding and retention of prose. In *The Psychology of Learning and Motivation*, ed. G. H. Bower, 11:1–39. New York: Academic. 323 pp.

Duncan, S., Fiske, D. W. 1977. *Face-to-Face Interaction: Research, Methods, and Theory*. Hillsdale, NJ: Erlbaum. 361 pp.

Fillenbaum, S. 1974a. *OR*: Some uses. *J. Exp. Psychol.* 103:913–21

Fillenbaum, S. 1974b. Information amplified: Memory for counterfactual conditionals. *J. Exp. Psychol.* 102:44–49

Fodor, J. A., Bever, T. G., Garrett, M. F. 1974. *The Psychology of Language*. New York: McGraw-Hill. 537 pp.

Forster, K. I. 1976. Accessing the mental lexicon. In *New Approaches to Language Mechanisms*, ed. R. J. Wales, E. Walker, pp. 257–87. Amsterdam: North-Holland. 296 pp.

Foss, D. J., Hakes, D. T. 1978. *Psycholinguistics*. Englewood Cliffs: Prentice-Hall. 434 pp.

Foss, D. J., Jenkins, C. M. 1973. Some effects of context on the comprehension of ambiguous sentences. *J. Verb. Learn. Verb. Behav.* 12:577–89

Frederiksen, C. H. 1975a. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognit. Psychol.* 7:139–66

Frederiksen, C. H. 1975b. Representing logical and semantic structure of knowledge acquired from discourse. *Cognit. Psychol.* 7:371–458

Garrett, M. F. 1975. The analysis of sentence production. In *The Psychology of Learning and Motivation*, ed. G. H. Bower, 9:133–77. New York: Academic. 317 pp.

Garrod, S., Trabasso, T. 1973. A dual-memory information processing interpretation of sentence comprehension. *J. Verb. Learn. Verb. Behav.* 12:155–67

Geers, A. E. 1978. Intonation contour and syntactic structure as predictors of apparent segmentation. *J. Exp. Psychol. Hum. Percept. Perform.* 4:273–83

Gentner, D. R. 1976. The structure and recall of narrative prose. *J. Verb. Learn. Verb. Behav.* 15:411–18

Gibbs, R. W. 1979. Contextual effects in understanding indirect requests. *Discourse Processes* 2:1–10

Glass, A. L., Holyoak, K. J. 1974–1975. Alternative conceptions of semantic theory. *Cognition* 3:313–39

Glenn, C. G. 1978. The role of episodic structure and of story length in children's recall of simple stories. *J. Verb. Learn. Verb. Behav.* 17:229–47

Glucksberg, S., Danks, J. H. 1975. *Experimental Psycholinguistics*. Hillsdale, NJ: Erlbaum. 233 pp.

Glucksberg, S., Hartman, D. E., Stack, R. A. 1977. *Metaphor comprehension is an automatic and parallel process*. Presented at Ann. Meet. Psychon. Soc., Washington DC

Glucksberg, S., Trabasso, T., Wald, J. 1973. Linguistic structures and mental operations. *Cognit. Psychol.* 5:338–70

Goodenough, D. R., Weiner, S. L. 1978. The role of conversational passing moves in the management of topical transitions. *Discourse Processes* 1:395–404

Grice, H. P. 1975. Logic and conversation. In *Syntax and Semantics: Speech Acts*, ed. P. Cole, J. L. Morgan, 3:41–58. New York: Academic. 406 pp.

Grimes, J. E. 1975. *The Thread of Discourse*. The Hague: Mouton. 408 pp.

Haberlandt, K., Bingham, G. 1978. Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *J. Verb. Learn. Verb. Behav.* 17:419–25

Halle, M., Bresnan, J., Miller, G. A., eds. 1978. *Linguistic Theory and Psychological Reality*. Cambridge: MIT Press. 329 pp.

Halliday, M. A. K., Hasan, R. 1976. *Cohesion in English*. London: Longman. 374 pp.

Harris, R. J., Monaco, G. E. 1978. Psychology of pragmatic implication: Information processing between the lines. *J. Exp. Psychol. Gen.* 107:1–22

Hasher, L., Griffin, M. 1978. Reconstructive and reproductive processes in memory. *J. Exp. Psychol. Hum. Learn. Mem.* 4:318–30

Haviland, S. E., Clark, H. H. 1974. What's new? Acquiring new information as a process in comprehension. *J. Verb. Learn. Verb. Behav.* 13:512–21

Hayes-Roth, B., Hayes-Roth, F. 1977. The prominence of lexical information in

memory representations of meaning. *J. Verb. Learn. Verb. Behav.* 16:119–36
Hersh, H. M., Caramazza, A. 1976. A fuzzy set approach to modifiers and vagueness in natural language. *J. Exp. Psychol. Gen.* 105:254–76
Hildyard, A., Olson, D. R. 1978. Memory and inference in the comprehension of oral and written discourse. *Discourse Processes* 1:91–117
Hollan, J. D. 1975. Features and semantic memory: Set-theoretic or network model? *Psychol. Rev.* 82:154–55
Homa, D., Omohundro, J. 1977. Search and decision processes in semantic memory. *J. Verb. Learn. Verb. Behav.* 16:383–94
Hupet, M., LeBouedec, B. 1977. The given-new contract and the constructive aspect of memory for ideas. *J. Verb. Learn. Verb. Behav.* 16:69–75
Hurtig, R. 1978. The validity of clausal processing strategies at the discourse level. *Discourse Processes* 1:195–202
Hutchinson, J. W., Lockhead, G. R. 1977. Similarity as distance: A structural principle for semantic memory. *J. Exp. Psychol. Hum. Learn. Mem.* 3:660–78
James, C. T., Hillinger, M. L. 1977. The role of confusion in the semantic integration paradigm. *J. Verb. Learn. Verb. Behav.* 16:711–21
Johnson-Laird, P. N. 1974. Experimental psycholinguistics. *Ann. Rev. Psychol.* 25:135–60
Johnson-Laird, P. N. 1977a. Procedural semantics. *Cognition* 5:189–214
Johnson-Laird, P. N. 1977b. Psycholinguistics without linguistics. In *Tutorial Essays in Psychology*, ed. N. S. Sutherland, 1:75–135. Hillsdale, NJ: Erlbaum. 173 pp.
Johnson-Laird, P. N. 1978. The meaning of modality. *Cognit. Sci.* 2:17–26
Kavanagh, J. F., Strange, W., eds. 1978. *Speech and Language in the Laboratory, School, and Clinic.* Cambridge: MIT Press. 511 pp.
Keenan, J. M., MacWhinney, B., Mayhew, D. 1977. Pragmatics in memory: A study of natural conversation. *J. Verb. Learn. Verb. Behav.* 16:549–60
Kintsch, W. 1974. *The Representation of Meaning in Memory.* Hillsdale, NJ: Erlbaum. 279 pp.
Kintsch, W. 1977. On comprehending stories. See Carpenter & Just 1977, pp. 33–62
Kintsch, W. 1979. Semantic memory: A tutorial. In *Attention and Performance VIII*, ed. R. S. Nickerson. Hillsdale, NJ: Erlbaum. In press
Kintsch, W., Bates, E. 1977. Recognition memory for statements from a classroom lecture. *J. Exp. Psychol. Hum. Learn. Mem.* 3:150–59
Kintsch, W., Green, E. 1978. The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes* 1:1–13
Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., Keenan, J. M. 1975. Comprehension and recall of text as a function of content variables. *J. Verb. Learn. Verb. Behav.* 14:196–214
Kintsch, W., Mandel, T. S., Kozminsky, E. 1977. Summarizing scrambled stories. *Mem. Cognit.* 5:547–52
Kintsch, W., van Dijk, T. A. 1978. Toward a model of text comprehension and production. *Psychol. Rev.* 85:363–94
Kintsch, W., Vipond, D. 1979. Reading comprehension and readability in educational practice and psychological theory. In *Perspectives on Memory Research*, ed. L.-G. Nilsson, pp. 329–65. Hillsdale, NJ: Erlbaum. 400 pp.
Klima, E. S., Bellugi, U. 1979. *Signs of Language.* Cambridge: Harvard Univ. Press. 262 pp.
Kripke, S. A. 1972. Naming and necessity. In *Semantics of Natural Language*, ed. D. Davidson, G. Harman, pp. 253–355. Dordrecht: Reidel. 769 pp.
Krulee, G. K., Fairweather, P. G., Bergquist, S. R. 1979. Organizing factors in the comprehension and recall of connected discourse. *J. Psycholing. Res.* 8:141–63
Krumhansl, C. L. 1978. Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychol. Rev.* 85:445–63
Labov, W. 1973. The boundaries of words and their meanings. In *New Ways of Analyzing Variation in English*, ed. C.-J. N. Bailey, R. W. Shuy, pp. 340–73. Washington DC: Georgetown Univ. Press. 373 pp.
Labov, W., Fanshel, D. 1977. *Therapeutic Discourse: Psychotherapy as Conversation.* New York: Academic. 392 pp.
Levelt, W. J. M. 1978. A survey of studies in sentence perception: 1970–1976. See Levelt & Flores d'Arcais 1978, pp. 1–73
Levelt, W. J. M., Flores d'Arcais, G. B., eds. 1978. *Studies in the Perception of Language.* Chichester: Wiley. 335 pp.
Lindig, K. D. 1976. *Contextual factors in shadowing connected discourse.* PhD thesis. Univ. Chicago, Chicago, Ill. 110 pp.
Lyons, J. 1977. *Semantics.* Cambridge: Cambridge Univ. Press. 2 vols. 897 pp.
Maclay, H. 1971. Overview. In *Semantics*, ed. D. D. Steinberg, L. A. Jakobovits, pp. 157–82. Cambridge: Cambridge Univ. Press. 603 pp.
MacLeod, C. M., Hunt, E. B., Mathews, N. N. 1978. Individual differences in the verification of sentence-picture relationships. *J. Verb. Learn. Verb. Behav.* 17:493–507
MacWhinney, B., Bates, E. 1978. Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *J. Verb. Learn. Verb. Behav.* 17:539–58
Mandler, J. M. 1978. A code in the node: The use of a story schema in retrieval. *Discourse Processes* 1:14–35
Mandler, J. M., Johnson, N. S. 1977. Remembrance of things parsed: Story structure and recall. *Cognit. Psychol.* 9:111–15
Marschark, M., Paivio, A. 1977. Integrative processing of concrete and abstract sentences. *J. Verb. Learn. Verb. Behav.* 16:217–31
Marslen-Wilson, W. D. 1975. Sentence perception as an interactive parallel process. *Science* 189:226–28
Marslen-Wilson, W. D., Tyler, L. K., Seidenberg, M. 1978. Sentence processing and the clause boundary. See Levelt & Flores d'Arcais 1978, pp. 219–46
Marslen-Wilson, W. D., Welsh, A. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognit. Psychol.* 10:29–63
Martin, R., Caramazza, A. 1979. Classification in well-defined and ill-defined categories: Evidence for common processing strategies. *J. Exp. Psychol. Gen.* In press
Massaro, D. W., ed. 1975. *Understanding Language: An Information-Processing Analysis of Speech Perception, Reading, and Psycholinguistics.* New York: Academic. 439 pp.
Massaro, D. W., Jones, R. D., Lipscomb, C., Scholz, R. 1978. Role of prior knowledge on naming and lexical decisions with good and poor stimulus information. *J. Exp. Psychol. Hum. Learn. Mem.* 4:498–512
McCloskey, M. E., Glucksberg, S. 1978. Natural categories: Well-defined or fuzzy sets? *Mem. Cognit.* 6:462–72
McCloskey, M., Glucksberg, S. 1979. Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognit. Psychol.* 11:1–37
McNeill, D., Lindig, K. 1973. The perceptual reality of phonemes, syllables, words, and sentences. *J. Verb. Learn. Verb. Behav.* 12:419–30
Medin, D. L., Schaffer, M. M. 1978. Context theory of classification learning. *Psychol. Rev.* 85:207–38
Mehler, J., Segui, J., Carey, P. 1978. Tails of words: Monitoring ambiguity. *J. Verb. Learn. Verb. Behav.* 17:29–35
Meyer, B. J. F. 1975. *The Organization of Prose and Its Effects on Memory.* Amsterdam: North-Holland. 249 pp.
Meyer, B. J. F., McConkie, G. W. 1973. What is recalled after hearing a passage? *J. Educ. Psychol.* 65:109–17
Meyer, D. E., Schvaneveldt, R. W. 1976. Meaning, memory structure, and mental processes. *Science* 192:27–33
Meyers, L. S., Boldrik, D. 1975. Memory for meaningful connected discourse. *J. Exp. Psychol. Hum. Learn. Mem.* 1:584–91
Miller, G. A. 1962. Some psychological studies of grammar. *Am. Psychol.* 17:748–62
Miller, G. A. 1978. Practical and lexical knowledge. In *Cognition and Categorization*, ed. E. Rosch, B. B. Lloyd, pp. 305–19. Hillsdale, NJ: Erlbaum. 328 pp.
Miller, G. A. 1979. Images and models, similes and metaphors. In *Metaphor and Thought*, ed. A. Ortony. Cambridge, England: Cambridge Univ. Press. In press
Miller, G. A., Johnson-Laird, P. N. 1976. *Language and Perception.* Cambridge: Belknap/Harvard Univ. Press. 760 pp.
Miller, G. A., Lenneberg, E., eds. 1978. *Psychology and Biology of Language and Thought.* New York: Academic. 294 pp.
Miller, G. A., Miller, K. 1979. Critical notice: John Lyons on semantics. *Q. J. Exp. Psychol.* In press
Morton, J. 1969. Interaction of information in word recognition. *Psychol. Rev.* 76:165–78
Morton, J., Long, J. 1976. Effect of word transitional probability on phoneme identification. *J. Verb. Learn. Verb. Behav.* 15:43–51
Newman, J. E., Dell, G. S. 1978. The phonological nature of phoneme monitoring: A critique of some ambiguity studies. *J. Verb. Learn. Verb. Behav.* 17:359–74
Oden, G. C. 1977. Fuzziness in semantic memory: Choosing exemplars of subjective categories. *Mem. Cognit.* 5:198–204
Olson, J. N., MacKay, D. G. 1974. Completion and verification of ambiguous sentences. *J. Verb. Learn. Verb. Behav.* 13:457–70

Ortony, A., Schallert, D. L., Reynolds, R. E., Antos, S. J. 1978. Interpreting metaphors and idioms: Some effects of context on comprehension. *J. Verb. Learn. Verb. Behav.* 17:465–77
Palermo, D. S. 1978. *Psychology of Language.* Glenview, Ill: Scott, Foresman. 261 pp.
Pezdek, K., Royer, J. M. 1974. The role of comprehension in learning concrete and abstract sentences. *J. Verb. Learn. Verb. Behav.* 13:551–58
Pisoni, D. B. 1978. Speech perception. In *Handbook of Learning and Cognitive Processes: Linguistic Functions in Cognitive Theory*, ed. W. K. Estes, 6:167–233. Hillsdale, NJ: Erlbaum. 311 pp.
Pollio, H. R., Barlow, J. M., Fine, H. J., Pollio, M. R. 1977. *Psychology and the Poetics of Growth: Figurative Language in Psychology, Psychotherapy, and Education.* Hillsdale, NJ: Erlbaum. 260 pp.
Potter, M. C., Valian, V. V., Faulconer, B. A. 1977. Representation of a sentence and its pragmatic implications: Verbal, imagistic, or abstract? *J. Verb. Learn. Verb. Behav.* 16:1–12
Poulsen, D., Kintsch, E., Kintsch, W., Premack, D. 1979. Children's comprehension and memory for stories. *J. Exp. Child Psychol.* In press
Ratcliff, R., McKoon, G. 1978. Priming in item recognition: Evidence for the propositional structure of sentences. *J. Verb. Learn. Verb. Behav.* 17:403–17
Reber, A. S. 1973. Locating clicks in sentences: Left, center, and right. *Percept. Psychophys.* 13:133–38
Rips, L. J., Smith, E. E., Shoben, E. J. 1975. Set-theoretic and network models reconsidered: A comment on Hollan's "Features and semantic memory." *Psychol. Rev.* 82:156–57
Rips, L. J., Smith, E. E., Shoben, E. J. 1978. Semantic composition in sentence verification. *J. Verb. Learn. Verb. Behav.* 17:375–401
Rochester, S. R., Martin, J. R. 1977. The act of referring: The speaker's use of noun phrases to instruct the listener. See Clark & Haviland 1977, pp. 245–69
Rosch, E. 1975. Cognitive reference points. *Cognit. Psychol.* 7:532–47
Rosch, E. 1978. Principles of categorization. See Miller 1978, pp. 28–48
Rosch, E., Mervis, C. B. 1975. Family resemblances: Studies in the internal structure of categories. *Cognit. Psychol.* 7:573–605
Rosenberg, S., ed. 1977. *Sentence Production: Developments in Research and Theory.* Hillsdale, NJ: Erlbaum. 323 pp.
Rosenberg, S., Lambert, W. E. 1974. Contextual constraints and the perception of speech. *J. Exp. Psychol.* 102:178–80
Rubenstein, H., Aborn, M. 1960. Psycholinguistics. *Ann. Rev. Psychol.* 11:291–322
Rubin, D. C. 1977. Very long-term memory for prose and verse. *J. Verb. Learn. Verb. Behav.* 16:611–21
Rubin, D. C. 1978. A unit analysis of prose memory. *J. Verb. Learn. Verb. Behav.* 17:599–620
Rumelhart, D. E. 1975. Notes on a schema for stories. See Bobrow & Norman 1975, pp. 211–36
Rumelhart, D. E. 1977a. Toward an interactive model of reading. In *Attention and Performance VI*, ed. S. Dornic, pp. 573–603. Hillsdale, NJ: Erlbaum. 777 pp.
Rumelhart, D. E. 1977b. Understanding and summarizing brief stories. See Clark 1977, pp. 265–303
Sachs, J. S. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. *Percept. Psychophys.* 2:437–42
Sachs, J. S. 1974. Memory in reading and listening to discourse. *Mem. Cognit.* 2:95–100
Sacks, H., Schegloff, E. A., Jefferson, G. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735
Schank, R. C., Abelson, R. P. 1977. *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures.* Hillsdale, NJ: Erlbaum. 248 pp.
Schenkein, J., ed. 1978. *Studies in the Organization of Conversational Interaction.* New York: Academic Press. 275 pp.
Schlesinger, I. M. 1977. *Production and Comprehension of Utterances.* Hillsdale, NJ: Erlbaum. 235 pp.
Schlesinger, I. M., Namir, L., eds. 1978. *Sign Language of the Deaf.* New York: Academic. 380 pp.
Searle, J. R. 1975. Indirect speech acts. See Grice 1975, pp. 59–82
Shaw, R., Bransford, J., eds. 1977. *Perceiving, Acting, and Knowing.* Hillsdale: Erlbaum. 492 pp.
Slobin, D. 1979. *Psycholinguistics.* Glenview, Ill: Scott, Foresman. 224 pp. 2nd ed.
Smith, E. E. 1978. Theories of semantic memory. See Pisoni 1978, pp. 1–56
Smith, E. E., Shoben, E. J., Rips, L. J. 1974. Structure and process in semantic memory: A featural model for semantic decisions. *Psychol. Rev.* 81:214–41
Staats, A. W. 1976. Social behaviorism and neo-psycholinguistics. *Die Neueren Sprachen* 2:127–41

Stein, N. L., Glenn, C. G. 1979. An analysis of story comprehension in elementary school children. In *Advances in Discourse Processing: New Directions in Discourse Processing*, Vol. 2, ed. R. O. Freedle. Norwood, NJ: Ablex. In press
Stein, N. L., Nezworski, T. 1978. The effects of organization and instructional set on story memory. *Discourse Processes* 1:177–93
Swinney, D. A., Hakes, D. T. 1976. Effects of prior context upon lexical access during sentence comprehension. *J. Verb. Learn. Verb. Behav.* 15:681–89
Thieman, T. J., Brewer, W. F. 1978. Alfred Binet on memory for ideas. *Genet. Psychol. Monogr.* 97:243–64
Thorndyke, P. W. 1977. Cognitive structures in comprehension and memory of narrative discourse. *Cognit. Psychol.* 9:77–110
Townsend, D. J., Bever, T. G. 1978. Interclause relations and clausal processing. *J. Verb. Learn. Verb. Behav.* 17:509–21
Tversky, A. 1977. Features of similarity. *Psychol. Rev.* 84:327–52
van Dijk, T. A. 1977. *Text and Context: Explorations in the Semantics and Pragmatics of Discourse.* London: Longman. 261 pp.
Walker, E., ed. 1978. *Explorations in the Biology of Language.* Montgomery, Vt: Bradford. 247 pp.
Warren, W. H., Nicholas, D. W., Trabasso, T. 1979. Event chains and inferences in understanding narratives. See Stein & Glenn 1979
Wexler, K. 1978. A review of John R. Anderson's *Language, Memory and Thought. Cognition* 6:327–51
Whitaker, H. A. 1976. Neurobiology of language. See Darwin 1976, pp. 121–44
Wickelgren, W. A. 1977. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41:67–85
Winograd, T. 1976. Computer memories: A metaphor for memory organization. See Cofer et al 1976, pp. 133–61
Wittgenstein, L. 1953. *Philosophical Investigations.* New York: Macmillan. 232 pp.
Woods, W. A. 1975. What's in a link: Foundations for semantic networks. See Bobrow & Norman 1975, pp. 35–82

READING AND UNDERSTANDING

Joseph Danks
Kent State University

Kathy Pezdek
California State College

Edited by Frank B. Murray
University of Delaware

1980
Published by
INTERNATIONAL READING ASSOCIATION
800 Barksdale Road Newark, Delaware 19711

Comprehension in Listening and Reading: Same or Different?

JOSEPH H. DANKS
Kent State University

"In order to read with comprehension, all the beginning reader has to do is pronounce the words correctly at a reasonable rate. His oral language comprehension processes will then suffice." Many researchers have tacitly assumed this proposition in their interpretation of research results and in their proposals for reading programs. Others have assumed the opposite, namely, that reading comprehension requires cognitive processes that are different in some major respects from those of listening. The fact that these assumptions are rarely made explicit should not lull us into overlooking them. Researchers have presented few data to support either assumption. Indeed, a cursory review of the literature reveals that data sufficient to resolve the issue are not available, primarily because of flaws in the rationales, designs, and procedures of the studies.

This chapter is divided into three sections. First, the two extreme positions are identified and elaborated, drawing from published writings of several well-known reading professionals. The implications for educational programs that follow from assuming each stance are identified. In particular, how one teaches reading and/or English and/or language arts in the later elementary school years is determined in large measure by the assumed relation between oral and printed communication systems. In the second section, methodological problems with empirical studies that bear on the issue are reviewed and illustrated. Finally, a task analysis is proposed that would direct research toward a valid comparison of listening and reading comprehension.

The discussion here is limited to what has been commonly called the comprehension aspect of reading in contrast to decoding. Word identification, which completes the

decoding phase, provides the grist for comprehension. However, when the dichotomy is pressed, it is difficult to maintain; for example, Ryan and Semmel (1969) describe the "downstream" influence that comprehension has on decoding. The decoding phases of listening and reading are necessarily different by virture of the different modalities of input and the amount of control the recipient has over the input flow. A listener has minimal control over the auditory signal which rapidly fades. A reader, on the other hand, has reasonably complete control over the amount and rate of visual input which is continuously available for reprocessing. If there is not a sharp break between decoding and comprehension, as there well may not be, then the different processing mechanisms at the decoding end may have a marked influence on the later comprehension stages. In general, this chapter focuses on comprehension and not decoding although this separation may be a fiction that ultimately cannot be maintained.

The Two Extremes

A Unitary Comprehension Process

There has been considerable discussion about the relation between a reader's processing of speech and print. Two conferences, reported in Kavanagh (1968) and Kavanagh and Mattingly (1972), were devoted to exploring this relationship. In the first conference, Alvin Liberman discussed several differences between listening and reading. However, all of his points focused on decoding speech and print into language, all of which presumably occurs prior to processing for meaning. As conveyed by the title of the report of the second conference, *Language by Ear and by Eye*, the emphasis there was also on the more peripheral aspects of the relationship, that is speech perception and reading implicitly defined as decoding. Few of the papers in that volume even hinted at the problem of comprehension and the similarities or differences as a function of mode of linguistic input. For the most part, the participants did not question, at least in the printed record of the conference, the assumption that once past word identification, the processes of comprehending speech and print do not differ.

The position that a unitary comprehension process is activated regardless of mode of input has been attributed in recent years to the proclamation of linguists that speech is primary and that print is a poor second cousin in the language family. The position was stated most decisively by Fries (1963, xv).

> Learning to read . . . is *not* a process of learning new or other language signals than those the child has already learned. The language signals are all the same. The difference lies in the medium through which the physical stimuli make contact with his nervous system. In 'talk,' the physical stimuli of the language signals make their contact by means of sound waves received by the ear. In reading, the physical stimuli of the same language signals consist of graphic shapes that make their contact with his nervous system through light waves received by the eye. The process of learning to read is the process of transfer from the auditory signs for language signals which the child has already learned, to the new visual signs for the same signals.

Psychologists interested in reading have frequently maintained a similar posture. For example, the research group at Cornell University who have been actively studying the reading process under the leadership of E. Gibson and H. Levin, have accepted a unitary comprehension process.

> Publications of the Gibson group suggest the model of the reader as "speaker, then hearer." The task of the reader is to translate graphemes into phonemes. This is the task particular to reading. After decoding to speech has been accomplished, comprehension processes associated with speech comprehension are brought into play [Wanat, 1971: 8-155].

In fairness it should be noted that Gibson (1972) in her keynote address to the conference reported in Kavanagh and Mattingly hinted that this position may be wrong. She commented that researchers have assumed that a reader's knowledge of grammar for oral language is transferred automatically and directly to the reading process but that this assumption may not be justified.

The assumption of a single comprehension process following initial decoding of either sound or print also has been

readily accepted by educators and educational psychologists. In many cases the assumption is not even made explicit and, when it is, it is noted in passing with the implication that a point so obvious scarcely bears mentioning. Note these examples from two prominent educational psychologists who spoke at the institute:

> The process of reading comprehension differs from the process of language comprehension only in the form of perceptual input. The latter works from oral input; the former from visual input [Goodman, 1966: 168].
>
> Performance in reading, at least after the basic decoding skills are mastered, is primarily an indicator of the general level of the individual's thinking and reasoning processes rather than a set of distinct and specialized skills [Thorndike, 1973-74: 135].

In sum, a frequently held position is that once the reader or listener has translated the particular form of input into an abstract coding, language is language. The language code is either an abstract code that is independent of the modality of input or a speech code to which print is decoded.

What effect does holding a unitary process view have on educational practice? Many school systems have accepted this position, albeit implicitly. Witness the fact that in many elementary schools, "reading" as a separate subject is taught only in the first two or three grades. At that point the teacher stops teaching reading and teaches language arts or English. The implicit assumption is that once the child has acquired decoding skills, training and practice with one form of language wil' ·ansfer to other forms. This educational practice is not ...ed on some amorphous notion of generalized transfer. Listening skills place an absolute ceiling on the level of reading skill a child may attain. As Bormuth (1972) has noted, "in the tradition and folklore of reading instruction, a student's ability to comprehend spoken language is regarded as an estimate of the upper limit of his capacity to learn reading comprehension skills" (p. 1134). Such an assumption demands that if education is to increase reading skills, oral language skills must first be increased beyond the desired level of reading. This strategy has been most evident in proposals for the remediation of supposed language deficits in speakers of dialects other than the dialect found in primers. Therefore, instead of teaching reading of prose, one should attempt to increase general verbal fluency, vocabulary, and even general knowledge. The only special training necessary with regard to print would be to emphasize reading speed and to introduce a variety of literature and other reading materials.

The assumption of a single comprehension process has also led to the language experience approach of teaching decoding. In the now popular technique, the teacher writes out a story told orally by a child so that he can read it later. The implicit reasoning behind the value of this exercise is that the child will learn to read easier and faster if he discovers the similarity between processing speech and print. Since the material is already meaningful to him, the comprehension process used for oral language can be more efficiently and smoothly activated while reading. In addition, the material is more relevant to the child's here-and-now, which also may facilitate comprehension. One cannot help but note from the vantage of a dual comprehensive process view, however, that when the teacher writes down what the child says, he or she implicitly admits that the language found in the child's primers is not the same as what the child himself speaks, and further that the language the child speaks is better for him to practice reading. If so, then the comprehension processes might not be the same for both speech and print. Let us now consider the dual process view.

Dual Comprehension Processes

As one might surmise from the previous section, few researchers in reading have claimed that the comprehension of speech and the comprehension of print involve totally different processes. However, some have questioned the assumption of a unitary process. Referring to "a student's ability to comprehend spoken language," Bormuth (1972: 1135) claimed that "although reading and listening abilities share some elements in common, they exhibit substantial differences, and ... we can no longer use listening abilities to estimate reading aptitude in the simple fashion we heretofore thought possible."

Weaver and Kingston (1971b) discussed the difference between listening and reading from the child's perspective of learning to do each. They mentioned several differences between learning oral language and learning to read print. One point in particular is germane to the present discussion. "The transfer of 'meaning' from spoken to graphic representation may not be the direct, simple operation commonly assumed" (1971b: 8-185). "Meaning" includes more than definitions of isolated words, but also the meanings of sentences integrated with the meanings of other sentences, context, and the reader's prior experience. Thus, in a word, "transfer of 'meaning'" is comprehension. Good oral language skills are clearly necessary for learning to read. However, the issue is whether oral comprehension skills are sufficient for becoming a skilled reader when combined with well-practiced decoding skills.

In his "Reflections on the Conference" reported in *Language by Ear and by Eye*, Miller (1972) noted a historical development that tangentially bears on the point that comprehension is not unitary. Writing did not originate as a more permanent form of speech. Rather writing, as pictographs, evolved independently of speech as an alternate form of communication. This development of writing was considerably later in man's evolution than speech. Only recently in mankind's history, say 3000 years ago, was writing associated more directly with speech when an alphabet was invented. It is invalid to rely wholely on diachronic arguments to explain ontogenetic and sychronic processes, but the historical development of writing vis-à-vis speech is suggestive that the two language processes may not be identical.

A final argument that there may be at least two comprehension processes—one for listening and one for reading—comes from the notion that speech and print are sufficiently different so as to demand different processing strategies. It is often assumed that the way we talk is considerably different from the language of what is read. What would be a "sufficient" difference to necessitate different comprehension processes? In what respects is print different from speech? Aristotle claimed that "the style of written prose is not that of spoken oratory" (1928, p. 1413b). Has knowledge on this point advanced in the intervening 2300 years? Many would claim that there are large differences in grammar, vocabulary, and style in what we write and in what we say. To the extent that differences are present, there well may be concomitant differences in the cognitive processes required for efficient extraction of meaning. Material produced for one channel is comprehensible when received in the opposite one, but processing may be more strained than when compatibility of production and reception modality is maintained.

Now consider what one would do in the classroom if the dual process view of comprehension were correct. Most importantly, once the child is able to decode one would not stop teaching reading. The teacher would teach reading specific skills that would lead to more efficient and faster reading. Among these skills would be previewing, outlining, maximizing the amount of material per fixation, and minimizing regressions. The teacher could also teach those comprehension skills that are peculiar to print. For example, how to analyze the structural properties of prose sentences and paragraphs, or how to follow the styles of thesis development and argumentation would be appropriate. Knowing specifically how speech and print differ would be important in designing these curricula. Clearly, teachers must not ignore speech and oral language skills: they are much too critical to relegate to haphazard development. Based on any real differences uncovered between speech and print, appropriate instructional programs could be developed.

A Compromise

The positions drawn here are extreme ones. Those who claim dual comprehension processes certainly would not be so extreme as to say that the two processes are totally unrelated systems triggered by different modes of input. Likewise, it is unlikely that those who hold a single process view would say that there are no differences whatsoever. Some intermediate position is more probable.

The processing that the listener-reader must perform on the input may be heavily influenced by the purpose for which he/she is attending to the material. Typical listening situations are conversing over coffee or cocktails, and listening to

the news on TV. Common reading contexts are reading the newspaper for information and a novel for entertainment. To what extent theoretical and experimental considerations apply to these typical situations must be carefully evaluated. The issue is more than the usual plaint about the generalizability of laboratory findings to the "real world," however, because there are so many real world contexts to be accounted for. Is it reasonable to assume that a single comprehension process can be defined for the multifarious real world contexts? Is listening comprehension in a conversation the same process as listening comprehension of a lecture? Or does the latter process have more in common with reading a text on the same topic? What about reading a play? Don't we hear the players speaking to us from the stage? There well may be commonalities in listening and reading comprehension, as well as some differences, but comparison of the wide variety of tasks that are included under the labels "listening" and "reading" will eventually be necessary. In the long run, an absolute answer to the question of whether listening and reading comprehension are unitary or not may not be possible because there is as much or more variability within each skill than between them. Before evaluating this possibility, let us review the studies that bear on listening and reading. Is the evidence adequate to identify in sufficient detail points of similarity and difference in listening and reading comprehension so that one can specify the underlying processes?

The Evidence

Superficially, the experimental design for resolving the question of whether listening and reading comprehension are the same or different is simple. Present "identical" material aurally (as speech) to one group of subjects and visually (as print) to another, then test their comprehension. If equal amounts and types of material are comprehended, then listening and reading comprehension processes are likely to be the same (unitary), but if different amounts are comprehended then the processes must be different (dual). However, there are numerous logical and methodological pitfalls in such an

experiment. These problems can be categorized as relating to presentation conditions, subject characteristics, language materials, and comprehension measures. As the research relevant to each of these four major classes of variables is reviewed, pitfalls are identified and designs that might be used to circumvent those particular problems are suggested.

Presentation Conditions

The primary manipulation for the comparison of listening and reading comprehension, is, of course, whether the material is presented as speech or in print. Yet there are certain physical properties of the presentation media as well as gross differences in listening and reading decoding that must be considered to insure comparability of the two conditions. The presence of suprasegmentals in speech undoubtably facilitates the listener's veridical understanding of the message. Although punctuation marks may provide some of the same type of information to a reader, it is doubtful that they can code anywhere near as much information as suprasegmentals can. The set of punctuation marks is far smaller than the range of stress, intonation, pitch, and speed possible in speech.

Under typical reading conditions a large segment of the material is before the reader at any point in time, while during listening the listener has access only to a relatively short segment amounting to no more than a few words or phrases. Since it is difficult, if not impossible, to provide a listener access to a larger segment of material, an investigator must focus on restricting the amount of material available to a reader. This restriction should be effected without causing the reader to alter substantially the processes he normally uses in reading. A limited amount of material, say a line or two of text, could be presented for a relatively brief period (Young, 1973), either at a fixed rate or subject paced. Too fast or too slow a presentation rate would reduce comprehension because too much or too little information would be available to process. The size of the segment presented and how it is segmented would also be likely to influence the fluidity of reading. Since subjects integrate at the end of clausal boundaries (Bever,

Garrett, and Hurtig, 1973; Jarvella, 1971), maintaining clausal units would facilitate comprehension.

A second difference between listening and reading presentation modes is the amount of time that the functional stimulus is available to the subject. Because of differences in typical listening and reading rates, time is likely to differ markedly unless specifically controlled. Typical silent reading rates are two to three times faster than normal speaking rates (100 to 150 words per minute for speaking versus 300 to 400 words per minute for reading). Thus, readers given the same length of time to read a passage that listeners require to hear that same passage may have time to rehearse or reread the material. Though many studies have equalized listening and reading times (Corey, 1934; Durrell, 1969; Horowitz, 1968; Horowitz and Berkowitz, 1967; Sticht, 1968), they failed to equalize functional listening and reading times. Since comprehension was tested with immediate recall, they compared memory in conditions with functionally different study times.

One solution to this problem has been the use of compressed speech. The most common technique to compress speech without increasing the pitch is to excise very short segments (shorter than a single phoneme) on a random basis. Numerous studies have shown that speech compressed to 275 to 300 words per minute, close to typical reading rates, can be understood with minimal or no loss of comprehension (Foulke, 1968; Rossiter, 1971). Compression to higher rates results in a sharp drop in comprehension although comprehension can be maintained at rates over 400 words per minute with training (Orr, Friedman, and Williams, 1965). Even elementary school children are able to comprehend compressed speech at rates of 225 words per minute (Woodcock and Clark, 1968). Jester and Travers (1966) tested college students for comprehension at listening and reading rates of 150 to 350 words per minute with a multiple-choice test. At the lower speeds (about 200 words per minute) listening was generally superior to reading; but at the higher speeds (about 300 words per minute) reading was superior. Not only does compressed speech permit the equalization of study times for reading and listening, but also the fact that the maximally efficient input rate is the same for listening and reading suggests that there may be some commonality in processing.

As a practical matter, educators have frequently considered whether there is transfer of listening comprehension training to reading, particularly in elementary school children. If listening and reading comprehension are handled by a single process, then training in one should produce equivalent improvement in the other. But if there are two distinct processes, or processes that overlap only marginally, then the transfer should likewise be marginal. The frequently obtained moderate correlation between scores on standardized listening and reading tests (cf. Duker, 1965; Kennedy, 1971) does suggest the possibility that improvements in listening skills may facilitate reading acquisition.

Reviewing five reviews of the literature on whether there is transfer from listening training to reading, Kennedy (1971) found three early reviews that concluded there was positive transfer. However two later ones, as well as her own review, concluded that there were only inconsistent or no effects of listening training on coding skills. Most of the studies reviewed took a rather limited view of potential transfer variables. Many of the studies (e.g., Lewis, 1952) trained general listening skills such as determining the main idea, noting details, and drawing conclusions and inferences. They then tested for general transfer to reading achievement. Not surprisingly, the results did not show clear transfer effects.

On a specific level, positive transfer does occur. Both Orr (1966) and Thomas and Rossiter (1972) found that reading practice paced by listening to compressed speech increased normal reading rates up to 350 words per minute. Further, the increase in reading rates was not accompanied by any loss in comprehension and the effects on reading were still present nine months after the conclusion of only ten days of training (Thomas and Rossiter, 1972).

Kennedy (1971; Kennedy and Weener, 1973) tested whether training on listening or reading comprehension would improve comprehension in the other mode. Third graders who were below grade level readers were trained for five 20-minute sessions on either a written (reading) or oral (listening) cloze

procedure. They were then tested on the Durrell Listening-Reading Series and on new cloze comprehension tests in both written and oral forms. Training on either the written or oral cloze produced equivalent positive transfer on both the written and oral cloze comprehension tests. The effects on the Durrell test were more complicated, however. Relative to an oral reading control group that received individualized attention, training on the written cloze improved reading comprehension but not listening comprehension, and practice on the oral cloze procedure had no effect on either.

While the equivalent cross- and within-mode transfer from cloze training to cloze test was consistent with a single comprehension process, the pattern of transfer to the Durrell tests was not consistent with that conclusion. The prediction training implicit in a cloze procedure may be a relatively restricted language skill that depends as much on nonlinguistic knowledge as it does linguistic rules. Thus, there was specific transfer of the skill regardless of modality. Further, if reading but not listening comprehension involves the prediction skill trained by the cloze procedure, as does not seem unreasonable, then the transfer pattern to the more general Durrell test is expected. Whatever the explanation, these transfer results complicate any conclusion of a unitary comprehension process.

Subject Characteristics

The most obvious variation in subjects relevant to listening and reading is age. Chronological age is not itself the critical variable, but variation in age is concomitant with several acquisition variables that interact with listening and reading. Children usually begin to speak (and to listen) and to read (and to write) at quite different ages and the rates of acquisition of oral language and reading also may be different. In addition, the typical context in which acquisition occurs is the home for speaking and listening and the school for reading and writing. Because of these age-related differences in the temporal and social context of acquisition, comparison of listening and reading at a single chronological age may not be a fair comparison. The linguistic and conceptual complexity of the material that can be processed easily may differ between listening and reading when tested at a single age. The selection of a subject at a particular age or a decision to systematically vary age must be carefully evaluated.

Although subjects from age six (first graders) to early twenties (college students and military personnel) have been tested in one experiment or another, most studies comparing listening and reading comprehension have concentrated on a single age or at best a limited age range. Considering only a sample of studies, elementary school children were tested by Kennedy and Weener (1973), Swalm (1971), Oakan, Wiener, and Cromer (1971); high school students by Thomas and Rossiter (1972); college students by Singer (1970), Abrams (1966), Corey (1934), Horowitze (1968), and Horowitz and Berkowitz (1967); and U.S. Army recruits by Sticht (1968). Because of the multitude of different materials and procedures used in testing subjects among these studies, no developmental conclusions can be drawn from what is basically a collection of unrelated cross-sectional studies.

One study (Durrell, 1969; Durrell, Hayes, and Brassard, 1969) attempted to compare listening and reading comprehension across a relatively wide age range, albeit for the purpose of validating a test instrument. Tests of listening and reading abilities were constructed in three levels covering grades one through eight and were designed to test both vocabulary and either sentence or paragraph comprehension. Since at each level the same language comprehension tests were used on carefully equated forms, the raw scores were purported to be comparable. In a large scale administration of the test to three to four thousand children at each grade, listening vocabulary was much superior to reading vocabulary at grade one (by a factor of two), but these vocabulary scores were comparable by grade eight. Sentence-paragraph comprehension also started out with listening comprehension better than reading comprehension by a factor of two, but comparable scores were achieved by grade six, and at grade eight reading comprehension was 12 percent superior to listening comprehension.

Durrell's rationale for constructing the listening-reading test was that, assuming a unitary comprehension process,

listening abilities set an upper limit on a child's potential to read. Thus, listening comprehension should always be superior, or at least equal, to reading comprehension. The vocabulary results were consistent with this rationale although at grades higher than those tested reading vocabulary might even be superior to listening vocabulary. However, the sentence-paragraph comprehension results were clearly inconsistent with the rationale. Durrell retreated from his original position to claim that "when comparing reading and listening, the higher score of either indicates a potential for the other" (1969: 456).

Although these results appear to support dual comprehension processes, the Durrell test is too fraught with practical problems to draw any firm conclusions. For one example, although the total time was constant for both listening and reading, this procedure can result in marked differences in actual time available for processing. For another, the items for the comprehension test were selected for their psychometric properties rather than for their ability to tap the underlying processing. In addition, some reading was required as part of the listening comprehension test, resulting in a possible confounding of listening and reading skills. Thus, although the test scores may have statistical comparability, they are probably based on very different scales and should not be directly compared.

No one has attempted to separate general developmental effects from those associated with specific training in reading, except in terms of cognitive readiness to begin reading (Elkind, 1969, 1974). Does the difference in absolute level of reading achievement of, say first and third graders, reflect only schooling, or might not some of the difference be attributable to general maturation or cognitive development? To separate completely these variables would require carefully matched samples of literate and illiterate subjects at all ages, a practical impossibility and perhaps a theoretical impossibility as well. A less extreme strategy to separate these age-correlated changes is to compare good and poor readers. The procedure for obtaining a proper match between good and poor readers is not simple. Further, one should test not only good and poor readers of the same chronological age, but also good readers of a younger chronological age with the same "reading age" as the older poor readers. Unfortunately, investigators do not always take pains to form their groups such that a difference in reading performance is the only distinguishing characteristic. For example, Swalm (1971) tested children in grades 2, 3, and 4 for listening and reading comprehension by a cloze procedure. Reading comprehension was better than listening for above-average readers, but the reverse was true for below-average readers; listening and reading comprehension scores were equivalent for average readers. However, the classifications probably reflected more general cognitive differences as well as differences in reading performance.

In contrast to Swalm, Oakan et al. (1971) matched good and poor readers on both age and IQ. They found that good fifth-grade readers comprehended standard text equally well when they listened to tape recordings of either another good reader or a poor reader. But, when they read it themselves, the good readers comprehended the standard text better than when they read a transcript of the poor reader reading aloud. Thus, for good readers, degradation of the input effected by filtering it through a poor reader was more disruptive of reading than of listening. In contrast, poor readers' listening comprehension was disrupted by the poorly organized input from another poor reader, but the already poor reading comprehension was not improved by supplementary training on word identification. Perhaps, in accordance with a unitary view of comprehension, one of the reasons that poor readers do not read well is that their listening comprehension skills, which may be necessary for good reading comprehension, are not well established as evidenced by the ease with which they were disrupted.

Another approach to the involvement of listening processes in reading comprehension comes from a study of deaf readers. Conrad (1971) compared reading comprehension of deaf and hearing children following both silent reading and reading aloud. The deaf children were divided into those who used articulatory coding on a short-term memory pretest and those who did not use articulatory coding. While the hearing children and the deaf articulators comprehended equally well in both reading conditions, the deaf nonarticulators were

disrupted by being required to read aloud. Although all of the deaf children could speak, none of their reading could have been mediated by acoustic coding since all were profoundly deaf from an early age. While some of the deaf children (the articulators) used their speech mechanisms to mediate reading, others evidently did not because they were disrupted when required to read aloud. Thus, speech skills may be functional for normal readers, but are not necessary. The results of the normal hearing children, plus other data presented by Conrad (1972), suggest that the favored mode of coding for normal subjects is phonological even when penalized for such coding.

Language Materials

How are speech and print alike? How are they different? If listening and reading comprehension involve the same processes, then written language and speech ought not to be too different in vocabulary, grammar, and style; otherwise, a unitary processing strategy might not be able to handle the differences. On the other hand, to the extent that speech and print are linguistically different, then one might expect to find differences in the processing of each. Differences in speech and print may result from speaker-writer differences because speech is usually feedback sensitive, while print is not. Evidence one way or the other cannot be considered conclusive. However, differences may be suggestive inasmuch as the materials set the task for the listener-reader.

From the widely scattered studies that have compared oral and written productions, the results, for the most part, have not been surprising. Oral productions, compared to written ones on the same topic, contain more words, phrases, and sentences; longer and more difficult words; more verbs; more ideas; more elaboration; and more repetition (DeVito, 1965, 1966; Driemann, 1962a; Horowitz and Newman, 1964; Portnoy, 1973). Most studies have also found a greater diversity of vocabulary in written than in oral production, usually as measured by a larger type-token ratio. This finding may be artifactual, however, resulting from a curvilinear relation between number of types and number of tokens, since more words (tokens) are usually produced in speech. Direct comparisons of type-token ratios from different sized samples would be appropriate only with a linear relation between the number of types and the number of tokens (Driemann, 1962a).

Introspective reports have indicated that the purpose of a communication, the subject matter, and the characteristics of the recipient affected the composition of written productions more than they affected oral ones (Driemann, 1962b). Written productions are more precisely composed. As a consequence, when production time was limited, written messages were more efficient as measured in ideas per words per unit time (Horowitz and Newman, 1964). But given unlimited time, the faster production rate of speech yielded relatively greater efficiency than handwriting (Horowitz and Berkowitz, 1964). Skilled typing and stenotyping increased the similarity of manual productions to oral productions, but there still were large differences (Horowitz and Berkowitz, 1964).

One surprising result was uncovered by DeVito (1965) when he compared the oral and written productions of ten highly skilled speaker-writers, speech professors at a major university. He found that their oral description of topics from their published papers did *not* differ from the printed version in sentence length defined either in terms of letters, syllables, or words. Further, the written passages contained a significantly *higher* proportion of simple sentences (one subject-verb clause) than did oral productions. Whether this finding is generalizable to other classes of speakers and writers, or is peculiar to these highly skilled writers, needs to be explored further.

Do these differences in production mode make a difference to listeners and readers? DeVito (1965) presented the oral and written productions described above to college students in written form as a cloze test. The cloze scores were nearly identical: both types of passages were equally comprehensible to undergraduates. Portnoy (1973) also used a written cloze procedure to assess comprehensibility of oral and written productions. College students both spoke and wrote paragraphs on "interesting experiences." Confirming DeVito, Portnoy found no overall difference in comprehensibility between production modes. However, there were individual differences among the producers. She identified about half as

"speakers" and the other half as "writers." The speakers' oral productions were significantly more comprehensible than were their written productions, and the writers' written paragraphs were more comprehensible than their oral ones, both as measured by the cloze procedure. There were no overall differences between speakers and writers, however. Speakers' and writers' productions differed on only one linguistic measure—writers used more long words.

A complete design would involve testing comprehension of both production types in both reception modes as well as using additional measures of comprehension. To the extent that listening and reading comprehension processes are different and that speakers and writers are sensitive to such differences, oral productions should be easier for listening and written productions easier for reading. None of the studies that have compared listening and reading comprehension attempted to control or manipulate the origin of the stimulus materials in terms of mode of production. Investigators have completely ignored this issue and most have used written materials produced for reading. These materials are read aloud, invariably by a skilled reader, for aural presentation.

A few studies have used orally produced materials. Based on normative frequencies of occurrence in children's oral language (Strickland, 1962), Ruddell (1965) constructed prose passages using either high or low frequency syntactic patterns. Fourth graders from the same school district from which Strickland drew her sample were tested for reading comprehension with a written cloze procedure. Those paragraphs composed of high frequency syntactic patterns yielded more accurate comprehension than did paragraphs with low frequency patterns. These results support a notion that reading comprehension is related to the syntactic patterns used in speech. This conclusion was verified in a modified replication of Ruddell by Tatham (1970) that used more carefully controlled item selection and a different measure of comprehension.

These two studies can be considered as no more than suggestive, however, because only a few selected syntactic patterns from the multitude of possible linguistic variables were manipulated. In fact, Ackerman (1974) found that syntactic complexity of written productions correlated significantly with progress in learning to read for first through third graders, but that syntactic complexity in oral productions correlated with learning to read only for third graders. The ideal study would compare both listening and reading comprehension of two types of material. One type would contain structures that have a high frequency in speech but a low frequency in writing. The other type would use those structures having a low frequency in speech but a high frequency in writing. Ruddell and Tatham tested only reading comprehension and only for materials that occurred with a high frequency in speech (frequency in writing was undetermined). This design is a minimal one for sorting out production mode effects from reception mode effects.

Comprehension Measures

That the measurement of comprehension is no simple matter is suggested by the considerable discussion by educators and by experimental psychologists (Carroll, 1971, 1974b; Carroll and Freedle, 1972; Farr, 1969). An important distinction when considering various measures of comprehension is whether one desires to measure some general comprehension ability of the listener or reader, the comprehensibility of the text for listeners and readers in general, or how much and/or what is comprehended by a specific subject from a specific passage. The response measures that have been used to assess these different types of comprehension have varied considerably.

Those concerned with individual differences in listening and reading comprehension have usually adopted standardized tests and psychometric analyses. In many of the studies reviewed by Kennedy (1971), transfer of listening training to reading was assessed with standardized tests. These tests claim to measure an individual's listening and reading ability at that point in time. However, in selecting a standardized test, an investigator cannot assume that it measures what it purports to measure, but must evaluate the task demands of each subtest (Carroll, 1974a). In most of the transfer of training designs, the emphasis was on tests of general listening and

reading skill. This rationale frequently assumes that listening and reading comprehension are rather homogeneous skills, or even a single skill. With respect to reading, this has been a much debated question (Davis, 1971; Thorndike, 1973-74). M. L. Clark (1972) presents an excellent review of the literature relating to this issue. If listening and reading comprehension refer to a unitary process, then there should be a high correlation between standardized tests of each. Across 23 studies reviewed by Duker (1965) and 12 additional ones reviewed by Kennedy (1971), correlations between scores on listening and reading tests ranged from .45 to .82 and the distribution was not skewed in either direction. However, even viewed with caution given the variety of tests, procedures, and subjects, there still remains a substantial proportion of variance that is unaccounted for and which may reflect differences between listening and reading comprehension processes.

Weaver and Kingston (1971a) have used standardized test data in a somewhat different psychometric analysis. If listening and reading comprehension involve the same processes, then test scores of different aspects of both should fall together when compared. In two separate studies, Weaver and Kingston applied a construct discriminant-validity analysis (Campbell and Fiske, 1959) to several measures of listening and reading comprehension. The measures were paired for construct similarity across oral and written language forms. The major discriminant of the tests, however, was between oral and written forms rather than among common constructs. There is apparently more commonality within a comprehension mode than between supposedly common constructs.

Experimental studies have typically used a measure of comprehension that relates directly to the specific material. Immediate free recall has sometimes been used to measure comprehension (Horowitz, 1968; Horowitz and Berkowitz, 1967; Larsen and Feder, 1940), but one is faced with the difficulty of scoring the recall. Whether one scores the protocols for verbatim repetition or for recall of "meaning" has clear implications for how one conceives of the comprehension process. People do not usually listen or read to retain what they hear or see word-for-word, but instead assimilate the meaning of the material to their prior knowledge (Bartlett, 1932; Bransford and Franks, 1971; Sulin and Dooling, 1974). Hence, placing demands for verbatim recall on the subject may alter his processing of the material, or at the very least underestimate the amount comprehended. Another recall method is to probe the subject with pictures (Tatham, 1970) or with questions, frequently of the multiple-choice type (Corey, 1934; Jester and Travers, 1966; Oakan et al., 1971; Sticht, 1968; Young, 1973). One is then faced with the question of what aspects of the text to probe. A criticism of both of these methods is that they require some delay, however brief, before recall. Although comprehension may occur prior to storing the material in memory (even this may not be always the case, particularly with a demand for verbatim recall: Glucksberg and Danks, 1969), the results may reflect retrieval differences rather than comprehension differences.

What is needed is a measure of comprehension that is taken simultaneously with input, so that as the subject is processing the input his comprehension is more-or-less continuously monitored. Unfortunately, most of the "on-line" measures of comprehension measure only the difficulty or complexity of processing at that point in the text rather than how much or what is comprehended. One exception to this generalization is the cloze procedure as used, for example, by DeVito (1965), Portnoy (1973), and Kennedy (1971; Kennedy and Weener, 1973). The better a subject is comprehending a passage, the greater the probability that he will be able to supply the correct word. The major problem with a cloze task is that it disrupts the reader's (and the listener's) normal processing. If he must stop every five to ten words and guess a word, then he cannot be said to be reading naturally.

A final issue with respect to the response measure used is whether the response mode is the same or different from the presentation mode and/or the production mode. To the extent that listening and reading reflect a unitary process, then there should be no cross-modal interactions. That is, whether the input comes from listening or reading should not interact with whether the response is oral or written. Horowitz (1968) did find an interaction between presentation mode and response

mode. Free recall in the opposite mode reduced both correctness and amount recalled. However, in a later study (Horowitz and Berkowitz, 1967), the interaction was not nearly so clear. A complicating factor in this rationale is that it assumes some sort of isomorphism between comprehension and production. Differences may be the result of a failure of production to reflect comprehension rather than true modality effects; but finding an interaction is suggestive of differences between the two modalities.

Auding and Reading: A Model and A Review

Sticht and his associates (Sticht, Beck, Hauke, Kleiman, and James, 1974; Sticht and James, in press) have considered explicitly the question of the relationship between auding and reading and have proposed a developmental model of all languaging processes. Unlike other investigators, they did not bury the assumption of a unitary comprehension process, but incorporated it as a central feature of their model, and then reviewed evidence for and against it. Since they arrived at a different conclusion about the available evidence, I discuss their model and the support they review in detail. Only because they were so explicit about their position is this discussion possible:

> The developmental aspect of the model is represented by the successive specialization of basic adaptive processes through interaction with the environment. Listening and looking develop out of the basic adaptive processes of hearing and seeing; uttering and marking grow out of basic motor movements. Then the languaging processes emerge—auding and speaking grow out of listening and uttering respectively, and finally reading and writing develop from looking and marking. All of these changes develop from the child's attempts to adapt to, interact with, and control his/her environment.
>
> The languaging processes—auding, speaking, reading, and writing—work in short-term memory and are connected to a cognitive content store in long-term memory by a single, unitary conceptualizing process.
>
> The receptive languaging components [auding and reading] serve to transform verbal or printed displays into nonlanguage conceptualizations which constitute the meaning of the message to the receiver. The conceptualizing process continually merges input from the languaging process with information from the cognitive content store to build the subjectively experienced, meaningful message.... Auding and reading are considered to be similar processes because both require the use of language and languaging; and because, with identical messages, both result in the formation of a single, mutual internal conceptualization.... To state it concisely, auding and reading differ primarily in the manner in which the individual receives the stimulus words; they are both similar in the sense that they are both receptive communication acts that require a central language and conceptualizing base [Sticht et al., 1974: 17-18, 68, 70].

The model leads to two basic assertions about the development of languaging competency:

> (a) competence in languaging by auding precedes competency in languaging by reading; (b) *when acquired*, reading utilizes the same cognitive content and languaging competencies that are used in auding, plus the competencies involved in searching the visual display and, at least initially, decoding print to speech [Sticht et al., 1974: 71, italics in original].

Sticht et al. derived four hypotheses from the model and these two basic assertions. They evaluated these four hypotheses in terms of the existing literature on auding (listening) and reading and find support for all four. This support led them to a strong conclusion in favor of their developmental model, in particular the unitary nature of auding (listening) and reading comprehension. "The confirmation of each of the four hypotheses provides evidence for the developmental model of reading. Reading is based upon, and utilizes the same conceptual base and languaging competencies used in auding, plus the additional competencies used in converting the visual display into an internal auditory display" (Sticht et al., 1974: 115-116). Let us consider the four hypotheses and the evidence they review.

Hypothesis 1. "Performance on measures of ability to comprehend language by auding will surpass performance on

measures of ability to comprehend language by reading during the early years of schooling until the reading skill is learned, at which time ability to comprehend by auding and reading will become equal" (Sticht et al., 1974: 71). Sticht et al. found 44 studies relevant to this hypothesis, of which they were able to examine only 31. When these studies were divided by grade level, they yielded 71 comparisons of auding and reading. Although Sticht et al. note several potential methodological problems with these studies, all comparisons were retained and weighted equally. Treating each comparison as a separate data point, they found that the percentage of comparisons in which auding was better than reading declined from 100 percent at age 3½ to about 20 percent at college/adult. On the other hand, the percentage of comparisons in which reading was better than auding increased from about 3 percent at age 6 to about 45 percent at college/adult. This interaction supports the hypothesis.

One problem with this analysis, however, is that, although the hypothesis is derived from the model, there are few plausible models that would not predict exactly the same results simply by virtue of the fact that reading is typically acquired later than auding. Imagine a completely separate, dual process model of auding and reading. Auding is acquired during the preschool years; reading is taught *de novo* beginning in first grade. Assume there is no dependence between the two skills. If so, auding would test better than reading in the early years, but as the child begins to acquire reading, the scores on the reading test would pick up and perhaps eventually pass the scores on the auding test in some cases. Nothing in the evidence for this hypothesis says anything about the underlying processes.

Second, the "box-score" methodology gives equal weight to both good and poor experiments (Gardner, 1966). One should accept the conclusions of one well-designed, carefully executed study rather than the total weight of many experiments, each of which is flawed. It is possible that the various faults of each experiment might cancel one another and that the "average" results would thereby be unaffected. However, the reverse is more likely. Investigators may have committed many of the same errors.

Hypothesis 2. "Performance on measures of ability to comprehend language by auding will be predictive of performance on measures of ability to comprehend language by reading *after* the decoding skills of reading have been mastered" (Sticht et al., 1974: 71-72, italics in original). Two hypotheses are embedded here: one is that tests of auding and reading comprehension will show a substantial correlation in older children and adults and another hypothesis is that this correlation will be low in children who are in the process of acquiring decoding skills.

Although Sticht et al. discovered no "ideal" test of this hypothesis, studies by Loban (see references in Sticht et al., 1974) in a [illegible] are a close approximation. Loban compared children longitudinally from grades four through eight on an oral vocabulary test and teacher ratings of oral language skill with the results on standardized reading achievement tests. The contingency coefficients between oral language and reading achievement increased from .36 in grade four to .52 in grade eight. Sticht et al. also tabulated 27 additional studies that reported a total of 125 correlations between auding and reading tests at various grades one through twelve and college/adult. The mean correlation across studies rose from .35 in grade one to a fairly stable plateau of .60 by grade four. This plateau is probably an asymptote close to a ceiling established by the reliabilities of the tests. Both aspects of the hypothesis thus were supported.

Since there were many differences among the studies in terms of tests used, administration conditions, response tasks, timing, and format the correlations are not comparable as interval data and the computed means are not valid descriptors of the psychological relationships. In addition, this analysis diluted good studies by mixing them with faulty ones. Each study should be evaluated and considered on its own, not as aggregated with diverse experiments.

Hypothesis 3. "Performance of measures of rate of auding and rate of reading will show comparable maximal rates of languaging and conceptualizing for both processes, assuming fully developed reading decoding skills" (Sticht et al., 1974: 72). The maximal reading rate with relatively complete [illegible]tion is about the maximum auding (listening)

rate it is possible to attain with compressed speech. Although there are undoubtedly individual exceptions, about 300 words per minute seems to be a common limit when complete comprehension is maintained. Sticht et al. attributed this limit to the bottleneck due to languaging and conceptualizing the information. Since these processes are the same for both auding and reading, the maximum possible rate should be the same, and indeed it is.

The conceptual support the confirmation of this hypothesis adds to the model is limited, however. A bottleneck at only one point in a long and complicated process would be sufficient to establish an upper limit on rate. Processes before and/or after a bottleneck might be the same or different for auding and reading, yet a single bottleneck in common could produce these results. Assume, for example, that the process of integrating word meanings with syntactic structure is the same for both listening and reading, but that lexical access of the mental dictionary is different because of the different modes of input. If syntactic-semantic integration is slow relative to lexical access, then the difference between the two lexical access processes would not be evident in the limits on comprehension rates. Thus, the specific languaging components that lead to the common maximum rate must be specified for this finding to provide support for the model.

Hypothesis 4. "Training in comprehending by auding of a particular genre (e.g., 'listening for the main idea') will transfer to reading when that skill is acquired. Conversely, once reading skill is acquired, new cognitive content learned by reading will be accessible by auding" (Sticht et al., 1974: 72). Sticht et al. identified 12 valid studies that measured transfer of auding training to reading, but no studies that tested for the reverse transfer of training in reading skills to auding. Ten of these reported significant improvement in (transfer to) reading and two showed no transfer.

The transfer design can be an effective analytic tool because specific components of the auding process can be trained to determine which ones transfer to reading comprehension and which do not. Using this rationale, Sticht et al. identified those factors that led to positive transfer. They concluded that "generally speaking, the studies reporting significant transfer are characterized by a fairly high level of correspondence between auding training and reading transfer measures" (Sticht et al., 1974: 112). This conclusion does not provide strong support for the model, however, because it attributes the positive transfer results to specific transfer related to the particular test instrument rather than to facilitation of the reading comprehension process itself.

In conclusion, the developmental model of auding and reading proposed by Sticht et al. may be correct when adequate data have been collected. It is consonant with the assumptions of numerous investigators and is compatable with conclusions derived from cognitive psychology and experimental psycholinguistics (cf. Glucksberg and Danks, 1975). However, the evaluation of the model was "overly ambitious" (Clark, 1975: 691). Some of the hypotheses derived from the model are not unique to the model, but are compatible with other assumptions about the relationship between listening and reading.

A Task Analysis

What conclusions can be drawn from the evidence? Although the data are varied and uneven in quality, neither the unitary nor dual view is apparently correct in the extreme. The comprehension process is not so robustly unitary nor so decidedly dual that one position rings clear no matter what test or design is used. There simply is not enough evidence from experiments that have a clear rationale, adequate procedure, and decisive results to support a strong conclusion on either side of the issue. In most of the studies, the dependent measure of comprehension was how much was comprehended, recalled, etc. Even if exactly the same results were obtained in the listening and reading conditions, there is no assurance that the processes producing those identical results were the same. Likewise, it would be possible to have the same basic comprehension process for both listening and reading and yet find different results caused by interactions with superficial performance factors. Experiment after experiment, all of which yielded the same basic pattern of results, would tend to be confirmatory. The key strategy is to design experiments that permit the isolation of various components. The compre-

[illegible] or processes can then be modeled as a function [illegible] tasks used to isolate the components.

H[illegible] an investigator proceed in such an instance? One cannot simply look for overall differences between listening and reading presentation modes and conclude that these differences reflect different comprehension processes. The underlying comprehension processes are complex cognitive events that are not amenable to direct observation. The solution is to systematically vary factors that apparently influence the comprehension process in some non-trivial way. Then, if listening and reading input modes show the same pattern of results, regardless of overall differences, the components influenced by that factor can be supposed isomorphic for listening and reading. If, on the other hand, the manipulated variable interacts with presentation mode, then a point of difference between the listening and reading processes has been uncovered.

Listening and reading conditions are variables that set a different task before the subject, at least in the intial processing stages. The question addressed in this chapter is whether the different tasks involved in listening and reading extend into the comprehension phase. One way to approach the problem is to analyze the task demands imposed by the particular situational conditions. Precisely this strategy has been followed at the decoding level. When necessary differences in listening and reading decoding processes are claimed to result from differences in the availability of the stimulus (transient for speech, durable for print), evaluation of the demands placed on the subject by the task are what lead to that conclusion. The availability of print can be reduced to approximately the duration of the speech stimulus, but in doing so the reader's task has been changed. However, even by examining his performance under these atypical conditions, we may gain insight into the reading process and the importance of a durable stimulus. Neither listening nor reading is a fixed process that is activated the same way in each situation. Listening and reading represent orientations that set the subject to devise strategies for processing speech and print inputs in a meaningful way. The question of listening and reading comprehension boils down to whether the same comprehension strategies are devised for each input modality.

The implicit task for the subject is defined not only by input mode but also by other situational factors. These factors must be evaluated in terms of whether they induce different comprehension strategies for listening and reading. Let us consider several variables as examples and see how they might be manipulated in this type of analysis. It would not be feasible to vary all these factors in a single experiment. Such an attempt would result in too complex a design to actually test. However, the primary interest is in how these variables interact individually with listening and reading.

Besides the task orientation implicit in listening or reading, the instructions to the subject affect the kinds of comprehension strategies the subject engages in. If he expects to be tested with verbatim recall of a paragraph, he will more likely engage in rote rehearsal than if he expects to be tested only for memory for gist. This difference in expectancy may very well interact with listening and reading presentation modes because the reading task is more amenable to a rote rehearsal strategy since print is continuously available. Unfortunately, many studies do not instruct the subject about what his comprehension test will be. It is left to his imagination or prior experience in such testing situations. Lacking instructions, subjects will formulate their own criterion of what information is important, perhaps on the basis of prior questions (Rothkopf, 1974; Rothkopf and Bisbicos, 1967), thus leading to large individual variation in comprehension strategies. The kinds of processing strategies induced by the particular comprehension tests should be analyzed so that the results can be interpreted in terms of the task demands.

Just as the expectations the subject has about what he will have to do with the information influences the processing strategy, so various aspects of the presentation itself also may affect his strategy. For example, if one varies the input rate, functions from slow to fast listening rates can be compared to functions from slow to fast reading rates. However, for a given

objective rate, the typical reader will have more functional processing time than will a listener. Relative rate may be more critical than absolute rate. If so, a comparison of the shapes of the curves will be more informative than comparisons at strictly equivalent objective rates. Even then, rate may interact with listening and reading. At very fast rates it may be easier for a reader to develop a relatively efficient skimming strategy, searching for the most useful information for constructing an integration of the passage. A listener, of course, does not have such an option available.

The other two classes of factors—language materials and subject characteristics—may also lead to variation in processing strategies. As an example of the former, listening and reading conditions may yield equal comprehension with familiar vocabulary, simple syntactic structure, and a linear conceptual organization of a paragraph. However, reading may have a distinct advantage when the vocabulary is unusual, the syntax complex, and the conceptual organization is hierarchical. These differences in language materials may lead listeners and readers to adopt different comprehension strategies. When reading a novel versus a critical commentary on that same novel, the processing strategies may differ. The same may hold for listening to a play versus an academic lecture on drama. Do the differences in processing strategy resulting from linguistic variation interact with listening and reading input modes? In some situations, there may be more commonality in comprehension strategies between listening and reading than there is within either listening or reading in other situations.

Finally, all of these factors may interact with age. Of all the factors considered here, age is probably the most critical outside the laboratory and it certainly is paramount in the schools. In a task analysis, there are two levels at which age-related factors can be influential. One is basic acquisition. Has the child acquired the information processing skills necessary to handle the information he is receiving? Obviously, a kindergartner who can only recognize letters cannot read. Likewise, a second-grade child who has not learned some of the more complex syntactic constructions (Chomsky, 1969, 1974; Palermo and Molfese, 1972) will not be able to comprehend them either by listening or by reading. At another level, however, a younger child may not have had sufficient experience devising strategies for efficient information processing. Young children may not know how or may not realize that it is important to devise processing strategies to meet the demands implicit in the task set before them. Since the other variables mentioned—listening/reading, comprehension tests, presentation rate, linguistic and conceptual complexity—define the task for the subject, skill in interpreting these demands and recognizing one's competencies and the available strategies may be major accomplishments for the child. (See Kreutzer, Leonard, and Flavell, 1975, for an examination of children's awareness of memorial strategies.)

The point of these examples is that it is possible to analyze task demands and to tap the underlying processes the listener/reader devises. If there are separate listening and reading comprehension processes, then interactions will result in different patterns of results. If there are no interactions, then perhaps listening and reading comprehension processes are the same. The unitary view does have a special problem, inasmuch as it is basically a null hypothesis. If the comprehension process were unitary, the best research strategy would be to show that any differences obtained between listening and reading were due entirely to decoding differences.

In conclusion, to draw this chapter full circle, a task and process analysis of listening and reading comprehension can provide a more precise notion of what should be attempted in the classroom with respect to teaching reading comprehension. If there are no fundamental differences between listening and reading, then a general language experience approach including both listening and reading would be sufficient. If, however, differences between listening and reading appear in one or another comprehension task, then training on that task would be appropriate. If there are vocabulary differences, then practice on vocabulary that is typically used in reading should be implemented. If there are syntax differences, then systematic introduction to the differing structures could be developed. If there are differences in the organization of ideas, then how written paragraphs are structured in contrast to conversation

or speeches should be added to the curriculum. In any case, suggestions more specific than such global conclusions as "one must teach reading comprehension" or "training in listening facilitates acquisition of reading" could be advanced. If one understands the underlying processes, then applied proposals can be not only more accurate, but also more specific. The research suggested here only answers "whether" and "how" listening and reading comprehension are related, but will not provide an answer as to the best program to implement that information. Additional experiments on how to transfer the knowledge gained through basic research to the classroom still will be needed.

Postscript

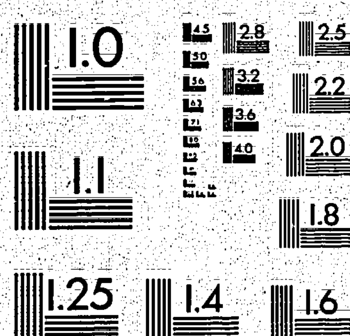
Several recent studies have raised the main question addressed in this chapter—whether the comprehension processes involved in listening and reading are the same or different—from the level of an assumption that is accepted as true to a hypothesis that can be tested empirically. These studies approach the question in a variety of ways and bear on a number of issues raised in this paper. However, with one exception to follow, these studies are not discussed, but are listed for the interested reader: Carver (1976), Cocking and Potts (1976), Goldman (1976), Neville and Pugh (1974, 1976-77), Perfetti and Goldman (1976), Sticht (in press), and Walker (1975-76).

An experiment reported by Mosenthal (1976-77) used the task analytic rationale for comparing oral and silent reading comprehension with listening comprehension. Mosenthal first assumed that H.H. Clark's (1969a, 1969b) theory of linguistic comprehension is essentially correct. Clark postulated three principles that operate in sentence verification tasks and which have been confirmed in numerous experiments by Clark and others (H.H. Clark, 1974). Mosenthal reasoned that testing whether or not Clark's principles operated identically in oral and silent reading as they did in listening would provide an adequate comparison of listening and reading comprehension processes. A straightforward statistical comparison is not strictly necessary although Mosenthal was able to perform such a comparison. Since Clark's principles describe components of the underlying comprehension process, confirmation of the principles and their identical manner of operation in listening and reading tasks would permit the conclusion that listening and reading comprehenion processes are identical. If the principles were not confirmed in both tasks, then the processing component described by that principle could be said to be different.

Using Clark's basic syllogistic reasoning task, Mosenthal tested second and sixth grade children for the operation of Clark's principles. In one experiment he compared oral reading with listening and in two additional experiments he compared silent reading with listening. In the first experiment, there was a significant decrement in the overall level of performance from listening to oral reading, but the pattern of results was the same for both modes. In the last two experiments, not only was the pattern of results the same, but there was no difference in the level of performance between listening and silent reading. Mosenthal concluded that silent reading and listening involved the same comprehension processes but that oral reading comprehension was different. Since Clark's principles were confirmed in general in the oral reading task, just as in the listening task, the proper conclusion is that the comprehension process is the same for oral reading as well.

Because Mosenthal used a task analytic rationale, the results of his experiments provide a valid comparison of the comprehension processes underlying listening and reading. One may reject the particular model of comprehension adopted by Mosenthal, but one cannot object to the rationale for the comparison. Based on the research reviewed here, it is unlikely that a direct comparison of listening and reading comprehension is possible. But an indirect comparison is possible as has been demonstrated by Mosenthal.

Acknowledgments. This chapter was written during the summer of 1974 and was revised the following summer. The postscript was added in the summer of 1977. I thank Rod Barron, John Carroll, Karen Cohen, Linnea Ehri, Karen Fischer, Sam Glucksberg, Kathy Pezdek, Betsy Ruddy, Ellen Ryan, Sandra Smiley, Tom Sticht, Tom Trabasso, and Dale Willows for their valuable comments on earlier drafts of this chapter. Preparation of this paper was


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

supported in part by Grant No. MH-21230 from the National Institute of Mental Health and by Grant No. NIE-78-0223 from the National Institute of Education.

References

ABRAMS, A.G. The relations of listening and reading comprehension to skill in message structuralization. *Journal of Communication*, 1966, *16*, 116-125.

ACKERMAN, M.D. Productive and receptive language of first-, second-, and third-grade children: Speaking, writing, and reading. American Psychological Association, New Orleans, 1974.

ARISTOTLE. Rhetorica. Translated by W.R. Roberts. In W.D. Ross (Ed.), *The works of Aristotle*, Vol. 11. London: Oxford University Press, 1928.

BARTLETT, F.C. *Remembering*. Cambridge: Cambridge University Press, 1932.

BEVER, T.G., GARRETT, M.F., and HURTIG, R. The interaction of perceptual processes and ambiguous sentences. *Memory and Cognition*, 1973, *1*, 277-286.

BORMUTH, J.R. Review of "Durrell Listening-Reading Series." In O.K. Buros (Ed.), *The seventh mental measurement yearbook*. Highland Park, N.J.: Gryphon Press, 1972.

BRANSFORD, J.D., and FRANKS, J.J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

CAMPBELL, D., and FISKE, D. Convergent and discriminate validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

CARROLL, J.B. Learning from verbal discourse in educational media: A review of the literature. ETS *Research Bulletin*, RB-71-61, 1971.

CARROLL, J.B. Psychometric tests as cognitive tasks: A new "Structure of Intellect." ONR Technical Report No. 4, Contract N00014-71-C-0117 NR 150 329, 1974. (a)

CARROLL, J.B. Research in language comprehension. Interdisciplinary Institute in Reading and Child Development, University of Delaware, 1974. (b)

CARROLL, J.B., and FREEDLE, R.O. (Eds.) *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.

CARVER, R.P. Optimal rate of reading. Unpublished paper, University of Missouri-Kansas City, 1976.

CHOMSKY, C. *The acquisition of syntax in children from 5 to 10*. Cambridge, MA: MIT Press, 1969.

CHOMSKY, C. (Presentation.) Interdisciplinary Institute in Reading and Child Development, University of Delaware, 1974.

CLARK, H.H. Influence of language on solving three-term series problems. *Journal of Experimental Psychology*, 1969, *82*, 205-215. (a)

CLARK, H.H. Linguistic processes in deductive reasoning. *Psychological Review*, 1969, *76*, 387-404. (b)

CLARK, H.H. Semantics and comprehension. In T.A. Sebeok (Ed.), *Current trends in linguistics, Vol. 12: Linguistics and adjacent arts and sciences*. The Hague: Mouton, 1974.

CLARK, H.H. Review of *auding and reading: A developmental model* by T.G. Sticht, L.J. Beck, R.N. Hauke, G.M. Kleiman, and J.H. James. *Contemporary Psychology*, 1975, *20*, 691.

CLARK, M.L. *Hierarchical structure of comprehension skills, Vol. I: Theoretical models and related research on the definition and measurement of reading and listening comprehension*. Hawthorn, Victoria, Australia: Australian Council for Educational Research, 1972.

COCKING, R.R., and POTTS, M. Psycholinguistics and the prediction of reading achievement. *Educational Testing Service Research Bulletin*, RB-76-20, 1976.

CONRAD, R. The effect of vocalizing on comprehension in the profoundly deaf. *British Journal of Psychology*, 1971, *62*, 147-150.

CONRAD, R. Speech and reading. In J.F. Kavanagh and I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge: MIT, 1972.

COREY, S.M. Learning from lectures and learning from reading. *Journal of Educational Psychology*, 1934, *25*, 459-470.

DAVIS, F.B. Psychometric research on comprehension in reading. In F.B. Davis (Ed.), *The literature of research in reading with emphasis on models*. New Brunswick, N.J.: Rutgers—The State University, 1971.

DeVITO, J.A. Comprehension factors in oral and written discourse of skilled communicators. *Speech Monographs*, 1965, *32*, 124-128.

DeVITO, J.A. Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs*, 1966, *33*, 73-76.

DRIEMANN, G.H.J. Differences between written and spoken language. *Acta Psychologica*, 1962, *20*, 36-57. (a)

DRIEMANN, G.H.J. Differences between written and spoken language. *Acta Psychologica*, 1962, *20*, 78-100. (b)

DUKER, S. Listening and reading. *Elementary School Journal*, 1965, *65*, 321-329.

DURRELL, D.D. Listening comprehension versus reading comprehension. *Journal of Reading*, 1969, *12*, 455-460.

DURRELL, D.D., HAYES, M.T., and BRASSARD, M.B. *Listening-reading tests*. New York: Harcourt, Brace, and World, 1969.

ELKIND, D. Reading, logic, and perception. *Educational Therapy*, 1969, *2*, 195-207. Reprinted in D. Elkind, *Children and adolescents*. New York: Oxford, 1970.

ELKIND, D. Cognitive development and reading. Interdisciplinary Institute in Reading and Child Development, University of Delaware, 1974.

FARR, R. *Reading: What can be measured?* Newark, Del.: International Reading Association, 1969.

FOULKE, E. Listening comprehension as a function of word rate. *Journal of Communication*, 1968, *18*, 198-206.

FRIES, C.C. *Linguistics and reading*. New York: Holt, Rinehart, and Winston, 1963.

GARDNER, R.A. On box score methodology as illustrated by three reviews of overtraining reversal effects. *Psychological Bulletin*, 1966, *66*, 416-418.

GIBSON, E.J. Reading for some purpose: Keynote address. In J.F. Kavanagh and I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge: MIT, 1972.

GLUCKSBERG, S., and DANKS, J.H. Grammatical structure and recall: A function of the space in immediate memory or of the retention interval? *Perception and Psychophysics*, 1969, *6*, 113-117.

GLUCKSBERG, S., and DANKS, J.H. *Experimental psycholinguistics: An Introduction*. Hillsdale, N.J.: LEA, 1975.

GOLDMAN, S.R. Reading skill and the minimum distance principle: A comparison of listening and reading comprehension. *Journal of Experimental Child Psychology*, 1976, *22*, 123-142.

GOODMAN, K.S. A psycholinguistic view of reading comprehension. In G.B. Schick and M.M. May (Eds.), *The fifteenth yearbook of the National Reading Conference*. Milwaukee: National Reading Conference, 1966.

HOROWITZ, M.W. Organizational processes underlying differences between listening and reading as a function of complexity of material. *Journal of Communication*, 1968, *18*, 37-46.

HOROWITZ, M.W., and BERKOWITZ, A. Structural advantage of the mechanism of spoken expression as a factor in differences in spoken and written expression. *Perceptual and Motor Skills*, 1964, *19*, 619-625.

HOROWITZ, M.W., and BERKOWITZ, A. Listening and reading, speaking and writing: An experimental investigation of differential acquisition and reproduction of memory. *Perceptual and Motor Skills*, 1967, *24*, 207-215.

HOROWITZ, M.W. and NEWMAN, J.B. Spoken and written expression: An experimental analysis. *Journal of Abnormal and Social Psychology*, 1964, *68*, 640-647.

JARVELLA, R.J. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 409-416.

JESTER, R.E., and TRAVERS, R.M.W. Comprehension of connected meaningful discourse as a function of rate and mode of presentation. *Journal of Educational Research*, 1966, *59*, 297-302.

KAVANAGH, J.F. (Ed.) *Communicating by language: The reading process*. Bethesda, Md.: NICHHD, 1968.

KAVANAGH, J.F., and MATTINGLY, I.G., (Eds.) *Language by ear and by eye*. Cambridge: MIT, 1972.

KENNEDY, D.K. Training with the cloze procedure, visually and auditorially, to improve the reading and listening comprehension of third grade underachieving readers. Unpublished doctoral dissertation, Pennsylvania State University, 1971.

KENNEDY, D.K., and WEENER, P. Visual and auditory training with the cloze procedure to improve reading and listening comprehension. *Reading Research Quarterly*, 1973, *8*, 524-541.

KREUTZER, M.A., LEONARD, C., and FLAVELL, J.H. An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development*, 1975, *40* (Serial No. 159).

LARSEN, R.P., and FEDER, D.D. Common and differential factors in reading and hearing comprehension. *Journal of Educational Psychology*, 1940, *31*, 241-252.

LEWIS, M.S. The effect of training in listening for certain purposes. *Journal of Communication*, 1952, *2*, 81-86.

MILLER, G.A. Reflections on the conference. In J.F. Kavanagh and I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge: MIT, 1972.

MOSENTHAL, P. Psycholinguistic properties of aural and visual comprehension as determined by children's abilities to comprehend syllogisms. *Reading Research Quarterly*, 1976-77, *12*, 55-92.

NEVILLE, M.H., and PUGH, A.K. Context in reading and listening: A comparison of children's error in cloze tests. *British Journal of Educational Psychology*, 1974, *44*, 224-232.

NEVILLE, M.H., and PUGH, A.K. Context in reading and listening: Variations in approach to cloze tasks. *Reading Research Quarterly*, 1976-77, *12*, 13-31.

OAKEN, R., WIENER, M., and CROMER, W. Identification, organization, and reading comprehension for good and poor readers. *Journal of Educational Psychology*, 1971, *62*, 71-78.

ORR, D.B. Recent research on reading and the comprehension of time-compressed speech. *College Reading Association: Proceedings*, 1966, *7*, 79-84.

ORR, D.B., FRIEDMANN, H.L., and WILLIAMS, J.C.C. Trainability of listening comprehension of speeded discourse. *Journal of Educational Psychology*, 1965, *56*, 148-156.

PALERMO, D.S., and MOLFESE, D.L. Language acquisition from age five onward. *Psychological Bulletin*, 1972, *78*, 409-428.

PERFETTI, C.A., and GOLDMAN, S.R. Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 33-42.

PORTNOY, S. A comparison of oral and written verbal behavior. In K. Salzinger and R.S. Feldman (Eds.), *Studies in verbal behavior*. New York: Pergamon, 1973.

ROSSITER, C.M. Rate-of-presentation effects on recall of facts and of ideas and on generation of inferences. *AV Communication Review*, 1971, *19*, 313-324.

ROTHKOPF, E.Z. Limitations in capacity to learn from written materials. Interdisciplinary Institute in Reading and Child Development, University of Delaware, 1974.

ROTHKOPF, E.Z., and BISBICOS, E. Selective facilitation effects of interspersed questions on learning from written material. *Journal of Educational Psychology*, 1967, *58*, 56-61.

RUDDELL, R.B. The effect of the similarity of oral and written patterns of language structure on reading comprehension. *Elementary English*, 1965, *42*, 403-410.

RYAN, E., and SEMMEL, M. Reading as a constructive language process. *Reading Research Quarterly*, 1969, *5*, 59-83.

SINGER, C.R. Eye movements, reading comprehension, and vocabulary as effected by varying visual and auditory modalities in college students. *Dissertation Abstracts International*, 1970, A31, 164.

STICHT, T.G. Some relationships of mental aptitude, reading ability, and listening ability using normal and time-compressed speech. *Journal of Communication*, 1968, *18*, 243-258.

STICHT, T.G. A model of auding-reading development with applications to beginning reading instruction and evaluation. In L.B. Resnick and P.A. Weaver (Eds.), *Theory and practice of early reading*, Vol. I. Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

STICHT, T.G., BECK, L.J., HAUKE, R.N., KLEIMAN, G.M., and JAMES, J.H. *Auding and reading: A developmental model*. Alexandria, VA HumRRO, 1974.

STICHT, T.G., and JAMES, J.H. Listening to language and looking at language. In S. Wanat, H. Singer, and M. Kling (Eds.), *Extracting meaning from written language*.

STRICKLAND, R.G. The language of elementary school children: Its relationship to the language of reading textbooks and the quality of reading of selected children. *Bulletin of the School of Education*, Indiana University, 1962, *38*.

SULIN, R.A., and DOOLING, D.J. Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology*, 1974, *103*, 255-262.

SWALM, J.E. Comparison of oral reading, silent reading, and listening comprehension assessed by cloze. *Dissertation Abstracts International*, 1971, A32, 3578-3579.

TATHAM, S.M. Reading comprehension of materials written with select oral language patterns: A study at grades two and four. *Reading Research Quarterly*, 1970, *5*, 402-426.

THOMAS, K.H., and ROSSITER, C.M. The effects of reading practice with compressed speech on reading rate and listening comprehension. *AV Communication Review*, 1972, *20*, 35-43.

THORNDIKE, R.L. "Reading as reasoning." *Reading Research Quarterly*, 1973-4, *9*, 135-147.

WALKER, L. Comprehending writing and spontaneous speech. *Reading Research Quarterly*, 1975-76, *11*, 144-167.

WANAT, S.F. Linguistic structure in reading: Models from the research of Project Literacy. In F.B. Davis (Ed.), *The literature of research in reading with emphasis on models*. New Brunswick, N.J.: Rutgers—The State University, 1971.

WEAVER, W.W., and KINGSTON, A.J. Assessment of oral-written language behaviors with the Campbell-Fiske analysis procedure. Paper presented at American Educational Research Association, New York, 1971. Reported and cited in Weaver and Kingston (1971b). (a)

WEAVER, W.W., and KINGSTON, A.J. Modeling the effects of oral language upon reading language. In F.B. Davis (Ed.), *The literature of research in reading with emphasis on models*. New Brunswick, N.J.: Rutgers—The State University, 1971. (b)

WOODCOCK, R.W., and CLARK, C.R. Comprehension of a narrative passage by elementary school children as a function of listening rate, retention period, and I.Q. *Journal of Communication*, 1968, *18*, 259-271.

YOUNG, R.Q. A comparison of reading and listening comprehension with rate of presentation controlled. *AV Communication Review*, 1973, *21*, 327-336.

6 An Interactive Analysis of Oral Reading

In A. M. Lesgold & C. A Perfetti (Eds.), Interactive Processes in Reading. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981. Pp. 131-153.

Joseph H. Danks and Gregory O. Hill
Kent State University

Models of reading and language comprehension have come to be dominated by the distinction between bottom-up and top-down sources of information and processing control (Danks & Glucksberg, 1980). Until recently, most information-processing models of reading have emphasized bottom-up processing (e.g., Gough, 1972; LaBerge & Samuels, 1974; Massaro, 1975). Such models attempt to formalize the intuitive observation that reading comprehension begins with the perception of print and ends with the construction of an abstract meaning representation. However, strictly bottom-up processing models have difficulty explaining some demonstrable aspects of reading performance. The crux of the difficulty is that readers and listeners frequently can anticipate parts of a linguistic message before bottom-up processing is completed. They habitually employ that ability to facilitate both perception and comprehension.

Evidence of anticipatory processing is found in two general classes of phenomena. One class results from the vagaries, ambiguities, and imperfections of typical human communication. The ease with which readers compensate for typographical errors, often without even noticing them, is an example of how the ability to anticipate an input provides a degree of insensitivity to minor linguistic violations (cf. Marslen-Wilson & Welsh [1978] for a similar effect in speech shadowing).

The other class of phenomena that poses difficulties for strictly bottom-up processing models is the facilitating effects of contextual information. Numerous demonstrations have shown that speed and accuracy of processing at one level are dependent on information from a more abstract level of representation (Rumelhart, 1977; Wildman & Kling, 1978-1979). Letter recognition is facilitated by lexical information; lexical access (word recognition) is faster when

syntactic, semantic, and factual information from sentences and paragraphs is available; semantic interpretation is aided by contextual information about the topic or theme; and paragraph comprehension is aided by general world knowledge. One or two of these contextual effects might be handled plausibly by modifications in a basically bottom-up model—for example, by the addition of feedback loops or by reordering processing stages. But the pervasiveness of top-down effects throughout all levels of representation indicates that bottom-up processing is only part of a more complicated picture.

There have been few purely top-down models proposed as alternatives to bottom-up models, and with good reason. Bottom-up effects can be readily demonstrated. Readers do not create the meaning of what they are reading wholly from prior knowledge. As a general principle, there is a greater reliance on less abstract information and bottom-up processing with difficult or unfamiliar material. However, if the text is not too difficult or unfamiliar, top-down information can facilitate processing greatly. The key to the benefits of top-down processing lies in the tremendous redundancy at all levels of analysis. Although redundancy tends to bog down any attempt at comprehensive bottom-up processing of all the information in the signal, it opens the door to a more efficient use of lower-level information.

The most realistic alternatives to bottom-up models are interactive models that envision both bottom-up and top-down directions of information flow (Danks, 1978). However, because of the traditional dominance of bottom-up models, interactive models have concentrated on describing the influence of contextual information. In order to demonstrate clearly the interaction between bottom-up and top-down processing, such models have focused on intermediate levels of processing where both types of processing have an opportunity to function. The most commonly chosen level is lexical access or word recognition. Visual-perceptual information and syntactic, semantic, and prior-knowledge information converge in lexical access. So it is a convenient point at which to demonstrate both bottom-up and top-down effects.

Selection of lexical access as the focus of study may not be the most felicitous choice, however, in that the interactions found there may not generalize to processing at other levels. There seems to be a fundamental change in the nature of comprehension at levels more abstract than lexical access. Words and letters can be thought of as self-contained units possessing a relatively enduring identity. It is meaningful to talk about word and letter "re-cognition" processes because we encounter those same units repeatedly. Beyond the level of words, however, a somewhat different picture emerges. Sentences and paragraphs do not have the relatively constant representations that we attribute to letters and words. Since almost every sentence that readers encounter is unique in their experiences, it does not make sense to talk about a sentence "re-cognition" process. The correspondence between the physical input and the mental representation is much less evident above the level of individual words.

STUDIES OF ORAL READING

Oral reading is a task used frequently in schools for teaching and evaluating reading skills (Durkin, 1978–1979), but we understand very little of the processing requirements of oral reading or how it relates to silent reading. The predominant requirement of oral reading is that each word be recognized and verbalized in serial order. In most oral reading, processing probably is carried beyond simple word identification, but higher-level comprehension is not generally the central focus. In contrast, silent reading does not require the serial recognition of every word and typically focuses on overall comprehension.

When oral reading is followed by a comprehension test—for example, recall, recognition, summarization, or question answering—there are two major tasks confronting the reader (Danks & Fears, 1979). One is to verbalize each word in succession. To accomplish this task, the reader must explicitly recognize each word. By this we presume that he or she locates the word in the mental lexicon and then uses the articulation information represented there to pronounce it. It is possible that a reader might use only lower-level information, such as grapheme-phoneme correspondences, spelling patterns, or syllabic structure, to pronounce each word without accessing it in the mental lexicon. However, since readers do not "read" pronounceable nonwords without hesitation, complete dependence on lower-level information is unlikely.

The oral reader's secondary task is to understand phrases, sentences, paragraphs, and themes of the text so as to be able to satisfy the demands of the comprehension test. The reader usually attempts to satisfy both the verbal performance and the comprehension tasks at the same time, although the press of the performance demand is greater. Analysis of oral performance provides an excellent opportunity to study the interaction of lexical access and text comprehension processes in a relatively natural reading situation.

Oral reading has the advantage of yielding a continuous on-line response that is roughly contemporaneous with the visual input. It is the reading counterpart to shadowing speech (cf. Lindig, 1976; Marslen-Wilson, 1975). In shadowing, the listener is presented with continuous speech and must produce the corresponding verbalization immediately while continuing to receive new input. In oral reading, the reader produces an oral rendition of the printed text shortly after the intake of the visual information. The main difference is that in oral reading the reader controls the intake of information, but in shadowing the listener does not. This difference corresponds to the general difference between listening and reading.

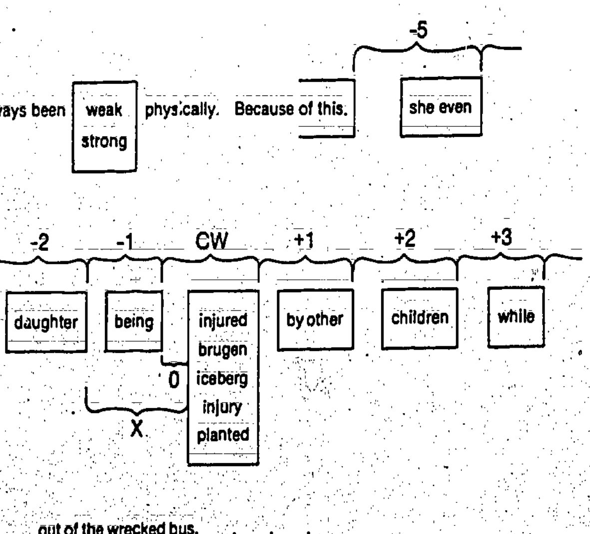
Rationale

In the experiments described in this chapter, we assessed how different types of textual information were used in oral reading by selectively violating or remov-

n from the text. If a particular type of information was used in r the oral production or the comprehension demand, there was a he oral performance at the moment when the altered information The basic procedure was to alter several critical words in a long g one or more types of textual information at each location. We readers' oral productions for disruptions at or around each critical

eral experiments, using readers at different levels of reading skill rade), we manipulated lexical, syntactic, semantic, and factual illustrated in a segment of text presented in Fig. 6.1. The story is hool girl who is severely injured when a train hits her school bus. In her mother has just heard about the accident and is worried about

The critical word (*injured* in the example) was replaced with a nonword (*brugen*). The nonword followed the rules of English tructure and was readily pronounceable. If the reader were relying pheme–phoneme correspondences to pronounce the word, there isruption of oral performance. But if the reader were using lexical vord recognition would be impossible, because the nonword could n the reader's mental lexicon. The lexical violation would disrupt



A portion of a story illustrating the different types of textual violations.

comprehension as well, because there was no syntactic or semantic information associated with the nonword.

2. *Syntactic + Semantic*. Both syntactic and semantic information was distorted by replacing the critical word with a word that was the incorrect part of speech and that was semantically anomalous as well (*iceberg* in Fig. 6.1). Here the word could be accessed in the mental lexicon, but it was inappropriate for the syntactic and semantic context of the sentence.

3. *Syntactic*. Syntactic information alone was distorted by retaining the root morpheme of the critical word but changing the inflection such that it indicated a part of speech that could not occur at that point in the sentence. In the example in Fig. 6.1, *injured*, a verb, was changed to *injury*, a noun. Although some semantic information is carried in the syntactic categories, most of the semantic information remains in the root. In this case, the reader could determine the meaning of the text relatively easily.

4. *Semantic*. To violate semantic information but not disturb lexical or syntactic information, the critical word was replaced with a word that was the correct part of speech but that was semantically anomalous; for example, *planted* replaced *injured*. Although the readers could determine grammatical structure, they had to concoct an implausible meaning. The best they could do was to imagine a very unusual circumstance in which the anomalous word could be interpreted metaphorically.

5. *Factual*. Factual information is what the reader accumulates from the proceeding text while reading a story. While reading, he or she constructs a representation of what the writer is conveying. New information is added to that representation. We violated factual information by introducing an inconsistency between the critical word and the preceding sentence. Unlike the other manipulations, neither the critical word nor the sentence containing it was altered. The sentence immediately before the sentence with the critical word was altered such that the critical word was factually inconsistent with the sense of the altered sentence. In the example in Fig. 6.1, the word *weak* in the preceding sentence was replaced with *strong*. The fact that her daughter was strong was inconsistent with the mother worrying about her being injured. There was nothing syntactically or semantically wrong with either sentence. They simply communicated inconsistent information.

All modifications were selected to assure that the readers would be unlikely to conceive of a continuation after the critical word that would eliminate the violation. For the factual manipulation, there did exist a plausible substitution for the critical word. Otherwise, the reader might sense something was amiss prior to reading the critical word. For example, if the daughter were strong, the mother might imagine her being safe, unharmed, helpful, or a heroine.

For purposes of analysis, five word units before and five word units after a critical word were identified (see Fig. 6.1). These were one or two words (rarely three words) that readers tended to pronounce together. Word units did not

y specific syntactic or semantic structures. The major dependent
frequency (later converted to a probability) of a major disruption
; the oral production of each word unit. Major disruptions were
s of at least 1 sec duration, substitutions, omissions, reversals,
ispronunciations, repetitions, and regressions. Only one disrup-
per word unit. A baseline probability of a major disruption was
control condition in which the original critical word was left
baseline probability was subtracted from each of the violation
interaction of violation conditions by word-unit position was the
was evaluated using both readers and critical word segments as

ionale was to look for disruptions in oral reading performance
1 of the violation types. If there were disruptions in oral perfor-
rs presumably were attempting to use the violated information to
itical word and/or to construct a meaning for the text. Fur-
lative position of the disruptions resulting from the different
indicate the order in which the information typically was used.
a bottom-up model, only lexical information was strictly neces-
oduction, although the reader would need to determine the
cture to read with appropriate prosody. If the reader were using
-phoneme correspondences to pronounce the words and did not
in the lexicon prior to uttering it, then the disruption from the
would occur after the critical word had been said. Syntactic and
ation would be used primarily to determine the meaning of the
disruptions would occur after the critical word at the clause
tions of factual information would be important only after the
lause had been determined. The reader would attempt to inte-
g of the clause with the representation of the preceding text and
unter the inconsistency.
:, semantic, and factual information would exert top-down in-
al access would depend on the particular interactive model
ctic, semantic, and/or factual information were used in lexical
ns from violating that information would occur at the same time
from lexical violations. Whether the disruptions occurred at the
nd had about the same extent would indicate the functional
e information sources. Violating two information sources, as in
mantic manipulation, would lead to different patterns of disrup-
n the interactions among the information sources.

g

ent with skilled readers used 12 paragraphs averaging about 125
a contemporary novel. One critical word was selected in each
ur versions of the story, lexical, syntactic + semantic, and

semantic violations were introduced at three critical words each. The remaining three critical words in each version were controls and were left unchanged. Across the four versions, manipulations were counterbalanced across critical words. Ten college students, who were unscreened for reading skills, read each version aloud. In order to ensure that they paid some attention to comprehension, they were required to give a brief summary of each paragraph immediately after reading it.

As shown in Fig. 6.2, the principal point of disruption resulting from lexical violations occurred at the critical word (word unit 0). Most of these disruptions were pauses as the readers hesitated before uttering the pronounceable nonword. Unable to locate that lexical item in their mental lexicons, they balked and sometimes had difficulty in pronouncing the nonwords just on the basis of phoneme–grapheme correspondences. The disruptions from the syntactic + semantic violation followed the disruption from the lexical almost perfectly. They both differed significantly from the control at the critical word and at word unit +1. Syntactic or semantic information or both were being used to locate the lexical item, so that removal of that information disrupted the readers. Since the patterns of disruption resulting from the lexical and syntactic + semantic violations were essentially the same, these information sources must have been

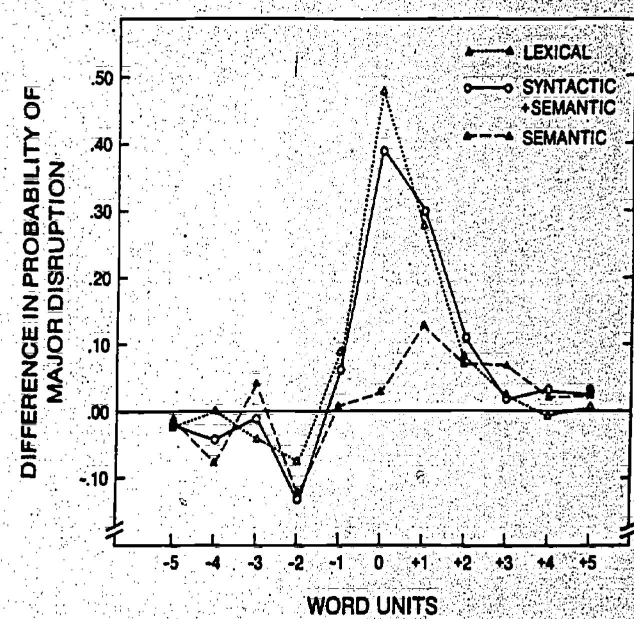


FIG. 6.2. Disruptions in skilled readers' oral reading performances resulting from lexical, syntactic + semantic, and semantic violations.

ame time, although they may have operated independently (Lazer-
75).
tion from the semantic violation did not occur until word unit +1,
was relatively much smaller. In reading this story, semantic infor-
ot being used for lexical access. The violation was discovered later,
clause integration occurred. When we considered the story more
realized why the semantic information may not have been used as
story was excerpted from a novel that was written in an abstract,
style. The semantic violations were anomalies that easily could
staken for intentional but incomprehensible metaphors. The readers
quite reasonable strategy of not giving high priority to semantic
ecause it was frequently figurative or anomalous.
experiment used a concrete, emotionally involving story (2171
about a high school girl who was severely injured when a train hit
s (see sample of text in Fig. 6.1). We dropped the lexical violations
ctual ones. The syntactic + semantic and the semantic violations
. However, given the general context of a concrete, literal story, the
rmation would be more useful in lexical access and the anomalies
iced. The other procedures and analyses remained unchanged ex-
tory was divided into four sections with four critical words (one of
n) in each section. Readers gave a summary after reading each
story.
yntactic + semantic and the semantic violations produced a disrup-
rd unit before the critical word (word unit −1; see Fig. 6.3). The
on from the syntactic + semantic violation was significantly larger
lting from the semantic violation alone, but the semantic violation
lasting effect, remaining significantly different from the control at
3 and +5. Since both syntactic + semantic and semantic violations
ptions before the critical word was produced, both syntactic and
rmation was being used during lexical access. Both violations were
e after the critical word was produced, indicating failures during the
n-integration task as well. The factual inconsistency was disruptive
critical word had been uttered (word unit +1). So factual informa-
nvolved in lexical access, but only in sentence meaning integration.
tic +semantic violation had a larger disruptive effect earlier than
tic violation. A violation in both of two independent information
d be noticed before a condition in which only one source was
he syntactic + semantic violation may have been a violation of two
knowledge sources (Siler, 1973–1974). The next experiment sepa-
actic and semantic violations and attempted to enhance the factual

experiment used the same story as the second, but several of the
ions were rewritten to make them even more inconsistent, and the

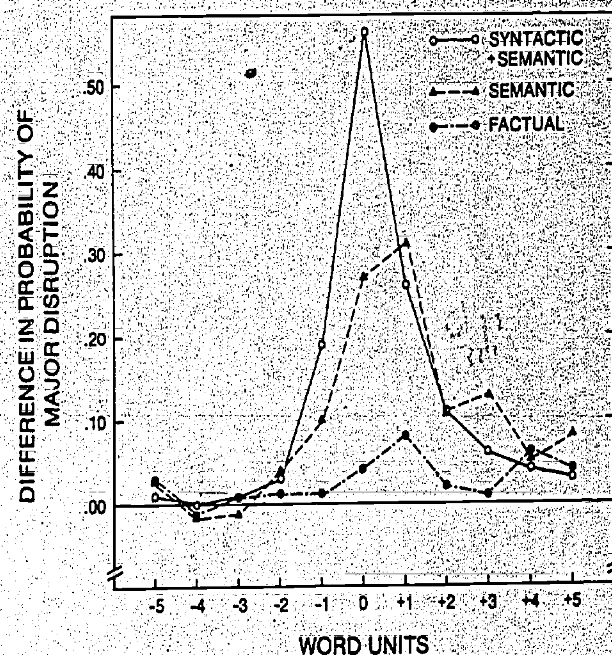
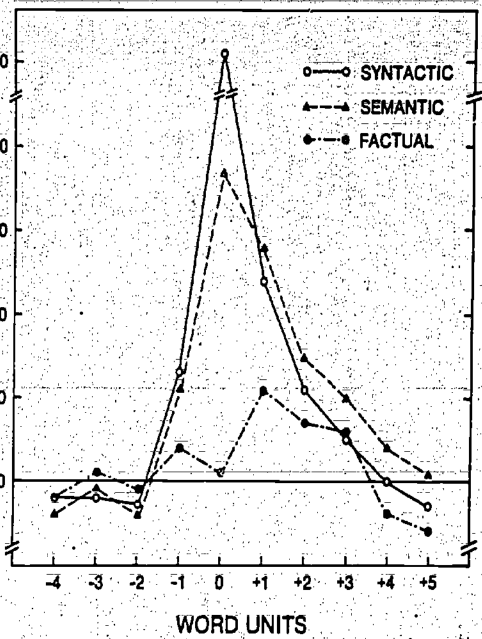


FIG. 6.3. Disruptions in skilled readers' oral reading performances resulting from syntactic + semantic, semantic, and factual violations.

syntactic + semantic violations were replaced with syntactic-only violations. The rest of the procedure and analyses remained the same except that the subjects answered 10 multiple-choice questions after each quarter of the text instead of giving summaries. Scores on the multiple-choice test were nearly perfect.

As shown in Fig. 6.4, the syntactic and the semantic violations produced very similar disruption curves. Both were significantly different from the control at word unit −1 and peaked at the critical word. The semantic violation produced a slightly longer disruptive effect (to word unit +3) than the syntactic (only to word unit +2). The biggest difference was in the magnitude of the disruption at the critical word. Most of the disruptions in the syntactic violation (54% of all disruptions) were restorations of the correct part of speech. Excluding restorations of the original critical word from the syntactic and semantic conditions, the proportions of disruptions at the critical word were virtually identical—.33 for syntactic and .34 for semantic. The restoration of the original critical word in the syntactic violation condition was a top-down effect resulting from syntactic constraints on the part of speech. Syntactic + semantic violations were never restored in the previous experiments. So the difference between syntactic +



. Disruptions in skilled readers' oral reading performances resulting from
, semantic, and factual violations.

semantic-only disruptions observed in the second experiment
lted from the combined effect of violations of two independent
ırces.

violation produced a larger effect than in the preceding experi-
till was first effective at word unit +1 and not at or before the
In contrast to syntactic and semantic information, factual consis-
have any apparent influence on lexical access of the critical word
ed in sentence meaning integration. Lindig (1976) obtained similar
ıal violations on speech-shadowing performance.

of lexical, syntactic, and semantic information disrupted oral pro-
ut the same point before the critical word was uttered. From the
lexical access, lexical information is bottom-up, and syntactic and
mation is top-down. Yet the pattern of their disruptive effects was
All three information sources were contributing to lexical access.
information source was violated, the normally automatic process
forcing a reliance on careful bottom-up processing to be sure what
ally printed.

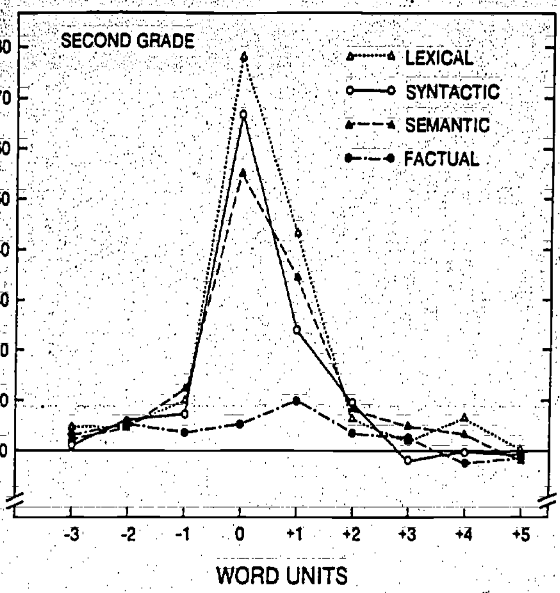
The occurrence of disruptions for several word units after the critical word represents more than a simple perseveration of the initial disruption, however. After the critical word had been uttered, the reader attempted to make sense of the inserted word. How could the clause be interpreted so that the word would not be inconsistent with the representation the reader was constructing for the sentence, paragraph, and story? This comprehension difficulty was evident with all violations, including the factual. Even though the disruption from the factual violation was much smaller, it occurred consistently for several words after the critical word.

Were the post-critical-word disruptions synchronized with the end of the clause? In preparing the stories, we did not attempt to control the location of the clause boundaries after the critical word. However, a post hoc analysis of the data in the last two experiments suggested that there may have been a peak of disruption at the clause boundary. Sometimes a peak at the clause boundary was the only disruption; sometimes it was in addition to one at the critical word; and sometimes it was missing entirely. However, the data were too few to draw strong conclusions about whether there were disruptions at the clause boundary.

Learning to Read

What is the interaction of information sources as a child is learning to read? A reasonable first hypothesis is that the child is paying most attention to bottom-up information, because that is where most instruction is focused and that is where the child is having the most difficulty. As bottom-up processing becomes more automatic (LaBerge & Samuels, 1974), children are gradually able to use more abstract information for lexical access and meaning integration. Alternatively, children initially might be overly dependent on context and prior knowledge simply because they lack proficiency in processing bottom-up information. As they gain skill in decoding, bottom-up information would become relatively more useful to them, and the balance between bottom-up and top-down processing would shift.

To investigate this question, the basic experimental paradigm used with skilled readers was adapted for children learning to read—second, fourth, and sixth graders. Stories were selected from primers one grade below the children's actual grade. The readabilities (Fry, 1968) of the stories were 1.6, 3.5, and 5.6 and the stories were 881, 1354, and 1617 words long. The stories were divided into four sections, and five critical words were selected in each quarter. Lexical, syntactic, semantic, and factual violations were developed for each critical word following the same criteria as for the skilled readers. Five versions of each story were constructed so that violations were counterbalanced across critical words and subjects. There were 50 children tested at each grade level, 10 on each version of the story. In order to ensure that the children paid some attention to



.5 Disruptions in second graders' oral reading performances.

ion, the children were asked three to four simple questions after quarter of the text. The scoring of major disruptions was the same as r skilled readers.

ption curves for the second, fourth, and sixth graders are shown in .6, and 6.7, respectively. Although there were some differences ree grades, the results were very similar. Lexical, syntactic, and lations all produced their largest disruptions at the critical word and, xtent, at word unit +1. A few of the conditions were significantly n the control as early as word units −2 or −1—namely, lexical in lexical in second and fourth at −1, syntactic in fourth and sixth at nantic in second and fourth at −1. Likewise, there were a few ffects at word units +2 and +3—namely, syntactic in second at +2 c in fourth at +2 and +3. But the dominant effect was at and after the critical word.

al violation produced small but significant disruptions in all three ord unit +1 in the second and fourth grades and at the critical word rade. Factual information may have been filling some useful role in s for the sixth graders, but not for the second and fourth graders. The nd the location of the children's factual disruptions indicated that

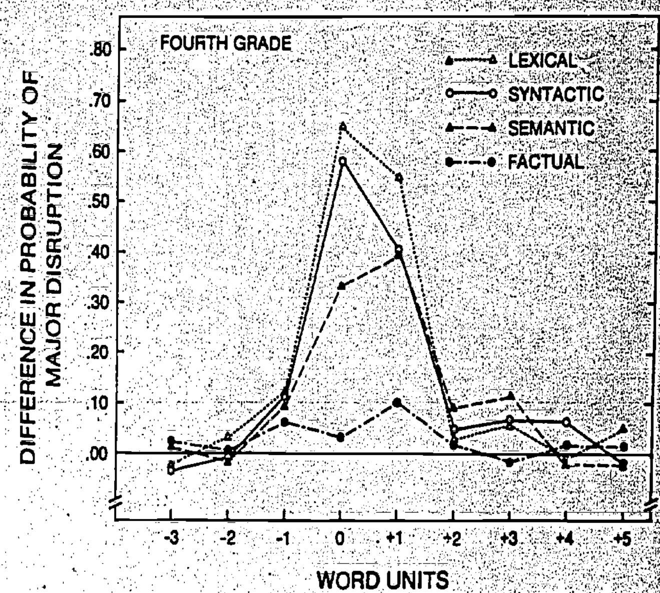
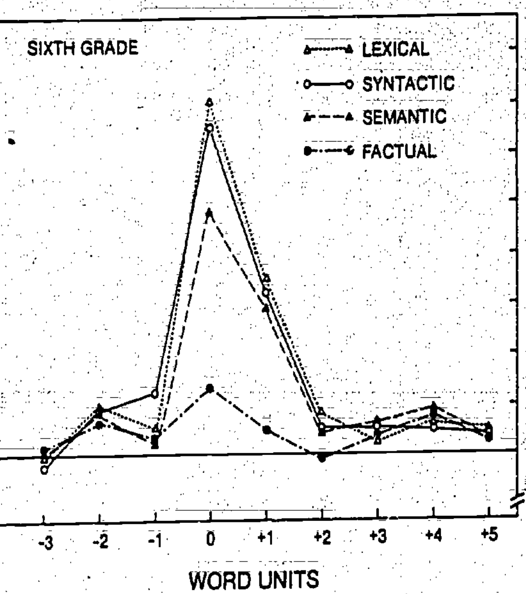


FIG. 6.6. Disruptions in fourth graders' oral reading performances.

resolving a factual inconsistency was more critical at meaning integration than lexical access.

The major conclusion was that syntactic and semantic information influence lexical access of the critical word as much as lexical information itself (cf Isakson & Miller, 1976; Miller, 1975). The magnitudes of the disruptions wer ordered from lexical violations producing the largest disruption, followed closel by syntactic and semantic violations. Perhaps these children were well along th way to being skilled readers, so that bottom-up processing was relatively autom tic, thus permitting top-down processes to operate (Isakson & Miller, 1976 Their reading rates indicated that this was not the case, however. Estimatin reading rates from the control condition, the second graders read at 12 syllables/minute, fourth graders at 161 syllables/minute, sixth graders at 18 syllables/minute, and college students at 270 syllables/minute. The children reading on the whole was not as fluent as that of skilled readers; they read mor slowly and haltingly.

In a comparison of the children's results with those of skilled readers, the were no major differences as a function of skill level. These particular childre may have been more highly skilled than one would expect from their grade leve (standardized reading-test scores were not available), but still they were n



Disruptions in sixth graders' oral reading performances.

t levels. The use of relatively easy, grade-appropriate stories may he children to use more skilled reading strategies. Also, some of inefficiencies may have been masked by their slow reading rates. reasons, there were no significant differences in the pattern of grades.

INTERACTIVE ANALYSIS OF ORAL READING

der an interactive analysis of oral reading and how such an analysis esults. There are two interactive models that are related to the one —Rumelhart's (1977) parallel-interactive model of reading and n and Welsh's (1978) direct lexical access model of speech perough there are convergences among the models, neither or Marslen-Wilson and Welsh's has been developed in the direcading.

ding Model

sk in oral reading is to verbalize each word of the text in succesnstant demand on the reader, one that cannot be ignored. There are

many points in a reading comprehension process at which the oral production might originate (Danks & Fears, 1979). Extraction of grapheme–phoneme correspondences might supply sufficient information for articulation. At the other extreme, a meaning representation of an entire sentence might be constructed first; a sentence production process might be initiated in the same way that sentences are produced in conversations. Neither of these extremes seems likely. The former would have readers articulate orthographically regular nonwords in context without any hesitation. However, readers do not do this (cf. disruptions from the nonword violation). The latter position would require that eye–voice spans encompass entire clauses, but eye–voice spans are not that large (Levin, 1979; Vázquez, Glucksberg, & Danks, 1977–1978).

A more plausible assumption is that oral production is initiated following lexical access, at least for familiar words. Once readers locate a word in their lexicon, they then have available the articulatory information necessary to pronounce it. There is no claim that lexical access is direct without any phonological mediation, only that initiation of an oral response does not occur until after the word has been accessed. The oral response does not originate in any phonological mediation that might be used in lexical access. When encountering an unfamiliar word, readers—especially children being taught with a phonics-based program—sometimes "sound out" the word slowly by breaking it into phoneme and/or syllable units. Then they repeat it, blending the parts together. In these cases, the first pronunciation is directed by letter–sound correspondences and the repetition, by the articulatory code accessed in the mental lexicon.

Lexical Access. Our model of lexical access is closely related to Marslen-Wilson and Welsh's (1978) direct access model. They proposed that a cohort of words is activated in the mental lexicon strictly on the basis of incoming perceptual information, which in their studies of speech shadowing was a phonemic analysis of the initial syllable of each word. Candidate words in the cohort then are evaluated against the continuing perceptual analysis and against syntactic and semantic constraints. Any word that is inconsistent with any constraint, whether it be perceptual or contextual, is deactivated. The elimination process continues until only one candidate remains in the cohort. That survivor is recognized as the target word.

Applying this model to reading, cohorts of words are activated by an initial visual analysis. However, we think that semantic and syntactic information, especially semantic, also can activate words for the cohort. In conversations, listeners frequently anticipate what speakers are going to say—as when a speaker pauses, apparently searching for a word, and the listener obligingly supplies it. Sometimes the listener anticipates correctly, sometimes not; but in either case the anticipation is present. We are not proposing a hypothesis-testing procedure in which contextually based predictions are produced and evaluated serially by comparison to perceptual information. Rather, we are proposing that many alter-

activated simultaneously, each of which is consistent with some
mation, either perceptual or contextual. These alternatives then are
parallel.
iation procedure differs from Marslen-Wilson and Welsh's as well.
ure is a process of elimination. Members of the cohort are tested for
with semantic and syntactic information, as well as with the results
uing perceptual analysis. Inconsistent items are eliminated from the
one is left. We propose a two-stage process. Initially, an elimination
members of the cohort for consistency with all available informa-
lly syntactic and semantic. This evaluation is fast and can be based
e perceptual information. If only one item remains in the cohort, it
However, if more than one item remains or if no item remains, then
slow, careful perceptual analysis is conducted to obtain relatively
er, syllable, and word information. Most of the time the elimination
ficient, but sometimes the reader must rely primarily on perceptual

rd has been located in the mental dictionary, information about how
that word is sent to a buffer where speech articulation is organized.
ion process is assumed to be nearly the same as spontaneous speech
he only difference is that in oral reading, subsequent syntactic and
rmation might be needed to specify some pronunciation or intona-
For example, whether the printed word *record* is pronounced with
nd-syllable stress may be determined by syntactic or semantic in-
er in the sentence. The use of subsequent syntactic and semantic
s not needed in natural speech production because the speaker
s whether *record* is a noun or a verb, just as he or she knows what
tern is appropriate for the message. In contrast, the relevant infor-
not be available to the oral reader until the end of a clause or
reader can scan quickly ahead in an attempt to locate the informa-
ect an interpretation using whatever information is available at that

d Sentence) Comprehension. At the same time as the reader is
e words, he or she is also selecting an appropriate meaning for each
y, 1979) and integrating those meanings in order to understand
ead. This step is necessary to satisfy the secondary task demand of
ecall or summary after the reading is finished. In some cases, the
forced to wait until reaching a clause boundary to execute the
ration (Carroll & Bever, 1976; Carroll, Tanenhaus, & Bever,
general, the reader attempts to integrate meanings word by word,
possible (Marslen-Wilson, Tyler, & Seidenberg, 1978; Tyler &
n, 1977).

As with lexical access, there are two processing options. Initially, the reader attempts a fast integration of each word's meaning with the representation that has been constructed up to that point. This process is efficient but requires the use of considerable top-down, contextual information about the meaning of what is written. With relatively easy or familiar material, the reader can succeed in performing this integration with minimal attentional and processing effort. Upon reaching the end of a clause, which is a conceptual as well as syntactic demarcation, the reader has in mind a coherent representation of the entire text up to that point.

Sometimes, however, word-by-word integration does not succeed. For example, in our semantic violation, the readers were unable to integrate the critical word. In more typical texts, there are ambiguities and vagaries to confound a reader's word-by-word integration. When such failures occur, a reader can postpone integration for a short time in hope that new information will resolve the difficulty. The individual meanings are stored in a temporary memory buffer. At the end of the clause or sentence, the reader is forced to integrate as best as possible because he or she cannot store and retain the meanings of individual words for very long. Since most clauses and sentences are conceptually integrated, the reader can reasonably expect a resolution at that point. At the end of a clause or sentence, the reader attempts to clarify the meaning of the clause by carefully analyzing its syntactic structure. In this respect, a careful syntactic analysis is a bottom-up process that provides definitive structural information specifying how to integrate the word meanings. If there is no resolution, the reader can suspend forward processing and regress in an attempt to find additional information or can plunge ahead, leaving a lacuna in the composite representation.

In summary, first the reader attempts a word-by-word integration of the meanings, based on prior knowledge and semantic expectancies about what is being communicated. If that process fails, the reader relies on a careful analysis of the syntactic structure ending at the clause boundary. The former process is fast and efficient, but risky. The latter is slow, but more likely to succeed.

In an oral reading task, how would these two processing strategies be evidenced? If word-by-word integration goes awry, oral performance would be disrupted at that point. For example, disruptions occurred both before and after a semantically anomalous word, one that could not be integrated, had been uttered. The disruption after the critical word could not have resulted from problems of lexical access because the word had to be accessed to be uttered. So a disruption after the critical word reflected difficulty with meaning integration. Attention appeared to be diverted from oral production to meaning integration. However, it is not clear that either oral production or meaning integration requires much attention in the normal case. Skilled readers can easily read aloud without attention, as any parent who reads stories to children can testify. Meaning integration

s relatively effortless. Yet somehow, difficulty with word-by-word distracted the reader's attention from oral production and resulted in a If word-by-word integration has failed and a careful syntactic analysis then the reader may have to pause at the clause boundary to allow tention) for the slower analysis to occur. In the experiments reported cation of the clause boundary was uncontrolled, so that disruptions at oundary were not observed in the primary analysis. However, a post s at the first clause boundary following each critical word indicated as a disruption at that point.

coherent representation of the entire text constructed? In word-by-ation, each word is integrated with a composite representation of the ip to that point, not just with the representation for that clause or a clause requires additional processing at the clause boundary in inderstood, then the reader has to adjust the meaning of the clause to osite representation. Such an adjustment is especially necessary if the he clause is inconsistent with another proposition already integrated entation. The factual inconsistency in our experiments blocked the a coherent representation. Elementary school children sometimes sistent propositions to remain in their mental representations 1979), but skilled, adult readers usually do not. If the inconsistency ficiently critical information for the understanding of a text, a reader egress, and attempt to solve it. Alternatively, he or she may alter erpretation of the prior existing portion of the inconsistency (Loftus, , Miller, & Burns, 1978) or the perception of the new information. ation is not crucial for understanding the whole text, the reader may attempt at resolution indefinitely. At recall, when that understanding e public, he or she may attempt a resolution, omit the inconsistent or repeat the inconsistency unaware. In any case, resolution of the not a pressing demand and can be postponed in hope of finding ater.

g Task Demands

analysis of oral reading, we can identify three implicit tasks for oral first is lexical access, which provides the information necessary to ral rendition. It also is the primary source of semantic information forming a representation of the meaning of the text, but it is the mand for oral production that drives lexical access. It cannot be s it might be in silent reading when an unknown word is -without producing a major disruption in oral production. Because acy of the demand, disruptions in lexical access appear as disrup-roduction before the critical word has been uttered.

The second implicit task is clause understanding and integration. Normally, this task proceeds word by word, but when that integration fails—as it would with our semantic violations—then more careful structural analysis must occur. There would be a disruption in oral production at the point within the clause where word-by-word integration failed. Attention would be diverted from oral production to meaning integration. The demand for meaning integration can be delayed to the end of the clause but no longer because of memory limitations. One result of a more careful analysis at the clause boundary is a disruption in oral production there.

The third implicit task is ensuring that the representation is coherent. Successful word-by-word integration satisfies this demand perforce. However, if an inconsistency remains at the end of a clause—as there would following a factual violation—the demand for coherency can be postponed to the end of the text or until an overt comprehension response, such as recall or summarization, is required.

What effect does the demand for an oral production have on these three implicit tasks? Some effects have been mentioned already. The competition for cognitive resources can result in less attention being allocated to integrating a meaning representation and to ensuring its coherence. For skilled readers, comprehension and oral production processes are sufficiently automatic that both can be carried out in parallel. Division of attention would be most evident when the comprehension process encounters difficulty. However, the sharing of cognitive resources may be facilitated by the fact that the mechanics of articulation slow down the intake of new information. The silent reading rates of skilled adult readers are approximately halved in oral reading. Many skilled readers read difficult texts aloud in order to assist understanding. If the rate becomes too slow, however, as in a child who is having difficulty with basic decoding skills, comprehension may be affected adversely. Individual word meanings would have to be retained too long in a temporary memory buffer and would be forgotten.

Having to read every word in sequence may alter the comprehension processes. In silent reading, skilled readers may not look at and access every word, especially highly predictable words (but for evidence to the contrary, see Zola, 1979; described by McConkie & Zola, Chap. 7, this volume). Skilled readers may not access the words in the order in which they appear on the page, either. Instead, they may jump ahead or regress in an attempt to optimize information intake. This strategy is analogous to that employed by many English-speaking students with an intermediate-level knowledge of German: Jump to the verb at the end of the clause before reading the rest of the clause.

Finally, auditory feedback from hearing oneself read aloud may change processing. Readers who spontaneously restored the correct part of speech to the syntactic violations may have paused after words because they realized that what

ered was not what was printed. Auditory feedback provides an back loop by which automatic oral production can be monitored.

on of Oral Reading Results

is analysis of oral reading explain the results of the experiments? nding is considered in terms of the analysis.

l, syntactic, and semantic violations disrupted oral production be- cal word was uttered (i.e., at word units −1 and 0). All three types on were used in lexical access. By definition, lexical violations ical access. Since the nonword was not in the lexicon, it could not Syntactic and semantic violations disrupted oral production before ord was uttered to the same extent as did insertion of a nonword. So rmation and semantic information were critical for lexical access as

l, syntactic, and semantic violations disrupted oral production after ord was uttered (i.e., at word units +1 and +2). Disruptions in s indicated that lexical, syntactic, and semantic information was construct a meaning of the sentence or clause. Semantic disruptions cal word tended to be somewhat larger than syntactic ones (cf. Fig. ing continuing attempts at meaning integration.

both syntactic and semantic properties were violated, disruptions the critical word was uttered and were more pronounced than when rmation alone was violated. Syntactic and semantic information nctionally independent knowledge sources, both of which were cal access. In terms of the model, violation of both information ed in the more rapid elimination of all members of the initial cohort access. Since the slower perceptually dependent process was re- arlier point, the disruptions also occurred earlier.

s frequently restored syntactic violations to the original and contex- iate part of speech. The part of speech was altered by changing the he root morpheme. There is a left-to-right scanning bias in reading Rayner, 1976), as well as in speech perception (Marslen-Wilson & . Initial analysis of the first letters or syllable would be consistent ographic, syntactic, and semantic constraints of a single lexical y, the original critical word. It would be the only word remaining in an early point in the elimination process. With the relatively early all but one lexical item from the cohort, perceptual processing of on of the printed word could be discontinued, and the original word red, restoring the altered inflection.

c violations in conceptually difficult text disrupted oral perfor- ter the critical word was uttered—that is, at word units +1, +2,

and +3 in Fig. 6.2. When the material was conceptually difficult, readers were unable to generate semantic expectations about upcoming words. They had few expectations with which to select items for the initial cohort, nor was there semantic information that could eliminate items from the cohort. After a lexical item was identified, primarily through perceptual and syntactic information, the reader could not integrate the meaning of the semantically anomalous word, resulting in a disruption after the critical word was uttered.

6. Factual inconsistencies between successive sentences disrupted oral performance only after the critical word had been uttered (word unit +1) and not before. Factual violations did not disrupt lexical access. There is nothing in the model that precludes factual expectancies from activating or eliminating items from the cohort. Perhaps the factual violations were not strong enough to produce inconsistent expectancies. Or perhaps factual information simply is not involved in lexical access. Factual inconsistencies did interfere with word-by-word meaning integration, however, although the disruption was relatively small.

7. In a post hoc analysis, some disruptions in oral performance occurred at the clause boundary immediately following the critical word. This pattern was most evident following factual violations. Assuming that this tentative finding is confirmed in future experiments, disruptions at the clause boundary indicated that readers were making another attempt to understand the clause with the violation. They may have reread the sentence silently; they may have regressed to the preceding sentence; or they may have engaged some covert problem-solving strategy. Language is flexible enough that most apparently inconsistent statements can be resolved eventually. This fact became obvious to us when constructing materials for the experiments. It was difficult to exclude the possibility of plausible alternative interpretations.

8. In the experiment with children, the same basic pattern of results was obtained at all grade levels. General processing strategies apparently were the same across differences in reading skill. There are several qualifications on that conclusion. First, reading skill differences were defined by grade, not by test scores. Good and poor readers at the same grade level might yield different patterns of results. Second, the stories were easy, somewhat below estimated reading-grade level. With difficult texts, differences in processing strategies might emerge as children acquire reading skill. Third, the youngest children were second graders in the second semester. Substantial changes in processing strategies may occur during the first 1½ years of learning to read (Miller & Isakson, 1978; but contrast with Lovett, 1979). In spite of these qualifications, the similarity of oral disruption patterns in Figs. 6.2 through 6.7 was striking. Processing differences may have existed, but the dominant impression was one of constancy.

These results are consistent with an interactive model of oral reading. Although we have not detailed all of the arguments and rationale, it does not seem

r any plausible bottom-up model to explain these results. Perhaps one leveloped (e.g., McClelland's [1979] cascade model), but we are look toward interactive models as a more promising avenue. Other models, such as Rumelhart's (1977) and Marslen-Wilson's Vilson et al., 1978; Marslen-Wilson & Welsh, 1978), are consistent sults which do not discriminate among these models in any significant e are many experimental demonstrations that require interactive which ours is one, but these demonstrations do not tell us very much the interaction works. The nature of the interaction among informas needs to be specified more precisely through experimentation.

ACKNOWLEDGMENTS

ents discussed here were conducted in collaboration with Lisa Bohn, Ramona Hall, and Ruth Miller. Their contributions to the research program are knowledged. Preparation of this paper and the reported experiments were Grant No. NIE-G-78-0223 from the National Institute of Education.

REFERENCES

, & Bever, T. G. Sentence comprehension: A case study in the relation of knowledge on. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Language* (Vol. VII). New York: Academic Press, 1976.

, Tanenhaus, M. K., & Bever, T. G. The perception of relations: The interaction of functional, and contextual factors in the segmentation of sentences. In W. J. M. Levelt lores d'Arcais (Eds.), *Studies in the perception of language*: Chichester, England: 8.

Models of language comprehension. *Polish Psychological Bulletin*, 1978, *9*, 183-192.

& Fears, R. Oral reading: Does it reflect decoding or comprehension? In L. B. P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 3). Hillsdale, N.J.: Erlbaum Associates, 1979.

& Glucksberg, S. Experimental psycholinguistics. *Annual Review of Psychology*, 391-417.

hat classroom observations reveal about reading comprehension instruction. *Reading Quarterly*, 1978-1979, *14*, 481-533.

eadability formula that saves time. *Journal of Reading*, 1968, *11*, 513-516; 575-578.

One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear* Cambridge, Mass.: MIT Press, 1972.

, & Miller, J. W. Sensitivity to syntactic and semantic cues in good and poor ders. *Journal of Educational Psychology*, 1976, *68*, 787-792.

& Samuels, S. J. Toward a theory of automatic information processing in reading. *sychology*, 1974, *6*, 293-323.

. The influence of highly variable spelling upon the reading performance of skilled Modern English. *Reading Research Quarterly*, 1974-1975, *10*, 583-615.

eye-voice span. Cambridge, Mass.: MIT Press, 1979.

Lindig, K. D. *Contextual factors in shadowing connected discourse*. Unpublished doctoral dissertation, University of Chicago, 1976.

Loftus, E. F. Leading questions and the eyewitness report. *Cognitive Psychology*, 1975, *7*, 560-572.

Loftus, E. F., Miller, D. G., & Burns, H. J. Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 19-31.

Lovett, M. W. The selective encoding of sentential information in normal reading development. *Child Development*, 1979, *50*, 897-900.

Markman, E. M. Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development*, 1979, *50*, 643-655.

Marslen-Wilson, W. D. Sentence perception as an interactive parallel process. *Science*, 1975, *189*, 226-228.

Marslen-Wilson, W., Tyler, L. K., & Seidenberg, M. Sentence processing and the clause boundary. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), *Studies in the perception of language*. Chichester, England: Wiley, 1978.

Marslen-Wilson, W. D., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 1978, *10*, 29-63.

Massaro, D. W. Primary and secondary recognition in reading. In D. W. Massaro (Ed.), *Understanding language*. New York: Academic Press, 1975.

McClelland, J. L. On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, 1979, *86*, 287-330.

McConkie, G. W., & Rayner, K. Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, 1976, *8*, 365-368.

Miller, J. W. Disruptive effect: A phenomenon in oral reading. *Reading Horizons*, 1975, *15*, 198-207.

Miller, J. W., & Isakson, R. L. Contextual sensitivity in beginning readers. *Elementary School Journal*, 1978, *78*, 325-331.

Rumelhart, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Siler, E. R. The effects of syntactic and semantic constraints on the oral reading performance of second and fourth graders. *Reading Research Quarterly*, 1973-1974, *9*, 583-602.

Swinney, D. A. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 1979, *18*, 645-659.

Tyler, L. K., & Marslen-Wilson, W. D. The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 683-692.

Vázquez, C. A., Glucksberg, S., & Danks, J. H. Integration of clauses in oral reading: The effects of syntactic and semantic constraints on the eye-voice span. *Reading Research Quarterly*, 1977-1978, *13*, 174-187.

Wildman, D. M., & Kling, M. Semantic, syntactic, and spatial anticipation in reading. *Reading Research Quarterly*, 1978-1979, *14*, 128-164.

Zola, D. *The perception of words in reading*. Paper presented at Psychonomic Society meetings, Phoenix, 1979.

Comprehension of Prose Texts During Reading[1]

Joseph H. Danks

Department of Psychology, Kent State University

Kent, Ohio 44242, U.S.A.

Presented at the

XXIInd International Congress of Psychology

Leipzig, 6-12 July 1980

1. The assistance of Lisa Bohn, Ramona Fears, Greg Hill, Jeff Hall, and Ruth Miller in the design and execution of the research reported in this paper is gratefully acknowledged. This research was supported by Grant No. NIE-G-78-0223 from the National Institute of Education.

Models of Reading Comprehension

What are the cognitive processes by which a reader arrives at an interpretation for a text? Research in the last couple of decades has focused on determining the nature of the mental representation that readers and listeners extract from verbal material. Recently, investigators have become more interested in specifying the cognitive processes by which the mental representations are constructed (Danks & Glucksberg, 1980). The shift in interest from representations to processes has altered the nature of the conceptual models advanced and the types of methodology used.

A major distinction that has come to pervade models of listening and reading comprehension, as well as cognitive models generally, is between bottom-up and top-down directions of information transfer and control (Danks, 1978; Danks & Glucksberg, 1980). Models that adopt a bottom-up orientation typically are composed of serial independent stages arranged in an inflexible order (e.g., Gough, 1972; Massaro, 1975). Comprehension is conceived as beginning with sensory analysis of linguistic input, proceeding through successive transformations, and terminating with a final representation that constitutes the meaning of the input. Bottom-up models have difficulty explaining how information from more abstract levels of analysis are able to facilitate (or hinder) processing at a lower level of abstraction (cf. Rumelhart, 1977, and Wildman & Kling, 1978-79, for reviews of these phenomena). In contrast, top-down models, which are usually cast as inter-

active models, incorporate the possibility of information from more abstract levels being used in the processing at lower levels. Since bottom-up processing is not precluded, these models depict construction of a representation using available information from any level. Processing proceeds in parallel on several different interacting levels. The process is adapted flexibly to meet the needs of the specific task and text.

Another conceptual distinction that has resulted from the shift in focus from representations to processes is the more precise specification of when a representation is formed. An initial hypothesis was that representations were formed clause by clause (Carroll & Bever, 1976; Carroll, Tanenhaus, & Bever, 1978). Meanings of individual words were integrated at the end of each clause into an autonomous representation of that clause's meaning. Then the clause meanings were integrated successively as the reader proceeded through the text. The clausal processing hypothesis was a natural outgrowth of the emphasis on representations. Since the final mental representation was organized from clause units, there was no need to specify precisely how or when clausal integration occurred.

The clausal processing hypothesis contrasts with a word-by-word integration hypothesis (Marslen-Wilson & Tyler, 1980; Marslen-Wilson, Tyler, & Seidenberg, 1978; Tyler & Marslen-Wilson, 1977). The listener or reader attempts to integrate each word's meaning with a comprehensive mental representation

as soon as he recognizes the word. He uses whatever lexical, syntactic, and/or semantic information is available at the time. The attempted integration does not always succeed; in which case, the listener/reader holds the word(s) in a short-term memory buffer until there is conceptual closure, usually at the end of the clause.

In order to differentiate between the clausal processing hypothesis and the word-by-word integration hypothesis, an on-line measure of comprehension, one that is taken at the same time as comprehension in proceeding, is required. Memory tasks assess only the final representation, and so provide limited information about the course of processing. The necessity of an on-line measure complicates interpretation because if one adds a behavioral task to a comprehension task, the comprehension processes may be altered by the presence of the secondary task. The naturalness of a secondary task is an important consideration in evaluating its impact. Is it one that readers perform at least occasionally in everyday life, or is it one that is devised specifically for the laboratory? More natural tasks are more likely to reflect typical comprehension processes.

In the experiments reported here, we used an oral reading task. Oral production occurs roughly simultaneously with the comprehension of the print input. Furthermore, it is a relatively natural task that is used extensively in most elementary school classrooms. Virtually every literate adult, especially parents, has had experience reading aloud. The

listening counterpart to oral reading is speech shadowing, in both cases the behavioral response is an oral rendition corresponding word for word to the input (Marslen-Wilson, 1975; Marslen-Wilson & Welsh, 1978). Oral reading is a more natural task than speech shadowing, however.

An oral reading task imposes two demands on the reader (Danks & Fears, 1979; Danks & Hill, in press). The more pressing demand is to produce an oral rendition of the printed text. The reader is in effect required to say each word in succession, and so must recognize each word by locating it in his mental lexicon. The second demand is to form a mental representation of the entire text. In these experiments, the reader was asked for a summary at several points during and after reading the text. The reader must use lexical, syntactic, and semantic information to construct an appropriate representation on which to base his summary. There is, however, an additional metacognitive task associated with the comprehension and construction of a representation. The reader must monitor the process to insure that the representation is meaningful, coherent, and internally consistent. The task of constructing a representation is usually considered to be the primary task in reading, but the additional requirement of oral production may supplant or interfere with comprehension.

Rationale and Method

There are several different types of information that are potentially available to a reader, i.e., visual, ortho-

graphic, lexical, morphological, syntactic, semantic, pragmatic, and factual. Which of these information sources are used by the reader, and what is the temporal order of use? In order to answer these questions, we distorted different information sources at several points in a story. We assumed that at the point in time when the reader attempted to use the distorted information, his normal flow of processing would be disrupted. Disruption of normal processing would result in a disruption in oral production. Thus, the pattern of when oral production was disrupted relative to the location of the distortions in the text would reflect the pattern of when the various information sources were being used.

The main text used with skilled readers was a story about a high school girl who was severely injured when her school bus was hit by a train. The story was adapted from a popular magazine and was 2171 words long with a readability of 7.8. In the sample shown in Figure 1 to illustrate the text manipulations, the girl's mother has just heard about the accident and is worried about her daughter.

Insert Figure 1 about here

Four types of information were distorted. (a) Lexical--a word was replaced with a pronounceable nonword, i.e., one that was orthographically regular in English, e.g., in Figure 1, the critical word injured was replaced with brugen. (b) Syntactic--the root morphemes of the critical word were retained, but the inflectional ending was changed so as to pro-

duce a syntactically unacceptable word for the sentence, e.g., injured was changed to injury. (c) Semantic--a word was replaced with a word that was semantically anomalous in the sentence, e.g., injured was replaced with planted. (d) Factual--the sentence preceding the one with the critical word was changed so that it was factually inconsistent with the critical word; neither the critical word nor the sentence containing it was altered, e.g., weak was changed to strong, which is inconsistent with the daughter being injured. (e) A Control condition in which the critical word was left unaltered provided a baseline for oral reading performance.

The manipulations were counterbalanced across 16 critical words in four different versions of the story, such that each manipulation occurred at each critical word an equal number of times across subjects. The story was read aloud by 40 college undergraduates into a tape recorder. The stories were divided into four sections, after which the subjects recalled as much of the story as possible. These recalls were to insure that the subjects were paying attention to what they were reading and were not systematically analyzed.

Two dependent measures were scored for three word units before and five word units after each critical word (see Figure 1). A word unit was one or two words that tended to be pronounced as a unit by the readers. The first dependent variable was whether a major disruption occurred in each word unit. A major disruption was defined as a substitution, mispronunciation, stammering, repetition, regression, reversal,

omission, or pause of at least one second. The second dependent variable was the production time for each word unit. The time from the end of one word unit to the end of the next was measured. The results of the production times mirrored almost perfectly the results from the major disruption data, so only the latter are reported here. The data were analyzed with an analysis of variance using both subject and critical word variance in the error term. The primary interest was in the interaction of text manipulations with word unit position. For purposes of presentation in the following figures, frequencies of major disruptions were converted to probabilities. The probabilities in the control condition were subtracted from the probabilities in each experiment condition.

Skilled Reading

The results from two experiments employing skilled readers as subjects are presented in Figure 2. One experiment used the text illustrated in Figure 1 and included syntactic, semantic, factual, and control manipulations. The other experiment used a different text with different subjects, but was functionally the same as the first experiment. The second experiment included lexical, syntactic-semantic, semantic, and control manipulations, but only the results of the lexical manipulation (vis-à-vis its control) are presented here.

Insert Figure 2 about here

The lexical, syntactic, and semantic violations produced

similar patterns of disruptions although there were some interesting differences among them. All three were significantly different from the control condition beginning at word unit -1 and peaking at the critical word (word unit 0). Disruptions from lexical and syntactic manipulations continued to word unit +2, and from semantic to word unit +3. The factual violation did not produce a significant disruption until after the critical word had been uttered (word unit +1), but maintained a small effect through word unit +3.

That the lexical violation disrupted oral reading performance before the critical word was uttered was expected. Although the nonwords could have been pronounced by reference to grapheme-phoneme correspondences, as they ultimately were, the normal strategy was to locate the pronunciation of the word through lexical access, but the nonword was not in the mental dictionary. The syntactic and semantic violations also disrupted oral reading performance before the critical word was uttered, just as soon as did a lexical violation. So syntactic and semantic violations were also disrupting the lexical access process. How can syntactic and semantic information be used in lexical access? If one assumes a strictly bottom-up model of lexical access, syntactic and semantic information are not involved because they are not available until after lexical access has occurred. However, under an interactive model, the reader forms expectancies about what parts of speech and what concepts are likely to occur next. He then uses that information to guide lexical access. When

there was a syntactic or semantic violation, the bottom-up perceptual information was inconsistent with the syntactic and semantic expectancies, so the reader took longer to locate the critical word.

The peak probability of a major disruption was considerably higher for the syntactic violation than for the semantic (also than for the lexical, but those conditions were from separate experiments). Most of the disruptions in the syntactic conditions (54%) were fluent restorations (substitution "errors") of the correct part of speech, e.g., the reader said injured instead of injury. Excluding these fluent restorations, the probabilities of a major disruption from syntactic and semantic violations were virtually identical, .33 for syntactic and .34 for semantic. Fluent restorations were also obtained by Marslen-Wilson (1975; Marslen-Wilson & Welsh, 1978) in a speech shadowing task. They provide further evidence that readers were using top-down information to facilitate lexical access.

All three manipulations produced significant major disruptions for two or three word units following the critical word. These disruptions resulted from the difficulties readers had integrating the information into an appropriate mental representation. Lexical, syntactic, semantic, and factual information were all used by the reader to construct a meaning for the phrase, clause, sentence, paragraph, and story. In addition to producing an oral rendition of the text, the reader also was expected to understand the story.

In an oral reading task, the comprehension demand is less pressing than that for oral production, but it still must be met. When the reader was unable to construct an appropriate representation, he might regress and reread some of the story either silently or aloud, or he might reconsider what he had just read while still continuing to read aloud. Whatever strategy he adopted, there would be a disruption following the critical word.

Especially interesting is the finding that the factual inconsistency produced a disruption shortly after the critical word was uttered. In order for a reader to recognize a factual inconsistency, information from one sentence had to be integrated with the representation being constructed for the following sentence. The factual violation in particular permitted a strong test of whether word meanings were integrated word by word, or whether integration occurred only at the clause boundary. In these experiments, the location of the clause boundary following the critical word varied freely. In a post hoc analysis of whether disruptions were synchronized with the location of the following clause boundary, there seemed to be disruptions at two locations--at or immediately following the critical word, or at the following clause boundary, or sometimes at both locations. These tentative findings have been confirmed in more recent experiments in which the location of the clause boundary has been explicitly manipulated.

Thus, readers were integrating each word's meaning into

the representation as soon as it was accessed in the mental dictionary. However, the violations blocked integration because the critical word was inconsistent with the representation already constructed. At the clause boundary, the reader again attempted to integrate the inconsistent information. The clause boundary was the appropriate place for the reader to attempt a second integration because clauses are usually conceptually integrated and there is a limit on how many and how long words can be retained in short-term memory.

In summary, the results of the experiments with skilled readers support an interactive model of oral reading. Word meanings are accessed through the interaction of bottom-up and top-down information, and are integrated word by word into a mental representation of the text, followed by an evaluation of the success of the integration at the clause boundary.

Learning to Read

How does this model of oral reading relate to children learning to read? What is the interaction among information sources as children become more proficient at reading? One hypothesis is that the child pays more attention to bottom-up information because that is where most reading instruction is focused and that is where the child is having the most difficulty. As bottom-up processing becomes automatic (LaBerge & Samuels, 1974), the child is gradually able to use more abstract information for lexical access and meaning integration. An alternative hypothesis is that the beginning

reader is overly dependent on context and prior knowledge simply because he lacks proficiency in processing print. As he gains skill in decoding, bottom-up information becomes more useful to him and the balance between bottom-up and top-down processing shifts.

To investigate this question, the basic experimental paradigm used with skilled readers was adapted for children learning to read--second (mean age = 8;0), fourth (mean age = 10;1), and sixth (mean age = 11;11) graders. Stories were selected from primers one grade below the children's actual grades. The readabilities (Fry, 1968) of the stories were 1.6, 3.5, and 5.6, and the stories were 881, 1354, and 1617 words long. The stories were divided into four sections and five critical words were selected in each quarter. Lexical, syntactic, semantic, and factual violations were developed for each critical word following the same criteria as for the skilled reader story. A sample segment from the second grade story is shown in Figure 3. Five versions of each story were constructed so that violations were counter-

---

Insert Figure 3 about here

---

balanced across critical words and subjects. There were 50 children tested at each grade level, ten on each version of the story. In order to insure that the children paid some attention to what they were reading, they were asked four or five simple questions after reading each quarter of the story. The scoring of major disruptions and production times was the

same as described for skilled readers. As with skilled readers, the production times mirrored the major disruption results.

The disruption curves for the second, fourth, and sixth graders are shown in Figures 4, 5, and 6 respectively. Although there were some differences across the three grades,

---

Insert Figures 4, 5, and 6 about here

---

the results were very similar. Lexical, syntactic, and semantic violations produced their largest disruptions at the critical word and to a lesser extent at word unit +1. A few of the conditions were significantly different from the control as early as word units -2 or -1, viz., lexical in sixth at -2, lexical in second and fourth at -1, syntactic in fourth and sixth at -1, and semantic in second and fourth at -1. Likewise, there were a few significant effects at word units +2 and +3, viz., syntactic in second at +2 and semantic in fourth at +2 and +3. The factual violation produced small but significant disruptions in all three grades--at word unit +1 in the second and fourth grades and at the critical word in grade six. The dominant disruptive effect of all four violations was at and immediately after the critical word was uttered.

Syntactic and semantic information influenced lexical access of the critical word as much as lexical information itself. The magnitude of the disruptions were ordered from lexical violations producing the largest disruption, followed closely by syntactic and semantic violations. The small size

and location of the children's factual disruptions indicated that resolving a factual inconsistency was more critical at meaning integration than at lexical access.

Comparing the children's results with those from skilled readers, there were no major differences as a function of skill level. Perhaps these children were well along the way to being skilled readers, so that bottom-up processing was relatively automatic, thus permitting top-down processes to operate. Their reading rates indicated that this was not the case, however. Estimating reading rates from the control condition, the second graders read at 123 syllables/minute, fourth graders at 161 syllables/minute, sixth graders at 181 syllables/minute, and college students at 270 syllables/minute. The children's reading on the whole was not as fluent as skilled readers; they read more slowly and haltingly. These particular children may have been more highly skilled than one would expect from their grade levels (standardized reading test scores were not available), but still they were not reading at adult levels. The use of relatively easy, grade-appropriate stories may have allowed the children to use more advanced reading strategies. Also some of the children's inefficiencies may have been masked by their slow reading rates. Whatever the reasons, there were no significant differences in the pattern of results across grades.

<u>Conclusions</u>

The results of these experiments support an interactive model of reading. There was considerable evidence that a

strictly bottom-up model would not be adequate. Perhaps a sophisticated serial model could be developed, say along the lines of McClelland's (1979) cascade model, but an interactive model appears more promising. Readers use all available information in constructing their mental representations. They integrate meanings word by word rather than leaving the construction of a sentence representation to the end of the clause. Furthermore, the interpretation of clauses is not autonomous, even though it may appear that way at times. Children learning to read use the same sorts of processing strategies as do skilled readers. Although they process the input more slowly and require texts of appropriate difficulty, the basic structure of the comprehension processes is the same for children and adults. Although the task used here was oral reading, the conclusions can be extended to reading generally. If the secondary oral production task had any effect on the reading process, it would have been to slow the process down and make it appear more bottom-up than it is in silent reading. Since this effect was opposite to what was obtained, the generality of the interactive model is enhanced.

References

Carroll, J.M., & Bever, T.G. Sentence comprehension: A case study in the relation of knowledge to perception. In E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception: Language and speech, Vol. VII. New York: Academic Press, 1976.

Carroll, J.M., Tanenhaus, M.K., & Bever, T.G. The perception of relations: The interaction of structural, functional, and contextual factors in the segmentation of sentences. In W.J.M. Levelt & G.B. Flores d'Arcais (Eds.), Studies in the perception of language: Chichester: Wiley, 1978.

Danks, J.H. Models of language comprehension. Polish Psychological Bulletin, 1978, 9, 183-192.

Danks, J.H., & Fears, R. Oral reading: Does it reflect decoding or comprehension? In L.B. Resnick & P.A. Weaver (Eds.), Theory and practice of early reading, Vol. 3. Hillsdale, NJ: Erlbaum, 1979.

Danks, J.H., & Glucksberg, S. Experimental psycholinguistics. Annual Review of Psychology, 1980, 31, 391-417.

Danks, J.H., & Hill, G.O. An interactive analysis of oral reading. In C.A. Perfetti & A.M. Lesgold (Eds.), Interactive processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Fry, E.B. A readability formula that saves time. Journal of Reading, 1968, 11, 513-516, 575-578.

Gough, P.B. One second of reading. In J.F. Kavanagh & I.G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass.: MIT Press, 1972.

LaBerge, D., & Samuels, S.J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Marslen-Wilson, W.D. Sentence perception as an interactive parallel process. Science, 1975, 189, 226-228.

Marslen-Wilson, W., & Tyler, L.K. The temporal structure of spoken language understanding. Cognition, 1980, 8, 1-71.

Marslen-Wilson, W., Tyler, L.K., & Seidenberg, M. Sentence processing and the clause boundary. In W.J.M. Levelt & G.B. Flores d'Arcais (Eds.), Studies in the perception of language. Chichester: Wiley, 1978.

Marslen-Wilson, W.D., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.

Massaro, D.W. Primary and secondary recognition in reading. In D.W. Massaro (Ed.), Understanding language. New York: Academic, 1975.

McClelland, J.L. On the time relations of mental processes: An examination of systems of processes in cascade. Psychological Review, 1979, 86, 287-330.

Rumelhart, D.E. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, NJ: LEA, 1977.

Tyler L.K., & Marslen-Wilson, W.D. The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 683-692.

Wildman D.M., & Kling, M. Semantic, syntactic, and spatial anticipation in reading. Reading Research Quarterly, 1978-1979, 14, 128-164.

Figure Captions

Figure 1. A portion of a story for college students illustrating the different types of textual violations.

Figure 2. Disruptions in skilled readers' oral reading performances resulting from lexical, syntactic, semantic, and factual violations.

Figure 3. A portion of a story for second grade children illustrating the different types of textual violations.

Figure 4. Disruptions in second graders' oral reading performances.

Figure 5. Disruptions in fourth graders' oral reading performances.

Figure 6. Disruptions in sixth graders' oral reading performances.

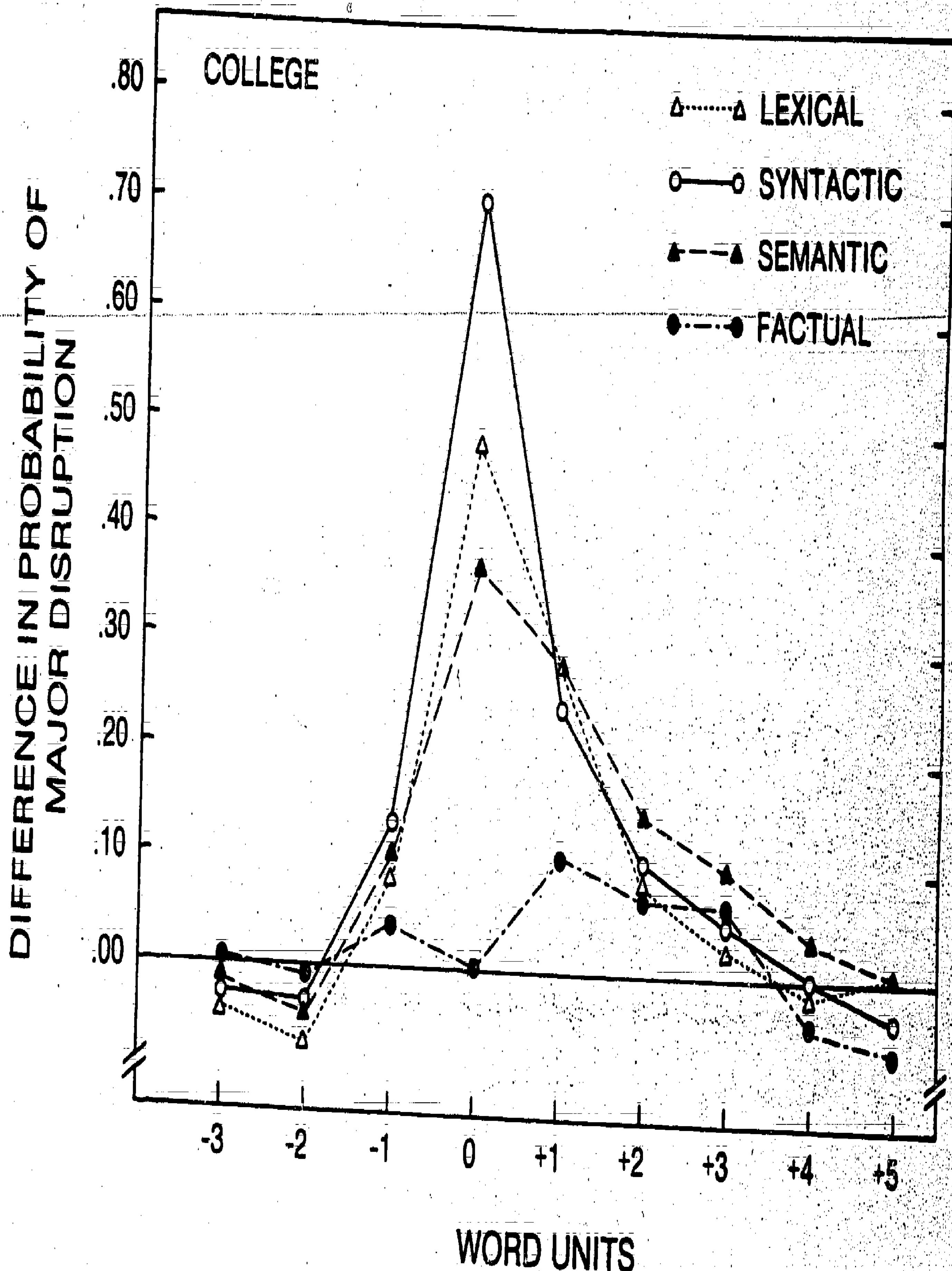
Her daughter had always been weak / strong physically. Because of this, -5 she even

-4	-3	-2	-1	CW	+1	+2	+3
imagined	her	daughter	being	injured brugen iceberg injury planted	by other	children	while

0

X

+4	+5	
trying	to get	out of the wrecked bus.


COLLEGE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS

-3

The Ducks found a fat / little stick. Because the stick

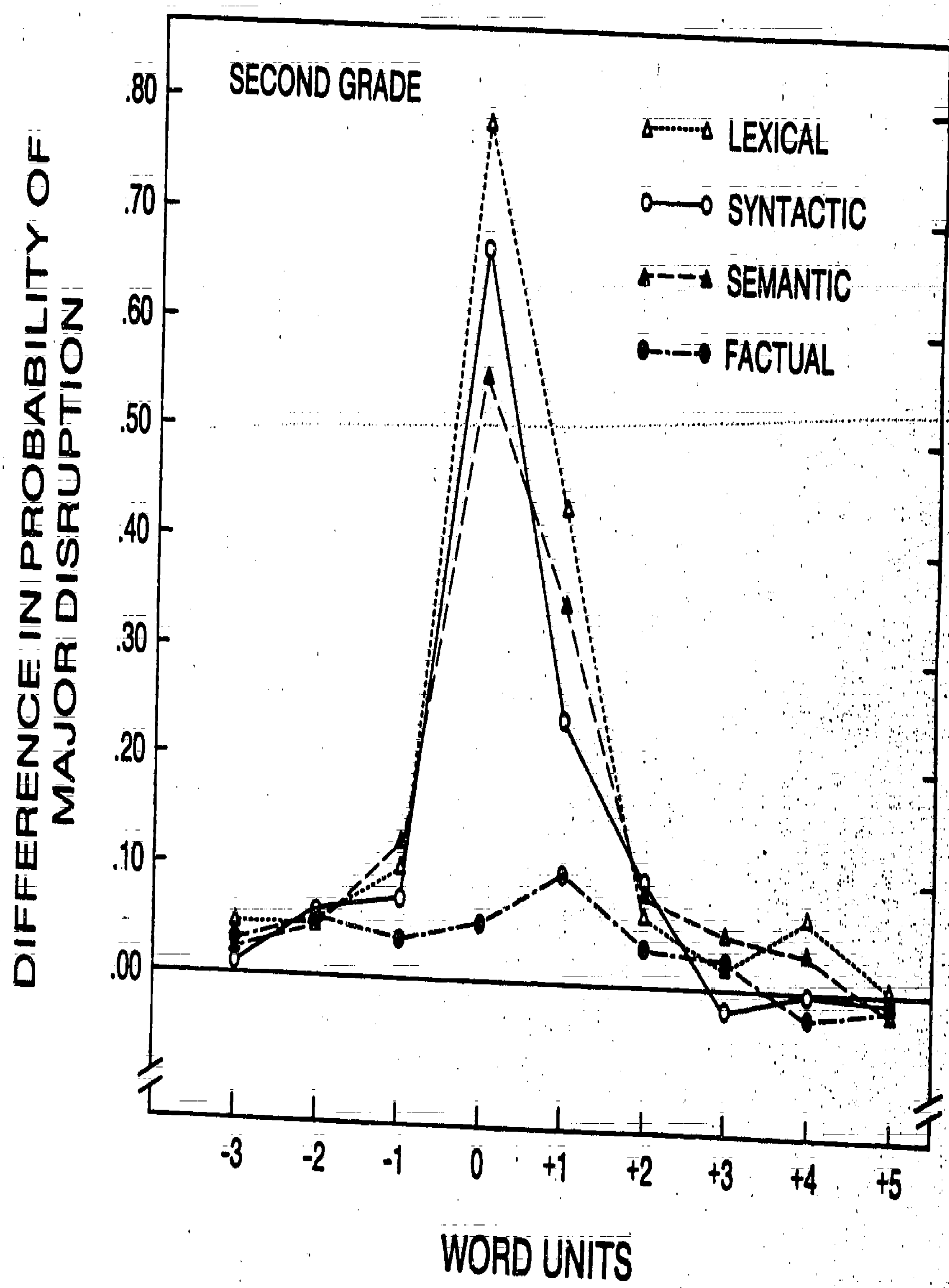
-2 -1 CW +1 +2 +3 +4 +5

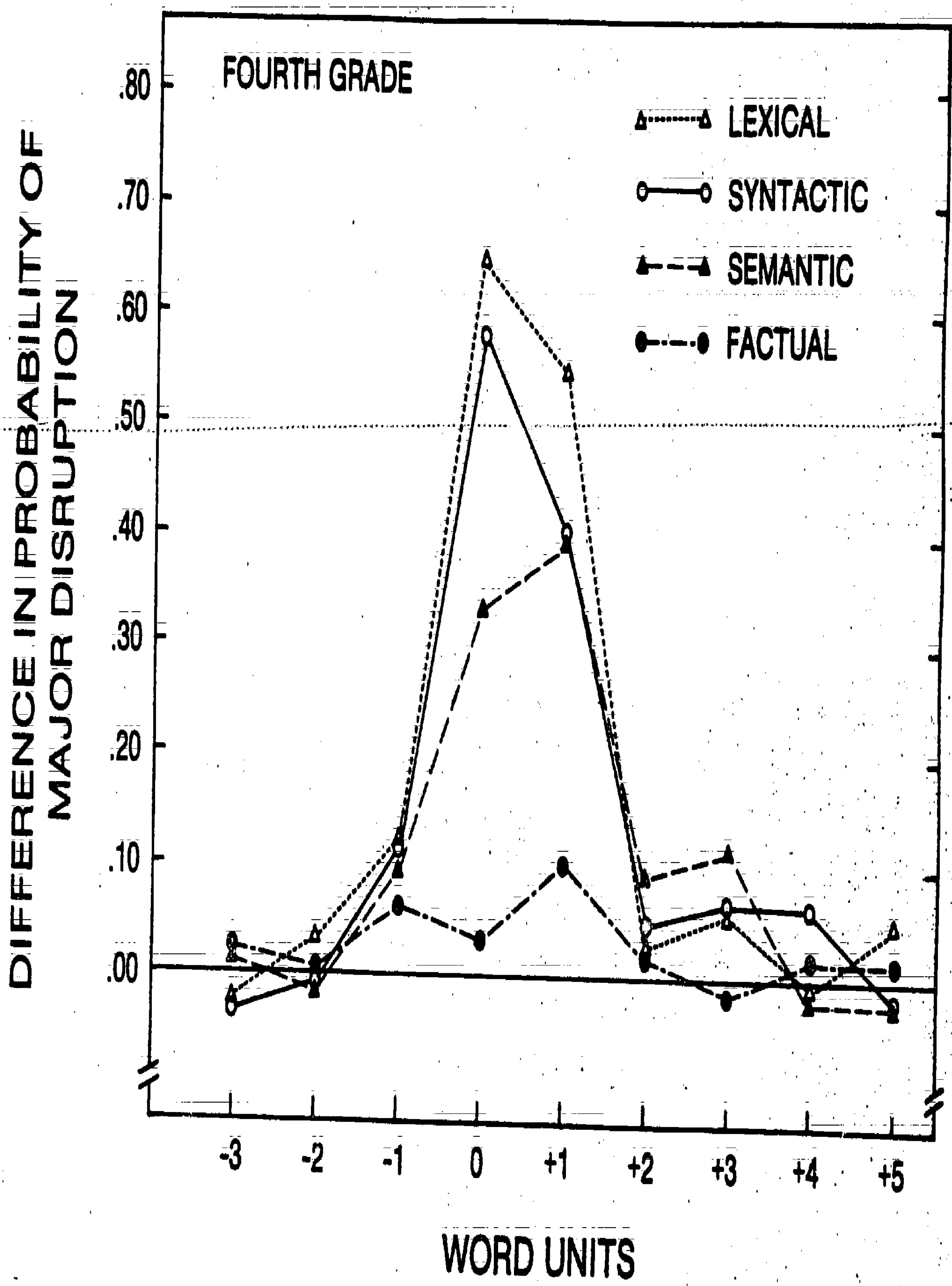
was so big / bis / biggest / mad the Turtle would be able

O

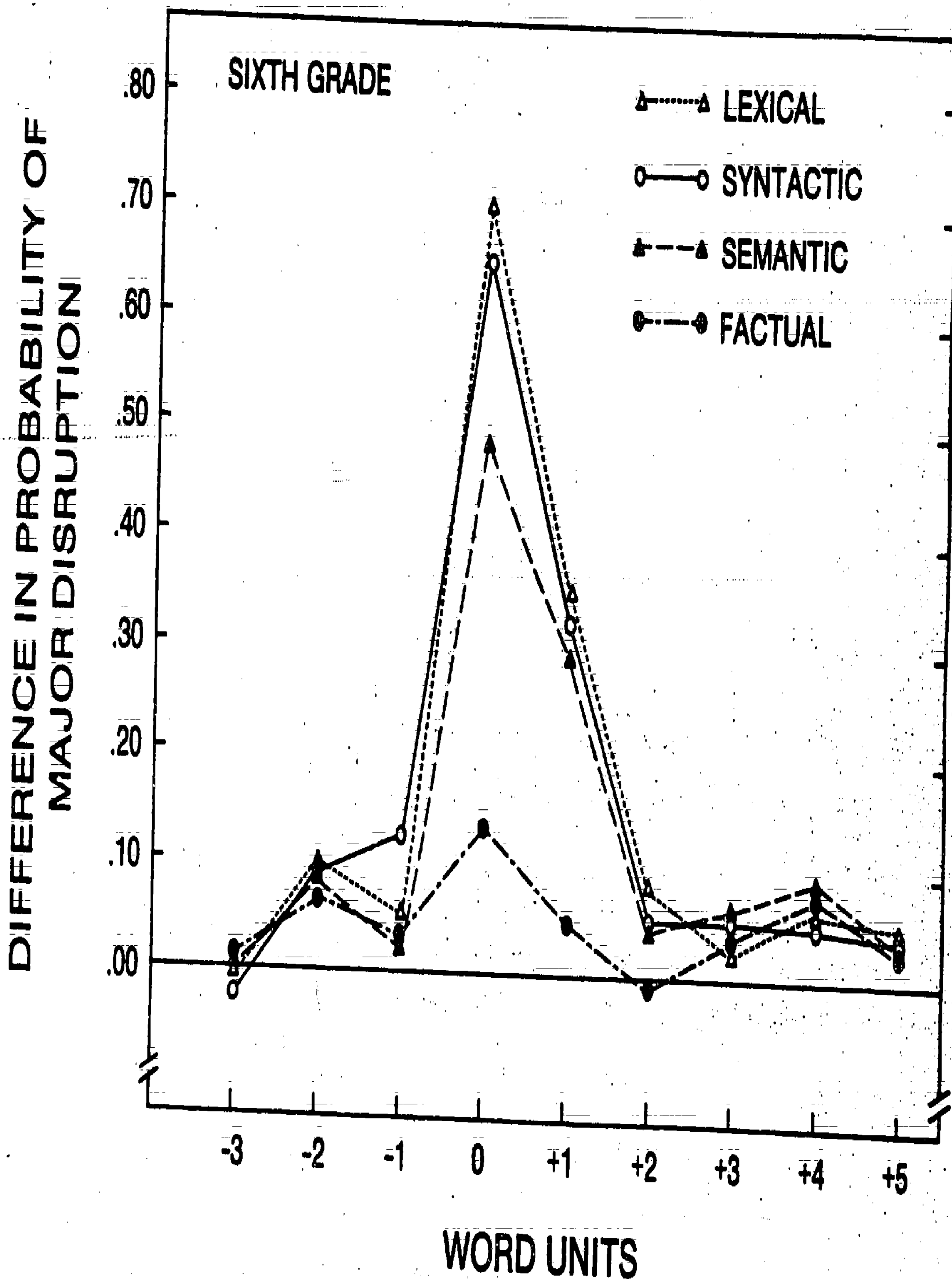
X

to hold on better. . . .


SECOND GRADE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS



FOURTH GRADE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS


SIXTH GRADE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS

Comprehension Processes in Oral Reading

Joseph H. Danks, Lisa Bohn, and Ramona Fears

Kent State University

In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), The process of language understanding. Chichester: Wiley, 1982, in press.

Mailing Address: Joseph H. Danks
Department of Psychology
Kent State University
Kent, Ohio 44242
U.S.A.

Interactive Models

Comprehension of language involves the transformation of speech or print into a mental representation of what the listener/reader thinks the speaker/writer intended. The representation arises from the interaction between the input and the knowledge base of the listener/reader. A major contemporary issue is how and when different types of information sources are used during comprehension. Many models of reading (e.g., Gough, 1972; LaBerge & Samuels, 1974; Massaro, 1975) have emphasized "bottom-up" processing in which different sorts of information are extracted in a strict serial order beginning with information closest to the physical input and ending with more abstract conceptual information.

Although a bottom-up orientation fits comfortably with our intuitive notions about comprehension, it has difficulty handling the robustness of human communication. Substantial distortions in the physical signal, for example, garbled speech or scrawled handwriting, commonly do not cause serious difficulty for listeners and readers. Listeners and readers rarely notice vague references and ambiguous phrasings. Apparently they interpret such constructions without realizing the presence of a potential problem. These casual observations of everyday phenomena have laboratory counterparts in which processing of lower-level information is facilitated by presentation in a larger context. Phonemes and letters are perceived better when presented in the context of a word or sentence (Reicher, 1969; Wheeler, 1970). Words are identified more accurately and rapidly in a grammatical, meaningful sentence than in ungrammatical or anomalous strings (Miller & Isard, 1963; Stevens & Rumelhart, 1975). Sentences that are vague and ambiguous out of context are readily interpreted in a story (Bransford & Johnson, 1973). Finally, interpretation of paragraphs is aided when the listener or reader

has sufficient knowledge about the topic to provide a schema for interpretation (Dooling & Lachman, 1971; Kintsch & Green, 1978). These phenomena result from more abstract information influencing decisions at a lower level of abstraction. Any adequate model of comprehension must account for such "top-down" effects.

The existence of top-down effects has been widely recognized (Danks & Glucksberg, 1980, Flores d'Arcais & Schreuder, this volume; Rumelhart, 1977; Wildman & Kling, 1978-1979), but precisely what they portend for models of comprehension is not widely agreed upon. A pure top-down model in which information passes serially from a meaning representation to input is implausible, if for no other reason than that the input itself does exert considerable control over the listener/reader's interpretation. One response to demonstrations of top-down effects in language comprehension was to add feedback loops such that information about decisions at more abstract levels of processing was returned to lower levels of processing. However, top-down effects are spread across the entire range of processing levels. They are not restricted to a single level, or even a few levels, of processing. So the number of feedback loops needed to account for all of the top-down effects would result in the interconnection of virtually every level of processing. In such a model, the notion of directionality of processing loses much of its force.

Several investigators have proposed interactive models of language processing (Danks & Hill, in press; Just & Carpenter, 1980; Marslen-Wilson, 1975; Rumelhart, 1977), but the details of the interactions are far from clear (cf. the papers in Lesgold & Perfetti, in press). Interactive models diminish the importance of unidirectionality of processing as a salient property of comprehension models and allow a more open system of how information is transferred

among components. Of necessity, some measure of directionality is retained because the listener/reader does start with a perceptual input and ends with a meaning representation. However, that directionality is not the overriding consideration. For example, a listener may impute meaning to a speaker's silence even though the listener has no input to process. So determining the directionality of information flow may not be as critical for building a model of language processing as had been assumed previously.

A problem with most interactive models is that they are too powerful; that is, they do not exclude any conceivable results. For example, most interactive models are capable of explaining bottom-up and t[illegible]ffects separately or in combination. Any model, and an interactive [illegible]fically, must be sufficiently precise that it can be falsified by some [illegible]eivable set of data. This requirement is complicated by the fact that language processing is extraordinarily flexible, so any model of language processing has to be very powerful. However, models cannot be allowed to grow too powerful or nothing is explained.

Properties of Interactive Models

In this section, several general properties of interactive models are described. Not all interactive models accept all of these properties, but each property can be identified in one or another interactive model.

1. Results of processing at any level are immediately and simultaneously available to all other levels. A model incorporating the most powerful variant of this property was proposed by Rumelhart (1977), in which all information resulting from processing at each level is deposited in a message center where it is available to all other processing levels. For example, as soon as a sequence of phonemes is identified, that information is available for use in lexical access. In some cases, a word might be identified before all its

phonemes are identified (Marslen-Wilson & Welsh, 1978). For another example, information about the discourse topic may aid identification of specific words (Foss & Ross, this volume).

The power of interactive models can be limited either by restricting the availability of information or by restricting the types of information specific levels are capable of processing. For example, only phonetic information might be used in lexical access to activate candidates for identification, but then syntactic, semantic, and thematic information could be used to select among candidates (Marslen-Wilson & Welsh, 1978). A variation of this property that permits retention of some bottom-up characteristics is to allow intermediate results from one processing level to be available to other levels before the first level has finished (cf. McClelland's cascade model, 1979, and Perfetti & Roth's reading model, in press). Thus, more interaction among levels of processing is permitted than with a strictly serial, autonomous model.

2. Processing proceeds at all levels in parallel constrained only by the availability of information on which to operate (Just & Carpenter, 1980; Marslen-Wilson, 1975; Rumelhart, 1977). Any given level of processing does not necessarily wait until processing has been completed at any other level. Any level is potentially active at any point in time.

One constraint on this property is that sometimes processing at one level cannot proceed without information from some other level. For example, construction of sentence meaning cannot proceed without identification of at least some of the words in the sentence. Listeners and readers may have a good idea about what speakers and writers are going to say and write, but they do not enjoy perfect prescience. A second constraint results from differences in rates of processing among processing levels. Some processes proceed automatically because they are so well practiced. For example, it is very difficult to

inhibit meaning access because determination of meaning is the primary goal of communication. In the Stroop task, the printed word interferes with naming the color of the ink. Lexical access of common words is so automatic that it is impossible for skilled readers to inhibit it. Likewise, in an oral reading task, sentence meanings may be constructed even though only oral production is required because sentences are understood faster than the words can be spoken (Danks & Fears, 1979).

3. Interactive models posit processing flexibility as a function of individual differences among listeners and readers, as a function of the listener/reader's purpose or task, and as a function of properties of the text (Danks, 1978). Comprehension, in either the listening or reading mode, is not a fixed, invariant process, but adapts to the specific situation. This adaptability represents the normal mode of processing. It is not something that happens just when processes have difficulty. Because there are so many things that can go wrong, the listener/reader never has the opportunity to develop one canonical process that can function effectively in most situations.

Some listener/readers are more skilled than others at specific types of processing. For example, some readers have excellent word recognition skills, and others are more adept at determining the meaning of sentences (Perfetti & Roth, in press). One obvious source of individual differences in processing skill is age. Children, in general, have different sorts of processing strategies than adults do (Bever, 1970), especially for reading (Schwartz, 1980). Skilled readers have developed some automatic processing strategies (LaBerge & Samuels, 1973). In contrast, children learning to read frequently have to attend more closely to each stage of processing.

Task differences are well recognized as contributing to differences in comprehension strategies (Aaronson & Scarborough, 1976; Danks, 1969; Frederiksen,

1975; Glucksberg, Trabasso, & Wald, 1973; Mistler-Lachman, 1972). Typically, one listens or reads with a purpose in mind. The purpose can direct attention to those processes that, when completed, will satisfy the task requirements. In addition, once task demands have been met at one level, processing at all levels may be terminated (Mistler-Lachman, 1972). However, there are limits to the amount of control that is possible because some processing levels proceed automatically or more rapidly than the processing at levels at which the task demands are met. For example, sentence comprehension usually proceeds more rapidly than does oral production in an oral reading task (Danks & Fears, 1979).

The type and structure of the discourse can alter processing strategies. With a difficult text, a reader may be more dependent on lower-level information being fed forward to higher levels of processing. With a simple text, in contrast, expectancies can be generated at abstract processing levels such that lower-level processing is facilitated or even short-circuited. In addition to variations due to text difficulty, processing changes according to the type of discourse. At a casual party, conversation may be processed only to the extent needed to make a polite reply, especially when the listener is more interested in a conversation across the room. With respect to print, reading a poem, a novel, a newspaper column, and a technical article yield different processing strategies, at least subjectively (Gibson & Levin, 1975). Most poetry keys on the sound properties of words, rhyme, rhythm, or word "color." Hearing a poem in one's mind's ear is crucial for understanding the poet's message, but hearing a novel would get in the way of understanding its theme. Technical articles require considerable conceptual processing; how the text sounds is not nearly so important.

4. An interactive model may describe interactions among types of information or interactions among processing components. First, consider inter-

actions among different types of information. A single processing component may have access to all types of information, but one type may be more useful than others. For example, perceptual information is more valuable for lexical access than is thematic information. In contrast, thematic information is more important for the construction of a story's macrostructure than any specific piece of perceptual information. If the information most typically used by a process has been distorted, violated, or is otherwise unavailable to the listener or reader, then another type of information may compensate for the deficiency (Stanovich, 1980). One reason for information unavailability is that something is wrong with the text. There may be physical distortions, or a story may be written in a vague, obscure, or metaphorical style (Bransford & Johnson, 1973). The lack of specific kinds of information can also be due to listener/reader factors. For example, children do not have as well developed story schemata as do adults (Poulsen, Kintsch, Kintsch, & Premack, 1980) and thus depend on bottom-up information to understand stories. Some less-skilled readers have sufficiently poor word recognition skills that their ability to comprehend sentences and stories is impaired (Perfetti & Lesgold, 1977). So a poor reader is more dependent on contextual information than is a good reader who has well developed word recognition skills (Stanovich, 1980). This compensatory mechanism gives a processing component considerable flexibility since it is not dependent on a single source of information.

A second way to formulate interactive models is through interactions among processing components. When processes make use of information resulting from other processes, the information-generating process influences the information-receiving process. If the information comes from a more abstract process, then the influence is top-down; if the information comes from a less abstract process, then it is bottom-up. Interactions among processes may be

complete or limited. If they are complete, then one process can influence another simply by generating information needed by the other. If interactions are limited, then there may be sectors of processing within which processes interact, but information transfer from one sector to another is restricted. For example, lexical access uses information from many levels of processing, but macroprocessing accepts information primarily from sentence comprehension.

An interactive model can be formulated in terms of what kinds of information interact in a single process or what kinds of processes influence other kinds of processes. These two types of formulations are complementary because different types of information are the result of different types of processing. So it is a question of whether the emphasis is on processing components or on types of information.

Processing Components

There are three major processing components in language comprehension: lexical access, sentence comprehension, and discourse understanding. Lexical access involves locating a lexical item in the mental dictionary and selecting an appropriate meaning. Bottom-up perceptual information, auditory and visual, is important for identifying a word; however, top-down contextual information, syntactic, semantic, textual, thematic, and factual, also influence lexical access. How lexical access might work in speech perception has been described by Marslen-Wilson and Welsh (1978). A cohort of potential words is activated by preliminary auditory analyses. In addition to continued processing of perceptual information, checks for consistency with contextual information are used to eliminate candidates from the cohort until a single item remains. Extending Marslen-Wilson and Welsh's model to lexical access in oral reading, Danks and Hill (in press) have suggested that top-down information may be used in the selection of the initial cohort as well as in eliminating candidates.

Forster (1976; cf. Bradley & Forster, this volume) has proposed a more strictly bottom-up model, in which a word is identified using only perceptual information and then is evaluated for contextual appropriateness. In interactive models, information from multiple sources converges in lexical access producing both bottom-up and top-down effects. Since lexical access provides the articulatory information that is needed to pronounce a word, any interaction among information sources will be reflected in oral reading and speech shadowing tasks.

In sentence comprehension, the listener or reader integrates the word meanings into a representation for the entire sentence. Syntactic structure is available to guide the integration, but how active a role it plays is not clear. There are two primary hypotheses about how the sentence comprehension component works. The clausal processing hypothesis (Carroll & Bever, 1976; Carroll, Tanenhaus, & Bever, 1978) proposes three autonomous steps in its simplest version. First, the meanings of words are stored in a working memory buffer as they are accessed. When the end of a clause is reached, a representation is derived for the clause. Finally, the representation of the clause is integrated with representations of prior clauses and with prior knowledge. The primary characterization of the clausal processing hypothesis is that it is serial, bottom-up, and autonomous. In contrast is the word-by-word processing hypothesis (Marslen-Wilson, Tyler, & Seidenberg, 1978; Marslen-Wilson & Tyler, 1980, this volume; Tyler & Marslen-Wilson, 1977). This listener/reader attempts to integrate each word's meaning into a comprehensive representation as soon as it is accessed. If immediate integration fails, the word's meaning is held in a memory buffer until integration is possible. Frequently, the most appropriate point at which to reattempt that integration is at the end of the clause. Interactive models imply a word-

by-word comprehension strategy as opposed to a clausal processing strategy. Because processing proceeds in parallel in all levels, each word is processed immediately to the maximal extent possible. For the clausal level, that means attempting to integrate the word's meaning into the composite representation at the time that it is accessed.

In discourse understanding, the listener/reader organizes the representations of individual sentences into discourse structures corresponding to the schemata of conversations, lectures, stories, and nonfiction prose. As sentences are comprehended, a discourse structure is constructed that is updated as additional information is received. At any point in time the representation is as complete as possible. One aspect of discourse understanding is to establish referential coherence for the discourse (Clark & Marshall, 1978). Listener/readers tie together the sentences of a discourse into a coherent representation, or text base (Kintsch & van Dijk, 1978), making bridges and inferences as needed (Carpenter & Just, 1977; Crothers, 1979; Warren, Nicholas, & Trabasso, 1979). In addition to local coherence, listener/readers derive a macrostructure, a schematic representation of the main ideas or gist of the text (Kintsch & van Dijk, 1978). Interactive models permit considerable processing flexibility in discourse understanding. Listeners and readers are able to adapt the processing strategies and knowledge they have available to the wide variety of texts that they encounter.

The emphasis in interactive models is on processes rather than on representations. In recent years, cognitive scientists generally have shifted away from attempting to specify representational structures to identifying the cognitive processes that produce those structures (Danks & Glucksberg, 1980). A change in the kinds of experimental tasks has attended this shift. Memory tasks provide information about mental structures and representations,

but little information about processes. Hence, investigators have developed other tasks, such as phoneme and word monitoring, speech shadowing, oral reading, and eye movement monitoring, that yield on-line measures of processing, dependent measures that are recorded concurrently with processing.

Experiments on Oral Reading Comprehension

Although oral reading is used frequently in schools to evaluate reading (Durkin, 1978-1979), its processing requirements are not fully understood (Danks & Fears, 1979). In contrast to silent reading, the dominant task demand in oral reading is that each word be uttered in serial order. Oral reading yields a continuous on-line response that is roughly contemporaneous with the visual input and with comprehension. It is the reading counterpart to speech shadowing. The main difference is that in shadowing the listener does not control the order and rate of input as the reader does, but this difference corresponds to the difference between listening and reading generally.

The oral reader's primary task is oral production. To accomplish this task, each word is located in the mental dictionary and the articulatory information is used to pronounce it. A reader potentially could use lower-level information, such as grapheme-phoneme correspondences, spelling patterns, or syllabic structure, to pronounce a word without accessing it. However, since readers do not read pronounceable nonwords unhesitantly, dependence on lower-level information is unlikely. When oral reading is followed by a comprehension test, the reader also needs to understand the phrases, sentences, paragraphs, and main ideas of the text. Information from several levels must be integrated to construct a reasonable interpretation. During oral reading, the reader is attempting to satisfy both the verbal performance and the comprehension demands at the same time, but the press of the perfor-

mance demand is greater. Analysis of oral reading performance provides an excellent opportunity to study lexical access, sentence comprehension, and discourse understanding in a relatively natural situation.

In the three experiments reported here, we investigated what kinds of information are used by the lexical access and sentence comprehension components. An oral reading task also permitted an estimation as to when the different types of information were being used. Specifically, the point in time when different types of information were processed was assessed by violating each information type. If that information were normally used in oral production or in comprehension, then oral performance would be disrupted because the normal interplay among processing components would be modified to compensate for the violation. Furthermore, the disruption would be temporally close to when the violated information was needed by the reader. Our basic method was to change several critical words in a story, such that one or more types of information was violated. We then analyzed readers' oral productions for disruptions near each critical word. The relative position of the disruptions resulting from the different violations indicated the order in which the information was typically used.

In three experiments, we manipulated various combinations of lexical, syntactic, semantic, and factual information. The first experiment established the basic pattern of results for syntactic, semantic, and factual (intersentential) information. The second experiment separated syntactic and semantic information types and replicated the results for factual information. The third experiment tested lexical, syntactic, and semantic information and added a global text factor, namely, the critical words were embedded in a difficult metaphorical story. The question raised by the text factor is whether information utlization changes when higher-level

information is obscured. Specifically, does the reader become more dependent on bottom-up processing strategies when top-down information is less available?

A bottom-up model predicts that disruptions resulting from having violated different types of information would be ordered from earliest to latest according to the level of abstraction of the violated information. The least abstract information, that was lexical information in our experiments, would produce the earliest disruption. Disruptions from syntactic and semantic violations would appear next, followed by disruptions from violating factual information, the most abstract in our experiments. Only lexical information is strictly necessary for oral production although a reader would need to determine the syntactic structure to read with appropriate prosody. So a lexical violation would produce a disruption near the critical word. Syntactic and semantic information would be used primarily at the end of the clause when the meaning of the clause is determined, so disruptions from these two violations would occur near the clause boundary. Violations of factual information would be important only after the meaning of the clause had been determined and the reader attempted to integrate the meaning of the clause with the representation of the preceding text.

Interactive models posit that several types of information are used in the same component, and so their violations would produce similar disruption patterns. Such models also permit more than one component to operate at the same time, so that different information types might be used at the same time by different components. As discussed earlier, part of the difficulty with interactive models is sorting out these two types of interaction. Both types of interactive models predict that violations of different information types results in disruptions that occur at the same time. Other aspects of the

data can be used to narrow the range of possible interactive models. For example, to the extent that the size and range of disruptions, as well as their timing, are similar, the involvement of only one processing component is more likely. The kinds of oral reading errors provide another source of evidence; for example, fluent restorations (cf. Experiment 2) reflect the operation of the lexical access component.

In general, interpretation of results depends on the relative positions of disruptions across violations, not on the absolute location of a disruption. So if violation of one information type produces a disruption before another violation type, utilization of the two information types is temporally ordered as well, regardless of the absolute positions of the disruptions. In some cases, however, the absolute position is interpretable; for example, disruptions that begin after a critical word has been uttered cannot reasonably involve lexical access. The absolute location of a disruption may reflect in part the eye-voice span, that is, the distance between where the eye is focussed and the word being uttered (Levin, 1979). But the relative positions of disruptions is not compromised because the size of the eye-voice span can be assumed to be relatively constant for all manipulations on average since all violations occurred equally often in each critical word segment. The size of the eye-voice span may vary systematically in other comparisons, such as in the results from different readers, for example, children at different levels of reading skill, or results from different texts, for example, the stories used in these experiments. In these cases, the effect of possible changes in the size of the eye-voice span must be considered.

Experiment 1

Method. A story about a high-school girl who was severely injured when a train hit her school bus was adapted from a popular magazine. The story

was rewritten to eliminate all conversation and any difficult or infrequent words. The final story was 2171 words long and its readability was 7.8 (Fry, 1968). It was divided into four sections of approximately equal length. Four critical words were selected in each section, separated by an average of 129 words. The critical sentences for all experiments are contained in the Appendix.

There were four manipulations of the critical words -- three types of violations plus a control condition. In the following sample segment, "Her daughter had always been weak physically. Because of this, she even imagined her daughter being injured by the other children while trying to get out of the wrecked bus," the critical word was injured. In the CONTROL condition, there was no change in the critical word or in the surrounding text. To produce the SYNTACTIC + SEMANTIC violation, both syntactic and semantic information were distorted by replacing the critical word with a word that was the incorrect part of speech and that was semantically anomalous. The critical word injured was replaced with iceberg. To violate SEMANTIC information without disturbing lexical or syntactic information, the critical word was replaced with a word that was semantically anomalous, but which was the correct part of speech, for example, planted replaced injured. Readers can still determine the grammatical structure of the sentence, but the meaning of the sentence is distorted. At best, they have to imagine some very unusual circumstances in which the anomalous word can be interpreted metaphorically. FACTUAL information is accumulated from the preceding text, so it was violated by introducing an inconsistency between the critical word and the preceding sentence. Unlike the other violations, neither the critical word nor the sentence containing it was altered. The sentence immediately preceding the sentence with the critical word was modified such that the critical word was factually incon-

sistent with the information of that sentence. In the example, the word weak in the preceding sentence was replaced with strong. The fact that her daughter was strong was inconsistent with the mother worrying about her being injured. There always existed a plausible substitution for the critical word; otherwise, the reader might sense something was amiss prior to reading it. For example, if the daughter were strong, the mother might imagine her being safe, unharmed, or helpful. All modifications were selected such that there was no plausible continuation following the critical word that would resolve the inconsistency.

Four versions of the story were constructed such that each violation occurred at each critical word segment in one version. There was one violation of each type in each section of the story. Each section of the story was typed starting on separate pages. Critical words did not occur in the top or bottom three lines on each page and were at least three words from the beginning and end of lines.

Subjects were 11 male and 29 female undergraduate students enrolled in general psychology courses at Kent State University. They received points toward their grades for their participation. All were native English speakers and were not screened for reading ability. Subjects for all experiments had not participated in any other oral reading experiment. Four experimental groups of 10 subjects were defined by the four versions of the story.

Subjects were tested individually. They were told that the purpose of the experiment was to examine the relationship between reading and comprehension. They were instructed to read each section aloud, and then to write a summary of it. They were given as much time as they needed to read and to summarize the story. The reading performances were tape recorded for later analysis. In order to provide some warm-up for the readers, the first critical word did not

occur until the bottom of the first page.

Analyses. In fluent speech, each word is not spoken with clearly distinguished beginning and ending sounds as it would be spoken in isolation. Ending sounds of one word are blended with the beginning sounds of the next making it difficult to mark precisely where one word ends and another begins. The result is that two words may be uttered as if they were one long word with no break separating them. At the other extreme, some words are pronounced with a break in the middle, depending on the particular configuration of phonemes. Because of these possibilities, the text surrounding each critical word was divided into word units in order to facilitate measurement of disruptions. Word units were specified by listening to several readers and dividing the text surrounding the critical words into groups that were pronounced as a unit. The most consistent phraseology across readers was adopted. Word units typically consisted of one or two words, rarely three words, and averaged 1.54 words long. They did not necessarily follow the syntactic structure of the sentence. Five word units before and five word units after each critical word were identified and were numbered from -5 to +5.

The primary dependent variable was the production time for each word unit. These times were measured by slowing the tape recorder to half-speed. An experimenter then pressed a key at the end of each word unit. A lab computer monitored the key presses and timed the latencies between them. Each interval included the production time for the word unit itself and any pause, hesitation, or filler words that preceded the word unit. Since the critical word physically changed between the control, syntactic + semantic, and semantic conditions, the production times for the critical word itself were not comparable. An additional latency was measured from the end of word unit -2 to the beginning of the critical word. The production time for word unit -1 was subtracted

from this additional latency. The difference corresponds to the time between the end of word unit -1 and the beginning of the critical word. This value was used as the production time for the critical word unit in all analyses.

It was impossible to have an experimenter who was blind to the experimental manipulations measure the production times because any English speaker would recognize the violations on hearing the taped protocols. In order to assess the extent of experimenter error in measuring the production times, inter- and intra-experimenter reliabilities were obtained. Seven randomly selected subjects' protocols were measured by a second experimenter. The latencies were correlated with those from the first experimenter, yielding an average correlation of .94. The same experimenter retimed four randomly selected subjects' protocols from Experiment 3 (reported below) about one month later. The average correlation between the two sets of measurements was .98. Finally, a spectrographic analysis of eight randomly selected subjects' protocols from Experiment 2 (reported below) was prepared. The relevant production times were measured from the sound spectrograms and correlated with the experimenter's timing; the average correlation was .91. Marking word boundaries on spectrograms is far from precise, but the source of the error is visual uncertainty in contrast to the auditory uncertainty of the experimenter's timing. Since the error sources are different, they would tend to attenuate the correlation. Thus, the procedure for measuring production times was reliable.

The production times were analyzed with a mixed analysis of variance. Groups of readers, as defined by the four versions of the story, was a between-subjects factor. Type of violation, word-unit position around the critical word, and section of the story were within-subject factors. Versions, violations, and segments were arranged in a Latin-square. This design permitted calculation of a quasi-F ratio (F'), in which both subjects (individual

readers) and language materials (critical word segments) were random factors contributing to a single error term (Clark, 1973). All reported effects were significant with $p < .05$. The three experimental means were compared to the control mean at each word-unit position using individual planned comparisons (Winer, 1971), based on the quasi-F mean square error term. The planned comparisons comprized less than 4% of the possible comparisons.

A second dependent variable was the probability of a major disruption at each word-unit position. Major disruptions were defined as pauses of at least one second duration, substitutions, omissions, reversals, stammerings, mispronunciations, repetitions, and regressions. In short, any deviation from fluent oral reading that indicated that the reader noticed a violation was scored as a major disruption. Only one disruption was tallied per word unit and the frequencies were converted to probabilities. The major disruption data were used to confirm the results of the production times and to provide qualitative information abc :t the disruptions. In Experiment 3, the correlation between production time and major disruption means was .93. Production times were lengthened by the major disruptions, as well as by a general slowing of oral production. Yet there is no reason to attempt to separate these influences because both reflect a disruption of the underlying cognitive processes.

Results. The mean production times as a function of violation type and word-unit position are presented in Table 1. Since word units differed markedly in physical size, only the differences in production times between the experimental and control conditions were interpretable. These differences are depicted in Figure 1. The effect of primary interest was the significant in: between violation type and word-unit position, F' (30, 240) =

5.029, $p < .001$. A difference of 62 msec. in mean production times was significant.

Insert Table 1 and Figure 1 about here

Both the syntactic + semantic and the semantic violations diverged from the control condition at word unit -1. The syntactic + semantic disruption peaked at the critical word, but the semantic disruption did not peak until word unit +1. Both disruptions continued to stay above the control through word unit +3 and also were significant at word unit +5. The factual inconsistency produced a much smaller effect, and was significantly different from the control only at word units +1 and +2. The mean probabilities of a major disruption confirmed the results of the production time analyses in all respects.

Since both syntactic + semantic and semantic violations yielded lengthened oral production times before the critical word was uttered, both syntactic and semantic information evidently was being used in lexical access because words had to be accessed before they could be uttered. Both violations also were disruptive well after the critical word was produced suggesting disruption of sentence comprehension. The factual inconsistency was disruptive only after the critical word had been uttered. So violation of factual information did not hinder lexical access, but it did affect sentence comprehension and story understanding.

The syntactic + semantic violation had a larger disruptive effect earlier than did the semantic, suggesting that the syntactic + semantic violation may have been a violation of two independent knowledge sources. If so, then confounding syntactic and semantic violations in a single manipulation can account for the greater disruption and the earlier peak of the

syntactic + semantic violation relative to the purely semantic violation. If two independent information sources were violated, the likelihood that a violation would be noticed in one of them is greater than if only one information source were violated.

Experiment 2

Experiment 2 was similar to Experiment 1. The story and critical words were the same. The syntactic + semantic violation was replaced with a syntactic only violation and the semantic violation was retained. This change permitted a direct comparison of whether syntactic information was used at the same time as semantic information and whether they were independ. .t sources of information.

Method. The story used in this experiment was identical to that used in Experiment 1 with the following exceptions. To violate SYNTACTIC information alone, the root morpheme of the critical word was retained, but the inflection was changed to that of a different part of speech. For example, the verb injured was changed to the noun injury. Although some semantic information is carried in syntactic categories, the reader could determine the intended meaning relatively easily. Several of the factual violations were rewritten so that the inconsistencies seemed more striking, at least to the intuitions of the investigators. Only four word units before the critical word were scored and the slightly altered word units averaged 1.60 words long.

The subjects were 17 male and 23 female undergraduate students enrolled in general psychology courses at Kent State University. All were native English speakers and were not screened for reading skill. The procedure was identical to that used in Experiment 1 with one exception. Instead of asking for summaries, multiple-choice questions were prepared for each section

of the story. These questions tested literal, factual information that was unrelated to the critical word segments. Readers were very accurate on these questions (over 95% correct) and there was no variation across sections or versions.

Results. The mean production times as a function of violation type and word-unit position are presented in Table 2. The differences in mean production times between the experimental and control conditions are shown in Figure 2. The critical interaction between violation type and word-unit position was significant, F' (27, 216) = 2.891, $p < .001$. / A 71 msec. difference in means was significant.

Insert Table 2 and Figure 2 about here

The syntactic and semantic violations both produced increased production times beginning at word unit -1 and continuing through word unit +3. The curves followed each other almost exactly, peaking at word unit +1 (the same as the semantic disruption in Experiment 1), except that the syntactic disruption declined faster at word units +2 and +3. The pattern of results from the probability of a major disruption analyses confirmed the production time results. The syntactic violation produced a major disruption at the critical word a large percentage of the time relative to other violations -- syntactic = 78%, semantic = 44%, and factual = 8%. Half of these syntactic major disruptions were restorations of the correct form of the base word with the syntactically correct inflection. For example, "injured" was uttered when injury was printed. Sixty-eight per cent of the restorations were fluent ones in that there was no pause or other disruption immediately before or during production of the restoration. If restorations are excluded from both the syntactic and semantic disruptions, the percentages of disruptions

at the critical word were virtually identical -- 33% for syntactic and 34% for semantic. The restoration of the original critical word in the syntactic violation condition was a top-down effect. Apparently, readers were attempting to make sense of the text, so the original critical word was anticipated and restored. Readers substituted the original critical word in their productions because that was the word they were anticipating and the first several letters of the printed word confirmed those expectations. This restoration effect is very similar to that obtained by Marslen-Wilson and Welsh (1978) in a speech shadowing task.

The disruption from the factual inconsistency was larger although it still was not as large as that produced by the syntactic and semantic violations. In comparison to Experiment 1, the factual disruption was spread over three word units (+1 through +3) instead of two and the peak was approximately 60% higher. Even though the factual violation produced a larger effect than in the preceding experiment, it still did not influence lexical access of the critical word. Its effect on sentence comprehension increased, but whether that effect occurred at the clause boundary or at an earlier point cannot be determined directly from these data.

Experiment 3

The story used in the preceding experiments was interesting and easy to understand in spite of the violations. It had good coherence among sentences and a macrostructure seemed easy to construct. How would processing change if the story were disjointed and if the events were strange and metaphorical? Many studies have shown that disordered and scrambled stories are difficult to understand and recall (Kintsch, Mandel, & Kozminsky, 1977; Mandler, 1978; Meyers & Boldrick, 1975; Stein & Nezworski, 1978). For this experiment, we selected unrelated paragraphs from a novel that described difficult to under-

stand happenings written in a highly metaphorical style. Because of this style, a reader would have difficulty constructing a coherent text base. In addition, it would be extremely difficult to construct a macrostructure incorporating all the paragraphs because they were not linked in the novel. So there would be less contextual information to aid lexical access and to guide sentence comprehension. In general, a reader would be more dependent on bottom-up processing to compensate for the reduced discourse information.

In addition to removing discourse information, we introduced a direct violation of lexical information by replacing the critical word with a pronounceable nonword. By definition, lexical access would be impossible because the nonword was not in the reader's mental dictionary. If the reader were using lexical information to pronounce the word, there would be a disruption in oral performance. But if the reader were relying solely on grapheme-phoneme correspondences to pronounce the word, there would be no disruption.

Method. Fifteen paragraphs averaging 124 words long were selected from a contemporary novel. The passages were not contiguous in the novel, but the temporal sequence was maintained. The paragraphs seemed to the investigators to be much more vague, metaphorical, and difficult to follow than the story used in the first two experiments. However, the readability of these paragraphs was 6.0 (Fry, 1968), nearly two grade levels less than the story. This contrast reflects the fact that standard readability formulae do not measure discourse and conceptual properties of texts (Kintsch & Vipond, 1979). The first three paragraphs were used for practice. In each of the remaining 12 paragraphs, one critical word was identified that was not near the beginning or the end of the paragraph. Four nouns, four verbs, and four adjectives were selected as critical words. Each critical word was changed to form a lexical, syntactic + semantic, or semantic violation, or it was left unchanged

as a control. For the LEXICAL violation, the critical word was replaced with a pronounceable nonword that followed the rules of English orthographic structure. In the following example, the critical word cool was replaced with brugen (lexical), flew (syntactic + semantic), or sharp (semantic):

When that was done, she pointed to his knee and removing the binding, spread the cool paste thickly on the swollen flesh.

There were four different versions of each paragraph representing each experimental manipulation. Critical words from other paragraphs were inserted to create the syntactic + semantic and semantic violations. The occurrence of critical words was counterbalanced such that no subject saw a specific critical word twice except in the control condition. The four versions of each paragraph were assembled into four presentation sets. Within each set, each manipulation occurred three times, once as a noun, verb, and adjective. The sets were complementary so that each manipulation occurred once in each paragraph across versions. The paragraphs were typed on separate pages and arranged so that the critical words did not occur near the beginnings or ends of lines.

The subjects were 13 male and 27 female undergraduate volunteers from general psychology classes at Kent State University. All were native English speakers and were not screened for reading ability. The procedure was the same as that used in the first two experiments except that the reader orally summarized each paragraph after reading it. The production times before the critical word were measured initially in a slightly different way than in the two preceding experiments (cf. Fears, 1978). The times and analyses presented here have been adjusted to conform to the previously described pattern. Otherwise, the analyses were the same as in the first two experiments.

Results. The mean production times as a function of violation type and word-unit position are presented in Table 3. The differences between the

production times of the experimental and control conditions are shown in Figure 3. The interaction between violation type and word-unit position was significant, F' (27, 160) = 3.899, $p < .001$. A difference of 59 msec. between means was significant.

Insert Table 3 and Figure 3 about here

Both lexical and syntactic + semantic violations lengthened production times one word unit before the critical word (word unit -1). The disruptions peaked at the critical word and continued through word units +2 and +4. Most of the disruptions from lexical violations were pauses as the readers balked before uttering the pronounceable nonword. Since the nonword was not in their mental dictionaries, they had to pronounce it solely on the basis of grapheme-phoneme correspondences. The curve for the syntactic + semantic disruption followed the curve for the lexical disruption almost perfectly, suggesting that syntactic and/or semantic information was used in lexical access. By [illegible] of the way the lexical violation was constructed, it disrupted the lexical access component. Since the syntactic + semantic violation produced an identical pattern of disruption, syntactic and/or semantic information apparently was being used by the same component. This result provides additional support for the conclusions from Experiments 1 and 2 that syntactic and semantic information was used in lexical access.

Production times at word unit -2 were significantly faster than the [illegible] in all three violation conditions. A possible explanation is that readers may have been processing the violation at this point since their eyes were several words ahead of their voices. They may have speeded up their oral productions in order to [illegible] the violation and to have more time to resolve it, but there [illegible] evidence for this speculation.

The disruption from the semantic violation did not occur until word unit +1 and was significantly smaller than the lexical and syntactic + semantic disruptions. (There also was a significant semantic disruption at word unit -3, but this disruption is unexplained because it was well in advance of any disruption obtained in any of our experiments.) In contrast to the first two experiments, the semantic violation did not produce a disruption until after the critical word had been uttered, so semantic information evidently was not being used for lexical access. The paragraphs were written in an abstract style using many figures of speech, so semantic violations were anomalies that easily could have been mistaken for metaphors intended by the author. Apparently, readers adopted the quite reasonable strategy of not giving high priority to semantic information for lexical access because it was not very informative. Lexical access used perceptual and syntactic information, which were not distorted by the difficulty a[illegible]veness of the story. The disruption produced by the semantic viola[illegible]fter the critical word was uttered indicates that semantic information was being used in sentence comprehension. Although semantic information may not have been useful for lexical access, it was essential for constructing a meaning for the sentence, so the semantic violation caused problems for the reader at that point.

In this experiment, the difficult text very likely shortened the eye-voice span relative to the first two experiments (Buswell, 1920; Morton, 1964). If such shortening occurred, the disruptions would have moved closer to the critical word, but their relative positions would not have been affected. Specifically, this shortening would not have changed the two primary results of this experiment, namely, the simultaniety of the lexical and syntactic + semantic disruptions and the delay of the semantic disruption until after the critical word was uttered.

Discussion and Conclusions

These experiments have provided information about how different kinds of information interact in the language processing components, particularly lexical access and sentence comprehension. Overall the results support an interactive model of language processing, but more importantly, they indicate something about the nature of the interactions.

Lexical access. Both bottom-up perceptual and top-down contextual information interact in lexical access. How violation of lexical information disrupted lexical access is evident: there was no mental dictionary entry for the nonword. According to Marslen-Wilson and Welsh's (1978) model of lexical access, a cohort of words would be activated corresponding to the initially processed portion of the nonword. For example, for the nonword brugen all known words beginning with bru- would be activated -- bruise, brunch, brunette, brunt, brush, brusque, brut, brutal, and brute. Continuing perceptual analysis quickly would eliminate all the candidates because there is no common English word beginning with brug- or even bru + a letter with a descender (g, j, p, q, or y). A check with syntactic and semantic information for contextual appropriateness also would eliminate all members of the cohort. Hence, the lexical violation would be discovered very quickly.

Syntactic and semantic information also were involved in lexical access, as demonstrated by the disruptive effects before the critical word was uttered. Furthermore, the pattern of disruption caused by the syntactic + semantic violations in Experiment 3 matched that from the lexical violations almost perfectly. If syntactic and semantic violations were disruptive of sentence comprehension only and were not involved in lexical access, the syntactic + semantic disruption would have been delayed at least slightly after the lexical disruption. Readers would not recognize that the syntactic-semantic

information was inconsistent with the rest of the sentence until the critical word had been accessed and the information made available. But the lexical and syntactic + semantic disruption curves were virtually identical suggesting that they affected either a common process, that process being lexical access, or processes operating at the same time.

The fluent restorations of the original critical word in the syntactic violation condition in Experiment 2 also supports the conclusion that syntactic information was involved in lexical access. The only way that the reader could have restored the original critical word was to have anticipated the part of speech from the preceding sentence context and then to have produced the syntactically appropriate ending for the critical word. Another piece of evidence for semantic involvement is the location of the semantic disruption in Experiment 3 relative to the preceding experiments. In the paragraphs used in Experiment 3, semantic information was not useful for lexical access and so semantic constraints apparently were suspended, shifting the disruption to after the critical word was uttered. When semantic information was useful, as in Experiments 1 and 2, its violation was disruptive before the critical word was uttered. This shift in when semantic information was used as a function of its utility indicates a flexibility of processing at the lexical level, a property represented in interactive models.

Another instance of processing flexibility was obtained when these experiments were replicated in Polish (Kurkiewicz, Kurcz, & Danks, in press). Inflectional information varies in usefulness to Polish and English readers. In Polish vis-à-vis English, most syntactic information is carried in suffixes and very little in word order. Violation of syntactic information by altering inflections produced a larger disruptive effect earlier in Polish readers than in English readers. Polish readers used inflectional information to a greater

degree in lexical access than did English readers.

Factual information was not used in lexical access because its violation was disruptive only after the critical word was accessed and uttered and not before or at the critical word. Although in principle it could have influenced lexical access by supplementing syntactic and semantic contextual information, it did not. The fact that factual violations did not influence lexical access across clause or sentence boundaries suggests that lexical access was clausally autonomous. These results are in contrast to those reported by Foss and Ross (this volume). They found that information from a preceding sentence facilitated lexical access as measured by a phoneme-monitoring task. An explanation of the contrasting results of the two experiments probably lies in the quite different relations between the information in the preceding sentences and the target words and in the different experimental tasks.

Sentence comprehension. Syntactic, semantic, and factual information were involved in sentence comprehension: violations of all three produced disruptions for several word units after the critical word was uttered. However, the results do not differentiate between the clausal processing and word-by-word hypotheses because the location of the clause and sentence boundaries following the critical word were uncontrolled. In other experiments (Danks, Bohn, End, & Miller, 1980), we have both controlled and manipulated the location of the clause boundary following the critical word. The results have been quite clear. Both semantic and factual violations produced small, but significant, disruptions between the critical word and the end of the clause, followed by a much larger peak of disruption immediately after the clause boundary. The disruptions resulting from semantic and factual violations were the same size. The only difference was that the semantic violation produced an additional peak at the critical word and the factual violation did not.

These more recent results support the word-by-word hypothesis of sentence comprehension. Violating either semantic or factual information produced a disruption while the clause was being read. As each word was accessed, the reader attempted to integrate its meaning into a larger representation of the text. This representation spanned more than just the immediate sentence because the factual violation, which involved the preceding sentence, produced a disruption as well. Since integration was not possible immediately, the reader had to hold the words in a memory buffer until the end of the clause. At that time, the increasing memory load and processing demands required a final attempt at integration. The end of a clause or sentence is a natural point for readers to resolve any problems they have had understanding a sentence. At the clause boundary, readers attempted to resolve the inconsistency, but given the nature of our violations, they were usually unsuccessful.

An interactive model of sentence comprehension provides the best account of these results. As words are accessed, each word's meaning is integrated into a global representation of the text to that point. The sentence comprehension component is not autonomous because the global representation spans more than the immediate sentence. Finally, sentence comprehension is a flexible process adapting to the difficulty of the text and the availability of information.

Discourse understanding. These experiments were not designed to evaluate discourse understanding, even though a story context was used and factual, intersentential information was violated. Most studies of discourse understanding have used memory tasks to assess whether the listener/reader has formed a coherent text base and a macrostructure representation (Kintsch, 1974; Meyer, 1975), but memory tasks do not provide much information about the process of story understanding. Several more recent studies have examined how

this component operates using on-line measures, such as eye fixations and sentence reading times. Just and Carpenter (1980) monitored eye fixations and were able to attribute substantial amounts of gaze time to macroprocessing, as well as to lexical and sentential components. Cirilo and Foss (1980) and Cirilo (1980) found that discourse understanding, as measured by reading times, varied in predictable ways based on Kintsch and van Dijk's (1978) processing model. Reading times increased for sentences where the referential antecedent of nouns occurred much earlier in the story, thus increasing the difficulty of establishing referential coherence. Reading times also increased when the sentence was high in the macrostructure, a main idea of the story. When the reader's task was altered, readers were flexible in adapting their processing strategies at the discourse understanding level to meet the task demands. These results support an interactive model of discourse understanding. Macrostructure and referential coherence influenced gaze durations for individual words and reading times for individual sentences. Processing strategies shifted flexibly with changes in the demands on the discourse processing component.

In summary, the results of the experiments reported here and other studies support interactive-type models for all three major components of language processing -- lexical access, sentence comprehension, and discourse understanding. Although we have considered the processing components separately for purposes of exposition, a comprehensive model will include all three components. We attempted to narrow the range of possible interactive models, but excessive explanatory power remains as their most salient fault.

Acknowledgments

We thank Richard Klich, Greg Hill, and Jeff Hall for their assistance in conducting and analyzing the experiments reported in this chapter. Experiment 3 was part of a master's thesis submitted by Ramona Fears to Kent State University. The research and the preparation of this chapter was supported by Grant No. NIE-G-78-0223 from the National Institute of Education.

References

Aaronson, D., & Scarborough, H. S. Performance theories for sentence coding: Some quantitative evidence. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 56-70.

Bever, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970.

Bradley, D. C., & Forster, K. I. Lexical access in on-line sentence processing. This volume.

Bransford, J. D., & Johnson, M. K. Considerations of some problems of comprehension. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Buswell, G. T. An experimental study of the eye-voice span in reading. Supplementary Educational Monographs, 1920 (Whole No. 17).

Carpenter, P. A., & Just, M. A. Reading comprehension as eyes see it. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Carroll, J.M., & Bever, T.G. Sentence comprehension: A case study in the relation of knowledge to perception. In E. C. Carterette & M. P. Friedman (Eds.), Handbook of perception: Language and speech, Vol. VII. New York: Academic Press, 1976.

Carroll, J.M., Tanenhaus, M. K., & Bever, T. G. The perception of relations: The interaction of structural, functional, and contextual factors in the segmentation of sentences. In W. J. M. Levelt & G. B. Flores d' Arcais (Eds.), Studies in the perception of language. Chichester: Wiley, 1978.

Cirilo, R. K. Referential coherence and text structure in story comprehension. Meetings of the Psychonomic Society, St. Louis, 1980.

Cirilo, R. K., & Foss, D. J. Text structure and reading times for sentences. Journal of Verbal Learning and Verbal Behavior, 1980, 19, 96-109.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language

statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 335-359.

Clark, H. H., & Marshall, C. Reference diaries. In D. L. Waltz (Ed.), Theoretical issues in natural language processing-2. New York: Association for Computing Machinery, 1978.

Crothers, E. J. Paragraph structure inference. Norwood, NJ: Ablex, 1979.

Danks, J. H. Grammaticalness and meaningfulness in the comprehension of sentences. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 687-696.

Danks, J. H. Models of language comprehension. Polish Psychological Bulletin, 1978, 9, 183-192.

Danks, J. H., Bohn, L., End, L., & Miller, R. Integration of sentence meanings in stories. Meetings of the Psychonomic Society, St. Louis, 1980.

Danks, J. H., & Fears, R. Oral reading: Does it reflect decoding or comprehension? In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading, Vol. 3. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

Danks, J. H., & Glucksberg, S. Experimental psycholinguistics. Annual Review of Psychology, 1980, 31, 391-417.

Danks, J. H., & Hill, G. O. An interactive analysis of oral reading. In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Dooling, D. J., & Lachman, R. Effects of comprehension on retention of prose. Journal of Experimental Psychology, 1971, 88, 216-222.

Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1978-1979, 14, 481-533.

Fears, R. Levels of processing in oral reading. Unpublished master's thesis. Kent State University, 1978.

Flores d'Arcais, G. B., & Schreuder, R. The use of syntactic and lexical

information during language processing. This volume.

Forster, K. I. Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), New approaches to language mechanisms. Amsterdam: North-Holland, 1976.

Foss, D. J., & Ross, J. R. Great expectations: Context effects during sentence processing. This volume.

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology, 1975, 7, 139-166.

Fry, E. B. A readability formula that saves time. Journal of Reading, 1968, 11, 513-516, 575-578.

Gibson, E. J., & Levin, H. The psychology of reading. Cambridge, Mass.: MIT Press, 1975.

Glucksberg, S., Trabasso, T., & Wald, J. Linguistic structures and mental operations. Cognitive Psychology, 1973, 5, 338-370.

Gough, P. B. One second of reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass.: MIT Press, 1972.

Just, M. A., & Carpenter, P. A. A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 87, 329-354.

Kintsch, W. The representation of meaning in memory. Hillsdale, NJ: Lawrence Erlbaum Associates, 1974.

Kintsch, W., & Green, E. The role of culture-specific schemata in the comprehension and recall of stories. Discourse Processes, 1978, 1, 1-13.

Kintsch, W., Mandel, T. S., & Kozminsky, E. Summarizing scrambled stories. Memory and Cognition, 1977, 5, 547-552.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In L.-G. Nilsson (Ed.), Perspectives

on memory research. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

Kurkiewicz, D., Kurcz, I., & Danks, J. H. Reading comprehension processes in Polish and English. Polish Psychological Bulletin, 1981, in press.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Lesgold, A. M., & Perfetti, C. A. (Eds.), Interactive processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Levin, H. The eye-voice span. Cambridge, Mass.: MIT Press, 1979.

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

Marslen-Wilson, W. D. Sentence perception as an interactive parallel process. Science, 1975, 189, 226-228.

Marslen-Wilson, W. D., & Tyler, L. K. The temporal structure of spoken language understanding. Cognition, 1980, 8, 1-71.

Marslen-Wilson, W. D., & Tyler, L. K. Processing structure for spoken language understanding. This volume.

Marslen-Wilson, W. D., Tyler, L. K., & Seidenberg, M. Sentence processing and the clause boundary. In W. J. M. Levelt & G. B. Flores d'Arcais (Eds.), Studies in the perception of language. Chichester: Wiley, 1978.

Marslen-Wilson, W. D., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.

Massaro, D. W. Primary and secondary recognition in reading. In D. W. Massaro (Ed.), Understanding language. New York: Academic Press, 1975.

McClelland, J. L. On the time relations of mental processes: An examination of systems of processes in cascade. Psychological Review, 1979, 86, 287-330.

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

Meyers, L. S., & Boldrik, D. Memory for meaningful connected discourse. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 584-591.

Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 217-228.

Mistler-Lachman, J. L. Levels of comprehension in processing of normal and ambiguous sentences. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 615-623.

Morton, J. The effects of context upon speed of reading, eye movements, and eye-voice span. Quarterly Journal of Experimental Psychology, 1964, 16, 340-354.

Perfetti, C.A., & Lesgold, A.M. Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Perfetti, C. A., & Roth, S. Some of the interactive processes in reading and their role in reading skill. In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates, in press.

Poulsen, D., Kintsch, E., Kintsch, W., & Premack, D. Children's comprehension and memory for stories. Journal of Experimental Child Psychology, 1979, 28, 379-403.

Reicher, G. M. Perceptual recognition as a function of meaningfulness and stimulus material. Journal of Experimental Psychology, 1969, 81, 275-310.

Rumelhart, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, NJ: Lawrence Earlbaum Associates, 1977.

Schwartz, R. M. Levels of processing: The strategic demands of reading comprehension. Reading Research Quarterly, 1980, 15, 433-450.

Stanovich, K. E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 1980, 16, 32-71.

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

Stevens, A. L., & Rumelhart, D. E. Errors in reading: An analysis using an augmented transition network model of grammar. In D. A. Norman & D. E. Rumelhart (Eds.), Explorations in cognition. San Francisco: Freeman, 1975.

Tyler, L. K., & Marslen-Wilson, W. D. The on-line effects of semantic context on syntactic processing. Journal of Verbal Learning and Verbal Behavior, 1977, 16, 683-692.

Warren, W. H., Nicholas, D. W., & Trabasso, T. Event chains and inferences in understanding narratives. In R. O. Freedle (Ed.), Advances in discourse processing: New directions in discourse processing, Vol. 2. Norwood, NJ: Ablex, 1979.

Wheeler, D. Processing in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Wildman, D. M., & Kling, M. Semantic, syntactic, and spatial anticipation in reading. Reading Research Quarterly, 1978-1979, 14, 128-164.

Winer, B. J. Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill, 1971.

Appendix

The critical word segments for the three experiments are listed in the order in which they appeared in the story. The critical word is italicized. The manipulations introduced to produce the violations are indicated in parentheses: the factual change is in the sentence preceding the critical word (Experiments 1 and 2) and the lexical (Experiment 3), syntactic + semantic (Experiments 1 and 3), syntactic (Experiment 2), and semantic (Experiments 1, 2, and 3) follow the critical word. The slashes separate word units used in scoring the protocols.

Experiment 1

1. Everyone in the Scott family was in high spirits because on Saturday they planned to fly to Mexico for a vacation (... was depressed that morning because on Saturday they had to attend Mary Jane's grandmother's funeral). / No one / in the entire / Scott / family / was more / excited (pumpkin, smoked) / than / Mary Jane, / an intelligent / and pretty / girl. /

2. Her daughter had always been weak (strong) physically. Because / of this, / she even / imagined / her daughter / being / trampled (iceberg, planted) / by other children / while / trying / to get / out of the wrecked bus. /

3. Sam Scott heard on the radio that morning (... went to work and did not know) that there had been a bus-train collision in / Lawrenceville. / Although / he was / relieved / that the / accident (subtract, delusion) / had not / occurred / in his / community / he felt / ....

4. There was enough time for the school bus to slow down

and yet it did not (The bus stopped at the stop sign and waited until the entire train had safely passed). A loud sound like a powerful explosion / could be / heard / as the / train / and bus / collided (family, mailed). / The train / hit / the bus / at the / midpoint. /

5. One boy had his left leg severed below the knee (One boy who was not injured was screaming hysterically). She / applied / a large / tourniquet / to stop / his / bleeding (freeze, writing) / leg. / Within minutes, / several ambulances / were on the / scene. /

6. They were unconscious and few had (They were fully conscious and had) either / wallets / or purses. / Immediate / identification / was / impossible (phonograph, athletic). / Grabbing / a felt-tipped / pen, / Dr. Carr / inscribed / ....

7. Mary Jane's name was first on a list of survivors (deceased). / She was / at that / moment / being / prepared / for / surgery (summarize, trailer). / From time / to time / a doctor / would enter / the cafeteria / ....

8. The doctor told the Scotts that the leg may have to be amputated (... that their daughter's leg may be saved) as soon as a specialist could be found / to perform / the surgery. / Vera / and Sam / were / saddened (rapidly, polished) / by this / possibility. / Vera / collapsed / but Sam / ....

9. Once her vital signs stabilized, she was moved (In spite of the doctors' efforts, her vital signs did not stabilize which meant that she could not be moved) / from her / hospital / room. / Mary / was / carried (piano, printed) / in her / bed / from the / intensive / care / ....

10. She was not aware of anything (She was aware of everything) that happened to her / after / the collision / of the bus / and the train. / She was / <u>unconscious</u> (baseball, classical) / from a / blow / to the / back / of her / ....

11. Mr. Scott was shakened, but calm (... and dazed) after all / that had / happened / that day. / He had / clearly / <u>comprehended</u> (umbrella, transported) / all / that the / doctors / had said. / The Scotts / ....

12. On occasion, parts of her body would jerk uncontrollably (Her whole body was motionless) as she / lay / in the / bed. / Her arm / and leg / <u>shook</u> (money, spoke) / violently. / Her eyes / were closed / as though / she were / ....

13. The next day her temperature shot up over six degrees and perspiration and dehydration accompanied her high fever (... returned to normal and the perspiration and dehydration accompanying her high fever were gone). As the hours / passed, / the nurses / continued / to change / her / <u>sweat-soaked</u> (accommodate, scuba-diving) / hospital / gowns. / For forty-eight / hours / her life / ....

14. Mary Jane became unusually more alert when she saw (Mary Jane rejected) / the special / drawing / board. / Immediately / she started / <u>copying</u> (divine, singing) / a picture / much / in the / style / of a / ....

15. She was allowed to have more visitors other than (Although Mary Jane was now feeling better she was only allowed to have visitors from) her immediate family. / Dozens / of her / classmates / came / to her / <u>bedside</u> (punish, teeth) / in the evenings. / To the / Scotts, / who had / ....

16. Ever since Mary Jane had been conscious and alert, she had been without exercise of any kind (..., she was careful to exercise her muscles everyday) while in the hospital. / As a / result, / her muscles / started / to / <u>waste</u> (fork, rejoice) / away. / Now / with the / help / of three / ....

<u>Experiment 2</u>

1. Everyone in the Scott family was in high spirits that morning because on Saturday they planned to fly to Mexico for a vacation (Everyone in the entire Scott family was depressed that morning because on Saturday they had to attend Mary Jane's grandmother's funeral). No one / in the entire / Scott / family / was more / <u>excited</u> (smoked, excitement) / than / Mary Jane, / an intelligent / and pretty / girl. /

2. Her daughter had always been weak (strong) physically. Because of this / she even / imagined / her daughter / being / <u>injured</u> (injury, planted) / by other / children / while / trying / to get out /..

3. Sam Scott heard on the radio that morning that there had been a bus-train collision in Lawrenceville (Sam Scott was not aware that there had been a terrible bus-train accident in Lawrenceville that morning). / Although / he was / relieved / that the / <u>collision</u> (collided, delusion) / had not / occurred / in his / community, / he felt / ....

4. There was enough time for the school bus to slow down and yet it did not stop (The bus driver saw the train just in time and miraculously stopped the bus) before / it reached / the tracks. / The bus / and train / <u>crashed</u> (crashing, mailed) / and a loud / sound / like a / powerful / explosion / ....

5. One boy had his left leg fractured below the knee (... was completely uninjured). She applied / a large / splint / to fix / his / broken (break, written) / leg. / Within minutes, / several / ambulances / were on / ....

6. Most of them were unconscious (conscious) and could not give (had given) their names to the / physician. / Immediate/identification / was not / possible (possibly, athletic)./ Grabbing / a felt-tipped / pen, / Dr. Carr / inscribed / ....

7. Mary Jane's name was first on a list of survivors (deceased). She was at that / moment / being prepared / for / surgery (surgically, trailer). / From time / to time, / a doctor / would enter / the cafeteria / ....

8. The doctor told the Scotts that the leg may have to be amputated (would be saved) as soon as a specialist could be found to perform / the surgery. / Vera / and Sam / were / depressed (depression, polished) / by this possibility. / Vera / collapsed, / but Sam / ....

9. Unfortunately (Fortunately), her vital signs did not (had) stabilized so it was impossible (possible) for the doctors and nurses / to move / her. / Mary Jane / was / carried (carrier, printed) / from the / intensive / care unit, / down the / corridor. /

10. After the collision of the bus and the train, she was not aware of anything (..., she was clearly aware of everything) / that had / happened / to her. / She was / unconscious (unconsciousness, classical) / from a / blow / to the / back / of her / ....

11. Mr. Scott was shakened but calm after all that had happened that day (... was so shakened and confused that he did

not understand any of the medical explanations concerning his daughter's condition). All that the doctors / had told / him, / he had / clearly / comprehended (comprehension, transported). / The Scotts / were / overwhelmed / with grief. / They realized / ....

12. On occasion parts of her body would jerk uncontrollably (Her whole body was completely paralyzed) as she / lay / in the bed. / Her arm / and leg / shook (shaky, spoke) / violently. / Her eyes / were closed / as though / she were / ....

13. The doctor then informed Mary Jane's parents that the leg would have to be removed (... that with more surgery, the leg would be able to be saved) / completely. / After / Mary Jane's / leg was / amputated (amputation, amplified), / she was / taken back / to her / room / in isolation / ....

14. Mary Jane became unusually more alert when she saw (Mary Jane rejected) the special / drawing / board. / Immediately, / she started / copying (copier, singing) / a picture / much / in the / style / of a / ....

15. She was allowed to have more visitors other than her immediate family (However, she still refused to have any visitors while she was bedridden). Dozens / of her / classmates / came / to her / bedside (bedridden, teeth) / in the / evenings. / To the Scotts / who had seen / her / ....

16. Ever since Mary Jane had been conscious and alert, she had been without exercise of any kind (..., she was careful to exercise her muscles everyday) while in the / hospital. / As a result, / her muscles / started to / waste (wasteful, rejoice) / away. / Now / with the / help / of three / ....

Experiment 3

1. ... / down / the dark / aisles / and sat / on the / moss (glurck, school, killed) / beneath / the great / trees; / and again / his mind / ....

2. ... / nostrils. / At first / the men swore. / They / struck (brugen, flew, brilliant) / and prodded / the camels. / Then / they / too / ....

3. ... / the grinding, / the screaming. / He / wriggled / in the / dry (kaysen, brilliant, iron) / straw / stuffing / his mouth / with it / to keep / ....

4. ... / crocodiles / were / shot / at, / the storks / flew (glurck, killed, streets) / into the air / in mass / panic. / The beat of / their wings / ....

5. ... / to his knee / and removing / the binding / spread the / cool (brugen, sharp, flew) / paste / thickly / on the / swollen / flesh. /

6. ... / town / and / entered it / and walked / through its / streets (kaysen, iron, sharp). / Faces / looked at / him / but he / did not / ....

7. ... / struck / something / hard / and he / felt a / sharp (glurck, cool, streets) / pain. / When he / looked up / the train / was / ....

8. ... / among / the dead. / Some of the / dead / had been / killed (kaysen, ran, school) / by the fire / and lay / faceless / in the / ....

9. ... / ships of / hard / fact, / of/hard / iron (brugen, moss, cool), / of coal, / machinery, / cargo / and / tonnage. /

10. ... / behind, / a black / headland / notched / the / brilliant (glurck, dry, moss) / sky. / Far ahead, / barely / visible / in the / ...

11. ... / from seven / to two / each / day / in the / school (brugen, streets, ran), / itself. / Later / they / went around / the town / ....

12. ... / came, / in wild / summer / thunderstorms. / The rain / ran (kaysen, struck, dry) / down / his clothing, / plastering, / the dust / in streaks / ....

Table 1

Experiment 1: Mean Production Times (msec.)

	Word-Unit Position										
Violation	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
Control	596	430	501	578	355	146	454	511	509	405	477
Syntactic + Semantic	549	449	469	609	621	586	791	700	636	448	553
Semantic	567	431	460	605	463	299	819	686	671	432	567
Factual	586	456	478	568	372	168	553	602	516	424	508

Table 2

Experiment 2: Mean Production Times (msec.)

	Word-Unit Position									
Violation	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
Control	410	540	620	355	140	500	565	470	425	550
Syntactic	440	545	595	435	415	890	690	555	455	580
Semantic	425	540	630	460	405	850	750	620	480	580
Factual	415	530	660	415	175	655	650	565	415	515

Table 3

Experiment 3: Mean Production Times (msec.)

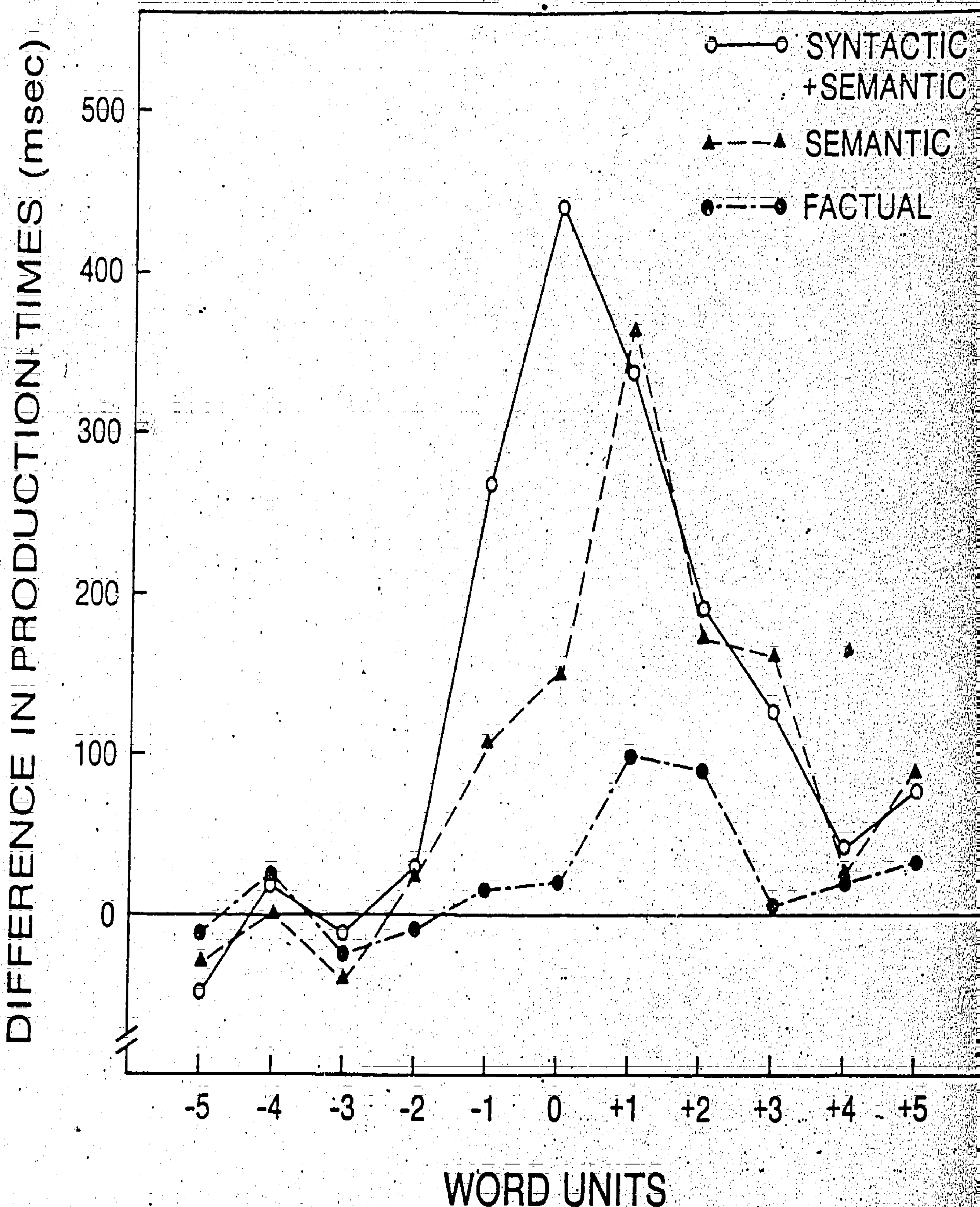
Violation	Word-Unit Position										
	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5
Control	520	520	485	535	405	267	450	640	495	460	435
Lexical	508	522	453	473	481	685	690	770	520	525	430
Syntactic + Semantic	528	506	433	441	478	632	785	745	540	535	455
Semantic	538	479	546	465	436	318	595	695	525	485	430

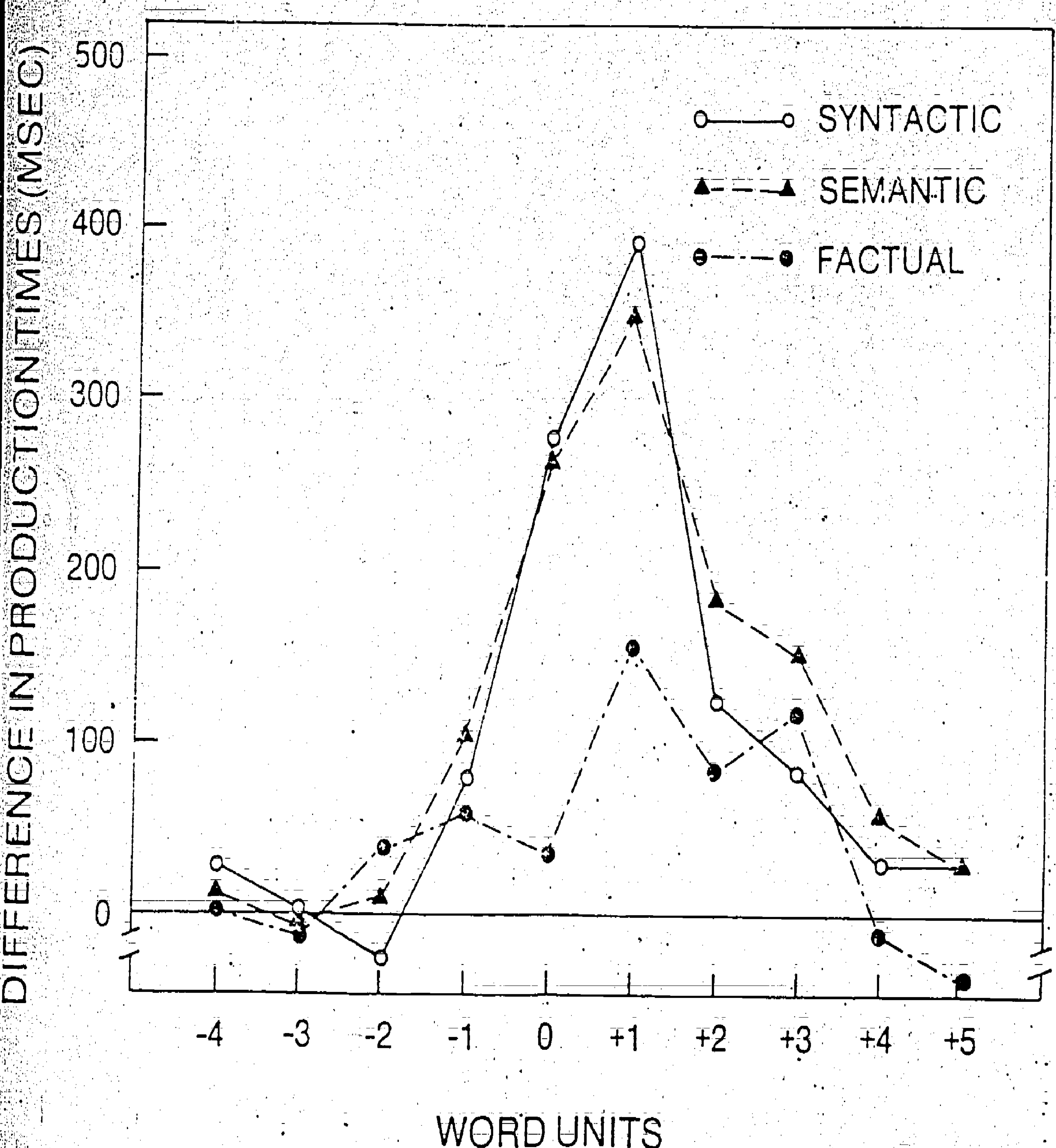
Figure Captions

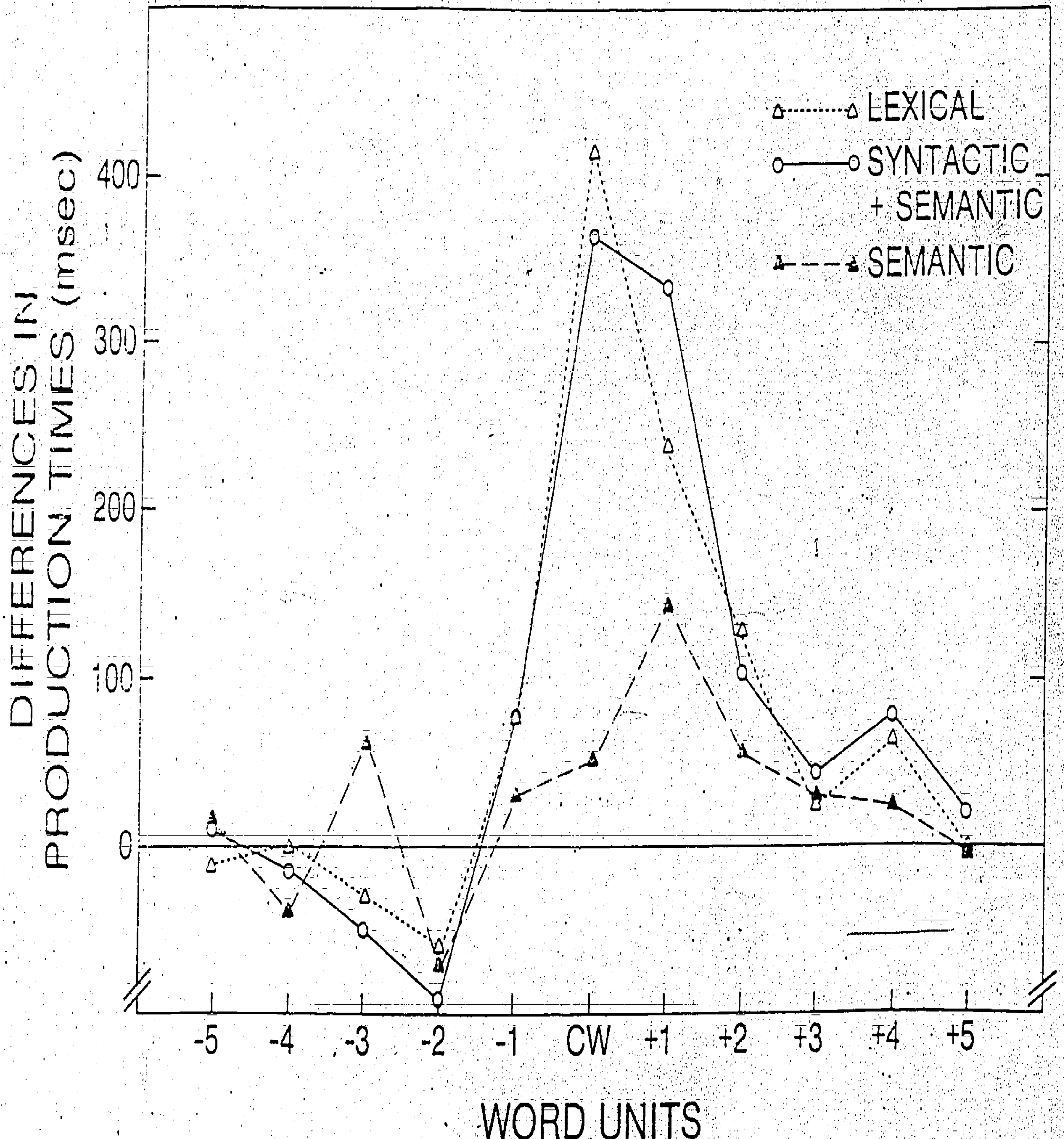
Figure 1. Experiment 1: Differences in production times.

Figure 2. Experiment 2: Differences in production times.

Figure 3. Experiment 3: Differences in production times.


SYNTACTIC +SEMANTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PRODUCTION TIMES (msec)
500
400
300
200
100
0
-5
-4
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS


DIFFERENCE IN PRODUCTION TIMES (MSEC)
500
400
300
200
100
0
SYNTACTIC
SEMANTIC
FACTUAL
-4
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS


LEXICAL
SYNTACTIC + SEMANTIC
SEMANTIC
DIFFERENCES IN PRODUCTION TIMES (msec)
400
300
200
100
0
-5
-4
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD UNITS

Integration of Sentence Meanings in Stories*

Joseph H. Danks, Lisa Bohn, Laural End, and Ruth Miller

Kent State University

Paper Presented at the

Meetings of the Psychonomic Society

St. Louis, November 13-15, 1980

This research was supported by Grant No. NIE-G-78-0223 from the National Institute of Education. Mailing address: Department of Psychology, Kent State University, Kent, Ohio 44242.

Abstract

Are meanings of sentences in stories constructed word by word or only at the end of a clause? Factual inconsistencies between sentences disrupted oral reading both at the inconsistent word and at the following clause boundary. Readers evidently attempted to integrate each word into a comprehensive representation of the story as soon as it was accessed.

What are the processes readers use to construct a representation of the texts they are reading? Three major components are involved (Danks, Bohn, & Fears, in press; Just & Carpenter, 1980). These components are not discrete, autonomous, or serially ordered, but operate interactively. In <u>lexical access</u>, readers locate words in the mental dictionary and select appropriate meanings. The information that readers use may be bottom-up perceptual information or top-down conceptual information. In the oral reading task used in the experiments reported here, readers gained access to the articulatory information needed to utter the printed word as a result of lexical access. In <u>sentence comprehension</u>, readers integrate the word meanings into a composite representation of the sentence using syntactic structure as a guide. There are two major hypotheses about how readers do this. Under a clausal processing hypothesis, word meanings are stored in a working memory buffer until the end of the clause. Then the meaning of the clause is computed, after which it is integrated with the representations of preceding sentences and with prior knowledge. Alternatively, a word-by-word integration hypothesis proposes that readers attempt to integrate each word's meaning as it is accessed. If immediate integration fails, then the word's meaning is held in a buffer until integration is possible, usually at the end of the clause. In <u>text understanding</u>, readers organize the text or story into a coherent text base or microstructure. Readers also develop a schematic representation of the main ideas or gist of the story, corresponding to the macrostructure (Kintsch & van Dijk, 1978).

These components use differentially various types of information to accomplish their tasks (Danks & Hill, in press). In

principle, all types of information are continuously available, but certain types of information are more useful than others for each component. What types of information are used by each component and when is each type of information used? Our rationale was to violate one or more types of information in the text. An oral reading task provided a continuous on-line indication of the status of the reading process. The point when the reading process was disrupted was reflected in the oral output.

The experiments used a 2000-word story about a high school girl who was injured in a school bus accident. We selected 24 critical words for manipulation. A portion of the text around one of the critical words is shown in Figure 1. In this example, the mother of the girl has just heard about the accident. The critical word is injured. The end of the clause (and the sentence) containing the critical word was four words after the critical word in all cases.

Three single violation types were introduced into the text. (a) A misspelled violation was created by misspelling the critical word such that it would be pronounced the same or very similarly to the critical word. The maximum possible misspelling that could still be pronounced correctly was used. In the example, injured was misspelled injerd. If readers "sounded out" the misspelled word, they could locate the intended word in their mental dictionaries. A spelling violation would disrupt primarily the lexical access component. (b) Semantic information was violated by replacing the critical word with a semantically anomalous word. In the example, planted was inserted for injured. Readers could devise an interpretation, but it would require considerable extra processing. The semantic violation would disrupt both lexical access and sentence

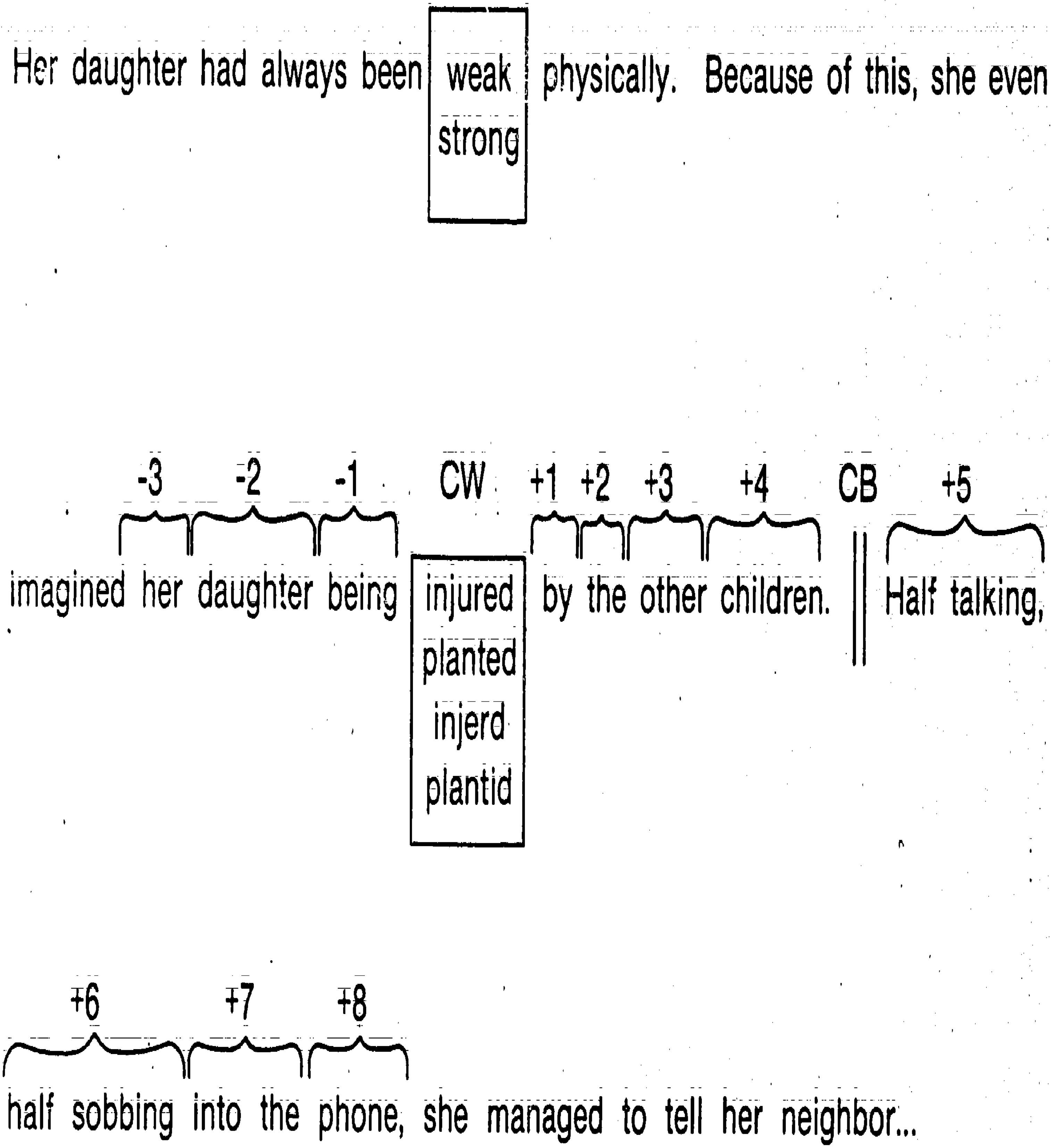


Figure 1. An example of a portion of text and manipulations.

comprehension components. (c) In the third type of violation, factual, neither the critical word nor the sentence containing it was changed in any way. The sentence preceding the critical sentence was changed so as to produce a factual inconsistency with the critical word. In the example, weak was changed to strong. The factual violation would disrupt the sentence comprehension and text understanding components more than it would lexical access.

Two additional violation types were produced by combining the misspelling with the semantic and factual violations. (d) In the misspelled + semantic violation, the semantically anomalous word was misspelled, e.g., planted was misspelled plantid. (e) In the misspelled + factual violation, the critical word was misspelled, e.g., injerd, and the preceding sentence was changed to produce the inconsistency, e.g., strong replaced weak. The multiple violations were intended to test for interactions among information types in the lexical access and sentence comprehension components. (f) A control condition in which there was no change in the critical word or the preceding sentence served as a baseline estimate of "spontaneously" produced oral reading errors. The story was constructed in six versions such that each violation type occurred four times in each version and equally often at each critical word across versions.

Sixty college undergraduates at Kent State University who were native English speakers but who were unscreened for reading ability read the story aloud. They were instructed to understand what they were reading and were given a cued recall test after reading the story. The reading aloud protocols were scored for major disruptions, i.e., substitutions, omissions, reversals, repetitions, regressions, insertion of extraneous words, pauses, stutterings, and stammerings.

Major disruptions were tallied for three words before the critical word and for eight words after (see Figure 1) and the frequencies were converted to probabilities. Because we were interested in disruptions of normal reading, the baseline probability of a major disruption at each word-unit position was subtracted from each of the five experimental conditions. The differences in mean probability of a major disruption are presented in Figure 2.

The misspelling violation produced a significant disruption that began one word before the critical word was uttered (word unit -1), peaked at the critical word, and continued through two words after the critical word (word unit +2). The semantic violation also was significantly different from the control beginning one word before the critical word (word unit -1), but continued through five words after the critical word (word unit +5). There were two peaks of disruption from the semantic violation -- at the critical word and just after the clause boundary at word unit +5. Since there was a disruption before the critical word was uttered in both conditions, both misspelling and semantic violations disrupted lexical access. The semantic violation produced a significantly greater disruption <u>before</u> the critical word was uttered than did the misspelling, indicating a top-down effect of semantic information on lexical access. This contextual effect did not extend across sentence boundaries, however, since the factual violation did not produce a disruption at or before the critical word.

Both semantic and factual violations produced a disruption peak immediately after the clause boundary (word unit +5). Furthermore, there was a significant disruption from the semantic and factual violations between the critical word and the clause boundary (word

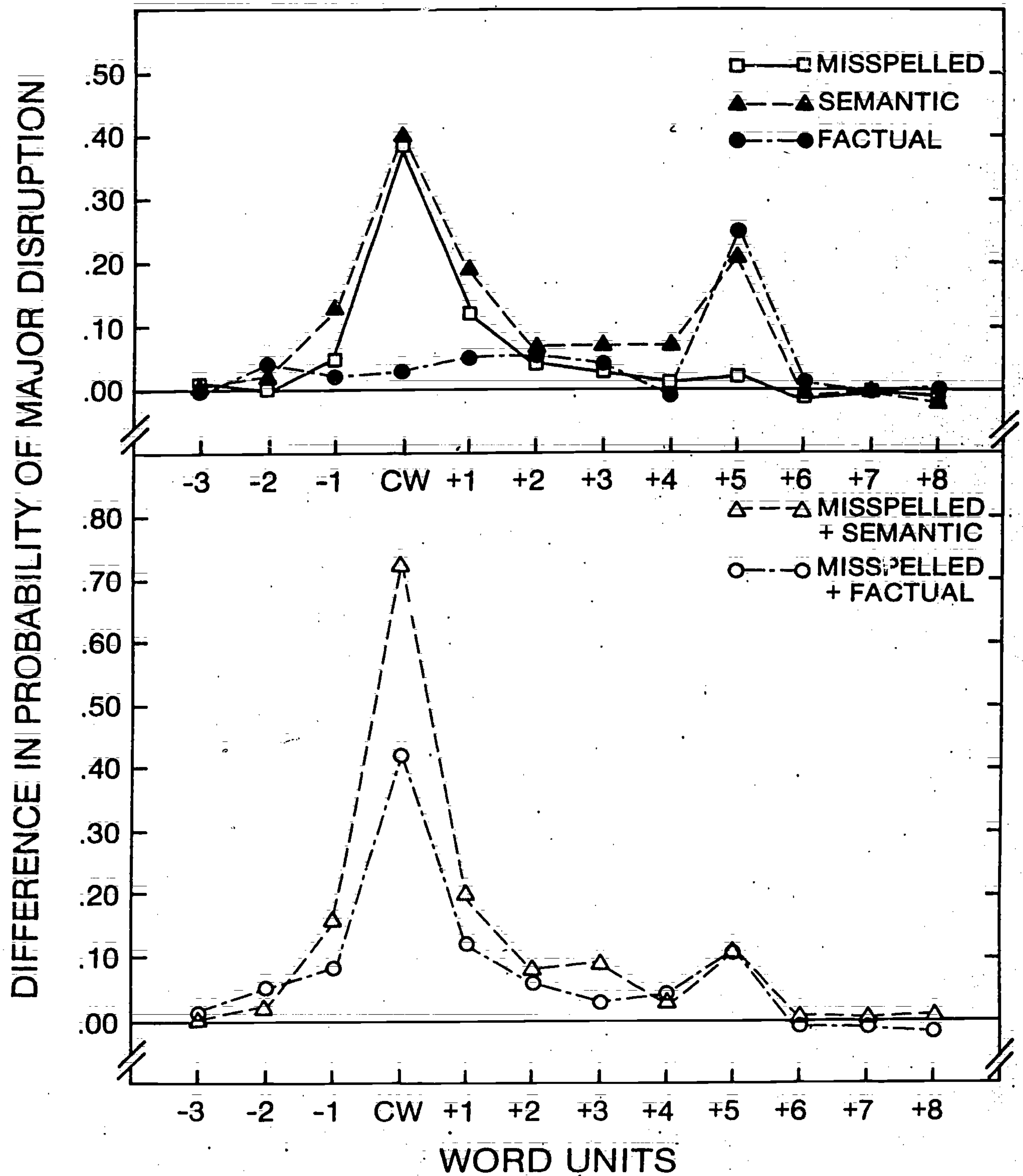


Figure 2. Mean probability of a major disruption.

units +1 through +4 for semantic and word units +1 through +3 for factual). This result supports the word-by-word integration hypothesis. Readers attempted to integrate the semantically anomalous word into a sentential meaning representation, but were unsuccessful. Then the reader attempted another resolution at the clause boundary producing the peak of disruption at word unit +5. The factual violation also disrupted sentence comprehension, but not lexical access. The disruption from the factual violation after the critical word and immediately after the clause boundary was the same size as the one from the semantic violation. Readers had difficulty integrating the critical word with the representation of the preceding sentence. This result with the factual violation is critical for two reasons. Firstly, the semantic disruption between the critical word and the clause boundary might be explained by perseveration of the lexical access disruption. Since the factual violation did not disrupt lexical access, a perseveration explanation is impossible for the factual disruption, and hence, is implausible for the semantic disruption. Secondly, in order for the factual violation to disrupt sentence comprehension before the clause boundary, readers had to be using information from the preceding sentence. Thus, sentence comprehension is neither clausally nor sententially autonomous.

The lower panel of Figure 2 shows the disruption curves for the two multiple violation types -- misspelled + semantic and misspelled + factual. When misspelling and semantic violations were combined, there was a significant disruption beginning one word unit before the critical word (word unit -1) and continuing through the third word after the critical word (word unit +3). When misspelling and factual violations were combined, there was a significant disruption

beginning two word units before the critical word (word unit -2) that continued through two words after the critical word (word unit +2). In both conditions, there was a small peak of disruption immediately after the clause boundary (word unit +5). The peak at the critical word was equivalent to the sum of the two peaks when spelling and semantic information types were violated separately. This result indicates that spelling and semantic information operated independently in lexical access. The lexical access disruption peak was also additive for misspelled + factual in that it was equivalent to the misspelled alone peak (the factual alone produced effectedly no disruption at the critical word).

The peak at the clause boundary in both cases was attenuated to about half the peak when semantic or factual information alone was violated. In sentence comprehension, there was an interaction of spelling with semantic and factual information. Our somewhat speculative interpretation of this result is that when there were two violations the reader gave up more easily attempting to resolve the multiple violations. This explanation involves metacognitive processes since it refers to how readers monitored their comprehension processes rather than to direct changes in the processes themselves.

After reading the story, readers were asked to recall the text surrounding the critical word, i.e., the sentence before the critical word, the sentence containing the critical word, and the sentence after the critical word. Readers were given a copy of the story with these three sentences blanked out at each critical word. The rest of the text provided cues as to what was missing. The recall protocols were scored in terms of the proportion of propositions recalled in each sentence. Both strict and lenient criteria of

whether a proposition was recalled were used. With the strict criterion all arguments had to be recalled; with the lenient only one argument had to be recalled along with the relation. The recall results are presented in Figure 3 as a function of whether there was a violation at the critical word (experimental) or not (control).

There was no difference in recall of the before sentence in the experimental and control conditions. However, the before sentence was manipulated only in the factual violation. The sentence before a factual violation was recalled better than one before a control condition with the strict criterion (but not with the lenient). This result suggests that readers mentally reviewed the before sentence when encountering a factual violation since the before sentence led to the inconsistency. This review resulted in better recall. The critical sentence was recalled less well when there was a violation present than when there was none. This depressed recall was in spite of the greater processing that the critical sentence received when a violation was present. Also, having a violation might have made the critical sentence stand out, a von Restorff effect, but this did not improve memory. Perhaps the depressed recall indicates that the critical sentence was less well integrated with the rest of the story when a violation was present and thus was less well recalled. The sentence after the critical sentence also was recalled less well when there was a violation in the critical sentence. The sentence after the critical sentence was never manipulated, so perhaps there was a continuing effort to resolve the violations in the critical sentence. If so, then the after sentence was processed at a reduced level. However, the after sentence was recalled less well after a misspelling violation as well as after the other types of

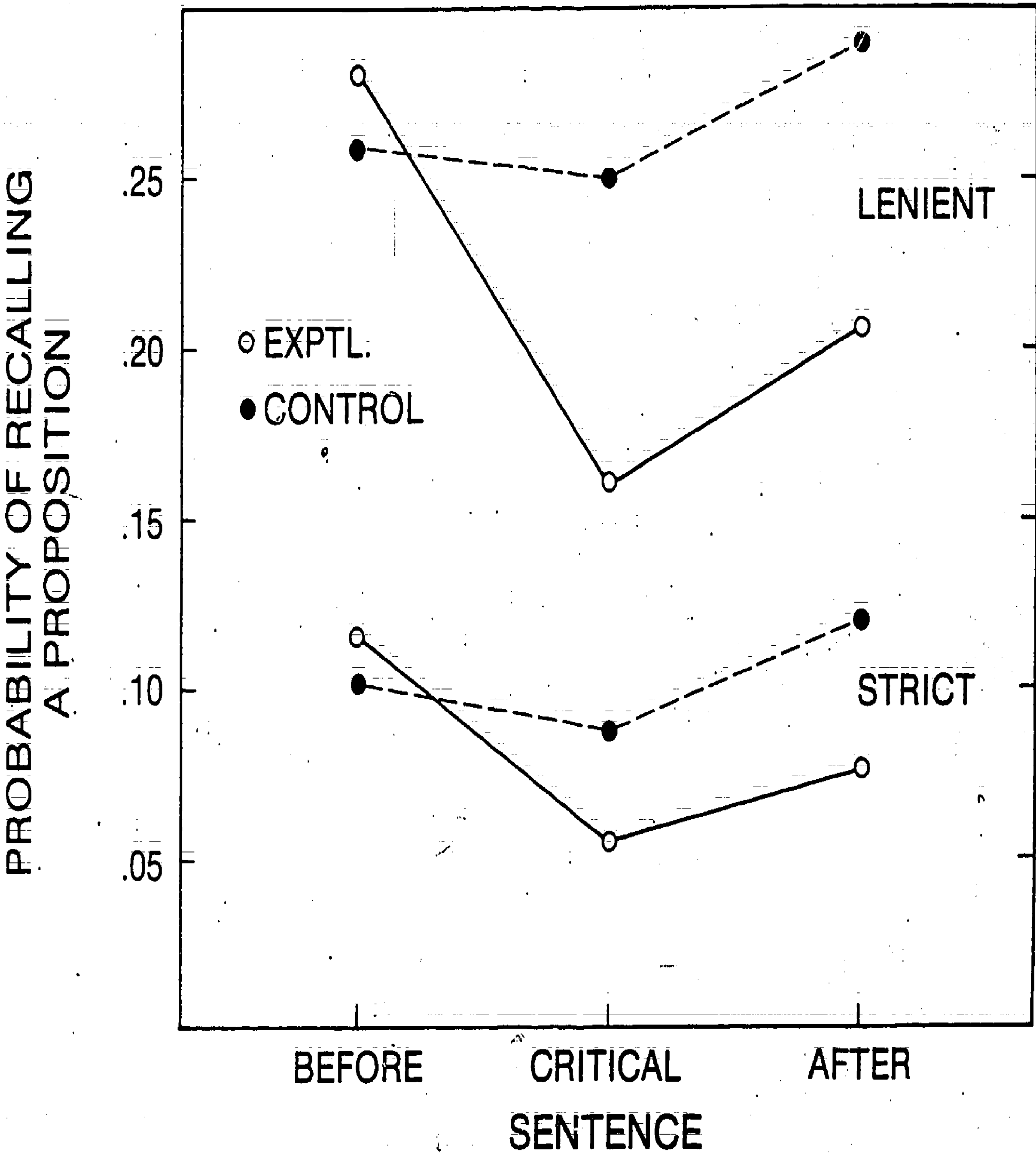


Figure 3. Probability of recalling a proposition.

violations. If the misspelling was resolved when the critical word was uttered, then it should not have had an effect on the processing of the after sentence.

Having established these results by controlling the location of the clause boundary following the critical word, we attempted a more subtle test by manipulating the location of the clause boundary. The portions of the critical sentence after the critical word were rewritten so that the clause boundary occurred either immediately after the critical word, two words after, three words after, or six words after. There were six different critical words for each clause boundary location, yielding 24 critical words in the story. The same six violations were implemented as in the experiment just reported, with appropriate counterbalancing so that each violation type occurred equally often with each clause boundary location.

The details of these results are too complicated to discuss in this paper; however, the basic pattern of results was replicated at each clause boundary location. Disruptions were obtained at the critical word with misspelling and semantic violations and disruptions immediately following the clause boundary with semantic and factual violations. The variability was much greater in this experiment than in the first. In the first experiment, the effects were replicated over 24 critical words whereas in this experiment they were replicated over only 6 critical words, so differences among the specific words emerged. In general, the basic pattern of disruptions was obtained and shifted in synchrony with the location of the clause boundary. The recall results also replicated those from the first experiment.

In conclusion, the lexical access component uses both spelling

and semantic information, but not intersentence factual information, to locate the word in the mental dictionary. Semantic information was relatively more important than spelling information as evidenced by the larger effect of the spelling violation before the critical word was uttered. With regard to sentence comprehension, the word-by-word integration hypothesis received support in contrast to the clausal processing hypothesis. Both semantic and factual violations produced disruptions before the clause boundary was reached, i.e., in the region between uttering the critical word and the end of the clause. There was a major attempt to resolve the violation at the clause boundary, as evidenced by substantial peaks there. The clause boundary is a natural stopping point to attempt such a resolution. If the reader continues to hold words in a limited capacity memory buffer beyond the clause boundary, the comprehension process might break down completely. With multiple violations, the information types did not interact in lexical access, but did in sentence comprehension. A misspelling simply added to the semantic and factual disruptions at the critical word, but attenuated the semantic and factual disruptions at the clause boundary.

In terms of recall, extra processing had a positive effect on the recall of the sentence before the critical word when there was a factual violation. In contrast, extra processing of the critical sentence itself was not sufficient to offset the reduction in recallability resulting from the difficulties of integrating the critical sentence with the rest of the story. The failure to resolve the violations carried past the clause boundary as indicated by the depressed recall of the sentence after the critical sentence when a violation was present.

References

Danks, J. H., Bohn, L., & Fears, R. Comprehension processes in oral reading. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), The process of language understanding. Chichester: Wiley, in press.

Danks, J. H., & Hill, G. O. An interactive analysis of oral reading. In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, NJ: LEA, in press.

Just, M. A., & Carpenter, P. A. A theory of reading: From eye fixations to comprehension. Psychological Review, 1980, 87, 329-354.

Kintsch, W., & van Dijk, T. A. Toward a model of text comprehension and production. Psychological Review, 1978, 85, 363-394.

Comprehension of Metaphors: Priming the Ground

Laurel J. End and Joseph H. Danks

Kent State University

Paper Presented at the

Midwestern Psychological Association Meeting

Minneapolis

May 6-9, 1982

This research was supported in part by Grant No. NIE-G-78-0223 from the National Institute of Education and by Grant No. SED-80-20159 from the National Science Foundation.

Much of the recent research on metaphor has been directed at discovering how metaphors are understood. Generally, theorists describe metaphors in terms of three elements; the topic or tenor, the vehicle, and the ground. The topic is the subject of the metaphor and the vehicle is the term that is used metaphorically. The ground is the relationship between the topic and vehicle from which the metaphorical meaning is derived. For example, in the metaphor Some roads are snakes, the topic is roads, the vehicle is snakes, and the ground is a conceptual relation like long, curvy and dangerous.

A central problem for metaphor comprehension is the nature of the ground. Two general classes of theories have been formulated: comparison and interaction theories. There are many variations of each type of theory, but some general similarities can be drawn. The comparison theory originated with Aristotle. He proposed that one word in a metaphor is replaced with another word that means the same thing. For example, in John is a fox, the word fox has replaced the word sly. The metaphor is understood when one compares a fox to John and discovers the common feature or attribute, which in this case is slyness. The underlying assumption is that the topic and vehicle share a number of features. In order to understand the metaphor the common features (ground) must be discovered by comparing the topic and vehicle (Billows, 1977; Ortony, 1979).

According to the interaction theory, the topic and vehicle interact to create the ground. There is no comparison of the topic and vehicle to find common features, rather, the elements interact to create the metaphorical ground. The resulting ground is a unique combination of the characteristics of the topic and vehicle, thus metaphor as embodied in the ground enables one to "see" the topic in a new or different way. As Black (1936) said of the difference between the comparison and interaction theories, "Looking at a scene through blue spectacles is different from comparing that scene with

something else" (p. 31, emphasis in original).

The distinction between the comparison and interaction theories is not clear-cut. Both are based on the notion that metaphorical meaning is derived from the relationship between the topic and vehicle. According to either theory, the ground can be conceptualized as ranging from a very specific entity, restricted to a single topic-vehicle combination to a more general, abstract entity.

Efforts to determine the nature of the ground have continued, but it is still not clear whether a ground is so specific that it is restricted to a single metaphor or whether a ground can be shared by a number of metaphors. We tested how metaphor-specific grounds are using a sentence priming task. Our rationale was that if several metaphors are based on a common ground then comprehension of one metaphor should prime another thereby facilitating comprehension. When the first metaphor is encountered, the comprehender must construct the relationship between the topic and vehicle in order to understand the metaphor. Once the ground has been constructed it is not necessary to construct it again, so the comprehension of subsequent metaphors based on the same ground should be facilitated.

Our first task was to construct a number of metaphors that seemed to share a common ground, at least to our intuitions. Ten groups of three to five metaphors were generated, totalling 42 metaphors. All metaphors were of the form <u>Some X are Y</u> and each group was based on a different ground. For example, one group consisted of the metaphors <u>Some roads are snakes</u>, <u>Some rivers are ribbons</u>, <u>Some subways are worms</u>, and <u>Some scarves are whips</u>.

In order to verify that the metaphors in each group actually had the same or similar meanings, we asked 30 subjects to sort the metaphors into ten groups. Each metaphor was typed on a file card. A standard from each group was placed on a table and subjects were instructed to place the remaining 32

metaphors under the standard that had the most similar meaning. Three groups of 10 subjects were presented with three different sets of 10 standards.

The results of the sorting task are presented in Table 1. The proportion of times each metaphor was sorted in its predetermined group is displayed in the first column. The three metaphors in each group which were clustered together most frequently were analyzed further. The proportion of times these three metaphors were clustered together is presented in the second column. The probability of clustering two metaphors in a specific group by chance was .10 and the probability of clustering three metaphors together by chance was .01.

As indicated in the table, clustering far exceeded the levels expected by chance. The proportion of times metaphors were clustered with a standard from the same group ranged from .55 to 1.00 with a mean of .78. Proportions for clustering all three metaphors together ranged from .40 to .90 with a mean of .67. Since subjects' clusterings were so consistent with our intuitive groupings, we concluded that our initial groups contained metaphors with the same or similar meanings.

For the priming task we used the three metaphors that were clustered together most frequently in eight of the ten groups. The two groups in which all three metaphors were clustered together less than an average of .63 were discarded. The eight triads of related metaphors then were embedded in a list of filler items. The fillers consisted of 24 filler metaphors and 48 literal sentences, all of the form <u>Some X are Y</u>. The filler metaphors were metaphorical statements that were unrelated to one another and to the triads of related metaphors. The literal fillers were literally true statements that were unrelated to one another and to the triads of related metaphors. The typicality of the literal fillers varied, with some of the sentences being fairly typical (e.g., <u>Some dogs are setters</u>) and others less typical

Table 1

Proportion of Times Metaphors Were Clustered in Their Predetermined Groups

Metaphor	Clustered in pre-determined group	All three clustered together
*Some roads are snakes	.80	
*Some subways are worms	.85	.63
*Some rivers are ribbons	.60	
Some scarves are whips	.20	
*Some clouds are cotton	.80	
*Some pillows are marshmellows	.90	.73
*Some skies are silk	.70	
*Some jobs are jails	.75	
*Some marriages are prisons	.90	.63
*Some drugs are handcuffs	.65	
Some hearts are closets	.17	
*Some rumors are diseases	.80	
*Some criminals are germs	.80	.63
*Some lies are cancers	.70	
Some prejudices are blindfolds	.23	
Some lives are ghettos	.13	
*Some stores are jungles	.60	
*Some buildings are mazes	.55	.40
*Some schools are zoos	.60	
Some homes are dungeons	.10	
*Some encyclopedias are goldmines	.90	
*Some words are jewels	.95	.90
*Some books are treasures	1.00	
Some friends are gems	.77	
Some ideas are diamonds	.90	
*Some stomachs are barrels	.70	
*Some mouths are canyons	.65	.50
*Some cheeks are balloons	.55	
Some legs are tree trunks	.53	
*Some desks are junkyards	.95	
*Some rooms are pigpens	.80	.70
*Some closets are warehouses	.75	
Some minds are swamps	.10	
*Some fogs are coats	.90	
*Some frosts are cloaks	.80	.80
*Some mists are veils	.85	
Some snowfalls are blankets	.67	

Table 1 (continued)

Metaphor	Clustered in pre-determined group	All three clustered together
*Some remarks are daggers	.85	
*Some jokes are spears	.90	.80
*Some tongues are knives	.90	
Some stares are slaps	.83	
Some smiles are razors	.83	

* indicates the metaphors used as standards for their groups

(Some vehicles are airplanes). No content words were repeated in any of the test sentences. In addition to the filler sentences, 18 practice sentences were generated, half of which were metaphors and half which were literal statements.

The presentation order of sentences was constructed as follows. Practice sentences were randomized and presented first. The filler items were randomly ordered and the triads of related metaphors were inserted randomly in the list of fillers. Each triad of related metaphors was presented as a block of three successive sentences. The only restriction on the placement of the related metaphor triads was that there be at least five fillers surrounding each triad. The number of fillers between triads ranged from five to twelve. Within this master list, six different versions were created by systematically rotating the related metaphors within each triad. So, if the three metaphors within each triad were ordered 1-2-3 in version 1, they were ordered 3-1-2 in version 2, 2-3-1 in version 3, 3-2-1 in version 4, 2-1-3 in version 5, and 1-3-2 in version 6. Each triad was presented in the same location in the list, only the order within the triad was varied.

Sixty undergraduate subjects (18 males, 42 females) were tested individually. Ten subjects were shown each of the six versions of the 114 sentence list. All of the sentences were typed in capital letters and reproduced in the center of a slide. The slides were presented via a rear-projection window in an I.A.C. chamber. Subjects were asked to indicate how easy or difficult it was for them to understand the sentences by pressing one of three response keys: the "D" key if the sentence was difficult to understand, the "E" key if it was easy to understand, and the "M" key if it was neither easy nor difficult, but somewhere in between (moderate). The position of the "E" and "D" keys was counterbalanced. The subjects were given a short rest halfway through the list.

In the final phase of the experiment, subjects were given a cued recall

task. They were presented with a randomized list of the 96 test items. Each sentence was typed with a blank space where the topic had been and subjects were asked to fill in the missing topic. For example, for the metaphor Some roads are snakes, subjects were presented with Some ________ are snakes, and were to provide the topic - roads.

The dependent measures were response latency, difficulty ratings, and correct recall probability. Response latency (in hundreths of a second) was measured from the onset of each slide until the response was made. Difficulty ratings were recoded from the response as E=1, M=2, and D=3. Recall was scored as the proportion correctly recalled (exact criterion) in each condition. The data were analyzed in two ANOVAs, using a quasi-F to test for significance across both subjects and item (Clark, 1973). The alpha level was set at $p=.05$ for all effects. Newman-Kuels were used for all individual comparisons. The first ANOVA compared the three different sentence types - literal filler, metaphor filler, and related metaphors. The second ANOVA tested for the effect of position (first, second, or third) within a triad using only the data from related metaphors. In order to have complete data on each metaphor at each position within the triad, subjects were matched across versions on the basis of their mean response latency to the literal fillers. These matched subjects were treated as a single subject in the ANOVA.

As indicated in Table 2, the filler sentences were responded to more rapidly, $F'(2,142)=38.16$, $MSe=17.81$, were subjectively easier to understand, $F'(2,140)=33.96$, $MSe=5.08$, and were better recalled, $F'(2,101)=32.79$, $MSe=1.65$, than either the related or filler metaphors. There were no significant differences between the filler and related metaphors in speed or difficulty of comprehension. However, the filler metaphor topics were recalled more often than were the topics of related metaphors.

These findings were expected given the nature of the sentences. The

Table 2

Mean Response Latencies, Difficulty Ratings, and Strict Recall Scores for Literal Fillers, Filler Metaphors, and Related Metaphors

	Sentence Type		
Dependent Variable	Literal Filler	Filler Metaphor	Related Metaphor
Response Latency	2.38	3.43	3.25
Difficulty Rating	1.08	1.60	1.53
Strict Recall	.79	.61	.47

literal fillers may have been easier to understand because the subjects were more familiar with their content or had actually encountered the ideas previously. On the other hand, subjects were less likely to have encountered the metaphors before because they are less common. Since there were no significant differences between the two types of metaphors, it appears that they were not perceived differently by subjects. Either type of metaphor, related or filler, was more difficult for subjects to understand and was recalled less often than were the literal sentences. This does not necessarily mean that metaphors are always more difficult to understand than literal expressions. The lack of familiarity with the metaphors may account for the longer response latencies . Ortony et al. (1978) found that familiar idioms were processed as quickly as syntactically and semantically comparable literal sentences. Also, Glucksberg, Gildea and Bookin (1982) found that comprehension of metaphors was so fast that it interferred with responding that the sentences were literally false. Perhaps more common metaphors would have required less processing than the relatively unique metaphors used in this experiment.

As indicated in Table 3, the metaphors in the first position were responded to more slowly, $F'(2,23)=12.71$, $MSe=2.24$, and were more difficult to understand, $F'(2,8)=29.80$, $MSe=.17$, than the metaphors in the second and third positions. There was no significant difference in the recall of topics across the three positions. Although there was a trend of decreasing response times and difficulty ratings as subjects progressed from the first to third position, the decrease between the first and second position was much larger than the decrease between the second and third position.

Priming was effective. The facilitation was large with only one priming metaphor. The first metaphor should have been the most difficult to understand if subjects had to search for or construct a ground. Once the ground was activated the second and third metaphors were understood more easily and

Table 3

Mean Response Latencies, Difficulty Ratings, and Strict Recall as a Function of Position Within Related Metaphor Triads

	Position in Triad		
Dependent Variable	First	Second	Third
Response Latency	3.53	3.15	3.10
Difficulty Rating	1.67	1.47	1.45
Strict Recall	.45	.47	.50

quickly. Since response times and difficulty ratings of metaphors in the second and third position were not significantly different, subsequent access to an already activated ground did not increase its effectiveness.

The recall data provided further evidence that the metaphors within the triads shared a common ground. The topics of related metaphors were recalled correctly less often than either filler metaphors or literal fillers. The majority of errors in recall of the related metaphors were confusions of topics within a related metaphor triad. Subjects confused topics within triads more than twice as often as they confused filler item topics (2.78 mean confusions per subject as compared to 1.28 mean confusions). Perhaps subjects confused the related metaphor topics more often because the topics within a triad could be interchanged without losing the shared meaning, or ground, of the metaphor.

The finding that metaphors could be consistently clustered on the basis of similarity of meaning, that priming was effective, and that topics within a triad were confused in recall indicate that metaphors share a common ground. Since grounds were restricted to unique topic-vehicle combinations, a conceptualization of the ground as a more abstract entity may be more accurate.

A number of researchers have proposed that grounds are fairly abstract. Grounds have been described as "conceptual bases" (Honeck, Riechmann & Hoffman, 1975), the interaction of several domains (Tourangeau & Sternberg, 1982; Verbrugge & McCarrell, 1977), and semantic fields (Glucksberg & Gildea, 1981; Keil, 1981). Glucksberg and Gildea (1981) have found that the comprehensibility of poor metaphors was enhanced when they were primed with their semantic fields. Additional research along these lines will enhance our understanding of the nature of metaphor and the comprehension process.

REFERENCES

Billow, R.M. Metaphor: A review of the psychological literature. Psychological Bulletin, 1977, 84, 81-92.

Black, M. More about metaphor. In A. Ortony (Ed.) Metaphor and Thought, Cambridge University Press, 1979.

Clark, H.H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research, Journal of Verbal Learning and Verbal Behavior, 1973, 12, 335-339.

Glucksberg, S. & Gildea, P. Understanding nonliteral speech: The used-car-lot theory of metaphor, Paper presented at the Symposium on Comprehension and Inference in Textual Information Processing, American Psychological Association, Los Angeles, 1981.

Glucksberg, S., Gildea, P., & Bookin, H. On understanding non-literal speech: Can people ignore metaphors? Journal of Verbal Learning and Verbal Behavior, 1982, 21, 85-99.

Honeck, R., Riechmann, P. & Hoffman, R. Semantic memory for metaphor: The conceptual base hypothesis, Memory and Cognition, 1975, 3, 409-415.

Keil, F. Semantic fields and the acquisition of metaphor. Paper presented at the meeting of the Psychonomic Society, Philadelphia, 1981.

Ortony, A. Beyond literal similarity. Psychological Review, 1979, 90, 227-234.

Ortony, A., Reynolds, R., & Arter, J. Metaphor: Theoretical and empirical research, Psychological Bulletin, 1978, 83, 919-943.

Tourangeau, R., & Sternberg, R.J. Understanding and appreciating metaphors, Cognition, 1982, 11, 203-244.

Verbrugge, R.R. & McCarrell, N.S. Metaphoric comprehension: Studies in reminding and resembling. Cognitive Psychology, 1977, 9, 494-533.

An Information-Processing Analysis of the Cognitive Processes Involved in Oral Reading*

Joseph H. Danks

Kent State University

PRESENTATION DRAFT: PLEASE DO NOT QUOTE WITHOUT PERMISSION.

American Educational Research Association

New York, 16 March 1982

* I thank Ruth Miller and Laurie End for their assistance with the research reported in this paper. The preparation of this paper and the research reported here were supported in part by Grant No. NIE-G-78-0223 from the National Institute of Education. Mailing address: Department of Psychology, Kent State University, Kent, Ohio 44242.

The main theme of this symposium seems to me to be that common educational problems and phenomena not only can be examined from many different perspectives, but must be so examined if we are to understand them. It is very likely that no one perspective has a corner on adequate analyses and solutions. My role in this symposium is to provide a cognitive science-psycholinguistic analysis of reading and to see how this analysis fits with the other perspectives represented here. To illustrate this orientation, I would like to analyze a specific reading situation that we have been studying the last several years--comprehension processes during oral reading (Danks & Fears, 1979; Danks & Hill, 1981; Danks, Bohn, & Fears, in press). I think that we have a reasonably good understanding of the basic process, but when we pushed our model in certain ways, we were relatively less successful. Specifically, we attempted to introduce a task demand manipulation via experimenter control. To telegraph my conclusion, one of the reasons that this manipulation was not as robust as we had anticipated, even though it seemed quite plausible at the outset, was that we were insensitive to some of the social interaction factors present in the experimental situation.

The basic question that we have asked is when during the comprehension process various kinds of information are utilized by the reader. In order to answer this question, we have used an oral reading task as an on-line measure of the difficulty of processing while a person is reading a story. We recognize that oral reading may be different in certain respects from silent reading, in fact we think that it is. However, oral reading occurs frequently in public schools today (Anderson, Shirey, & Mason, 1981) and is something that most literate people do relatively easily. Thus, it is a convenient task with which to study reading in a relatively natural situation.

We ask the reader, either a child or an adult, to read a story aloud while we tape record the performance. The general rationale is that we look at the types of oral reading errors and the locations of those errors as

indicative of the kinds of information, or misinterpretation of information, that might produce those errors. However, instead of waiting for readers to produce errors spontaneously, we insert various sorts of violations into the text in order to increase the likelihood that the reader will commit an error of interest.

For the stories used with children, which are the primary focus in this paper, we selected stories from primers with readabilities of 1.6, 3.5, and 5.6 (Fry, 1968). Since we intended to use these stories with second, fourth, and sixth graders, respectively, they should not be too difficult for the children to read in terms of their overall reading level. We identified 20 critical words for manipulation with the implicit restrictions that they be reasonably spaced throughout the story and that they be susceptable to the sorts of manipulations that we planned to introduce. They varied over several different parts of speech and several different positions within the sentences. For example, a portion of text from the second grade story is shown in Figure 1 and reads "The Ducks found a fat stick. Because the stick was so big the Turtle would be able to hold on better." Each critical word, big in the example, was violated in one of four different ways:

(a) In a lexical violation, the critical word was replaced with a pronounceable nonword of about the same lenght and shape. In the example, the critical word big was replaced with the nonword bis.

(b) In the second type of violation, a syntactic violation, the morphology of the critical word was altered, usually by changing the inflectional suffix, such that the syntactic structure of the sentence was violated. Where possible we attempted to change completely the part of speech, but often we did not have sufficient flexibility to do so. So some of the syntactic violations were within the same part of speech. In all cases, however, the reader could still figure out the meaning of the critical sentence even though it was syntactically ill-

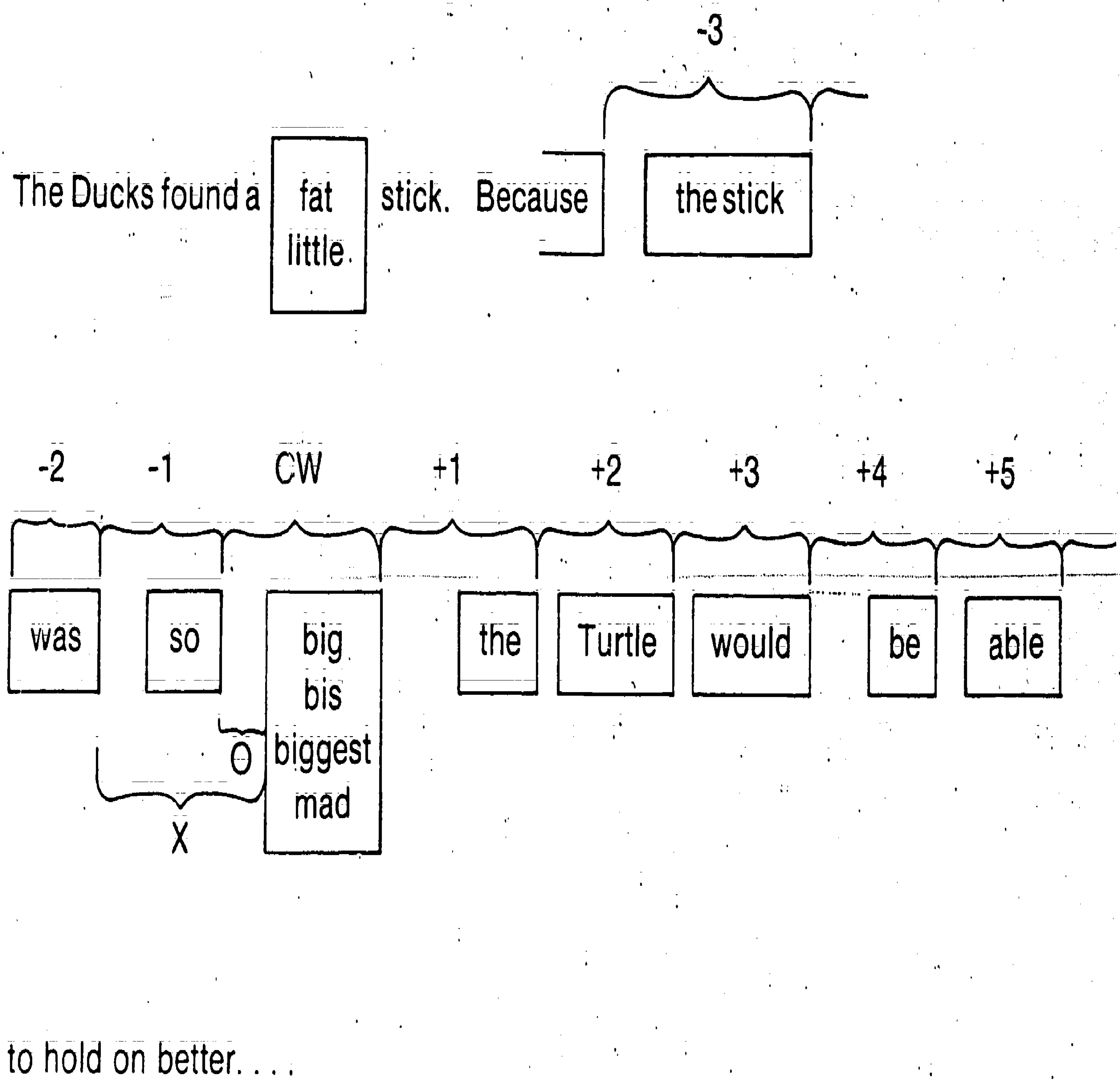

-3
The Ducks found a
fat
little
stick.
Because
the stick
-2
-1
CW
+1
+2
+3
+4
+5
was
so
big
bis
biggest
mad
the
Turtle
would
be
able
O
X
to hold on better. . . .

Figure 1

formed. In the example, the syntactic violation was formed by substituting biggest for big.

(c) The third type of violation was a semantic violation. Here a semantically anomalous word was introduced, that is, one that did not immediately make sense, but which was the same part of speech as the original critical word. In the example, mad was substituted for the critical word big. It was sometimes possible to make sense out of these semantically anomalous words, but it required extra processing to do so.

(d) The fourth type of violation was a factual violation. The critical word was unchanged, but the sentence preceding was changed such that a violation of fact was produced at the critical word. In the example, little was substituted for fat in the preceding sentence producing a factual inconsistency with the critical word big. Both of the sentences were fully grammatical, acceptable, and made sense in the story. However, they were factually inconsistent with each other, specifically at the point of the critical word.

(e) The fifth manipulation was a control condition in which the critical word and the preceding sentence both were left unaltered. This condition furnished a baseline for estimating the number of "spontaneous" errors that would be produced without our violations.

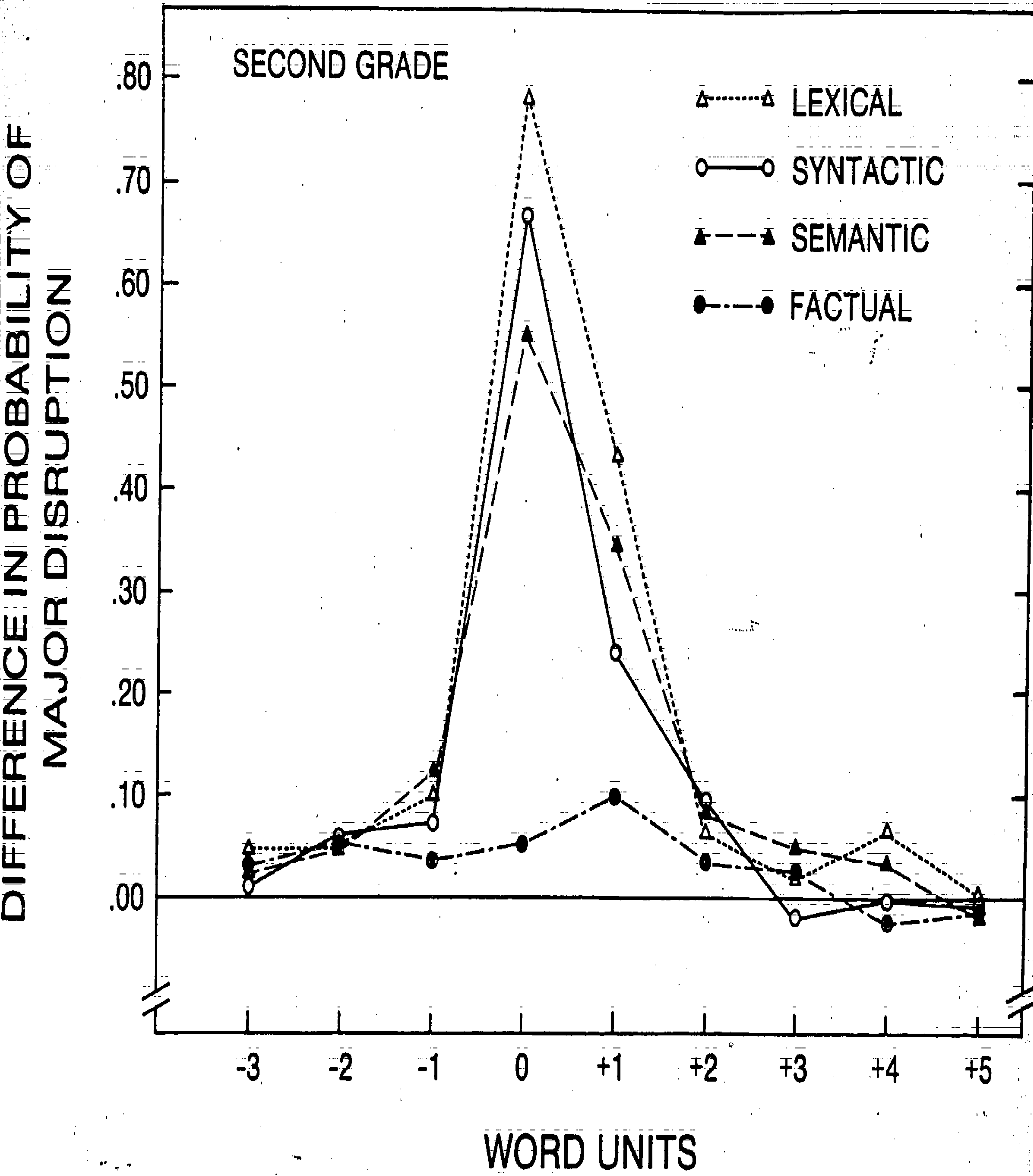
The stories were composed in five versions such that each violation occurred equally often at each critical word and an equal number of times in each version. Each version was read by 10 children at the designated grade levels, so that we had a completely counterbalanced design. School records and standardized test results were not available to us, but the children apparently were reading at or near grade level, or a little above, as estimated by their teachers. Our impressions from listening to the tapes confirmed the teachers' judgments. We did not have children who read either well above or well below grade level. In order to insure that children were comprehending the story, they were asked

several simple, literal questions after reading each quarter of the story.

The children's oral reading protocols were scored by dividing the text around each critical word into word units--three before each critical word and five after--as shown in Figure 1. In general, these word units were single words, although some word units were larger because the children tended to pronounce them together as a unit. At each word unit position we tallied whether a disruption occurred. A major disruption was defined as a substitution, omission, reversal, mispronunciation, stammering, regression, repetition, and pauses longer than those occurred normally for that reader. Only one disruption was scored per word unit per trial per reader. For presentation purposes, I have subtracted the probability of a disruption occurring in the baseline condition from the four experimental conditions in order to obtain a better estimate of the effect produced by the violations. The disruption curves for the second, fourth, and sixth grade readers are shown in Figures 2, 3, and 4, respectively.

The dominant impression that one gets from these figures is that the three sets of curves were very similar. There were some differences, but these tend to be relatively minor. The lexical, syntactic, and semantic violations all produced their largest disruptions while the critical word was being uttered or immediately before (word unit 0) and to a lesser extent in the word unit following (word unit +1). A few of the conditions were significantly different from the control as early as word units -2 and -1, that is before the critical word itself was even uttered, but there was no systematic pattern to these early disruptions. Likewise, there were a few significant disruptions at word units +2 and +3, but these were few. The dominant effect of the lexical, syntactic, and semantic violations in all three grades was that there was a peak of disruption at the critical word that continued for one word unit after.

The factual violation produced smaller, but significant disruptions in all three grades. In the second and fourth grades, this disruption occurred at



SECOND GRADE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS

Figure 2

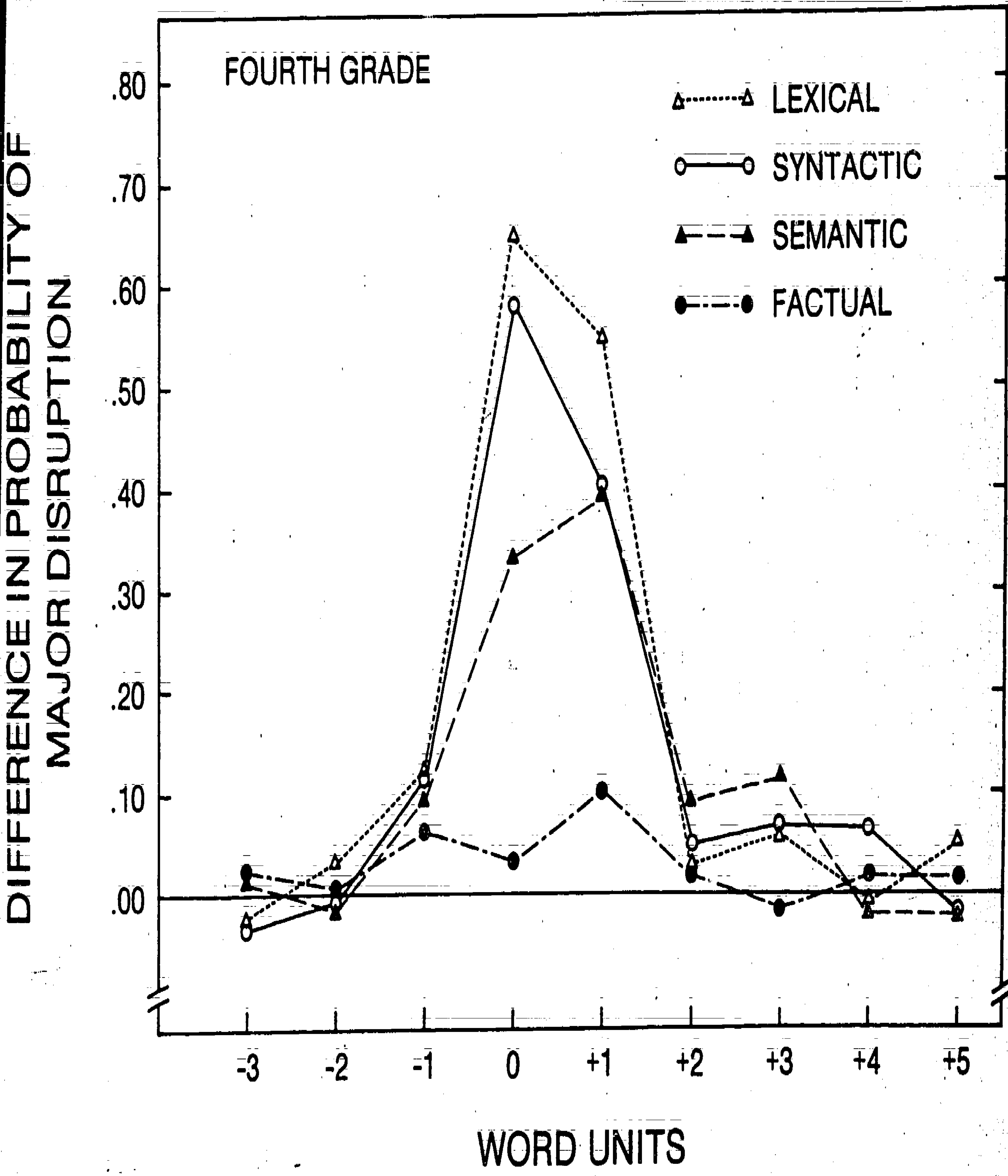

FOURTH GRADE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS

Figure 3

word unit +1, after the critical word had been uttered and after the other three violations had started their disruptions. In the sixth grade, the factual disruption occurred at the critical word. A possible interpretation of this change in the factual disruption is that as children begin to emphasize comprehension over decoding in the upper elementary school years, factual violations produce disruptions at an earlier point in the reading process. In the sixth grade, children processed the most abstract level of information earlier than did children in the lower grades.

This interpretation was belied, however, by the results from college readers, as shown in Figure 5. The factual violation did not produce a disruption until one word unit after the critical word had been uttered, a result that has been replicated in several other studies with skilled adult readers. So the "early" disruption peak from the factual violation in sixth graders was an anomaly, probably a result of specific story manipulations. Note that the lexical, syntactic, and semantic violations produced peaks of disruptions at the critical word and word unit +1 in college readers just as they did in the younger readers. Although all three of these violations were significantly different from the control at word unit -1, the dominant pattern in skilled adult readers was quite similar to that obtained with less skilled children readers.

There were some interesting qualitative differences in the disruptions produced in the syntactic condition. Readers at all grade levels tended to replace the syntactic violation with the original critical word, that is, they corrected the syntactic violation. This substitution accounted for 32%, 44%, 38%, and 54% of the second, fourth, and sixth graders', and college students' major disruptions, respectively. Many of these restorations were fluent in that there was no other disruption associated with the correction. The readers corrected the violation fluently and smoothly without pausing, hesitating, repeating, etc., many times without even realizing what they had done. This

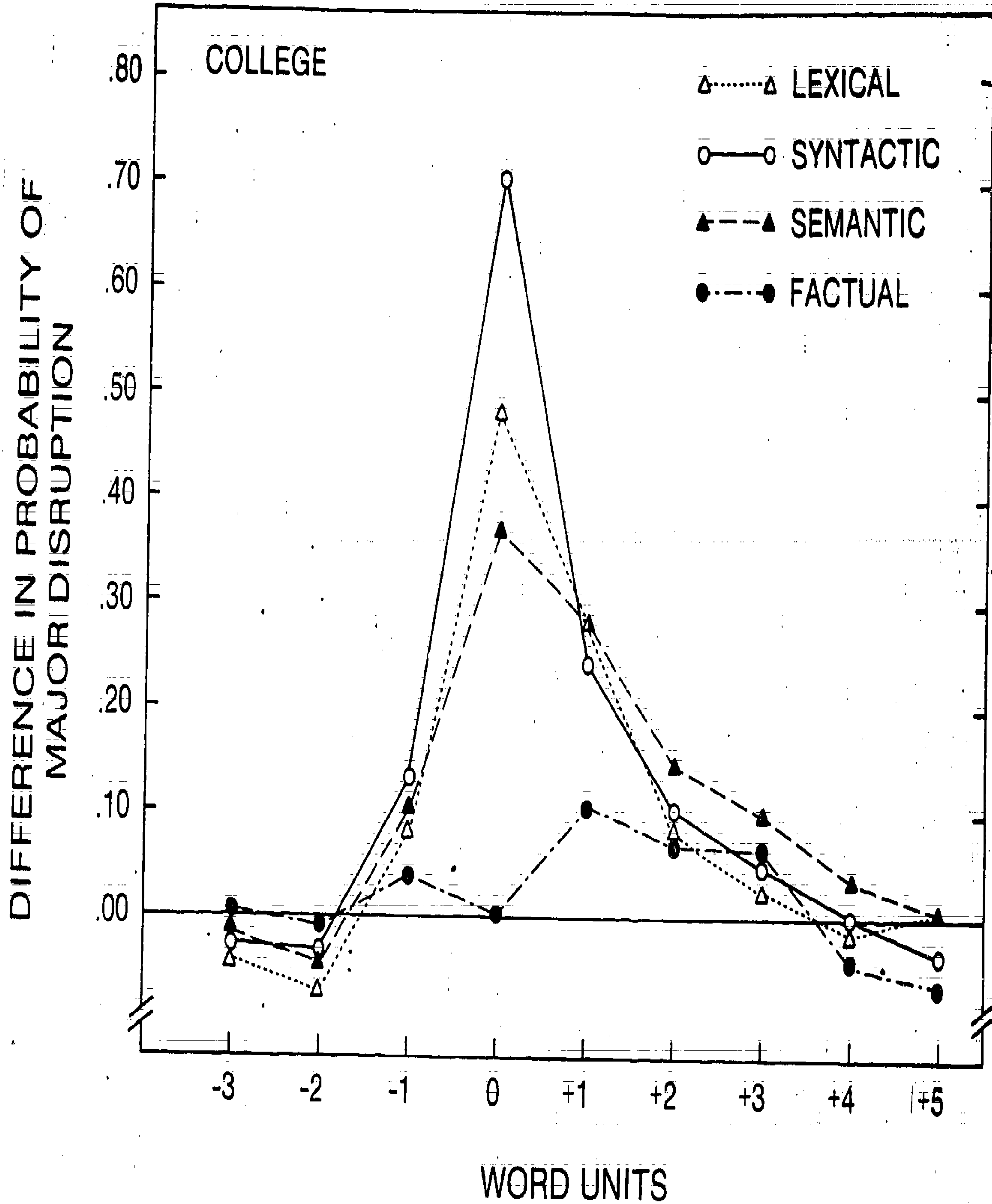

COLLEGE
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
0
+1
+2
+3
+4
+5
WORD UNITS

Figure 5

result suggests that there was considerable top-down processing since the sentence structure constrained the part of speech of the critical word and the readers added the appropriate inflection after "reading" only the stem of the critical word.

The model we have used to interpret these results emphasizes two components (actually there are additional components for which we have evidence, but these two are critical here). One component is lexical access, the process by which words are located in the mental dictionary. Not only is semantic and syntactic information located there, but also the articulatory information needed to render the word orally. Lexical access operates both with bottom-up information, as represented in the lexical violation, as well as with top-down information, as represented in the syntactic and semantic violations. The second component of the model is sentence integration. Here the meaning of the sentence is constructed by integrating the word meanings as well as information from preceding sentences. Integration of information from one sentence to the next begins before the clause or sentence boundary is reached (Danks, Bohn, End, & Miller, 1980). That is, each clause or sentence is not understood automously before integration with the information from preceding sentences. There is an on-going word-by-word integration of information as each word is accessed.

In this experiment, syntactic and semantic information were used in lexical access, nearly as much as the bottom-up lexical information was. The lexical violation did not produce an earlier disruption than did the syntactic and semantic information. Lexical information was critical for lexical access. Lexical access was blocked by the pronounceable nonword because it did not exist in the readers' mental dictionaries. If the syntactic and semantic information were used only after lexical access, say only in sentence integration, then one would expect the syntactic and semantic disruptions to have occurred after the lexical disruption. But in virtually every case, the syntactic and

semantic violations produced disruptions just as soon as did the lexical violation. Additionally, the restorations in the syntactic condition supported this conclusion that there was substantial use of top-down information in lexical access. Factual information, in contrast, seemed to be involved in a later stage of processing since its disruptions tended to occur after the disruptions from the other violations.

Why then were the children not more different from the adult readers, and why was there not a progression in the processing structure as the children became more proficient readers? Perhaps the children readers were well along the way to being skilled readers. This is partly true, but only in part, as indicated by the children's reading rates. The second grade children read about 120 syllables per minute, fourth graders about 160 syllables per minute, sixth graders about 180 syllables per minute, and the college students read about 270 syllables per minute. There were substantial differences in the reading rates and the children were not reading as fluently as well. Our impressions from listening to the tapes were that the children read more haltingly and less smoothly. Yet the patterns of disruptions resulting from the different violations were very similar across this relatively wide range of skill levels. Our interpretation of this lack of differences was that when reading relatively easy stories children can engage automatic processes which permit them to perform at an optimal level.

Now that we think we understand what processes were affected by the violations and we know what pattern of disruptions to expect, what changes in the pattern of disruptions are produced by the introduction of task demands similar to those imposed on children in typical classrooms? One reported phenomenon (Durkin, 1978-79) is that teachers supervising children in round-robin reading tend to focus attention on the oral rendition by correcting any deviations from the text and by giving relatively little attention to whether the children understand what they

read. Teachers frequently correct the children for mispronunciations and other sorts of oral reading errors, but there is little discussion of what a story means and how it can be interpreted. This emphasis on pronunciation, as opposed to an emphasis on comprehension, probably affects the sorts of processing children engage in while reading (Pehrsson, 1974).

We attempted to simulate these emphases in two groups of second grade readers. One group received pronunciation-emphasis instructions and a second received instructions that emphasized comprehension. In the pronunciation-emphasis condition, the children were instructed to read the story very carefully and accurately as if they were reading to a blind child. No mention was made of comprehension. During the reading of a practice story, ev[illegible]iation error was corrected by the experimenter, no matter how small, [illegible], the experimenter was hypercritical of their performance. Then the [illegible]er was presented with the experimental story. Although the experimenter did not interrupt or correct the children while they were reading the experimental story, she did reemphasize accurate oral production during short breaks between sections of the story.

In the comprehension-emphasis condition, the children were told to pay attention to the content of the story because they would be asked questions about the story when they were finished. The reading aloud was mentioned almost as an afterthought. While reading the practice story, they were not corrected in any way. After they were finished, they were asked some very difficult questions about the story. The experimenter pressed them for answers and urged them to pay more attention to what they were reading. The child was required to go back to the story and find the answer to any questions that he/she could not answer. The children then were given the experimental story and during the breaks the comprehension orientation was reemphasized.

In addition to the task emphases, we also manipulated the difficulty of

the story. Half of the readers in each task group were presented with the second grade story as in the first experiment. This story was relatively easy for the second grade readers. The other half of each task group were given the original fourth grade story. This story was relatively difficult, though not quite frustrating, for second graders. The practice story (readability = 1.6) was the same for all groups.

Twenty-five second graders were tested in each of the four conditions defined by the interaction of task emphasis and text difficulty. Five readers were given each version of the stories. Exactly the same stories were used as in the first experiment, including the same four violations (lexical, syntactic, semantic, and factual) and control. After reading the entire experimental story, all children were asked a standard set of questions that centered around the critical segments that had been manipulated, as well as giving a straight free recall of the story. The scoring and analyses of the protocols were the same as in the first experiment.

In general, we view children as having limited cognitive resources to devote to the tasks implicit in oral reading. In all cases, an oral rendition of the story was required. The task emphases changed the relative amount of cognitive resources allocated to oral production. With an easy story, the oral production task can be satisfied relatively easily leaving some resources for comprehension regardless of task emphasis. A difficult story, in contrast, cannot be processed quite so automatically to yield an oral rendition, leaving fewer resources to be applied to comprehension. The distribution of cognitive resources is indexed by the relative sizes, locations, and patterns of the distributions caused by violating different types of information.

Our expectations were that the disruptions produced by the violations in the stories would differ as a function of the task emphasis and the difficulty level of the story because of the differential demand on and allocation of available cognitive resources (Stanovich, 1980). Specifically, an emphasis on

pronunciation would focus attention on the less abstract information, such as lexical and syntactic, producing relatively larger and perhaps earlier disruptions than would semantic and factual violations and than would a comprehension emphasis. The syntactic information was considered closely related to pronunciation because syntactic information would be needed to produce an appropriate intonation, stress, etc. The factual and semantic violations would produce less of an effect because the children would not have to pay attention to the content of what they were reading. In contrast, the comprehension emphasis would show an increase in the disruption from semantic and factual violations as compared to the pronunciation emphasis because the reader would be attempting to understand the story. Lexical and syntactic information also are relevant to understanding the story, so their disruptions probably would not be diminished under a comprehension emphasis.

The difficulty manipulation also would affect differentially the size and pattern of the disruptions. With a difficult story, children would have to devote relatively more attention to lexical and syntactic information in order to read aloud, thereby reducing the attention they could devote to more abstract information. Thus, we expected the lexical and syntactic violations would be relatively more disruptive in difficult stories than would semantic and factual violations. With easy stories, the children could process information at all levels because the less abstract information would not consume a disproportionate amount of attention.

Furthermore, we expected task emphasis and text difficulty to interact with each other. With the easy story, children might be able to comprehend the story even while emphasizing accurate pronunciation, so that semantic and factual violations might have some effect even with a pronunciation emphasis. However, with a difficult story, children would have difficulty producing an oral rendition regardless of emphasis, so that comprehending the story would receive relatively

less attention, and thus the semantic and factual violations would have a relatively smaller effect, even with a comprehension emphasis.

The results are presented in the next four figures (Figures 6, 7, 8, and 9) using the same format that was used to present the results of the first experiment. The dominant impression that one receives from the four figures is one of similarity. There were no salient differences in the patterns of disruptions resulting from the task emphasis and text difficulty manipulations. The curves were somewhat more variable than those of the first experiment because they were based on half as many subjects.

In all conditions the lexical, syntactic, and semantic violations produced peaks of disruption at the critical word with slightly smaller disruptions one word unit after. Some of the disruptions were significant one or two word units before the critical word as we had obtained in the first experiment. However, the major differences came from the factual violation. There was a significant disruption from the factual violation in only two conditions, namely, the pronunciation-easy story and the comprehension-difficult story conditions. In both cases, the factual disruption was significantly different at the critical word. But there was no factual disruption present in the comprehension-easy story condition, a condition where we most expected to obtain one. If the children were disrupted by the factual violation in the easy story with a pronunciation emphasis, they must have been able to process the easy story at all levels of information. That being the case, they surely should have been able to process factual information in the same story when comprehension was emphasized. We have no explanation for this anomaly.

Although there were no obvious differences in the overall pattern of disruptions, a more finely grained analysis uncovered some significant differences in the direction we had expected. Disruptions at or immediately after the critical word (word units 0 and +1) reflected primarily disruptions to lexical

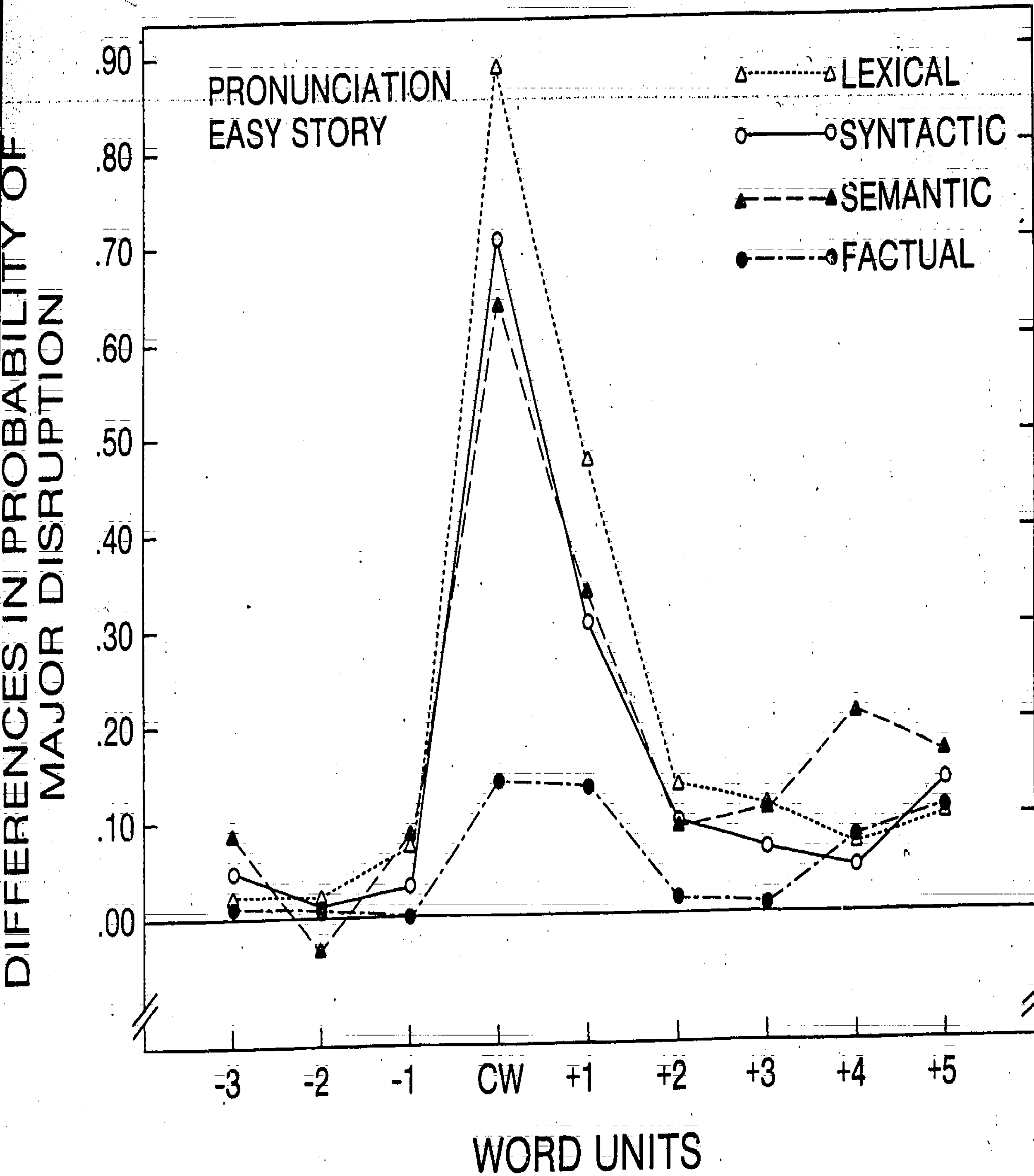

PRONUNCIATION
EASY STORY
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCES IN PROBABILITY OF MAJOR DISRUPTION
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD UNITS

Figure 6

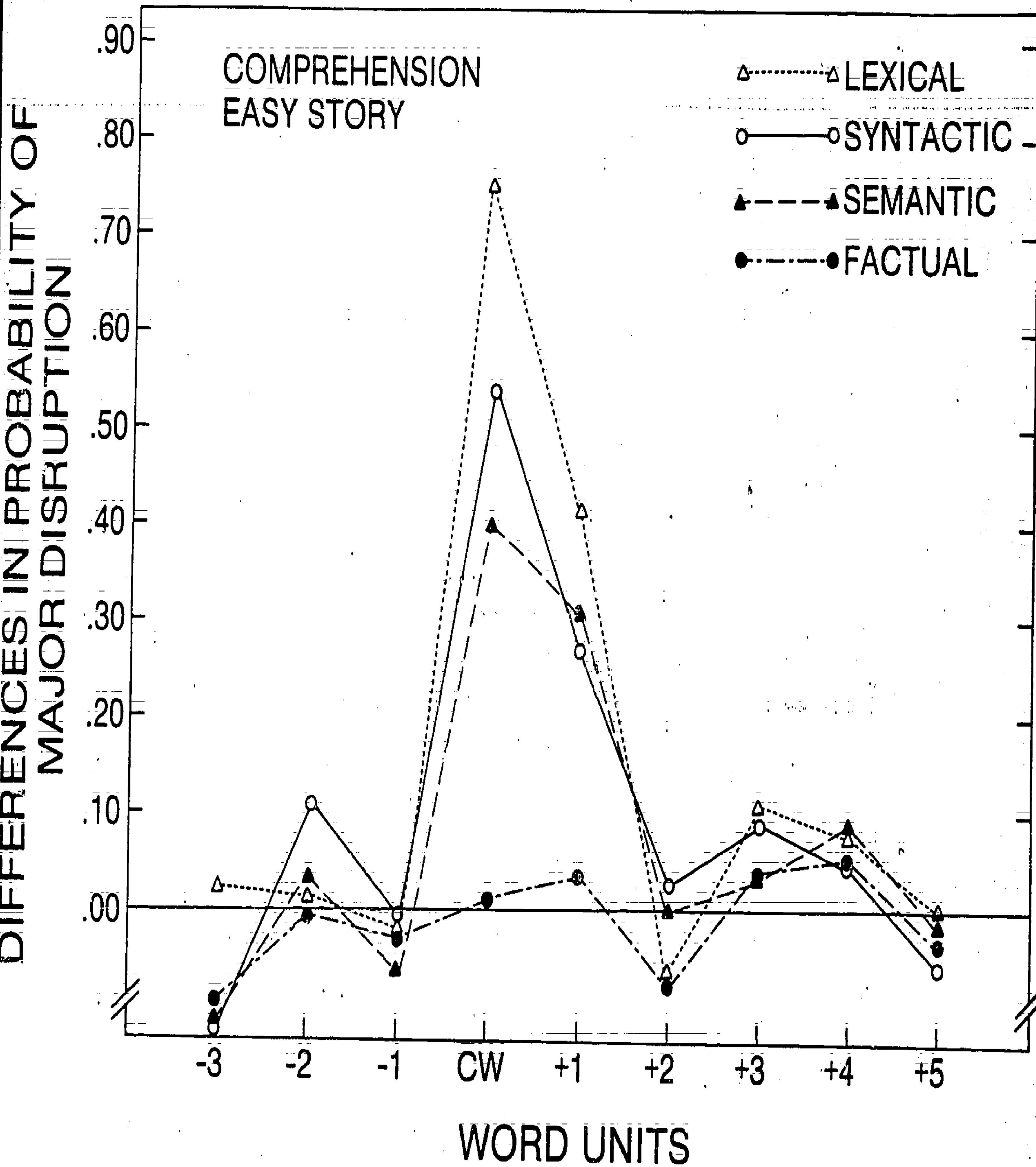

COMPREHENSION
EASY STORY
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCES IN PROBABILITY OF MAJOR DISRUPTION
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD UNITS

Figure 7

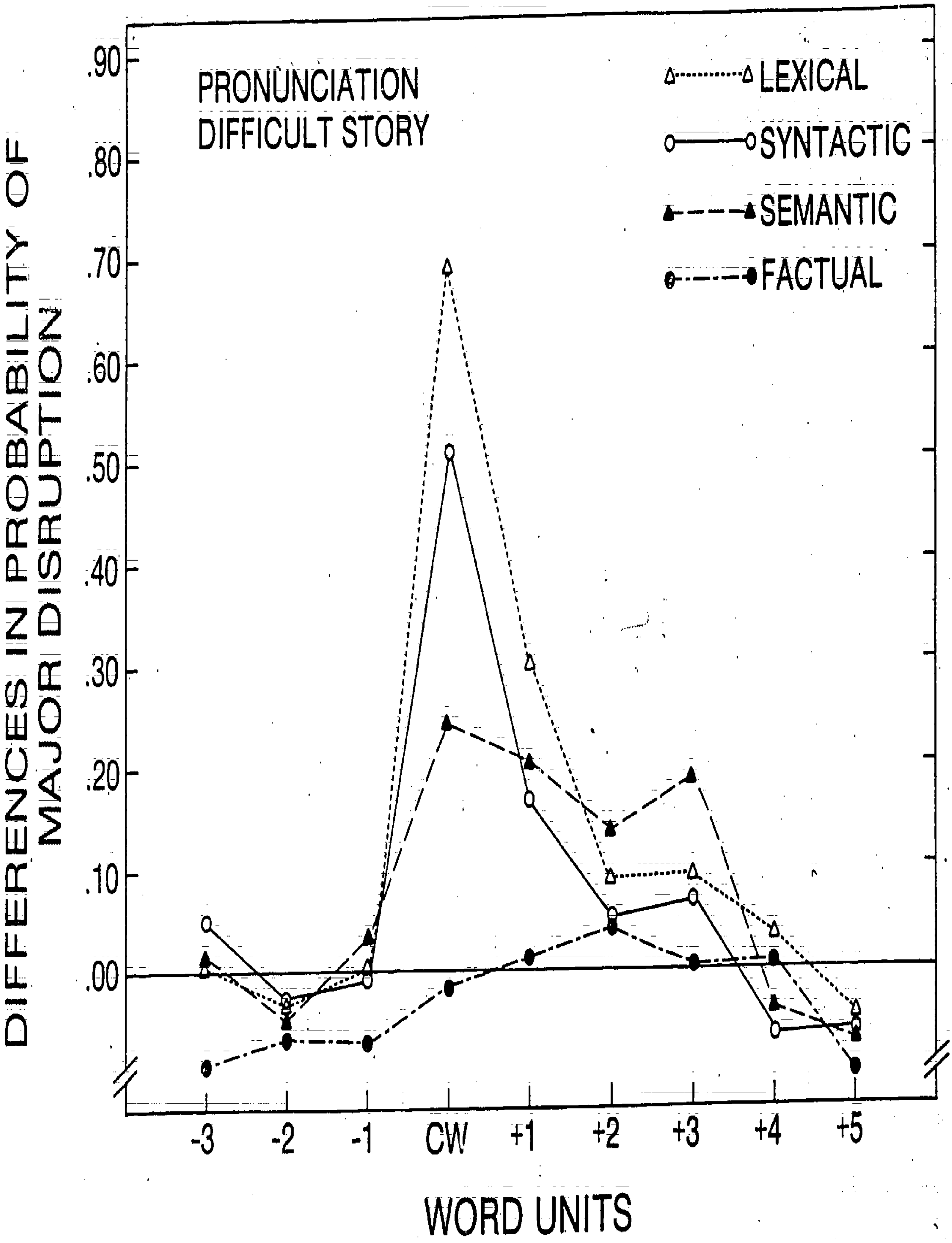

PRONUNCIATION
DIFFICULT STORY
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCES IN PROBABILITY OF
MAJOR DISRUPTION
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD UNITS

Figure 8

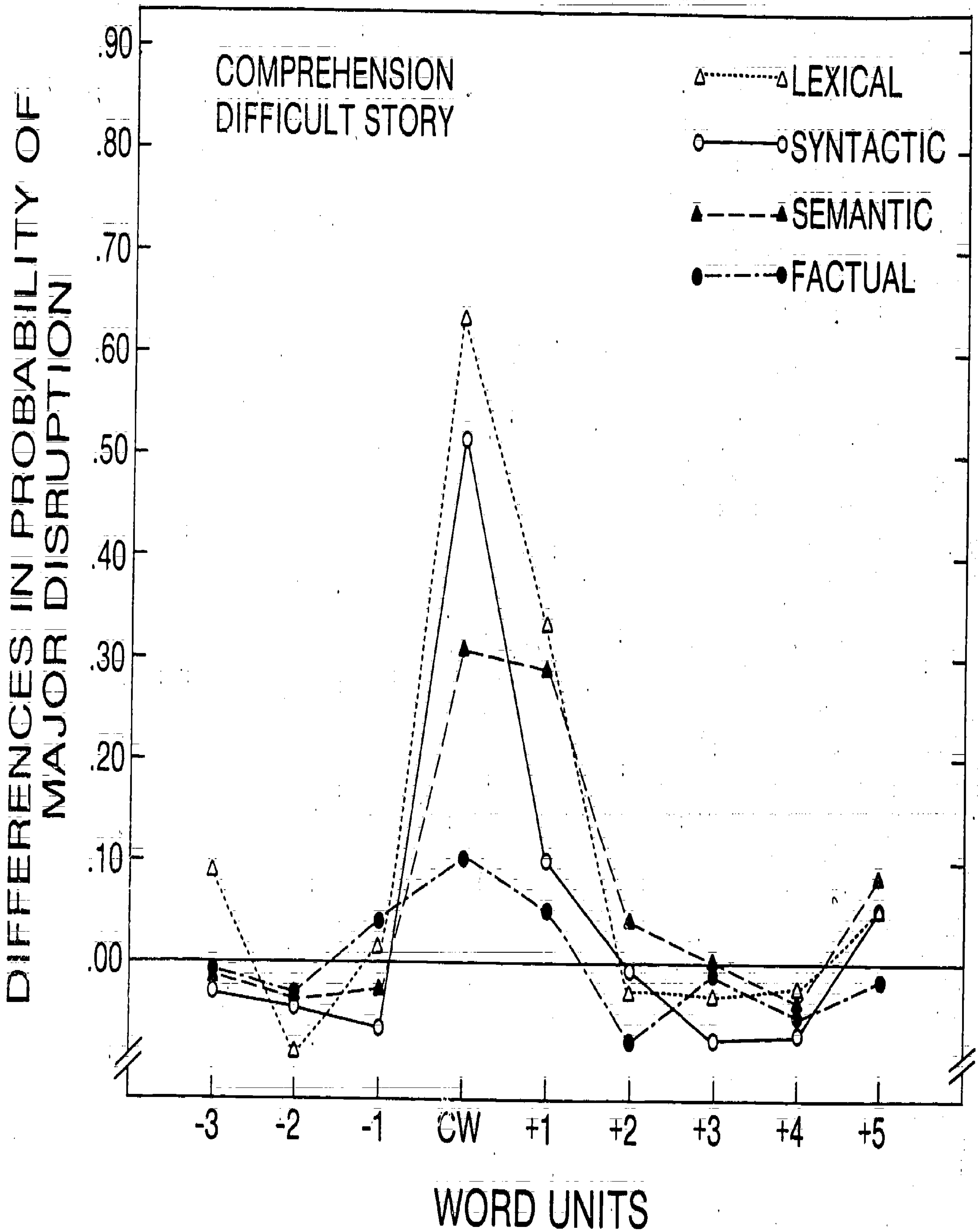
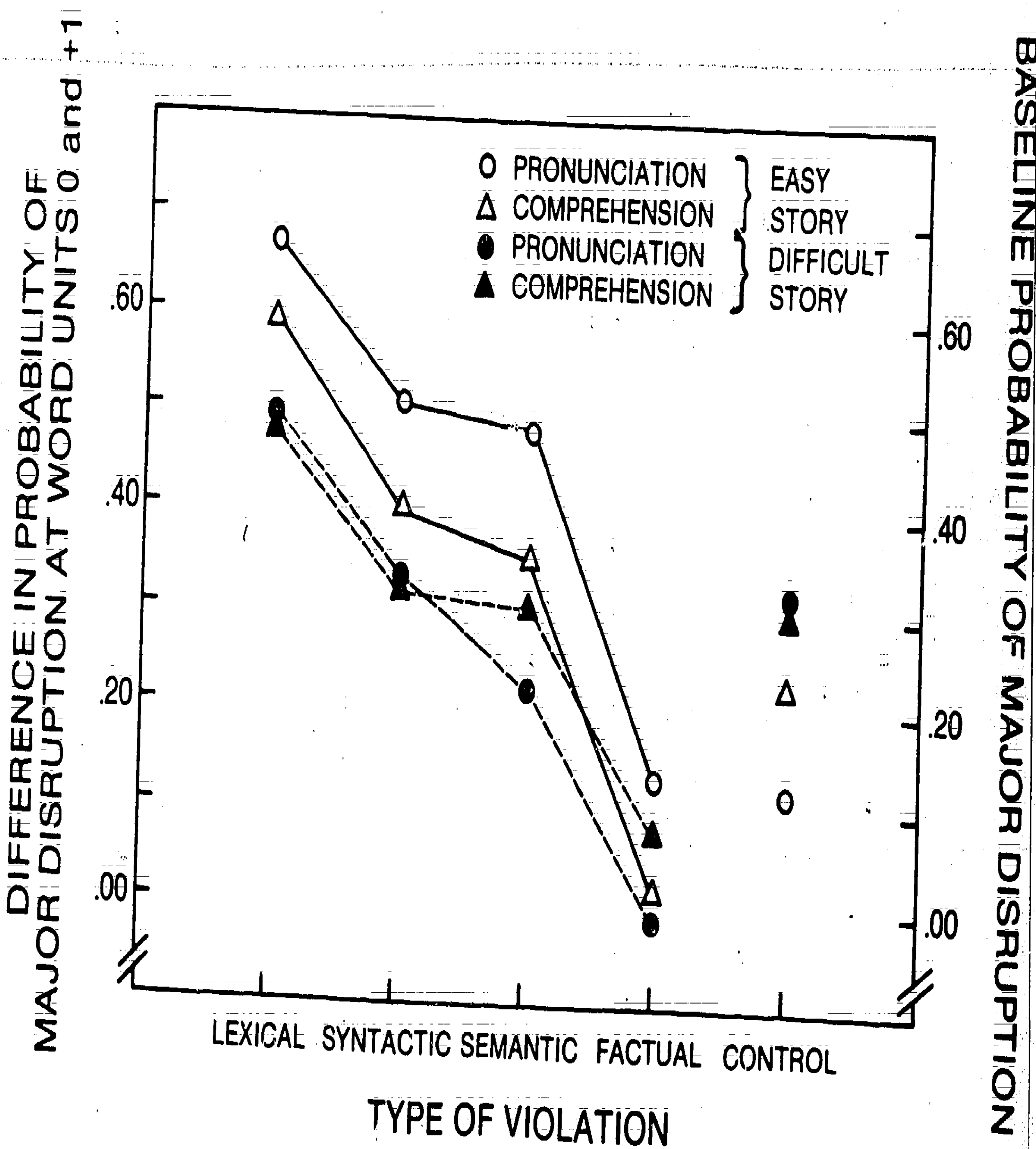

COMPREHENSION
DIFFICULT STORY
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
DIFFERENCES IN PROBABILITY OF MAJOR DISRUPTION
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD UNITS

Figure 9

access and to the attempted integration of that word's meaning into the representation of the story. Since lexical access is necessary for oral production (because the articulatory program must be accessed) and also is involved in comprehension processes, effects at this point should reflect differences in both task emphasis and text difficulty. In Figure 10, the average magnitude of the disruptions at word units 0 and +1 is presented as a function of the violation types. On the right, the baseline probability of a disruption is presented. In the baseline-control condition, there were fewer disruptions with the easy story than with the difficult one, as one would expect. There also were fewer disruptions with the pronunciation emphasis, but only for the easy story. With the difficult story, there was virtually no difference between comprehension and pronunciation emphases. These results support the analysis that, when the story was difficult, readers were less able to vary their processing to meet task emphases.

With respect to the disruptions (difference scores between control and violation conditions), there was a regular decrease in the size of the disruption with increasing abstractness of the information violated--from lexical to syntactic to semantic to factual. With the easy story, the pronunciation and comprehension curves were nearly parallel with the comprehension disruptions being significantly less than the pronunciation disruptions. Why did a comprehension emphasis produce smaller disruptions than a pronunciation emphasis? In the control condition there were more errors under the comprehension emphasis. These disruptions reflected primarily disruptions of lexical access because the most immediate demand in both conditions was the production of an oral rendition. That demand was not stressed in the comprehension emphasis, but it still was present. However, it was stressed by the pronunciation emphasis. So when there was a violation that derailed the lexical access process, it was more disruptive to the pronunciation group because of the greater emphasis to produce a perfect



DIFFERENCE IN PROBABILITY OF MAJOR DISRUPTION AT WORD UNITS 0 and +1
BASELINE PROBABILITY OF MAJOR DISRUPTION
PRONUNCIATION
COMPREHENSION
EASY STORY
PRONUNCIATION
COMPREHENSION
DIFFICULT STORY
.60
.40
.20
.00
LEXICAL
SYNTACTIC
SEMANTIC
FACTUAL
CONTROL
TYPE OF VIOLATION

oral rendition.

With the difficult story, however, there was no difference between the pronunciation and comprehension emphases at the two lower levels of violations--lexical and syntactic--because both groups were near the limits of their abilities to produce an oral production. At the two more abstract levels of violation--semantic and factual--the comprehension group was disrupted more than was the pronunciation group. The pronunciation group plunged ahead being relatively less affected by the semantic and factual violations since that information was less relevant to the pronunciation task. The comprehension group was trying to understand the difficult story, so the semantic and factual violations disrupted that comprehension process. Although these effects were small, they were reliable. These more detailed results, thus, supported the analysis of the reading process as it operated in these experiments. The effects simply were not as robust as expected.

Restorations of the original critical word when there was a syntactic violation followed a pattern expected from the conceptual analysis of the conditions. With the easy story, a greater percentage of the disruptions were restorations under the comprehension emphasis (41%) than with the pronunciation emphasis (26%). With the pronunciation emphasis, readers were more careful to read what was printed, whereas with the comprehension emphasis, the linguistic context exerted a top-down influence that led the reader to restore the syntactic form of the critical word. In the difficult story, this difference disappeared: the percentage of disruptions that were restorations was virtually the same under the comprehension emphasis (35%) as under the pronunciation emphasis (36%). With the difficult story, readers had to devote more cognitive resources to less abstract types of information so there was less opportunity for the top-down linguistic context to influence lexical access.

Turning now to the results of the recall test, the answers to the compre-

hension questions that were asked of all children immediately after they finished reading the story, the percentage of questions answered correctly was significantly greater for the easy story (78%) than for the difficult story (57%). However, there was virtually no difference between the two task emphases in amount recalled (pronunciation emphasis = 69%, comprehension emphasis = 66%). Whatever the immediate task emphasis, all readers were comprehending and remembering the story well enough to answer the questions. The failure to find a difference between the task emphases was especially surprizing because the recall test was expected by the comprehension group, so they were presumably remembering intentionally. The recall test was a surprize for the pronunciation group because they had not been given any questions after the practice story, so their recall was incidental to the oral production task. There also was no hint of any interaction between the task emphasis and text difficulty conditions. This result also was unexpected. The pronunciation group recalled as much of the easy story as did the comprehension group (77% and 79%, respectively), a result consistent with the notion that with an easy story children could both produce an oral rendition and comprehend the story at the same time. However, the difficult story should have been sufficiently difficult to occupy most of their resources in producing an acceptable oral rendition, without having much left over for comprehension. But even with the difficult story, the pronunciation group recalled as much as did the comprehension group (56% and 59%, respectively). Perhaps the comprehension questions were not sufficiently sensitive, yet they were sensitive enough to yield a difference between story difficulties.

In summary then, with regard to the lexical access and sentence comprehension processes, the effects were consistent with our model of how reading comprehension processes operate. The pronunciation emphasis had its primary effect on lexical access by focussing attention on the information types that most facilitated producing an oral rendition of the story, namely, lexical and syntactic information.

The comprehension emphasis, in contrast, tended to focus relatively more on semantic and factual information that would be used to construct a representation of the story. The adjustments to processing were not very robust, but were evident only in detailed analyses of the reading aloud protocols. Regardless of the task emphasis, the second graders were able to retain a representation of the text.

How then are we to interpret this relative lack of robustness? We think that there are two primary explanations as to why we did not obtain larger effects from the task-emphasis manipulations. One possibility is that the task-emphasis manipulations were too short-term to affect the reading process. The children were in the experiment about 30 minutes, so the experimenter had a relatively brief period of time in which to implement the task emphases, in contrast to the hours and hours of classroom instruction. The other side of this explanation is that children simply are not very flexible with their reading strategies. They are unable to vary their reading processes in any appreciable ways to meet specific task demands. Their reading processes change, but only with lots of practice and continued instruction. So we should not have been surprised that the short-term manipulations did not have much impact on the ingrained processes that children had acquired through classroom instruction.

A second possible explanation is that in the experimental situation the implicit social demands overwhelmed the specific manipulations. In all conditions, we used a one-on-one examination of the children's oral reading by an adult examiner. The child was under implicit pressure to perform for the experimenter, just as the child performs in many similar situations for the teacher. Performing well in reading implicitly means reading aloud accurately and understanding what is being read. The implicit task has components very similar to the manipulations, but the relative weighting of these subtasks has been set by the classroom teacher's use of oral reading and not by the manipulations. From the child's

perspective, then, the task-emphasis manipulations were relatively minor aspects of what the child perceived as an intensive individual examination of his/her reading skill. Of course, the experimenter was friendly, cordial, and tried to put the children at ease; she tried to emphasize that the experiment was non-evaluative of individuals and the results would not affect their grades. But it seems evident in retrospect that the children still perceived this situation evaluatively, as an examination, despite demurrals to the contrary. So we think that a major reason for the lack of robustness of the task-emphasis manipulations was our insensitivity to the children's perceptions of what the experimental situation was all about. These perceptions are not changed easily, and perhaps not at all. We should be pleased with the results that we did get.

What solutions are there to this problem, assuming that we are not satisfied with the current evidence of how flexible and adaptable (or inflexible and non-adaptable) children's reading processes are? One solution is to identify teachers with differing orientations toward reading comprehension and oral reading specifically. Then using experimental stories like ours, one could determine the extent to which children adapt their reading strategies to satisfy the teacher's implicit task emphasis over the period of a year. The testing could be done in the classroom as part of the regular instructional program, rather than in separate experimental situations, thus reducing the more intense one-on-one evaluation demands.

This approach does not solve all of the problems, however, because we have lost the ability to control the manipulation of the task emphasis. What kinds of teachers adopt or develop the different orientations? What kinds of children flourish or flounder under the different orientations? A better solution is a research program in which multiple research orientations interact (Danks, 1982). Neither a naturalistic nor an experimental approach can provide all of the answers. But naturalistic studies can suggest hypotheses that are testable in

experiments and experimental results require naturalistic observations to be properly interpreted and validated. When dealing with complex social-cognitive systems, and the educational process certainly is one, we should not restrict ourselves to only a few research strategies, but converge on the issues from many angles.

References

Anderson, R. C., Shirey, L. L., & Mason, J. M. Aspects of teaching method, materials, and child characteristics that make a difference in round robin reading. Psychonomic Society, Philadelphia, 1981.

Danks, J. H. Communication without cross-talk. (Review of Children's oral communication skills edited by W. P. Dickson.) Contemporary Psychology, 1982, in press.

*⇒ Danks, J. H., Bohn, L., & Fears, R. Comprehension processes in oral reading. In G. B. Flores d'Arcais & R. J. Jarvella (Eds.), The process of language understanding. Chichester: Wiley, in press.

Danks, J. H., & Fears, R. Oral reading: Does it reflect decoding or comprehension? In L. B. Resnick & P. A. Weaver (Eds.), Theory and practice of early reading, Vol. 3. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.

Danks, J. H., & Hill, G. O. An interactive analysis of oral reading. In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.

Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1978-79, 14, 481-533.

Fry, E. B. A readability formula that saves time. Journal of Reading, 1968, 11, 513-516; 575-578.

Pehrsson, R. S. V. The effects of teacher interference during the process of reading or how much of a helper is Mr. Gelper? Journal of Reading, 1974, 17, 617-621.

Stanovich, K. E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 1980, 16, 32-71.

* Danks, J. H., Bohn, L., End, L., & Miller, R. Integration of sentence meanings in stories. Psychonomic Society, St. Louis, 1980.

Polish Psychological Bulletin
1981, Vol. 12(1), 25—31.

DANUTA KURKIEWICZ
IDA KURCZ

Institute of Psychology
University of Warsaw

JOSEPH H. DANKS

Department of Psychology
Kent State University

READING COMPREHENSION PROCESSES IN POLISH AND ENGLISH

Data obtained in an English-language experiment, in which subjects had to read aloud a story containing either syntactic, or semantic, or factual inconsistencies, were compared with the results of a parallel Polish-language experiment. The general pattern of disruptions in the reading process was very similar in the two languages, except for the syntactic condition, where — probably due to the specific role of word endings in Polish — the magnitude of the disruptions was much larger, and substitution errors were much less frequent, in Polish than in English. The ultimate conclusion is that reading comprehension processes are functionally the same in Polish and English and are best represented by an interactive model.

Language comprehension in general and reading comprehension specifically involve the construction of a meaning representation for a linguistic input. A physical stimulus, speech or print, undergoes a series of transformations as a listener or reader constructs a meaning. Psychological studies of comprehension have focused on the form of the final meaning representation instead of on the processes by which that representation is constructed. The nature of the comprehension process is not self-evident from a rational or linguistic analysis (cf. Fodor, Bever & Garrett, 1974). Intuitive models usually posit a "bottom-up" interpretation of an input, i.e., a unidirectional series of processing stages as the physical input is transformed into a meaning representation. However, these interpretive models seriously understimate how much more abstract information facilitates meaning construction (Danks & Glucksberg, 1980; Rumelhart, 1977).

The paper is a result of collaboration that has developed under an exchange agreement between the University of Warsaw and Kent State University. The Polish language experiment is part of an M. A. thesis presented by the first author to the University of Warsaw, supported by Grant W.2.1.-W-266. The English language experiment and the preparation of this report were supported in part by Grant No. NIE-G-78-0223 from the National Institute of Education to Kent State University.

Requests for reprints should be directed to either Joseph H. Danks, Department of Psychology, Kent State University, Kent, Ohio 44242, USA, or to Ida Kursz, Instytut Psychologii UW, Stawki 5/7, 00-183 Warsaw, Poland.

For example, phonemes and letters are easier to understand when presented in the context of a word, words are easier to understand in a sentence, and sentences are easier to interpret in a large context such as a story or paragraph. Such influence is called "top-down" because higher, more abstract information facilitates (or sometimes hinders) lower level processing.

Since both top-down and bottom-up effects occur, an adequate model of language comprehension must permit both directions of information flow. Interactive models accommodate this requirement most easily (Danks, 1978). They typically allow several sources of information to be active at any given time while a meaning representation is constructed more or less continuously. While interactive models are most acceptable as a class, there still exists a possibility that an adequate bottom-up model can be devised. Forster (1979) has proposed that various sources of information are processed in parallel but independently, and that the final integration (meaning computation) is achieved at the end of some portion of material, after a clause boundary, for example. There is evidence, however, that meaning is constructed more or less continuously during text prosessing (Marslen-Wilson, Tyler & Seidenberg, 1978).

There are several different types of information available to the listener or reader during comprehension: physical — the actual speech or print; lexical — information stored with the representation of a word in the mental dictionary; morphological — the various morphemes contained within a word; syntactic — how a word fits into the grammatical structure of a sentence; semantic — how a word's meaning contributes to the meaning of a sentence; textual — how the meaning of a sentence fits with the schema for the story or text being processed; and factual — how a story or text and its components relate to a reader's knowledge about the world.

In the experiments discussed here, we focused on how syntactic, semantic, and factual[1] sources of information contributed to a reader's understanding of a story, and how reading comprehension processes might be different in Polish and English. The basic rationale was to violate or distort one of these information sources at specific points in a story. Subjects then read the story aloud. When they encountered a violation, their comprehension processes would be disrupted at the point in time that they attempted to use the violated information to understand the story. Because increased attention would have to be devoted to resolving the distortion, there would be less attention available for producing the speech corresponding to the text. Thus, the shift in attention would cause a disruption in the reader's oral production at the point in time when he attempted to use the violated information.

METHOD

Material

A story selected from an American magazine had been used in previous experiments using this procedure (Danks & Hill, in press). The story was translated into Polish with some adjustments, e.g., names of characters, to make it appropriate for Polish readers. Sixteen words (out of about 2000 words in the story) were identified as critical words for manipulation. In the story, a school girl is injured in a bus accident. In the following example, the girl's mother has just heard about the accident and is worried about her daughter. "Her daughter had always been weak physically. Because of this,

−5 −4 −3 −2

/ she even / imagined / her / daughter /

[1] In these experiments, we did not differentiate between violations of textual and factual information. In any case, both are more abstract than either syntactic or semantic information.

−1 being / 0 injured / +1 by other +2 children / +3 while / +4 trying / +5 to get / out of the wrecked bus".[2] The Polish translation was "Córka jej zawsze była −5 słaba −4 fizycznie. Z tego powodu / nawet / wyobra- −3 żała / sobie, / −2 że córka / −1 została / po- 0 trącona / +1 przez inne / +2 dzieci / +3 w chwi- +4 li / gdy usiłowała / +5 wydostać się / z uszkodzonego autobusu." The critical words were *injured* and *potrącona*.

For the *syntactic* manipulation, the base morpheme(s) of the critical word was retained, but the ending was altered such that the word was the incorrect part of speech for that point in the sentence. For example, the verbs *injured* and *potrącona* were changed to the nouns *injury* and *potrącenie*. Although the same formal manipulation was performed in both Polish and English, word endings function very differently in the two languages. In Polish, the part of speech (as well as some other properties) of most words can be determined directly from the word ending; e.g., *-ek*, as in *korek*, indicates a masculine noun. The endings are similar to English endings such as *-tion*; however, while most content words in Polish have such endings, only a few English words are so marked. So the syntactic manipulation was much easier to execute in Polish than in English. In addition to endings that mark part of speech, virtually all content words in Polish also have inflections which indicate the syntactic function of the word in the particular sentence. For example, *kobieta* is nominative case, *kobiety* is genitive, and *kobietą* is accusative.[3] While English has a few inflections, e.g., singular/plural, genitive in some cases, and verb tense, it depends primarily on word order to indicate syntactic function. In Polish word order is relatively much freer, serving primarily a pragmatic function. We expected that, since word endings, including inflections, are relatively more important for determining syntactic structure in Polish, introducing a syntactic violation by altering a word ending would be more disruptive for Polish readers than for English readers.

For the *semantic* manipulation, the critical word was replaced with another word of the same part of speech, but one which was semantically anomalous in that sentence. For example, *injured* was replaced with *planted*, and *potrącona* was replaced with *posadzona*. For the *factual* manipulation, the critical word was not altered. The sentence preceding the critical word was altered so that an inconsistency was created between that sentence and the critical word. In the example, *weak* and *słaba* were changed to *strong* and *silna*, respectively. The daughter being strong is inconsistent with the mother being worried about her being injured by the other children. There were no apparent differences between Polish and English in implementing the semantic and factual manipulations, so we expected no differences in the timing or size of the disruptions resulting from them.

Design and procedure

Sixteen critical words were identified and manipulated as described. Four versions of the story were prepared such that each critical word occurred once in each manipulation, including a control condition in which the text was not altered in any way. The stories were read aloud by skilled readers — 40 undergraduate students at the University of Warsaw and 40 at Kent State University, 10 readers per version. In order to insure that sub-

[2] An explanation of the numbers and slashes in the example follows in the next section; see also Figure 1.

[3] Violation of inflections, which is practically impossible to accomplish in English, has been introduced in a second experiment with the Polish text. The results of that experiment will be reported in a later paper.

jects were attempting to comprehend what they were reading, the subjects gave a summary of what they had read following each fourth of the story.

For scoring purposes, the portions of the sentence preceding and following each critical word were segmented into five word units. These word units are marked by slashes and are numbered from the critical word in the example sentences above. The word units were one word long except when readers

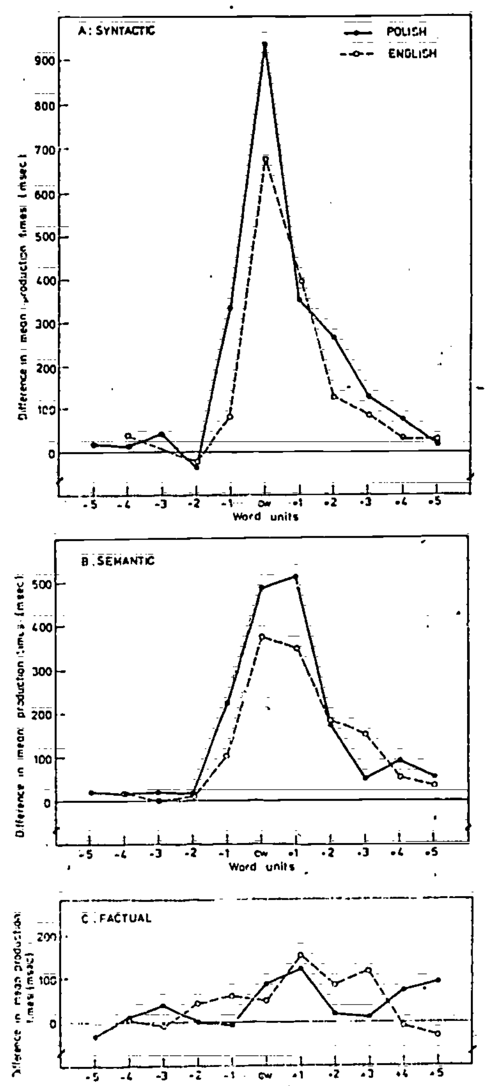


Fig. 1. Mean increase in production times of syntactic, semantic, and factual manipulations (panels A, B, and C, respectively) over the control condition as a function of word unit position for the Polish and English language experiments.

tended to pronounce two words as a single unit, e.g., *she even* and *w chwili*. Two dependent variables were scored for each word unit — production times and major disruptions. Production time was the time it took a reader to pronounce a word unit, timing from the end of the previous word unit to the end of the one being scored. Any pauses or breaks in oral production were included in the production time of the following word unit. Major disruptions were any disruption or error that indicated that the reader was having a problem comprehending the text. The types of major disruptions were: mispronunciations, repetitions, substitutions, omissions, reversals, stutterings, hesitations, and pauses. In general, the results of the major disruption analyses confirmed the results obtained from the production times, so only the latter are presented except where additional information was obtained from major disruptions. All results are presented relative to the control condition. The absolute production times and major disruptions were not of interest. All reported effects were significant at the .05 level of significance.

RESULTS AND DISCUSSION

The mean production times from the Polish and English language experiments are shown in Figure 1 as a function of word unit position. The difference times (after subtracting the control condition times) for the syntactic, semantic, and factual manipulations are shown in panels A, B, and C, respectively.

Syntactic manipulation

In both the Polish and English language experiments, there was a significant disruption due to the syntactic manipulation beginning at word unit −1 (see Figure 1, panel A). The magnitude of the disruption was larger in Polish than in English. At the point of initial disruption (word unit −1) the disruption was much greater in the Polish experiment than in the English one, and the peak was higher in the Polish experiment. So although the timing of the disruption was the same in both experiments, the overall magnitude of the disruption due to the syntactic manipulation was greater in Polish than in English.

The disruptions due to the syntactic manipulation were qualitatively different in Polish and English as well. In the English experiment, 50% of the major disruptions at the critical word were restorations of the original form of the critical word, e.g., *injured* was restored from the printed *injury*. Many of these restorations were without pause until after they had been uttered, then the reader realized too late that what he had said was not what was printed. If the English readers were not paying much attention to the ends of words, they would perceive *injur-*, which was syntactically and semantically consistent with a possible word, and would say the syntactically correct response, *injured*. Thus, top-down information was used for lexical access of an appropriate word.

In the Polish experiment, only 9% of the major disruptions in the syntactic condition were substitution errors (which would include fluent restorations). If Polish readers paid more attention to the ends of words, they would be more likely to notice that the ending in the syntactic conditions was inconsistent with the sentence structure. Sometimes this recognition occurred before they uttered the word. In which case, they would hesitate or pause (46% of the major disruptions in Polish *vs.* 33% in English). Sometimes this recognition did not occur until they had already started to pronounce the word, which in turn would cause them to stutter or stammer (9% of the major disruptions in Polish *vs.* 7% in English). Much of the time Polish readers were hesitating, pausing, or stuttering before uttering the syntactically incorrect ending. Polish and English readers pay attention to different parts

of the word because the two languages distribute the linguistic information differently.

The quantitative and qualitative differences between the disruptions in Polish and English are consistent with the differences between Polish and English in how information in word endings contributes to the specification of syntactic structure. Word endings are not as important in English as they are in Polish for determining the part of speech or for indicating how a word functions syntactically in a sentence. Readers of English focus more attention on the beginnings of words because the beginnings carry the most relevant information. Syntactic information in English is carried primarily by word order, the extraction of which is a natural consequence of left-to-right reading strategies. In Polish, in contrast, the reader must pay close attention to the word endings because they provide the syntactic information necessary to determine the syntactic structure of the sentence. Skilled readers of Polish evidently have adopted the attentional strategy of paying close attention to the ends of words as well as to the beginnings. Since the syntactic manipulation distorted the part of speech information in the word endings, this violation had a larger disruptive effect earlier. Word order, which was not disturbed, still provided the required information to English readers even when the part of speech was changed by the word ending. The word order information was still available to the Polish readers as well, but it was not nearly as useful to them.

Semantic manipulation

The introduction of a semantic anomaly produced a disruptive effect beginning at word unit -1 and peaking at word unit $+1$ in Polish and at critical word in English (see Figure 1, panel B). The shape of the disruptive effect curve was similar in Polish and English. In the later word units, the disruptive effect fell off faster in Polish than in English. The fact that the two curves were nearly the same shape indicates that semantic information was functioning in a similar way for both Polish and English readers. There is no apparent difference between Polish and English in this aspect of semantics, nor in how the semantic manipulation was instituted. So the similarity of semantic effects was in line with our expectations.

Factual manipulation

The factual violation produced a significant disruption in the Polish experiment at word units 0 and $+1$ (see Figure 1, panel C). In the English experiment, the significant disruption began only at word unit $+1$, but continued through word unit $+3$. In addition, the magnitude of the disruption was larger in Polish, but this difference was probably not meaningful for comparing the use of factual information by Polish and English readers. The factual manipulation is difficult to compare closely because subtle differences in phrasing can enhance or attenuate the effect quite easily. That small differences in phrasing can make large differences in interpretation was observed while constructing the materials for the experiments. The important point is that the factual inconsistency did produce a significant disruption of about the same magnitude at about the same point in the sentence in both language experiments. There is nothing about the nature of Polish and English languages that would lead us to expect any large differences in how factual information functions in reading comprehension.

CONCLUSIONS

In spite of the salient differences in magnitude of the disruptions between the Polish and English experiments,

the pattern of results was very similar for Polish and English. The similarities in pattern indicate that the underlying reading comprehension processes are likewise very similar. A top-down influence of higher-order information on lexical access was quite evident in the disruptive effect of the semantic and syntactic manipulations before the critical word was uttered. The factual inconsistency produced a significant disruption after the other two, indicating that factual information was more critical for sentence and story integration than for lexical access. The overall pattern of results in both languages supported an interactive model of reading comprehension (Danks, 1978; Danks & Hill, in press).

This similarity in the pattern of results should not obsure the differences that were obtained in the syntactic condition. Certainly such differences indicate that Polish readers [illegible] the text somewhat different[illegible] the English — speaking rea[illegible] ever, the inferred difference [illegible]s-sing can be attributed to differ[illegible]s in Polish and English languages. The nature of Polish requires that the reader pay close attention to word endings. When the syntactic manipulation distorted the information in those endings, it produced a pronounced disruption immediately. Since in English more syntactic information is conveyed by word order than by word endings, English readers do not have to pay such close attention to the ends of words and so were not quite so disrupted by the syntactic manipulation. The basic comprehension process remained the same, however. In both languages, the reader used syntactic information for lexical access. The difference lies in where that information was located: in the print — at the ends of words, or in word order. Thus, the primary conclusion stands: reading comprehension processes are functionally the same in Polish and English and are best represented by an interactive model.

REFERENCES

Danks J. H., Models of language comprehension, *Polish psychol. Bull.*, 1978, 9, 183–192.

Danks J. H. & Glucksberg S., Experimental psycholinguistics, *Ann. Rev. Psychol.*, 1980, 31, 391–417.

Danks J. H. & Hill G. O., An interactive analysis of oral reading, in: C. A. Perfetti & A. M. Lesgold (eds.), *Interactive processes in reading*, Hillsdale, N. J., Lawrence Erlbaum, in press.

Fodor J. A., Bever T. G. & Garrett M. F., *The psychology of language: An introduction to psycholinguistics and generative grammar*, New York, McGraw-Hill, 1974.

Forster K. I., Levels of processing and the structure of the language processor, in: W. E. Cooper & E. Walker (eds.), *Sentence processing*, Hillsdale, N. J., Lawrence Erlbaum, 1979.

Marslen-Wilson W., Tyler L. K. & Seidenberg M., Sentence processing and the clause boundary, in: W. J. M. Levelt & G. B. Flores d'Arcais (eds.), *Studies in the perception of language*, Chichester, Wiley, 1978.

Rumelhart D. E., Toward an interactive model of reading, in: S. Dornic (ed.), *Attention and performance VI*, Hillsdale, N. J., Lawrence Erlbaum, 1977.

(Received June, 19

A Comparison of Reading Comprehension Processes in Polish and English

Joseph H. Danks and Ida Kurcz

Kent State University, USA — University of Warsaw, Poland

Running Head: Reading Comprehension in Polish and English

Mailing Address for Proofs: Joseph H. Danks
Department of Psychology
Kent State University
Kent, Ohio 44242
USA

Experiments 1 and 2 were parts of master's theses submitted to the University of Warsaw by Jerzy Tokaj and Danuta Kurkiewicz, respectively. Preliminary results from Experiment 1 have been reported previously by Kurkiewicz, Kurcz and Danks (1981). The collaboration on this research and the preparation of this paper was supported by exchange agreements between the University of Warsaw and Kent State University and between the Polish Academcy of Sciences and the National Academy of Sciences (U.S.A.). Support also was received from Grant W.2.1. (W-266) from the Polish Academy of Sciences and Grant No. NIE-G-78-0223 from the National Institute of Education (U.S.A.). Requests for reprints should be sent to JHD, Department of Psychology, Kent State University, Kent, Ohio 44242, U.S.A.

Abstract

Two experiments in Polish replicated several conditions in English experiments on reading comprehension processes. In Polish syntactic information is represented primarily by morphology and in English primarily by word order. Lexical, syntactic + semantic, within-syntactic class, between-syntactic class, semantic, and factual information were violated in stories. Native speakers read the stories aloud and the protocols were scored for increases in production time around violations. The disruptions in oral reading caused by the syntactic violations were relatively larger and occurred earlier in Polish than in English. The semantic and factual violations produced broader disruptions in English than in Polish. Reading comprehension processes varied to meet the cognitive demands imposed by how the available information was represented. Polish readers adopted a focused strategy and English readers a more diffused strategy.

Resume

En deux expériences conduites en langue polonaise, on a reproduit plusieurs conditions d'expériences conduites en langue anglaise sur les processus de la compréhension de la lecture. En polonais, l'information syntactique est donnée principalement par la morphologie, alors qu'en anglais, c'est surtout par l'ordre des mots. Dans des récits, des informations lexiques, syntactiques + semantiques, dans-la-même-classe-syntactique, entre-deux-classes-syntactiques,

sémantiques, et positives ont été déformées. Des sujets de langue maternelle polonaise ont lu les récits à haute voix et les protocoles ont été évalués en fonction de l'augmentation du temps de production aux endroits affectés par les déformations. Les perturbations de la lecture à haute voix provoquées par des déformations syntactiques ont été relativement plus importantes et ont eu lieu plus tôt en polonais qu'en anglais. Le déformations sémantiques et positives ont provoqué des perturbations plus importantes en anglais qu'en polonais. Les processus de la compréhension de la lecture ont varié en fonction des exigences cognitives imposées par la façon dont les informations disponibles ont été représentées. Les lecteurs polonais ont adopté une stratégie concentrée alors que celle des lecteurs anglais était plutôt diffusée.

Many recent investigations have viewed language comprehension as an interactive process in which the comprehender utilizes many sources of information to arrive at an appropriate interpretation of the speech or print (Danks and Glucksberg 1980; Just and Carpenter 1980). One result of comprehension being interactive is that it is also quite flexible (Danks 1978), that is, there is no single integrated comprehension process that is applied uniformally each time linguistic input is encountered. Rather comprehenders (listeners and readers) adapt their comprehension strategies to the situation. We read novels for pleasure in a different way than we read technical articles. We do not process cocktail party small talk in the same way as dramatic dialogue. The comprehension process changes to meet the cognitive demands of the situation. Comprehenders also differ in the cognitive skills that they have available to meet these demands. Children who are just learning how to read and listeners who are learning a second language have fewer cognitive skills so that they must allocate their cognitive energies differently than do skilled adult readers and native speakers. The flexibility of language processing then is represented in the interaction between the cognitive demands of the situation and the cognitive skills of the comprehender.

What sorts of variations in the comprehension situation lead to differences in the cognitive demands placed on the comprehender? If comprehenders expect a verbatim recall

test they process the text differently than if they are to solve a problem (Frederiksen 1975). Comprehension is different when comprehenders are required to produce a continuation of the read text than when required to recall the information (Mistler-Lachman 1972). Emphasizing oral performance in reading results in a different comprehension process than when comprehension is emphasized (Danks 1982a, 1982b; Pehrsson 1974). Many of the apparent differences that have been found between listening and reading modalities are the result of different demands that the two modalities of presentation place on listeners and readers (Danks and End in press; Rubin 1980).

The linguistic structure of the text itself can place different sorts of demands on the comprehender. Stories that follow the story grammar for a particular culture are more easily and completely understood than those that do not (Kintsch and Green 1978; Mandler 1978; Stein and Glenn 1979). Although there may be universal properties of folk tales such as the scientific values of the story characters that are understood easily across cultures (Bobryk and Dobrowolski 1981; Propp 1958), text difficulty also can affect processing. If the input material is quite difficult, either in vocabulary or syntax, the comprehender may have to resort to a lower level of processing in order to salvage some information from the text (Danks 1982a).

Finally, differences in language structure can affect the structure of the comprehension process. Some languages, like English, encode syntactic information primarily in terms of word order. Some languages, like Polish, Russian, and Finnish, encode most of the syntactic information in suffixes. One would expect then that comprehenders of these languages would be more attentive to word endings than would English language comprehenders. The primary purpose of the experiments reported here was to explore what effect such differences in language structure would have on comprehension processes.

Many studies of language comprehension, especially reading comprehension, have measured subjects' memory for texts (cf. Bobryk and Dobrowolski 1981; Danks and Glucksberg 1980). However, such data provide only indirect information about the process by which the representation was constructed. In order to gain a fuller understanding of comprehension processes, many investigators have turned to on-line measures of comprehension, such as phoneme and category monitoring (Foss and Lynch [illegible]; Marslen-Wilson and Tyler 1980), eye movement monitoring (Just and Carpenter 1980), reading time (Kintsch and Keenan 1973; Cirilo and Foss 1980; Cirilo 1981), and speech shadowing (Marslen-Wilson 1975). We have used oral reading performance as an on-line indicator of the comprehension process.

Two tasks are implicit in an oral reading task (Danks, Bohn and Fears in press; Danks and Fears 1979; Danks and Hill 1981). One continuing on-line demand is to produce an oral rendition of the print about the same time as it is fixated. Variation in the speed of oral production and the kinds of oral reading errors provide quantitative and qualitative indications of the difficulty and nature of the comprehension process at that point in time. In this sense, the oral reading task is very similar to a speech shadowing task except that the input is print rather than speech; the output is a relatively contemporaneous oral rendition of the input. Satisfying the oral production demand requires lexical access of the printed word. By locating the word in the mental dictionary, readers have access to the articulatory information needed to say the word.

The second demand implicit in oral reading is to understand what is being read. As we have used the task, readers are asked to give a summary of the story every couple of pages. Consequently, they must attempt to construct a representation of the story as they are reading it. The comprehension demand is not on-line, so readers can delay understanding if they encounter some difficulty. But completing a gist representation from verbatim code cannot be delayed for long because of short-term memory constraints. The comprehension demand induced by a request for a summary requires that sentence and discourse processing be completed through the construction of a macrostructure.

The contribution of different types of linguistic information to lexical access and sentence comprehension components was investigated in several experiments (Danks et al. in press; Danks and Hill 1981). The rationale was to introduce violations of different types of linguistic information at specific points in a story. The effect of these violations was measured by delays in oral production times and disruptive errors. The relative positioning and shape of the disruption curves resulting from each type of violation indicated how the different types of information were processed. Violating lexical (inserting a pronounceable nonword), syntactic (changing the part of speech), and semantic (inserting an anomaly) information disrupted oral production just before the critical word was uttered. So top-down syntactic and semantic information were used in lexical access just as was lexical information itself. Violating factual information (making two successive sentences factually inconsistent) did not yield a disruption of oral production until after the critical word had been uttered. So factual information was not used in lexical access, but was involved in sentence and discourse processing.

The purpose of the experiments reported here was to compare whether the comprehension processed revealed by the pattern of oral reading disruptions would be altered by differences in the cognitive demands imposed by differences in language structure. Several of the conditions and experiments previously reported for English-language readers were replicated with Polish readers.

Polish is a Slavic language that differs in many ways from English, which is Germanic (cf. Fisiak, Lipińska-Grzegorek and Zabrocki 1978; Schenker 1973). In addition to obvious differences in pronunciation, spelling, and vocabulary, Polish and English differ in their syntactic structure. In English syntactic information is indicated primarily by how words are ordered in sentences. The position of a given word in relation to other words in a phrase or clause indicates how the word functions syntactically. In Polish the syntactic function is marked by the morphological structure of the word itself, not by its sequential relation to other words. Most words have specific suffixes that indicate the part of speech. Additionally, within each part of speech, inflections organized in declensions and conjugations further specify syntactic function by differentiating among genders, numbers, tenses, and cases. For example, in Alicja dała książkę przyjacielowi, "Alice gave a book to a friend," the -a suffix on Alicja indicates feminine, singular, nominative case, the -ła on dała indicates past, third-person singular, feminine subject, the -ę on ksiazkę indicates feminine, singular, accusative case, and -owi on przyjacielowi indicates masculine, singular, dative case. If the friend had been a girl, then the ending would have been -ce, przyjaciółce (note stem variation as well).

Both Polish and English are subject-verb-object (SVO) languages, but in Polish the basic SVO word order can be altered quite readily for stylistic and pragmatic purposes,

such as to fulfill the given--new contract (Haviland and Clark 1974). If a speaker wished to emphasize the friend in the above example, przyjacielowi could be moved to the first position in the sentence with no other change. Likewise, any other word in the sentence could be placed in first position, and almost any other ordering of the four words is also possible. The literal meaning of the sentence would remain the same; only the pragmatic emphasis would change. In English, of course, very few changes in word order are possible that do not also change the meaning of the sentence. But English has very few inflectional endings that indicate syntactic function.

Intrasentence concordance rules, such as subject-verb agreement, are more numerous and more constraining in Polish than in English. In Polish, the subject noun must agree with the verb in number, person, and gender, and a noun and adjective must agree in number, gender, and case. The crucial role of suffixes in marking functions is observed in all language use situations: speaking, listening, writing, and reading. Distinctive pronunciations of the suffixes constitute a major source of dialectal differences between social classes and geographical regions.

These differences lead to an asymmetry in listening/reading and speaking/writing between Polish and English. Since the word itself contains more explicitly marked syntactic information in Polish than in English, a reader/listener finds more information in isolated words in Polish than in

English. For example, a Polish reader knows that Alicja is a subject of a sentence, whereas an English reader does not know whether Alice is a subject, direct or indirect object, or object of a preposition without seeing where it is located in a sentence. So the English reader is dependent on the linguistic context for word-order information, whereas the Polish reader can concentrate more directly on each individual word. In contrast, a Polish speaker must be more aware of the other words in the sentence than is an English speaker in order to insure the concordance constraints. For example, the adjective and noun, mała książkę, "small book" (feminine, singular, accusative case), must agree in number, gender, and case, so the speaker must select the noun before the correct ending can be determined for the adjective. So if the speaker had selected zeszyt, "notebook" (masculine), instead, then the adjective would have to be changed to mały. At the level of morphology, the Polish speaker is more dependent on linguistic context than is the English speaker.

Thus, we expect that Polish readers have developed reading strategies of attending not only to the semantic content of each word, as an English reader would, but also to the word ending (cf. Przetacznikowa and Kielar 1981). English readers would pay relatively less attention to the endings of words and relatively more to their position in the sentence. Introducing a syntactic violation by changing the ending of a word so that it is a different part of speech,

we expected a relatively larger disruption of oral reading performance in Polish than in English. We also expected that the spread of disruption in English readers would be relatively wider than in Polish since English readers would attempt to use contextual information to resolve the violations. Polish readers would have relatively narrower patterns of disruptions because they focus more on isolated words than do English readers. So Polish readers would have a more focused strategy, while English readers would have a more diffused strategy.

We report here two experiments in Polish that were close replications of previously reported experiments in English (Danks et al. in press; Danks and Hill 1981). The manipulations, procedures, and analyses were kept as similar as possible although there were enough unavoidable differences to prevent direct statistical comparison. Hence, only the patterns of results are compared.

Method

Materials. A story was selected from an American magazine about a high school girl who was injured severely when a train hit her school bus. The story was just over 2000 words long and had a readability of 7.5 (Fry 1968). Sixteen critical words spaced throughout the story were selected for manipulation. The story was translated into Polish with some adjustments to make it appropriate for Polish readers, e.g., names of characters, geographical names, and a few culture-dependent details were changed. A portion of the

story surrounding one of the critical words is shown in both Polish and English in Figure 1. The critical word is potrącona in Polish, injured in English. The girl's mother has just heard about the accident and is worried about her daughter. Lexical, between-synactic, within-syntactic, semantic, syntactic + semantic, and factual information were violated in the critical words.

--- Insert Figure 1 about here ---

(a) Lexical. The lexical manipulation replaced the critical word with a pronounceable nonword, that is, one that followed the orthographic rules of Polish or English. In the example in Figure 1, injured was replaced with brugen, and potrącona was replaced with pomerana. In Polish, the syntactic information in the suffix -a was appropriate to the sentence context, in this case, the suffix indicated verb participle, or adjective, nominative case, feminine, singular. Because of the importance of morphological cues associated with Polish words, we attempted to manipulate only the lexical information so that the violation would be equivalent for Polish and English readers. In Polish, as in English, the syntactic relationships in the sentence were not distorted, although the syntactic structure in English was not reinforced since syntactic inflections were not added to the nonsense words. The "word" stem had no meaning and could not be located in the mental dictionary. Thus, in both languages, if readers were relying solely on phoneme-grapheme correspondences to render

the print as speech, there would be no disruption of their oral productions with the lexical violations. However, if the readers were accessing aritculatory information in their mental dictionaries in order to pronounce the word, there would be a disruption before they uttered the word. Because the cognitive demands of the manipulations were similar in Polish and English, we expected no differences between Polish and English readers in this condition.

(b) Between-Syntactic. In the English language experiments, the syntactic information was violated by changing the ending such that the part of speech of the critical word was changed to one clearly inappropriate for the sentence. For example, the verb past participle injured was changed to the nominal form injury. Syntactic information might have been distorted in English by altering the word order, but the morphological manipulation was used to avoid involving more than the critical word.

Given how syntactic information is represented morphologically in Polish, two variants of the English syntactic manipulation were possible. In both variants the ending was changed, but in one case the part of speech was altered and in the other there was a syntactic error within the same part of speech. In the former, which we called "between-syntactic," the word ending was changed such that the part of speech changed, for example, the past participle potrącona was changed to a noun potrącenie, "injury."

The aim of the between-syntactic violation was to determine to what extent syntactic information is involved in lexical access and sentence comprehension. If syntactic information is involved in lexical access, its violation should produce a disruption before the critical word was uttered just as with the lexical violation. If it is involved in sentence comprehension, then the disruption would continue for some distance after the critical word was uttered. Although some semantic information is carried in syntactic categories, most of it comes from the root morpheme. So the reader could determine the meaning intended for the sentence even when the syntactic information had been altered. Because of the heavy use of morphological information in Polish, we expected that the between-syntactic violation (and to a lesser extent the within-syntactic; cf. below) would have a greater effect somewhat earlier and to a greater extent on Polish readers than on English readers.

(c) Within-Syntactic. The second type of syntactic violation that is possible in Polish, we called "within-syntactic." In this case, the word endings were changed within the same part of speech, but gender, number, case, and/or tense relations were violated. For example, the verb participle potrącona, which marks feminine gender, was altered to potrącony, still a verb participle, but one that marks masculine gender. In general, syntactic manipulations were easier to construct in Polish than in English, but the

within-syntactic is especially difficult in English, being limited to number and tense changes.

We compared the within-syntactic change in Polish to the between-syntactic change in English, but one must be aware of the different types of syntactic information being manipulated in the two languages. We expected that the within-syntactic manipulation would yield disruptions similar to the between-syntactic ones, except that they might be somewhat smaller and less extensive. Parts of speech are broader syntactic categories than those formed by within-syntactic information. The broader the class of information violated, the greater would be the size and extent of the disruption.

(*d*) *Semantic*. Semantic information was violated by inserting into the sentence a semantically anomalous word that was the appropriate part of speech. For example, *injured* was replaced with *planted* and *potrącona* was replaced with *posadzona*, "planted." Since syntactic information was not distorted, the reader could determine the grammatical structure of the sentence, but a semantic interpretation was much more difficult if not impossible. At best only an implausible or metaphorical circumstance could be imagined. Since there were no apparent differences between Polish and English in implementing the semantic manipulation, we did not expect any major differences in the pattern of disruptions.

(e) Syntactic + Semantic. Both syntactic and semantic information were distorted by replacing the critical word with one that was the incorrect part of speech and which was semantically anomalous as well. For example, injured was changed to iceberg, and potrącona was replaced with posadzenie, "the act of planting." In constructing the Polish violations, the morphological changes of the between-syntactic violations were added to the roots used in the semantic violations. Hence, linguistically the Polish syntactic + semantic violations were more accurately a sum of the semantic and between-syntactic violations than were the English.

With this manipulation we investigated whether syntactic and semantic information were additive in their effects. Since we expected a difference between Polish and English with the between-syntactic violation, but not with the semantic, the relation between syntactic and semantic information could be assessed more accurately by comparing the effects across languages.

(f) Factual. As readers progress through a story they accumulate factual information about what is happening in the story and construct an event sequence using their knowledge about how things happen in the real world (Warren, Nicholas and Trabasso 1979). This knowledge, which we called "factual" information, was violated by changing the sentence preceding the one with the critical word such that a factual inconsistency was produced when the critical word

was encountered. The critical word itself was not altered and both sentences were syntactically and semantically normal. In the example, *słaba* was changed to *silna*, and *weak* was changed to *strong*. If the daughter were strong, the mother would not necessarily worry about her being injured by the other children.

We expected no difference between Polish and English readers with regard to the factual violation because processing this information is beyond any linguistic differences. Although there are discourse differences between Polish and English with respect to textual structure, the factual representation of events should be similar.

(g) *Control*. There also were control conditions in which there was no change in the critical word or in the preceding sentence. This condition provided a baseline for normal oral reading times.

These conditions were combined in two experiments in Polish. In Experiment 1, lexical, within-syntactic, syntactic + semantic, and control were manipulated. In Experiment 2, between-syntactic, semantic, factual, and control were manipulated. The English-language experiments have been reported in detail in Danks et al. (in press; cf. also Danks and Hill 1981). In addition to control conditions in each experiment, syntactic, semantic, and factual were included in one experiment; lexical was included with syntactic + semantic and semantic in another; and syntactic + semantic was included with semantic and factual in a third experiment.

Because of some differences in wordings around the critical words and even the story itself (the second English-language experiment used a different story), the English comparisons for the between-syntactic, semantic, and factual conditions were taken from the first English experiment, lexical from the second, and syntactic + semantic from the third. These experiments provided the closest available English-language comparisons for the two Polish experiments.

In all of the experiments, four versions of the story were constructed. Each of the four modifications occurred four times in each version. The manipulations were counter-balanced across the 16 critical words and across versions such that each manipulation was applied once to each critical word. The stories were divided into four approximately equal sections, each of which contained four critical words. In each sentence, each of the four manipulations occurred once. The stories were typed such that the critical words did not occur near the beginnings or ends of lines, nor near the tops or bottoms of pages.

Subjects. The subjects in both Polish experiments were students at the University of Warsaw. Their native language was Polish. The subjects in Experiment 1 were 35 women and 5 men, who were residents of University dormitories and none of whom were psychology majors. In Experiment 2 the subjects were 40 women, majoring in psychology, but who were unfamiliar with the specific research and its objectives. The subjects in each of the English experiments were 40

undergraduates enrolled in general psychology at Kent State University.

Procedure. The procedure was identical in all of the experiments. Subjects were tested individually. They were told the experiment was about reading comprehension and were asked to read the story aloud into a tape recorder. After each of the four sections, they were asked to recall as much as possible about the story. The purpose of asking for recalls was simply to insure that the readers tried to understand what they were reading. The summaries were not scored for accuracy. Nothing was mentioned about the manipulations of the text.

Scoring and analyses. In order to facilitate scoring around the critical words, the text was divided into five word-units before and five word-units after each critical word. (In the first English experiment only four word-units were scored before the critical word.) These are indicated in Figure 1. Since the pronunciations of some words are physically joined, it is difficult, if not impossible, to know precisely where one word ends and another begins. The specification of word-units attempted to reflect this coarticulation.

The primary dependent variable was the production time for each of the word-units before and after the critical word as well as the time for the critical word itself. These times were obtained by slowing the tape to half-speed and pressing a key at the end of each word-unit. The time

between key presses was measured on a digital clock and divided by two to obtain real time measures.

Reliabilities of the measurement of production times were obtained by having two experimenters measure eight subjects in each Polish experiment. The correlations between the experimenters were .99 for Experiment 1 and 0.92 for Experiment 2. In the second English experiment, similarly measured production times were compared with measurements taken from speech spectrographs for eight subjects. The correlation was .91. Thus, the procedure used to measure production times was reliable.

A second dependent variable was the kind of oral reading errors caused by the violations. These were such things as substitutions, omissions, additions, mispronunciations, regressions and hesitations. The probability of making an error confirmed the production time results as well as provided a qualitative indication of the readers' oral performances in the neighborhood of the critical words.

The production times in all experiments were analyzed with a four-way Latin-square analysis of variance with versions as a between-subject variable and violations, sections, and word-unit positions as within-subject variables. All effects were tested against subject and critical-word error variances using a quasi-$\underline{F}$ (Clark 1973). The effect of primary interest was the interaction between violations and word-unit positions. Planned comparisons were conducted at each word-unit position between each violation condition and

the control. All reported effects were significant with a probability of .05 or less.

Results

Overall results. The mean production times for each manipulation at each word-unit position are presented in Table 1. In both Polish experiments, the main effect of violation type and the interaction of violation type with word-unit position were significant: Experiment 1, violation type, F (3, 24) = 25.00, violation type by word-unit position, F (3, 240) = 10.15; Experiment 2, violation type, F (3, 24) = 38.53, violation type by word-unit position, F (3, 240) = 8.91. As shown in the last column of Table 1, all violation types were significantly different from their controls.

--- Insert Table 1 about here ---

We now discuss each of the violation types by evaluating the points at which the violations produced significantly longer production times. These differences are presented in Figures 2, 3, 4, and 5. The English-language results are included in the figures for comparison.

Between-Syntactic. The production time disruption curves for the between-syntactic violations are presented in Figure 2. In Polish the between-syntactic violation produced a significant disruption beginning at word-unit -1 that continued through word-unit +2 with a peak during the critical word. In contrast, a between-syntactic violation in English did not have an effect until the critical word,

one word-unit after the Polish effect began. Polish readers were disrupted earlier than were English readers by the between-syntactic violation. This result was as expected based on the fact that Polish suffixes are more informative than are English. Polish readers apparently attended more to the word endings, so their reading comprehension processes were disrupted more by a violation present in those endings.

--- Insert Figure 2 about here ---

The kinds of oral reading errors committed at the critical word were qualitatively different. In English, 50% of the oral reading errors at the critical word were restorations of the original form, e.g., injured was uttered instead of injury. Of these restorations, 68% were fluent in that there was no hesitation or other disruption immediately before or during the critical word. (This result is consistent with the fluent restorations obtained in a speech-shadowing task by Marslen-Wilson 1975). English readers apparently used top-down contextual information to locate the critical word in their mental dictionary such that identification of the original critical word occurred before the printed form was completely processed. In Polish, only 9% of the oral reading errors were substitutions of any sort, including restorations. If Polish readers were attending closely to the ends of words, then they would be more likely to notice the syntactic violations before top-down contextual information could restore the original

critical word. Thus, there were more pauses immediately before the critical word in Polish than in English (46% vs. 33%) as Polish readers hesitated before uttering what was printed. The differences in the types of oral reading errors support the conclusion that Polish readers attacked the printed word differently than did English readers.

Within-Syntactic. The disruption curve for the Polish within-syntactic violation is shown also in Figure 2. The differences in production times resulting from the within-syntactic violation in Polish differed from the control only at the critical word and at word-unit +1, with a peak at the critical word. This pattern was clearly different from the between-syntactic curve in Polish in that it was much smaller in height and breadth. In comparison with the English between-syntactic disruption it was one word-unit shorter.

The within-syntactic violation violated fewer linguistic distinctions than did the between-syntactic. Since Polish readers needed to make fewer repairs to determine what was meant by the sentence, we expected that the within-syntactic violation would be less disruptive. A more recently completed English-language study (Danks 1982b) used a somewhat different design to compare between- and within-syntactic violations in English. In contrast to Polish, the between- and within-syntactic disruption patterns were very similar. The within-syntactic violation produced the same magnitude of disruption in Polish as both between- and

within-syntactic did in English. English readers apparently treated all syntactic violations more or less equivalently, whereas Polish readers were sensitive to the degree of linguistic violation. These later English results also confirm that it is reasonable to use the English between-syntactic disruption curve for comparison with the Polish within-syntactic.

The number and pattern of oral reading errors was very similar for Polish within-syntactic and English between-syntactic. However, they were distributed quite differently. In Polish, only 15% were substitutions including restorations of the critical word, in contrast to 50% in English. But there 49% hestitations at the critical word in Polish and only 33% in English. As with the between-syntactic violation, Polish readers were sensitive to the printed form of the critical word so that the within-syntactic violation was not restored.

The linguistic difference between the between-syntactic and within-syntactic violations in Polish also was reflected by the number of letters changed for each type of violation. An average of 3.8 letters were changed to produce the between-syntactic violation and only 1.8 letters were changed for the within-syntactic. In the recent English-language experiment, the average number of letters changed was 3.4 for the between-syntactic and 2.3 for the within-syntactic violations. However, a difference between the Polish and English violations of less than one letter would

not appear to be large enough to explain the result that there was no between-within-syntactic difference in English and a very large difference in Polish.

Semantic. The disruption curves for the semantic violation are shown in the upper panel of Figure 3. In Polish the semantic violation produced a difference in production times beginning at word-unit -1 that continued through word-unit +2. In English, the differences were significant beginning at the critical word and continuing through word-unit +3. The difference at word unit -1 missed being significant by 9 msec. Both Polish and English curves had peaks at the critical word and word-unit +1.

--- Insert Figure 3 about here ---

The pattern of the semantic disruptions were nearly the same in Polish and English (although the English was not significant at word-unit -1), but the English distribution continued for one word-unit longer. The Polish curve was higher than the English, but the shapes were quite similar. The slightly greater breadth of the English disruption supported the linguistic analysis that English readers were more dependent on linguistic context to determine the sense of a sentence, whereas Polish readers focused a bit more on individual words.

Syntactic + Semantic. The production time disruption curves for the syntactic + semantic violations are shown in Figure 4. Syntactic here means a between-syntactic violation for both languages. Both Polish and English syntactic

+ semantic violations produced significant differences from the control conditions beginning at word-unit -1 with peaks at the critical word. The English disruption was significant through word-unit +3, but the Polish was significant only through word-unit +2.

--- Insert Figure 4 about here ---

The peaks of both curves were considerably higher, especially for Polish, than either the between-syntactic or semantic peaks separately, but were not as high as their sum. Such a result would be expected when two related sources of information were violated because the reader would have more difficulty in making the necessary repairs. But the information sources were not completely independent because the disruptions were not additive, or possibly the processes using each information type operated in parallel. The Polish curve was not as broad as the English, similar to the syntactic curves, supporting the conclusion that Polish readers were more focused on individual words and sensitive to word endings.

Lexical. The production time disruption curves for the lexical violation are shown in Figure 5. Both Polish and English disruption curves from the lexical violation were significantly different from the control beginning at the critical word and continuing to word-unit +1 in Polish and to word-unit +2 in English. The curve was a bit sharper, or more peaked, in Polish than in English. As with the other cases of sharpening, we attribute this difference to the

fact that Polish words are more self-contained in marking syntactic information, so the Polish reader can focus on individual words more directly. Also in Polish the lexical violation explicitly preserved syntactic information through appropriate morphology. In English it was only implicit in the word order. This difference, one which results naturally from the different modes of linguistic representation, may have contributed to the sharpening in Polish.

---Insert Figure 5 about here ---

For both Polish and English, the lexical disruption was two word-units shorter than the corresponding semantic disruption. The semantic violation was an obvious anomaly that had to be resolved. The lexical violation was a nonsense word with no established conventional meaning. Thus, once readers realized that it was not a real word, or at least one that they did not know, they were free to infer an appropriate meaning from the sentential context. Since the semantic violation was a real word with definite meaning to be integrated with the other meanings in the sentence, what the sentence meant as a whole could not be determined easily. More of the sentence had to be processed in an attempt to resolve the semantic anomaly; whereas with the nonsense word, readers could devise a plausible meaning for the word and continue.

Factual. The production time disruption curves resulting from the factual violations are shown in the bottom panel of Figure 3. The factual disruption in Polish was

significantly different from the control when averaged across all word-unit positions (cf. Table 1). However, it was not significantly different at any single word-unit position although the increase at word-unit +1 missed being significant by only 15 msec. The English curve was significantly different at word-units +1 and +3 (but not at word-unit +2). The fact that the English curve continued longer than the Polish reflects a strategy of English readers in which they attempt to use information following the critical word to attempt to resolve the factual inconsistency more than do Polish readers.

No involvement of factual information in lexical access has been found for English readers (Danks et al. in press; Danks and Hill 1981). There evidently was none for Polish readers either because the factual disruption was delayed relative to the lexical violation and the uttering of the critical word itself.

Other experiments (Danks 1982b) have shown that factual violations are most disruptive at the clause boundary following the critical word. It is possible that during translation the clause boundaries were altered significantly, but such was not the case. The clause boundary was located an average of 1.81 word-units after the critical word in the English text and 1.69 word-units after in the Polish text. It is also possible that the distance between the inconsistent material in the preceding sentence and the critical word was altered during translation. This possibility also

did not occur: in English the distance averaged 7.75 words and in Polish 7.06 words. Based on our intuitions about the factual violations gained through constructing the stories, we think that factual violations would be the most susceptible to incomparabilities due to the subtlies of translation. However, we have been unable to document any basis for that intuition.

Discussion

Based on these results we can identify two complementary reading strategies -- a focused strategy and a diffused strategy. Although a focused strategy was used more by Polish readers and a diffused by English, the use of each strategy was controlled by more than just language differences. With a focused strategy, the reader attends to a relatively narrow portion of the text, perhaps only a few letters. The "perceptual window" itself probably does not change in size (McConkie and Zola 1981), but attention is more sharply focused. The reader would adopt such a strategy in situations where information is concentrated in a relatively narrow portion of the text. In Polish, syntactic information is concentrated in the letters at the end of each word. Polish readers need information about the case relations of nouns and adjectives, properties of verbs, and government relations between adjectives, nouns, verbs, adverbs, and prepositions. The syntactic information can be determined solely from the word ending without even accessing the word in the mental dictionary. While learning to

read, Polish readers would develop a focused strategy because most syntactic information is concentrated. This strategy is not absolute nor rigidly used all of the time, but rather is a habit adopted for its heuristic value. English readers would tend not to use such a strategy because linguistic information is distributed more broadly in English text, but in appropriate circumstances, English readers might find a focused strategy valuable.

With a diffused strategy, in contrast, the reader assimilates information from a much broader portion of the text, perhaps spanning several eye fixations. Such a strategy would be functional when the information in the text is distributed over several words. Such is the case with English syntax. It is based primarily on the ordering of word classes. The syntactic information associated with each word in the mental dictionary must be related to that of the surrounding words. So in English, in contrast to Polish, processing syntactic information would be facilitated by a diffused strategy.

The largest differences between the Polish and English results were from the syntactic violations. Polish readers were very sensitive to the violation of syntactic information. There was a higher between-syntactic peak in Polish, fewer restorations of the critical word, and a sensitivity to whether between- or within-syntactic information was violated. Polish readers attended more closely to the word endings and the information encoded there, reflecting their

general adoption of a focused strategy. English readers in contrast showed a more diffused strategy.

Determining the meaning of a sentence requires the integration of the word meanings. Since several words have to be accessed before such integration can begin, a diffused strategy is functional for processing semantic information. This is the case for both Polish and English readers because the distribution of semantic information is quite similar, especially with the close translation of the story used in these experiments. A diffused strategy does not require that semantic processing be delayed until the clause boundary, however, thereby precluding word-by-word processing (Marslen-Wilson, Tyler and Seidenberg 1978). It is only that several words are needed before a meaningful representation can be constructed.

The semantic violation should reveal the diffused strategy in both Polish and English readers, and it did because the shapes of the curves were similar. Although the peak of the Polish disruption was higher than in English, more striking was the similarity of the flattened peaks on both curves; plateaus were formed between the critical word and word-unit +1. This flattening was not produced by any other violation and reflected a diffused strategy in both Polish and English readers in specific response to the violation of semantic information. Polish readers did not show a focused strategy here because appropriate syntactic information was present. Rather they were faced with the same problem as

were the English readers, namely, how to integrate the anomalous semantic information with the remainder of the sentence. Hence, a diffused strategy was evident in both language groups.

The results of the syntactic + semantic violation also were consistent with this view. Both strategies were evident because both between-syntactic and semantic information were violated. The between-syntactic and semantic violation produced a larger, more focused disruption in Polish readers and the semantic violation produced a longer, more diffused disruption in English readers.

Comparison of the lexical disruptions is more problematic than the other violations because the violations were implemented in somewhat different ways in Polish and English. In Polish completely appropriate syntactic endings were affixed to the nonsense words. In English, no syntactic endings were added, but word order still provided some syntactic information. However, providing the syntactically appropriate endings in Polish may have caused Polish readers to believe that the nonsense words were real, albeit infrequent, words. The syntactic endings made them look more like real words than did the English ones. If so, then Polish readers may have persisted longer searching for the word in their mental dictionaries. Even in English, having a string of nonsense words marked with appropriate syntactic endings makes them seem more like a sentence and easier to learn (Epstein 1961, 1962). Such an effect probably would

be even stronger in Polish. A second difference between the lexical manipulations was that the English results came from an experiment that used a very different story than the one translated into Polish. This other story was abstract, metaphorical, and quite difficult to follow, so that the disruptions related to meaning violations (including lexical) were attenuated (Danks et al. in press).

Even with these qualifications, the disruption curves from the lexical violations showed some aspects of the focused and diffused strategies. The higher peak indicated that Polish readers were relatively more focused in their attempts to locate a meaning for the nonsense word in their mental dictionaries. English readers appeared to be more diffused in searching for information about the nonsense word's meaning in the rest of the sentence since their disruption was larger.

The disruptions from the factual violations also illustrated a focused strategy for Polish readers and a diffused strategy for English readers. There was an increase in production times in Polish near the critical word indicating a more focused strategy. The English disruption was later reflecting the gathering of information from a broader portion of the sentence. Thus, even with the factual violation, which involved integration of information between two sentences, focused and diffused strategies were evident.

We have emphasized the differences between Polish and English in our discussion, but one should not overlook the similarities in the two sets of results. All violations produced disruptions in both languages. Furthermore, the general shapes of the disruption curves were quite similar, e.g., the flattening of the semantic disruptions. Within each language, the ordering of the magnitudes of the disruptions, from syntactic + semantic being the largest to factual being the smallest, was virtually identical. The Polish disruptions were generally higher, but that was not true in every case, e.g., the within-syntactic and factual disruptions. These general similarities then provide support for the general interactive model of reading comprehension (Danks et al. in press; Danks and Hill 1981) using Polish, a language that represents syntactic information quite differently than English. The linguistic differences led to predictable differences in results based on the model.

A central property of this model is that readers use information -- of whatever type -- as soon as it becomes available to construct a representation of the text. In Polish and English, syntactic information becomes available at different times because of how it is encoded. So Polish and English languages present different sorts of problems for readers. They make different demands on the cognitive processing system. Differences in processing are explicable in terms of the differences in cognitive demands. Adams

(1980) and Stanovich (1980) have argued that readers attempt to compensate for inadequate processing skills by emphasizing those skills that they have. In our case, it was not differences in processing skills that resulted in the different strategies between Polish and English readers, but the differential availability of linguistic information.

Polish readers did not employ an exclusively focused strategy nor did English readers employ an exclusively diffused one, but all readers adapted their reading strategies to the information available. Thus, skilled readers of Polish and English have developed somewhat different reading strategies in the course of learning to read. These strategies represent a general orientation toward reading comprehension and are not completely fixed. The strategies can be modified to match the needs of specific situations, e.g., the tendency toward a diffused strategy in Polish readers when they encountered a semantic violation. These differences represent an example of the flexibility of processing capabilities. The basic structure of the reading process appeared to be quite similar for Polish and English readers, involving lexical access, sentence comprehension, and discourse understanding components. However, how these processes operated depended on the information available to the components and the form that the information took in print.

The linguistic differences between Polish and English lead to predictable differences in two other areas. The tendency to be sensitive to syntactic morphemes also has been observed in Polish studies on word association (Kurcz 1966, 1967). Since all Polish nouns and adjectives must be marked for gender, number, and case, even isolated stimulus words, especially adjectives, must be so marked. Typical word association studies use nominative case and singular number for stimuli, and usually masculine gender. However, in one study (Kurcz unpublished), different groups of subjects received adjectives marked for different genders, e.g., <u>biał-</u>/-<u>y</u>/-<u>a</u>/-<u>e</u>/m/f/n/, "white." Of the word association responses, 95% to 99% agreed with the stimulus in gender. These percentages combined paradigmatic (another adjective) and syntagmatic (a noun) responses. Isolated Polish words are more specific than English. For example, <u>białe</u> does not refer to just anything white, but something white of neuter gender. So <u>białe</u> could modify <u>wino</u> ("wine," neuter), but not <u>śnieg</u> ("snow," masculine). In Polish word association studies (Kurcz 1966, 1967), there was a stronger tendency toward syntagmatic responses than in American norms (Russell and Jenkins 1954): 41% in Polish versus only 18% in American norms, in which paradigmatic responses predominate (82%, cf. Rosenweig 1961). The syntactic markers required of Polish words even in isolation provoke more syntagmatic responses, concordant in gender, number and case.

Another area in which linguistic differences between Polish and English have proven illuminating is early language acquisition. Studies of children learning English have found that young children are especially sensitive to word order; inflections are learned much later (Brown 1973). However, studies of children learning Polish (Shugar 1977; Smoczyńska in press) have found that young children in the two-word stage use many syntactic inflections appropriately. For example, they acquire the dative relatively early because it is used in giving and receiving: Da dzidzi, "give baby," (-i is a correct dative ending). In contrast to English children, Polish children pay relatively less attention to word order. Polish children adopt language acquisition strategies that are appropriate for the way syntactic information is encoded in Polish. Thus, Smoczyńska (in press) has argued that there are very few universal operating principles (Slobin 1973), but many language specific strategies for language acquisition.

These studies as well as the ones reported here illustrate the necessity for testing psycholinguistic models in several languages. Most psycholinguistic research in the past couple of decades has been on English. However, many of the conclusions derived from that work may be language specific. With additional comparative studies, we can construct more general models of language processing.

References

Adams, M. J., 1980. Failures to comprehend and levels of processing in reading. In: R. J. Spiro, B. C. Bruce and W. F. Brewer (eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bobryk, J., and R. Dobrowolski, 1981. The psychosemantic representation of texts constructed according to Greimas's model of narrative structure. Polish Psychological Bulletin 12, 33-40.

Brown, R., 1973. A first language: The early stages. Cambridge, Mass.: Harvard University Press.

Cirilo, R. K., 1981. Referential coherence and text structure in story comprehension. Journal of Verbal Learning and Verbal Behavior 20, 358-367.

Cirilo, R. K., and D. J. Foss, 1980. Text structure and reading times for sentences. Journal of Verbal Learning and Verbal Behavior 19, 96-109.

Clark, H. H., 1973. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior 12, 335-359.

Danks, J. H., 1978. Models of language comprehension. Polish Psychological Bulletin 9, 183-192.

Danks, J. H., 1982a. An information-processing analysis of the cognitive processes involved in oral reading. American Educational Research Association, New York.

Danks, J. H., 1982b. Letters, words, sentences, and text: Some studies of reading comprehension. Midwestern Psychological Association, Minneapolis.

Danks, J. H., L. Bohn, and R. Fears, in press. Comprehension processes in oral reading. In: G. B. Flores d'Arcais and R. J. Jarvella (eds.), The process of language understanding. Chichester: Wiley.

Danks, J. H., and L. End, in press. Processing strategies for listening and reading. In: R. Horowitz and S. J. Samuels (eds.), Comprehending oral and written language. New York: Academic Press.

Danks, J. H., and R. Fears, 1979. Oral reading: Does it reflect decoding or comprehension? In: L. B. Resnick and P. A. Weaver (eds.), Theory and practice of early reading, Vol. 3. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Danks, J. H., and S. Glucksberg, 1980. Experimental psycholinguistics. Annual Review of Psychology 31, 391-417.

Danks, J. H., and G. O. Hill, 1981. An interactive analysis of oral reading. In: A. M. Lesgold and C. A. Perfetti (eds.), Interactive processes in reading. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Epstein, W., 1961. The influence of syntactic structure on learning. American Journal of Psychology 74, 80-85.

Epstein, W., 1962. A further study of the influence of syntactic structure on learning. American Journal of Psychology 75, 121-126.

Fisiak, J., M. Lipińska-Grzegorek, and T. Zabrocki, 1978. An introductory English-Polish contrastive grammar. Warszawa: Państwowe Wydawnictwo Naukowe.

Foss, D. J., and R. H. Lynch, 1969. Decision processes during sentence comprehension: Effects of surface structure on decision times. Perception and Psychophysics 5, 145-148.

Frederiksen, C. H., 1975. Effects of context-induced processing operations on semantic information acquired from discourse. Cognitive Psychology 7, 139-166.

Fry, E. B., 1968. A readability formula that saves time. Journal of Reading 11, 513-516, 575-578.

Haviland, S. E., and H. H. Clark, 1974. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior 13, 512-521.

Just, M. A., and P. A. Carpenter, 1980. A theory of reading: From eye fixations to comprehension. Psychological Review 87, 329-354.

Kintsch, W., and E. Green, 1978. The role of culture-specific schemata in the comprehension and recall of stories. Discourse Processes 1, 1-13.

Kintsch, W., and J. M. Keenan, 1973. Reading rate as a function of the number of propositions in the base structure of sentences. Cognitive Psychology 5, 257-274.

Kurcz, I., 1966. Inter-language comparison of word association responses. International Journal of Psychology 1, 151-161.

Kurcz, I., 1967. Polskie normy skojarzeń swobodnych na 100 słów z listy Kent-Rosanoffa. Studia Psychologiczne 8.

Kurkiewicz, D., I. Kurcz, and J. H. Danks, 1981. Reading comprehension processes in Polish and English. Polish Psychological Bulletin 12, 25-31.

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

Marslen-Wilson, W. D., 1975. Sentence perception as an interactive parallel process. Science 189, 226-228.

Marslen-Wilson, W. D., and L. K. Tyler, 1980. The temporal structure of spoken language understanding. Cognition 8, 1-71.

Marslen-Wilson, W. D., L. K. Tyler, and M. Seidenberg, 1978. Sentence processing and the clause boundary. In: W. J. M. Levelt and G. B. Flores d'Arcais (eds.), Studies in the perception of language. Chichester: Wiley.

McConkie, G. W., and D. Zola, 1981. On the nature of perceptual processing during reading. In: A. M. Lesgold and C. A. Perfetti (eds.), Interactive processes in reading. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Mistler-Lachman, J. L., 1972. Levels of comprehension in processing of normal and ambiguous sentences. Journal of Verbal Learning and Verbal Behavior 11, 615-623.

Pehrsson, R. S. V., 1974. The effects of teacher interference during the process of reading or how much of a helper is Mr. Gelper? Journal of Reading 17, 617-621.

Propp, V., 1958. Morphology of the folktale. International Journal of American Linguistics 24, Part 3.

Przetacznikowa, M., and M. Kielar, 1981. Comprehension of sentences with simple syntactic structure in children aged 3 to 13. Polish Psychological Bulletin 12, 17-24.

Rosenzweig, M. R., 1961. Comparison among word association responses in English, French, German, and Italian. American Journal of Psychology 74, 374-360.

Rubin, A., 1980. A theoretical taxonomy of the differences between oral and written language. In: R. J. Spiro, B. C. Bruce, and W. F. Brewer (eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Russell, W. A., and J. J. Jenkins, 1954. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff word association test. USN Technical Report No. 11, University of Minnesota.

Schenker, A. M., 1973. Beginning Polish, Vol. 2 (Revised ed.). New Haven: Yale University Press.

Shugar, G. W., 1977. Early semantic relations and their surface expression: A study of Polish in give/take transactions. Conference on Psycholinguistics, Saltzburg.

Slobin, D. I., 1973. Cognitive prerequisites for the acquisition of grammar. In: C. A. Ferguson and D. I. Slobin (eds.), Studies of child language development. New York: Holt, Rinehart & Winston.

Smoczyńska, M., in press. Acquisition of Polish. In: D. I. Slobin (ed.), The cross-cultural study of language acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Stanovich, K. E., 1980. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly 16, 32-71.

Stein, N., and C. Glenn, 1978. An analysis of story comprehension in elementary school children. In: R. Freedle (ed.), Discourse processing: Multidisciplinary perspectives. Norwood, N.J.: Ablex.

Warren, W. H., D. W. Nicholas, and T. Trabasso, 1979. Event chains and inferences in understanding narratives. In: R. Freedle (ed.), Advances in discourse processes, Vol. 2. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Table 1

Mean Production Times as a Function of Violation Type and Word-Unit Position in the Two Polish Experiments

Violation Type	Word-Unit Position											
	-5	-4	-3	-2	-1	CW	+1	+2	+3	+4	+5	Mean
Experiment 1												
Control	528	530	507	567	461	517	595	465	478	495	454	509
Within-Syntactic	586	562	522	599	517	1266*	827*	519	453	547	499	627*
Syntactic + Semantic	518	497	512	659	764*	2179*	1193*	697*	638	611	634	890*
Lexical	621	552	489	567	629	1853*	913*	572	589	456	609	714*
Experiment 2												
Control	606	551	541	631	463	523	631	484	569	554	554	555
Between-Syntactic	625	564	590	581	885*	1688*	1068*	814*	701	621	544	789*
Semantic	630	570	552	651	737*	1164*	1264*	700*	613	643	600	739*
Factual	565	565	588	631	459	625	799	499	556	621	555	586*

*Experimental violation type was significantly different from its control at $p < .05$ using a MS_e comprised of both subject and critical word variances.

Figure Captions

Figure 1. A sample portion of the Polish and English texts surrounding the critical words potrącona and injured. The word-unit positions are indicated.

Figure 2. Differences in mean production times between the syntactic violations and their controls.

Figure 3. Differences in mean production times between the semantic and factual violations and their controls.

Figure 4. Differences in mean production times between the syntactic + semantic violation and and its control.

Figure 5. Differences in mean production times between the lexical violation and its control.

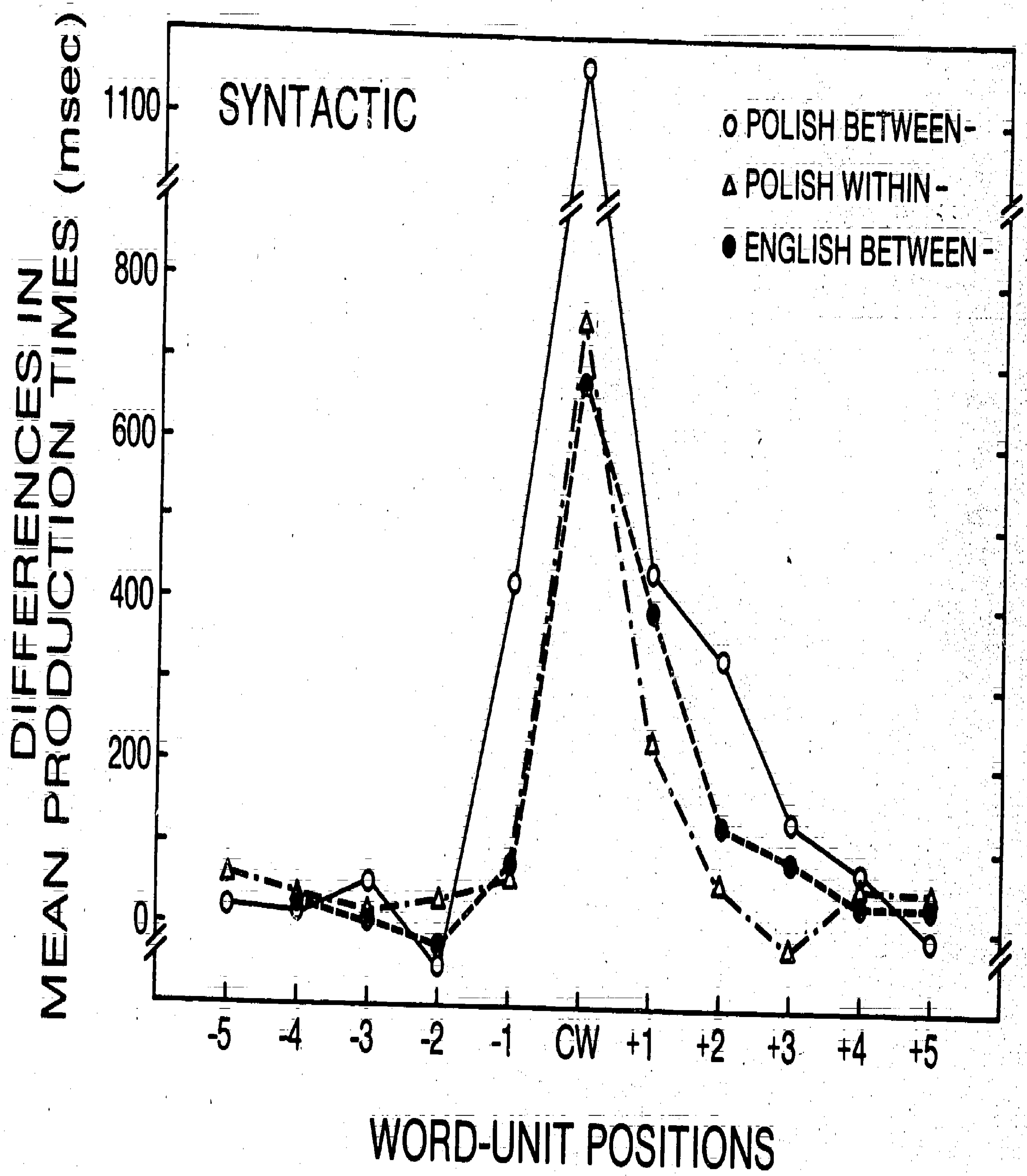
Córka jej zawsze była słaba fizycznie. Z tego powodu
Her daughter had always been weak physically. Because of this,

-5
nawet
she even

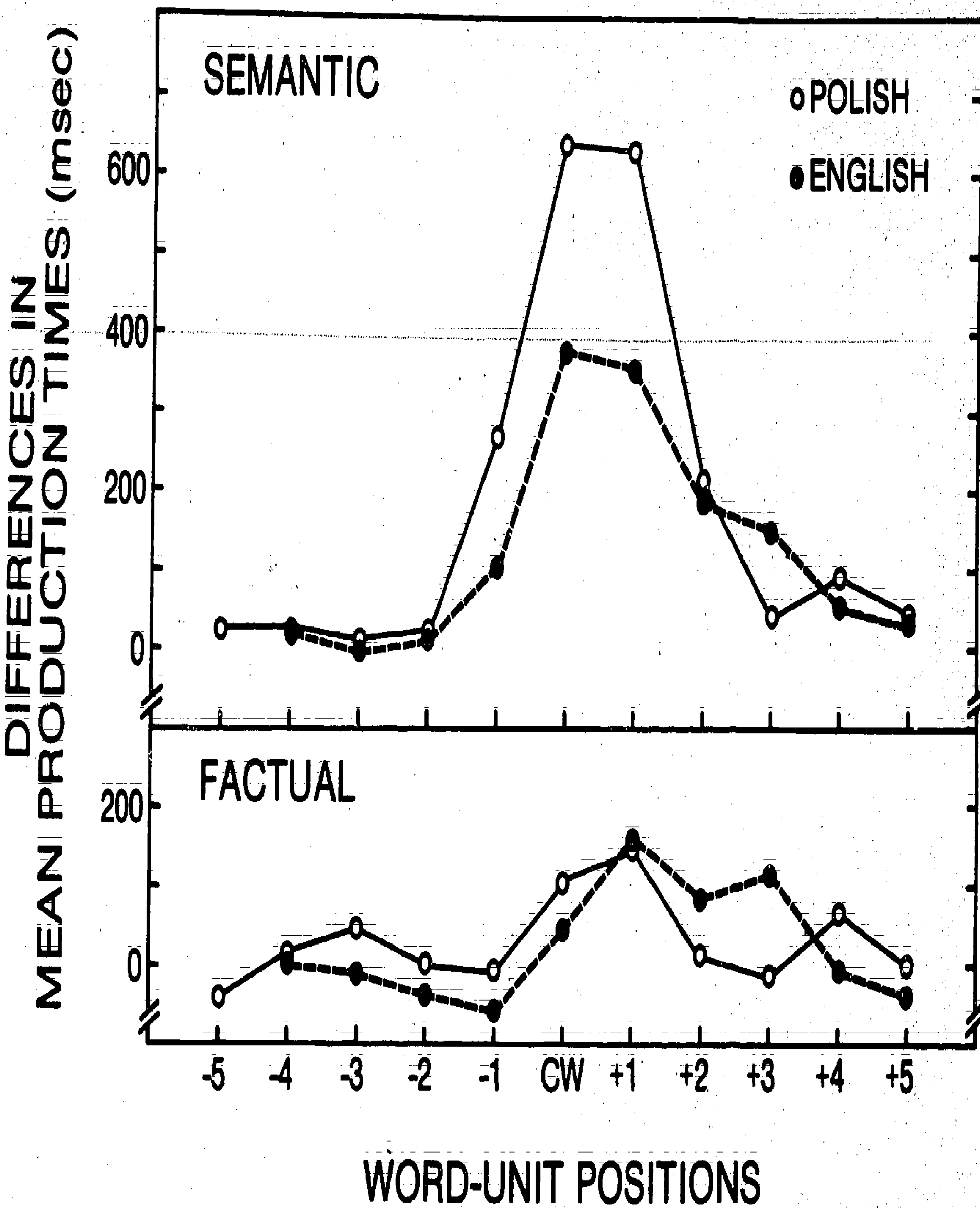
-4	-3	-2	-1	CW	+1	+2	+3
wyobrażała	sobie,	że córka	została	potrącona	przez inne	dzieci	w chwili
imagined	her	daughter	being	injured	by other	children	while

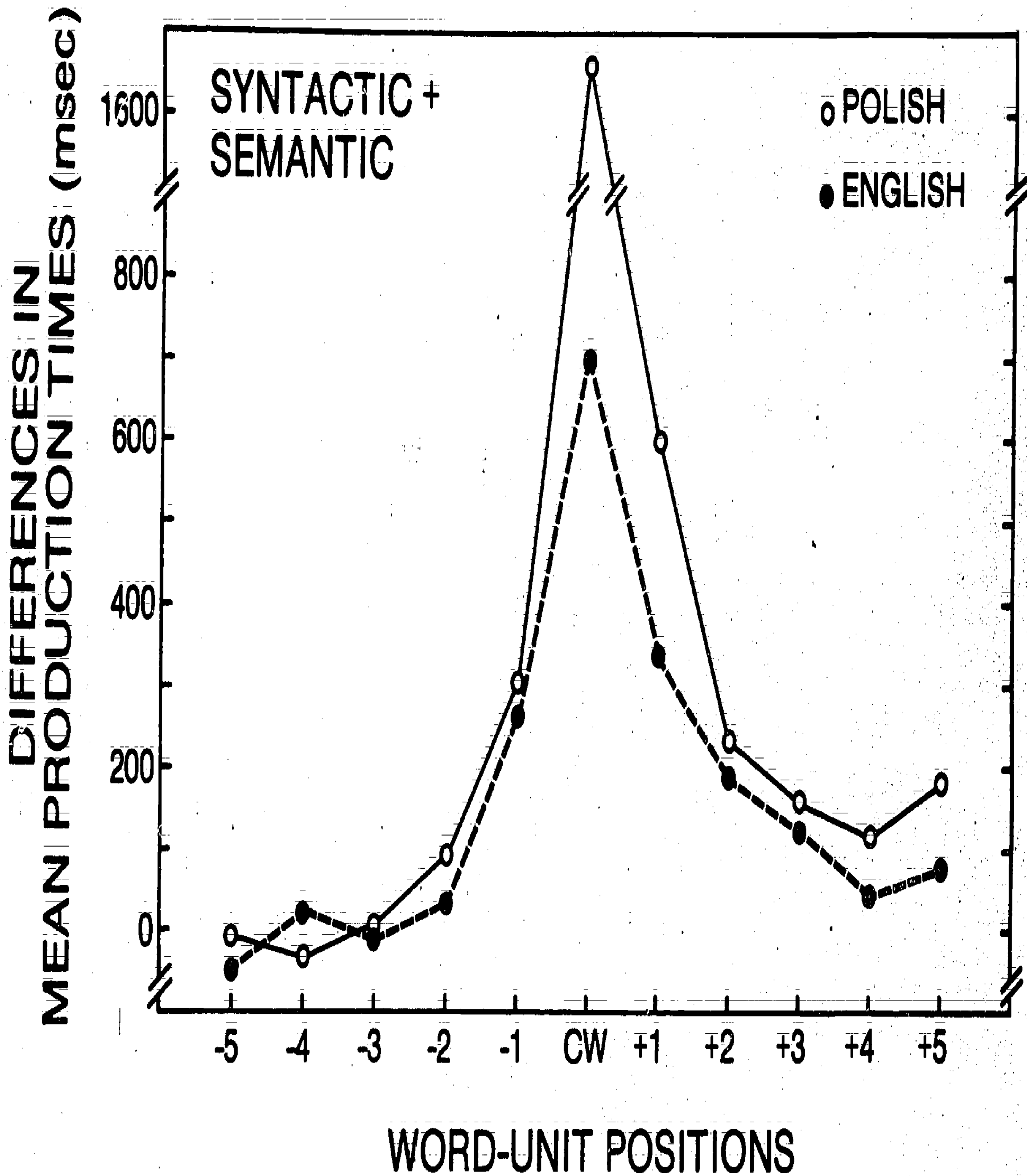
+4	+5
gdy usiłowała	wydostać się
trying	to get

z uszkodzonego autobusu. . . .
out of the wrecked bus. . . .

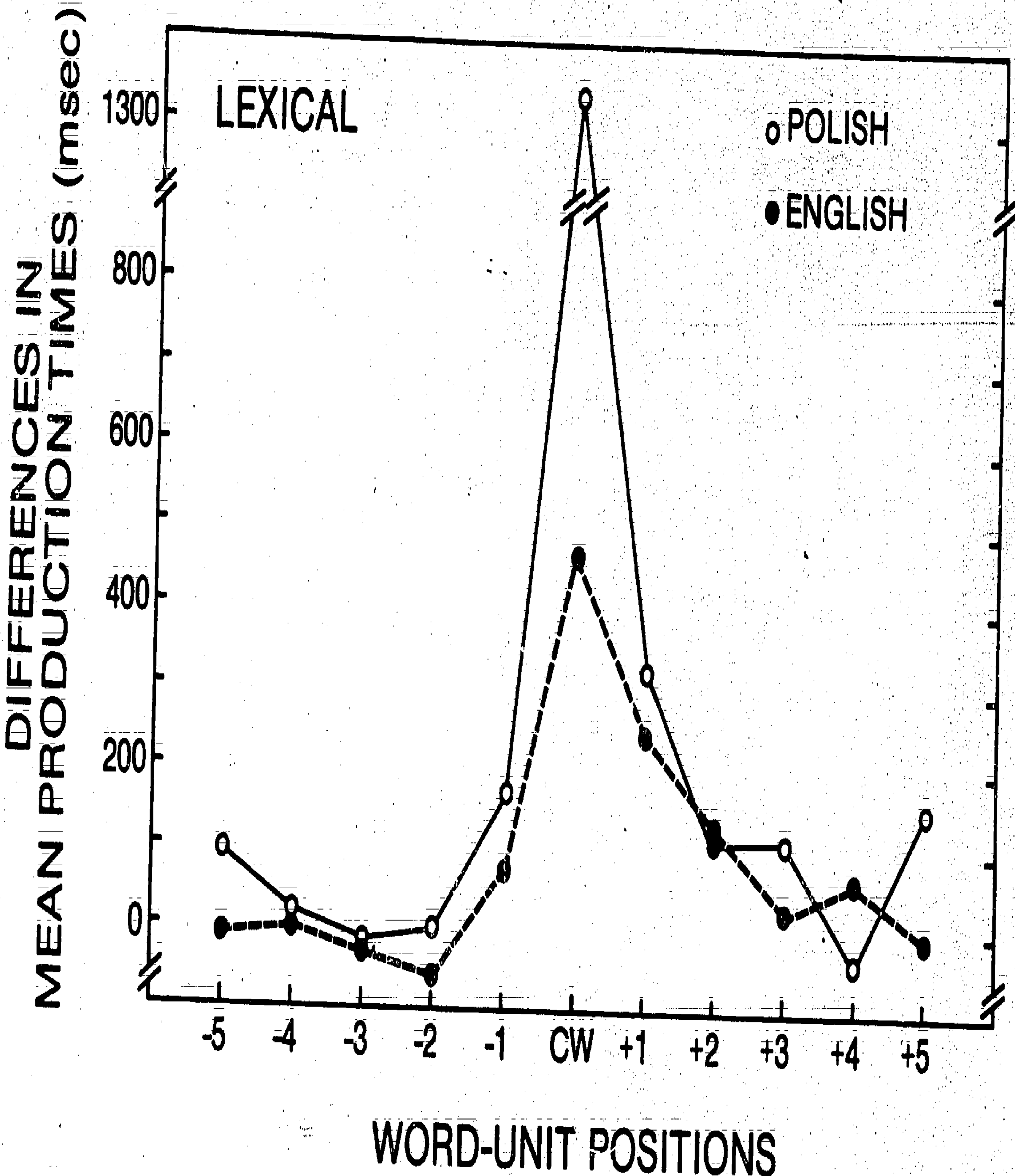


DIFFERENCES IN MEAN PRODUCTION TIMES (msec)
SYNTACTIC
o POLISH BETWEEN-
Δ POLISH WITHIN-
● ENGLISH BETWEEN-
1100
800
600
400
200
0
-5
-4
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD-UNIT POSITIONS


SEMANTIC
POLISH
ENGLISH
FACTUAL
DIFFERENCES IN
MEAN PRODUCTION TIMES (msec)
600
400
200
0
200
0
-5
-4
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD-UNIT POSITIONS



DIFFERENCES IN
MEAN PRODUCTION TIMES (msec)
SYNTACTIC +
SEMANTIC
o POLISH
● ENGLISH
1600
800
600
400
200
0
-5
-4
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD-UNIT POSITIONS



LEXICAL
o POLISH
● ENGLISH
DIFFERENCES IN MEAN PRODUCTION TIMES (msec)
1300
800
600
400
200
0
-5
-4
-3
-2
-1
CW
+1
+2
+3
+4
+5
WORD-UNIT POSITIONS

Memory and Metamemory Processes: Levels of Processing and Cognitive Effort in the Retention of Prose

Elizabeth A. Bohn
Kent State University
Kent, Ohio 44242
(216) 672-2166

Memory and Metamemory Processes

Mailing Address:

Elizabeth A. Bohn
Department of Psychology
Kent State University
Kent, Ohio 44242

Memory and Metamemory Processes: Levels of Processing and Cognitive Effort in the Retention of Prose

Elizabeth A. Bohn

Topic: Memory-Prose
Memory-Metamemory

Footnote: This paper is based on a dissertation submitted to Kent State University. Special thanks goes to my Committee Chairman, Joseph H. Danks, for his guidance and comments in the preparation of this draft. I also thank the members of my committee, David Riccio, Steven Levin, Roy Lilly, Robert Chabot, and John Gordon for their help and encouragement.

ABSTRACT

Memory and metamemory processes were compared by introducing an idea unit ranking task to prose text adapted to depth of processing and cognitive effort theories. Depth of processing was operationalized through fluency, concreteness, and category-relatedness instructions. Cognitive effort was manipulated through text organization: matched-to-task (easy), scrambled (difficult), and narrative organizations. Proportion recalled (memory) and proportion estimated recall (metamemory) for each idea unit level comprised the primary dependent measures. Depth of processing was successfully extended to both prose memory and metamemory. Cognitive effort was not, since easy matched and difficult scrambled organizations were equivalent. Recalling passages provided additional information to increase the accuracy of metamemory estimations. Metamemory (estimation) patterns in general echoed those of memory (recall).

We are often aware of our capacities and limitations in gathering information. We may know, for example, that it is easier to remember three unrelated items than to remember ten, or that it is easier to multiply three digit numbers with paper and pencil than in our heads, or that our oldest son is more capable in Arithmetic while our youngest is better in English. Our knowledge about our own and others' cognitive processing capabilities can be broadly defined as metacognition (Flavell, 1979). According to recent taxonomies (Brown, 1977; Flavell, 1979; Flavell, 1981; Paris, 1978), metacognition includes both our internal representation of abilities (metacognitive knowledge) as well as our ongoing monitoring of cognitive processes (metacognitive processing). In contrast to theories and research in cognitive information processing which describe the ongoing flow of processing from input to output, metacognition examines a person's understanding and monitoring of this information as it influences his/her and others' cognitive performances.

Metamemory, or knowledge about one's own memory, has received the greatest attention in recent research, however. Metamemory, like metacognition, consists of both metamemory knowledge and memory monitoring (Wellman, 1977). Metamemory knowledge entails general facts known about the characteristic properties of memory, for example, that a person can remember short lists better than long ones. In contrast, the ongoing internal judgments, attributions, and assessments that a person makes about specific items in memory is called memory monitoring. Memory monitoring, in the form of estimations about memory, is the type of metamemory assessed in the present study.

Memory monitoring and its predictive utility for recognition success was first examined by Hart (1965, 1967). Following an analogue of Brown and McNeill's (1966) tip-of-the-tongue experiments, Hart asked adult subjects to estimate from unrecalled items, those they believed they could recognize. He found subjects were able to predict both recognition failure and success quite accurately for unretrieved items. By far, the bulk of memory monitoring, as well as metamemory research in general, has been developmental in nature, focused mostly on differences between children and adults, on production versus mediation deficiencies, and on demonstrations of how teaching effective mnemonic strategies can yield increments in performance in problem solving (Brown, 1977; Brown & DeLoache, 1977; Brown & Smiley, 1977; Flavell, 1981; Keniston & Flavell, 1979; Markman, 1981; Yussen & Paquette, 1978).

Tha majority of literature in meta-processing has yet to connect itself explicitly with the vast theoretical foundations in cognitive information processing. Few researchers have noted the possible connections between the memory theories which guide their investigations and metamemory. Only one study to date (Yussen & Paquette, 1978) has attempted such a comparison by using the constructivist paradigms of Bransford and Franks (1971) and Paris and Carter (1973). However, their interest was in developmental differences between children and adults' predictions of unrelated versus related sentences, not in theoretical explorations of metamemory per se. One further study (Cavanaugh & Borkowski, 1980) has attempted multiple memory-metamemory comparisons, but it did not test

specific cognitive theories by comparing and manipulating relevant variables.

Traditional memory and metamemory explanations are not necessarily diametrically opposed. Rather, differences in recall could be the result of both mechanisms. Separate analyses of memory and metamemory processes would clarify the role each plays in a general characterization of memory. This investigation made use of such separate analyses with the assumption that metamemory and memory phenomena are not isomorphic. Sometimes there is reciprocity between them and sometimes not. Memory processing could influence metamemory experience and the monitoring of memory input. In turn, experiences and monitoring could result in control over ongoing memory processing.

Metamemory, in the form of memory monitoring was conceived within a processing framework. In this scheme, metamemory is a higher-order flexible process, one which monitors and controls lower-order memory processes. In this sense, metamemory becomes an executive processor (Brown & DeLoache, 1977) that functions as an intentional regulator and generator of lower-order routines (Brown, 1977). The central processor not only receives and evaluates memory input, but in addition controls the flow of that information through the cognitive system.

The interdependence and divergence of memory and metamemory processes was examined in this study using prose materials. In prose, main themes abstracted from text could influence the feeling of knowing and ongoing memory monitoring of the

text itself. Baker (1979), for example, examined adult subjects' comprehension monitoring of text which contained various confusions. Although subjects reported the existence of few confusions, Baker discovered that poor recall of confusions was due to subjects' rapid application of cognitive strategies to resolve them. In this case, failure in metamemory (reporting the existence of confusions) was due to rapid resolution in cognitive processing.

The intention of the research reported here was twofold: First, existing cognitive theories were used as methodological tools to explore memory-metamemory connections. By manipulating factors derived from these theories in a memory-metamemory framework, metamemory processing can be investigated and some of its characteristics established. If the manipulated theoretical variable shows identical results in memory and metamemory, not only is the theory extended to include metamemory processing, but in addition new information is learned about metamemory itself. If different results are registered in memory and metamemory measures, then these results delineate limits of the manipulated theory as well as showing qualitative differences in memory-metamemory processes. Although hundreds of theoretical cognitive variables potentially could be manipulated in this way, two were chosen because of the possible interrelationships between them: levels of processing and cognitive effort. Manipulation of these theoretical variables represents a first attempt to flesch out the metamemory processing system.

In addition, a second consideration motivating this research

was to introduce these two established variables to prose material. Although metamemory for prose has been studied (Baker, 1979; Brown & Smiley, 1977) and although levels of processing has been manipulated in prose (Schallert, 1976), simultaneous examination of both levels and effort on prose memory and metamemory has not been tested.

Recall of Prose: Depth of Processing versus Cognitive Effort

Depth of processing, formally introduced by Craik and Lockhart (1972), refers to a series of hierarchic cognitive processes applied to stimulus input, ranging from "shallow" structural analyses to "deeper," more elaborate semantic operations. Although recent research has contraindicated some of its notions (Baddeley, 1978; Craik & Tulving, 1975; Nelson, 1977), empirical results in general have supported the hypothesis of semantic levels of analyses leading to better retention than shallow perceptual levels for words (Craik & Tulving, 1975) and for sentences (Mistler-Lachman, 1974, 1975).

For prose text, Dooling and Christiaansen (1977) discussed the implications of depth of processing for a constructivist orientation. The fact that main ideas lead to better recall than details in prose carries an implicit levels flavor.

One major study (Schallert, 1976) has examined prose recall under different levels of processing. Schallert manipulated depth of processing by having subjects examine ambiguous text under different task instructions. Subjects counted either four-letter words or pronouns, rated the degree of ambiguity of the text, or read the text for later recall. She found

that the number of idea units recalled varied directly with depth of processing. Shallow tasks (counting words and pronouns) led to worse recall than deep tasks (rating for ambiguity and intentional recall).

One difficulty with Schallert's manipulations was that she varied both the units processed as well as the depth of processing. "Unit" refers to the input chunk of processing specified by the task (letter, word, sentence, or paragraph), and "depth" refers to the type of processing operation applied (structural, phonemic, or semantic). In Schallert's experiment, units and depth were confounded, such that subjects receiving shallow instructions (counting) examined word units, and subjects receiving deep instructions (ambiguity ratings) examined whole text. Processing operations at different depths should be applicable to units of a variety of sizes. A continuum of operations can be applied at any chunk unit, be it word, sentence, or paragraph.

The present research chose an intermediate-sized, more abstract unit for prose, an idea unit (Johnson, 1970). Idea units are parts of or whole sentences which express only one complete idea. Idea units were written as complete sentences in the present experiment. Subjects ordered all sets of idea units according to a specified criterion. Hence, all idea units had to be processed equally. The present experiment extended the generality of the depth of processing framework by analyzing sentence idea units in text.

These idea units were to be analyzed according to one

of three task instruction criteria: fluency (shallow), concreteness (intermediate), or category-relatedness (deep) processing. Category processing has been found to be an effective orienting task with words, enhancing recall over phonemic and rhyming tasks (Craik & Tulving, 1975). Intermediate rhyming tasks, however, have been problematic since task focus is directed only at word endings (Nelson, 1977). For sentence idea units in text, varying task depths were constructed intuitively, so as to be applicable to this larger unit of analysis. Instead of processing words according to their relatedness to a category, idea units in text were processed in terms of the text's category topic, which itself was highly related to the text's thematic title. In the same way, an intermediate depth was defined by directing processing towards idea units' concreteness in the world. A shallow depth was determined by a criterion that focused processing on fluency characteristics of idea units. Whereas the fluency task examined how easily sentence idea units flowed in speech and reading, concreteness tasks examined how easily the idea units were visualized or connected to the real world. In contrast, deep category-relatedness tasks required the greatest degree of semantic processing--relatedness to its category topic. Subjects sorted all sentence idea units into three groups within each task depth--low, medium, and high (fluency, concreteness, or category-relatedness). Hence idea units were scaled within each particular task instruction depth.

To ferret out the effects of depth of processing on text

comprehension and recall, cognitive effort was introduced by varying idea unit organization. Effort has been defined as "the amount of the available processing capacity of the limited capacity central processor utilized in performing an information-processing task" (Tyler, Hertel, McCallum, & Ellis, 1979, p. 608). In Tyler et al.'s experiment, the effort notion was tested within a depth of processing framework, using anagram and sentence completion materials as representative of shallow and deep levels respectively. Within each level, low effort (easy) and high effort (difficult) materials were manipulated. Cognitive effort was monitored through a secondary tone detection task. Both effort and task depth produced differential recall. High effort, difficult materials increased recall over low effort, easy materials. Deeper levels of processing showed an advantage in recall over shallow levels. General results depicted no interaction between effort and depth of processing. Probe (tone) reaction time reflected differences in effort but not in depth of processing.

Tyler et al.'s procedure split effort into easy and difficult materials within each depth. Analogously, the present study varied effort within each task depth through how the idea units were organized in presentation to subjects. Easy organizations were those that matched the task criterion. For example, a levels task in which units were ordered in terms of the fluency of their expression was termed "easy" if the input idea units were already pre-organized from most to least fluent. "Difficult" effort was required when the idea units were presented

in random scrambled orders because they bore no obvious relation between the order and task criterion. A third organization presented the text in its original narrative order. How much effort narrative order might require was uncertain because of contrary influences. More effort might result because the subjects would have to break down the narrative organization so that the task criterion could be met. On the other hand, less effort might be needed because initial reading would be easier (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975; Stein & Nezworski, 1978).

In sum, different idea unit organizations were high or low in cognitive effort depending on the task. Difficult, high effort, scrambled organizations resulted when subjects had to sift through disorganized information in order to meet the task requirements. Easy, low effort, matched organizations resulted when idea unit ordering matched the task requirements. Matched organizations required little effort because they provided an external ordered structure directing subjects to process the task-relevant information efficiently. The amount of cognitive effort depended on the match between text organization and task requirements.

Both cognitive effort and depth of processing were hypothesized to affect recall. A depth of processing interpretation predicts better recall with deeper task instructions: the fluency-of-expression task would yield the least amount recalled and the category-relatedness task would result in the greatest amount recalled. A cognitive effort interpretation

predicts enhanced recall with difficult idea unit organizations: scrambled organization should result in a greater amount recalled than matched-to-task organizations.

Metamemory Estimation: Depth of Processing versus Cognitive Effort

The primary question the current study addressed was how depth of processing and cognitive effort affect metamemory estimations of recall and their relative error. Although past investigations have shown adults to be accurate in estimating their performance (Yussen & Paquette, 1978; Wellman, 1977), little research has examined the memory variables which influence their estimation accuracy. Many cognitive variables have been identified in information processing, but few have been explored within the context of metamemory. Yet metamemories are themselves cognitive processes. The present research represents a first approximation directed at uncovering relevant theoretical mechanisms in metamemory processing. Specifically, degree of cognitive effort and depth of processing were expected to influence metamemory-based estimations as well as actual recall. Since both variables have an effect on memory performance, do they affect metamemory performance as well? Would this effect be identical or divergent from assessed recall?

A depth of processing main effect was expected in recall, such that deeper, category-relatedness processing would enhance recall over shallower, fluency-based processing. Metamemory would be considered sensitive to depth of processing if estimations of recall duplicated the recall results. If a depth effect was present in recall and recall estimations indexed no differences due to depth, then metamemory would be considered

insensitive to depth. Conversely, if differences in depth of processing were absent in recall, yet present in estimations, then metamemory monitoring would be in error. If such were the case, subjects would have believed erroneously that depth of processing contributed to recall, when in fact it did not. This type of metamemory insensitivity is qualitatively distinct from the case where recall differences are present, but differences in estimates are not. Finally, these two types of erroneous metamemory can be distinguished from a third case, wherein both recall and estimation differences are present, but with qualitatively different patterns. For example, a category-relatedness task might produce the greatest recall, but a concreteness task might produce the largest estimations.

In the same way, differences in effort organizations could emerge in recall, but not in estimations of recall. Then, metamemory would be considered insensitive to variations in cognitive effort. The same types of divergences between recall and estimations described for depth of processing apply to the cognitive effort manipulation, as well as the idea unit level manipulation.

Different patterns emerging from memory and metamemory imply qualitatively distinct processing. Subjects would not be just insensitive to their own recall, but their beliefs about their memory processes would be in error as well. If a particular pattern were present in recall estimates (for example, an interaction between effort and depth), but absent in recall, subjects apparently would be in error in believing that one variable (depth of processing) influenced another

(cognitive effort).

Metamemory error refers to how far estimations of recall are from actual assessed recall. Over- or underestimates of recall indicate differential error in the monitoring process. What causes subjects to inaccurately monitor their memory? They may over- or underemphasize the cognitive effort expended while processing the passage, they may assume that one level of processing was deeper or shallower than it actually was, or they may inaccurately assess the degree of their cognitive effort when processing at a particular depth. In a variety of ways exist in which subjects may err in memory monitoring. Differential error of estimations allowed a clearer assessment of the informational basis used for metamemory estimations and how these estirations interacted with actual memory processing mechanisms.

METHOD

The purpose of the present study was to compare memory and metamemory processes in the context of prose materials. Two theoretical variables, depth of processing and cognitive effort, were operationally defined and manipulated. Depth of processing was introduced through fluency, concreteness, or category-relatedness task instructions. Cognitive effort was manipulated through matched-to-task (easy), scrambled (difficult), and narrative organizations. Separate groups of subjects read three passages and ranked each set of sentences in a passage into three groups: low, medium, or high, fluency, concreteness, or category-relatedness. Half of the subjects estimated the amount they would remember before they recalled

the passages. The remaining subjects first recalled the information and then estimated how much they believed they recalled.

Materials

Three articles ("Windpower," "TV," and "Genetic Research") were adapted from *Time* (1980). Each passage consisted of 30 idea units constructed by the experimenter so that each sentence could stand alone without pronouns or anaphoric references to other idea units. The units were written in the original narrative order, one idea unit per line, and numbered 1-30.

Prior to the main experiment, the investigator randomly scrambled the 30 sentence idea units within each passage. These three scrambled passages were given to an independent group of 102 raters drawn from a pool of General Psychology students attending Kent State University in fall semester, 1980. Three separate groups of 34 subjects each sorted the units into three idea unit levels of either fluency, concreteness, or category-relatedness. All groups were given all three passages. The mean rank across subjects for each sentence idea unit within each passage and task instruction depth was calculated. The idea units in each passage then were ordered from high to low in terms of their mean ranking for each instructional condition. This procedure resulted in three unique orderings per passage: from most to least fluent, from most to least concrete, and from most to least category-related.

These three orderings served as input materials for the

main experiment's matched (easy effort) condition. Each ordering was matched to its particular task instruction depth. Scrambled (difficult) order was the same in both preliminary and main experiments and was determined randomly. Narrative order was derived from the original passages as constructed by the previously discussed experimental criteria. Spearman Rank Order correlations among orders were computed to determine similarity of organizations. The orders were unrelated with an average intercorrelation of .004 and with a range of .185 to -.213.

Experimental booklets and instructions. Two experimental booklets were prepared, one for the instructional ranking task and one for the recall and recall estimation tasks.

In the first booklet, a practice passage appeared on the first two pages after the cover sheet, followed by task-specific instructions. The final six pages consisted of the three experimental passages arranged according to one of the three experimental organizations (scrambled, matched-to-task, or narrative). Passage order was counterbalanced in each condition.

Task instructions were incidental, only explaining the orientation of the task and the idea unit ranking procedure. Fluency instructions indicated that idea units ranged in fluency of expression for reading and speaking. Fluency was said to refer to the flow of sentences. Concreteness instructions stated that idea units ranged in terms of how easy they were to connect to the real world. Concreteness instructions

referred to how easy the idea units were visualized or conceptualized. Category-relatedness instructions explained how some idea units were more related than others to a particular topic. Subjects in the category instruction condition read passages headed by their respective topic categories ("Windpower," "TV," "Genetic Research"). An example was given of the task criterion for each instruction set.

The second experimental booklet had two alternate forms: (a) a prediction estimation page followed by three blank pages for recalling the three experimental passages, or (b) three blank pages for recall followed by a postdiction estimation page. For the prediction estimation condition, instructions directed subjects to predict the number of idea units they believed they could recall from each of their own rank levels (low, medium, high) for each passage ranked (first, second, third). This page was followed by three blank pages in which subjects were requested to write down what they could remember in any order (that is, not necessarily by idea unit level), for the first, second, and third experimental passages, consecutively.

In order to investigate effects of prior recall on estimation, additional groups of subjects received the recall and estimation tasks in reverse order. In this condition, a postdiction task was administered. Postdiction instructions required subjects to estimate how many idea units per level they believed they actually did recall. Both estimation and recall instructions asked subjects to estimate and recall

idea units in terms of their own words.

Procedure

Subjects were run in separate groups by task instructional conditions. The experimenter passed out the first booklet and read the instructions to the subjects. Subjects were oriented to the nature of the task (fluency, concreteness, or category-relatedness) and then were told to read through the set of sentence idea units for the practice passage. The idea unit ranking procedure was explained. First, subjects chose a third of those idea units at a specified level (least fluent, least concrete, or least category-related) and drew a line through them. Then the procedure was repeated for the next level (in which a circle was drawn around 10 idea units intermediate in fluency, concreteness, or category-relatedness), leaving the remaining 10 idea units untouched (that is, the 10 most fluent, concrete, or category-related). This procedure was repeated for the three experimental passages.

After the ranking task, the passage booklets were collected. Then the recall estimation booklets were passed out. For prediction estimation, subjects were asked to predict their recall for each rank level within each passage for all three passages. Then they recalled each of the passages in order of presentation. Instructions for postdiction asked subjects to first recall the passages and then to estimate the number of units from each rank level for each passage they believed they actually had recalled in gist.

Subjects

Subjects in the main experiment were drawn from a pool of General Psychology students attending Kent State University in the spring semester of 1981. Subjects served in the experiment as partial fulfillment of a course requirement. Nine groups of 25 subjects each received a particular organization by task instruction condition in the prediction estimation group. Nine comparable groups of 25 subjects also were tested in the postdiction condition. The total number of subjects participating in the main body of the experiment was 450.

Scoring and Analyses

Recalls were scored by the investigator who was blind to the particular experimental condition of each subject. Because the recall task was incidental, protocols were scored by lenient criteria. If the major theme of an idea unit was present in the protocol, regardless of verbatim accuracy or detail, it was considered as recalled. Recall of idea units was expressed as proportion of units recalled as a function of each subject's own ranked idea unit level. To assess scoring agreement, an independent rater scored 144 passage recalls. Across idea unit levels, 91% of the raters' judgments were either in total agreement or one idea unit away.

Results were analyzed in a 3 X 3 X 2 X 3 X 3 mixed analysis of variance with three levels of task instructions (fluency, concreteness, category-relatedness), three levels of text organization (matched, scrambled, narrative), two levels of recall estimation (prediction, postdiction), three levels of idea units (high, medium, low), and three levels of passages

("Windpower," "TV," and "Genetic Research"). Task, organizational effort, and prediction/postdiction testing were between-subjects' factors. Passage topic and rank levels of idea units were within-subjects' factors.

Three dependent variables were analyzed: (*a*) the principal memory measure was proportion recalled as a function of idea unit level; (*b*) the principal metamemory measure was proportion estimated recall as a function of idea unit level; and (*c*) a derived error measure was estimated minus actual recall. All post hoc comparisons were conducted using Newman-Keuls. All effects reported as significant are with $p \leq .05$ unless indicated otherwise.

RESULTS AND DISCUSSION

Recall

Both concrete and category-relatedness instructions, although not differing themselves, produced significantly better recall than did fluency instructions, F (2, 432) = 31.421, MS_e = .071, as shown in Table 1. The means for scrambled and

Insert Table 1 about here

matched organizations were not significantly different. Task instructions and organization did not interact, F (4, 432) < 1, MS_e = .071. Previous findings that recall of isolated word lists was enhanced by deeper levels of task processing (Craik & Lockhart, 1972; Craik & Tulving, 1975) also apply to prose texts. However, greater effort did not enhance recall, as had been found in past research with words (Tyler

et al., 1979). Difficult scrambled organizations were equivalent to easy matched organizations. Only narrative organizations improved recall. The difference between prediction/postdiction testing was not reliable, $\underline{F}$ (1, 432)<1, MS_e =.071. So, metamemory estimations did not influence amount recalled.

For idea unit level, high level units (most fluent, concrete, or category-related) were recalled better (.203) than medium level units (.171), and low level units produced the worst recall (.121), $\underline{F}$ (2, 864) = 146.838, MS_e = .016. However, this main effect was qualified by several higher-order interactions. The idea unit level pattern differed depending on task instructions, $\underline{F}$ (4, 864) = 17.077, MS_e = .016, as shown in Figure 1. Both concreteness and fluency instructions showed linear decreases in amount recalled, but the slope was steeper under concreteness instructions. Category-relatedness instructions showed equivalent recall for high level and medium level units, both producing greater recall than low level units.

- - - - - - - - - - - - - - - -

Insert Figure 1 about here

- - - - - - - - - - - - - - - -

There were significant differences among passages in the amount recalled, $\underline{F}$ (2, 864) = 26.153, MS_e = .023, and passages interacted with task instructions, $\underline{F}$ (4, 864) = 3.600. MS_e = .023, and idea unit level, $\underline{F}$ (4, 1728) = 9.991, MS_e = .013. However, both two-way interactions entered into a triple interaction of task, idea unit level, and passage topic, $\underline{F}$ (8, 1728) = 9.650, MS_e = .013. The patterns of results across

idea unit levels were replicated in all three passages for concreteness and fluency instructions. However, category-relatedness instructions produced variable results across passages. Category-relatedness showed greater sensitivity to passage topic than either concreteness or fluency instructions. Concreteness and fluency instructions replicated a linear idea unit ordering over passages. Category-relatedness instructions, however, lacked this stability and were more dependent on the content of the passages.

It was expected that category-relatedness instructions would result in better structured recall. Intuitively, they were presumed to induce the deepest semantic level processing. Relating idea units to category meanings seemed deeper than to concreteness in the real world. Recall was not better structured under category-relatedness instructions, however. These instructions were related to processing content of the passages, and recall reflected sensitivity to differences in that content.

Estimations

Concreteness instructions did not differ from category-relatedness instructions, both producing reliably greater estimations than the fluency task, $\underline{F}$ (2, 432) = 5.786, MS_e = .192, as shown in Table 2. The effect of text organization reflected a similar pattern for estimates as for recall (narrative > scrambled = easy), but this effect was not statistically significant, $\underline{F}$ (2, 432) = 2.928, $\underline{p}$ = .053, MS_e = .192.

- - - - - - - -

Insert Table 2 about here

- - - - - - - -

Prediction/postdiction interactions, absent in recall, were present in estimations. Prediction/postdiction interacted with task instruction, $\underline{F}$ (2, 432) = 3.071, MS_e = .192. As shown in Table 2, concreteness instructions resulted in higher recall estimations for postdictions than predictions, whereas fluency and category-relatedness instructions resulted in slightly lower postdictive estimations. Subjects processing text according to its concreteness and who had already recalled the passages tended to estimate their recalls as producing more idea units than those who had not had the opportunity to recall text.

Subjects' recall estimates were highly sensitive to idea unit level, $\underline{F}$ (2, 864) = 57.460, MS_e = .029; all ordered pairwise comparisons between levels (high = .279, Medium = .231, and Low = .211) were significantly different. This pattern was the same as the one obtained in actual recall. The main effect of idea unit level was qualified, however, by an interaction with prediction/postdiction estimation, $\underline{F}$ (2, 864) = 5.070, MS_e = .029, as depicted in Figure 2. Postdictions showed a linear ordering of estimations as a function of idea unit level. However, in prediction testing, only high level

- - - - - - - -

Insert Figure 2 about here

- - - - - - - -

idea units were differentiated clearly. Medium and low levels did not differ. Recalling passages prior to estimation produced the same idea unit orderings to those of actual recall. When passage information was recalled, further memory feedback corrected divergent idea unit level estimations. A task instruction by idea unit level interaction, present in recall, was not significant in estimations, $\underline{F}$ (4, 864) = 2.027, $\underline{p} < .09$, MS_e = .029. The linear patterns was present in all tasks.

In contrast to recall results, there were no significant effects involving passage topic. All the significant effects in recall estimation were replicated across passages and none were compromised by an interaction with passages.

In summary, reliable effects in metamemory estimates were produced as a function of instruction, idea unit level, and interactions of prediction/postdiction tests with task and with idea unit level. The effects of text organization, passage topic, and interactions with passages were absent.

Relative Error (P-A)

Main effects of between-subjects' variables of prediction/postdiction estimations, task instruction, and text organization were absent in this analysis, indicating subjects' estimations were equally accurate for these variables. A passage topic main effect was present $\underline{F}$ (2, 864) = 9.871, MS_e = .040, and passage interacted with organization, $\underline{F}$ (4, 864) = 2.423, MS_e = .040, and idea unit level, $\underline{F}$ (4, 1728) = 6.747, MS_e = .020. A triple interaction between passage, task instruction, and idea unit level was also present, $\underline{F}$ (8, 1728) = 6.599, MS_e =

.020.

More importantly, for idea unit levels, medium level idea units were the most accurately assessed, followed by high level units and then low level units, $\underline{F}$ (2, 864) = 7.861, MS_e = .040. However, this main effect was qualified by the interaction with prediction/postdiction testing, $\underline{F}$ (2, 864) = 3.573, MS_e = .040 (see Figure 3). Postdiction estimates produced relatively equivalent accuracy across idea unit level,

---

Insert Figure 3 about here

---

but prediction estimates were relatively less accurate for high and low level idea units and more accurate for medium level idea units. Idea unit levels also interacted with task instruction, $\underline{F}$ (4, 864) = 5.116, MS_e = .040. The shallow fluency task, produced no differences in accuracy by idea unit level, but the concreteness and category-relatedness tasks did. Concreteness instructions showed greater overestimation of low level idea unit, but category-relatedness instructions resulted in greatest overestimation of both highest and lowest level units. Deeper tasks produced differences in error for idea unit structure.

Subjects tended to overestimate their recall in all experimental conditions. Perhaps, the greater number of passages read, the greater the overestimation. Another reason for overestimation could have been that recall instructions were lenient. Stringent instructions might have resulted

in underestimations of recall. In the present experiment recall instructions were lenient, so subjects may have responded by overestimating their recall. Finally, overestimation could have been due to internal scaling differences. Subjective mental scales of memory may differ from experimentally defined recall measures.

Comparisons of Recall and Estimation Measures

The main effects of task instruction and idea unit level were reflected clearly in both memory and metamemory measures (refer to Table 3). These two factors were the most salient parts of the experimental instructions. Subjects were told explicitly to attend to the task orientation and to the sorting procedure (idea unit levels). Experimental instructions did not call attention to text organization. Although organization affected recall, this pattern failed to reach significance in the metamemory measure.

\- - - - - - - -

Insert Table 3 about here

\- - - - - - - -

Memory itself was not affected by whether a prediction test was administered before recall. In contrast, metamemory was sensitive to whether estimates were given predictively or postdictively. In the prediction condition, estimates produced in response to task instructions showed no differences. However, after recalling passages, subjects elevated their estimates when given concreteness instructions. Feedback from recall in the concreteness condition raised estimates above those from category-relatedness instructions. Hence, the

postdictive estimation ordering more closely duplicated the recall pattern. In addition, prediction/postdiction estimations interacted with idea unit level. Whereas predictive estimates failed to differentiate between medium and low level idea units, postdictive estimates did, and approximated the recall results more accurately. Generally, recalling passages affected metamemory processing for the two strongest variables, task instruction and idea unit level. In the present study, memory influenced metamemory processing, not the reverse.

Differences in patterning between memory and metamemory processing were found in the task instruction by idea unit interaction (see Table 3). The primary difference was in the metamemory system's failure to pick up a fine-grained effect, although the pattern of means was similar. The cause of this pattern difference was found in the task-idea unit-passage triple interaction. Examination of this pattern showed memory to be sensitive to the variable passage-specific memory effects of idea unit level in the category-relatedness task. In contrast, metamemory was not sensitive to passage content; the linear effect of idea unit level was present in all three passages. Both idea unit level and task were replicated across all three passages in metamemory, but not in the memory measure.

To compare the relative strengths of effects, omega^2 was computed for all relevant results in memory and metamemory measures. Omega^2 examines proportion of the total variance accounted for by each treatment effect (in this case total within-subject variance). Hence, this statistic indexes the

the strength of effects obtained from $\underline{F}$ ratios (Hays, 1973). In the context of metamemory processing, omega^2 was used as an index of sensitivity to memory variables. Comparisons of omega^2 between recall and estimated recall are presented in parentheses in Table 3.

Overall, the proportion of accounted for variance decreases from memory to metamemory levels. In all cases, when a significant effect was present at the memory level, it was attenuated at the metamemory level.

The relative magnitude of significant recall effects, ordered from most to least, showed task instruction to be the strongest variable, followed by idea unit level and text organization, respectively. In estimations, these same variables were reduced in strength. Significant estimation effects of task instructions and idea unit level were not as strong when compared to recall variables. However, task instruction and idea unit level, the strongest effects in recall, were also the strongest effects in estimations. The weaker text organization recall effect was not reliable in estimations. Prediction/postdiction interactions, present only in the estimation measure, showed relatively nonexistent effects, as shown by omega^2.

The general summary of results revealed differences in recall, and recall estimates, as a function of task instruction, text organization, passage topic, idea unit level, and prediction/postdiction estimations. Estimation patterns did not reflect those of recall in all cases. Obtained effects at a metamemory level weakly echoed those already present

in memory. Stronger effects in memory (as determined by omega^2) were more likely to be reflected in metamemory. When all measures are compared, it was possible to assess the relative patternings of each, clarifying the relationship of memory-metamemory connections. The description of these connections depicted a pattern of echoing from memory to metamemory levels. Metamemory traced similar patterns of main effects. But the echoing in metamemory was weaker, suggesting a second-order, more abstractive relation registered in metamemory.

GENERAL DISCUSSION

Memory Processes

Task orientation, idea unit level, text organization, and passage topic produced reliable differences in recall. Depth of processing, an effect commonly found with words (Craik & Tulving, 1975), has been extended to sentence idea units in text. Similar to Schallert's (1976) work with task instructions, processing to deeper levels of analysis enhanced performance. Unexpectedly, concreteness instructions either equalled or exceeded category-relatedness instructions. To relate subordinate idea units to a category heading seemed "deeper" than to relate them to their concreteness in the real world. Concreteness instructions did not seem to necessitate semantic operations, since subjects need only visualize the sentences. Category-relatedness instructions, in contrast, were believed to require subjects to make use of internal semantic representations. However, concreteness instructions appeared to induce a relationship to an even deeper, more

elaborate, schematic base--the real world, which is more inclusive than any particular category subset of that world. Both instructional sets could be interpreted as relational to semantic representations, whereas fluency instructions lacked this semantic component. The fluency orientation directed processing towards the surface representation of the sentences--their choppiness and flow, rather than towards the sentences' underlying ideas.

In contrast to Tyler et al.'s (1979) results, cognitive effort failed to have an effect: Difficult scrambled and easy matched organizations produced equivalently low recall. Only original narrative organization enhanced recall. It is possible that matched and scrambled organizations did not cause different levels of cognitive effort. In Tyler et al.'s experiment, an independent index of effort was used to assess amount of central processing capacity. Longer latencies of tone detection (the secondary measure) were assumed to reflect greater utilization of the central processor, and hence, greater effort. Because of the size of the subject sample in the present experiment, such a secondary measure was not feasible. It is uncertain then, whether scrambled passages indeed induced greater effort in processing compared to matched passages. Further, narrative order could have conceivably produced the greatest effort. It is recommended that future research monitor processing capacity with secondary measures.

Higher idea units were recalled better regardless of sorting on. The ordering of idea unit levels was stable

across passages for concreteness and fluency instructions, but tended to vary across passages for category-relatedness instructions. It is interesting to note that idea unit level work has been researched only as a function of a thematic importance tasks (Johnson, 1970; Brown & Smiley, 1977; Brown, 1977), itself analogous to the category-relatedness task used in the present study. These results showed idea units ranked as most important to the main theme of the narrative are recalled better than units ranked as less important. In the present experiment, other task directives (fluency, concreteness, and category-relatedness) produced similar mean recall orderings of idea unit levels as those obtained in thematic importance. Hence, the linear effect of idea unit level occurs across a variety of tasks, not solely for a thematic importance task. This result suggests that thematic importance rankings reflect a deeper level of task orientation.

Further, narrative organization produced greater recall independent of the task instruction criterion by which idea units were sorted. If the idea unit effect were dependent on narrative structure alone, then when this structure was altered through scrambled or matched organizations, the idea unit effect should have been eliminated, producing a flat function. This did not occur. The idea unit effect was present in all text organizations, but as defined by task, not narrative structure. Such a result runs counter to previous interpretations of thematic importance as inherently reflecting the narrative structure of passages. In contrast, the present results strongly suggest that the monotonic ordering of idea

units, from most important to least, is a result of general task-oriented processing, and not due to narrative structure. When underlying narrative structure was disrupted by changes in text organization, the idea unit effect still remained, although the total amount of recall decreased.

The rankings of idea units according to task criteria also showed that idea units were processed primarily according to task, not according to organizational structure. If idea unit rankings were dependent on different organizations, then idea unit rankings would be uncorrelated. Subjects would have rank the same idea units differently when presented in different organizations. However, correlations between rankings obtained with differing organizations were relatively high. Average correlations ranged from .74 between easy and scrambled, .76 between easy and narrative, to .79 between scrambled and narrative organizations. Hence if a particular unit was ranked high in narrative organizations, it also tended to be ranked high for other organizations. In contrast, idea unit rankings were uncorrelated across different task instructions. Idea units ranked high in one task were not necessarily ranked high in other tasks. The correlation of fluency ranks with concreteness ranks was -.03, fluency with category-relatedness was -.10, and concreteness ranks with category-relatedness was -.13. High-ranked idea units that were recalled well in a fluency task were different from high-ranked idea units that were recalled well in category-relatedness or concreteness tasks. Idea unit recall clearly was dependent on task, not organization.

All these results point to an interpretation of the linear idea unit effect as one based on task-dependent processing. Although narrative structure enhances recall overall, it does not determine the shape of the recall function (in terms of idea unit level).

Metamemory Processes

Task instruction depth and idea unit level were reflected clearly in metamemory, as in recall. That estimations registered differences in these two variables extends depth of processing and idea unit level to salient cognitive variables for metamemory. Since idea unit level scaled estimations within each task depth, from least to most fluent, concrete, or category-related, idea unit level results also showed a monitoring of task orientation in general. In contrast, text organization and passage topic were not monitored reliably in metamemory, although they did indeed affect recall. Similarly, estimations were not sensitive to higher-order recall interactions among these variables. Metamemory-based estimations tended to match actual task-directed recall more closely when subjects estimated directly after recalling the passages than before.

Omega^2 (Table 3) depicted estimation effects as diminished in strength, compared to recall. The description of metamemory processing lies, not within any particular memory effect, but within the relative changes in manipulations of memory processes. When metamemory reflects the same relative patterning in memory, it is considered sensitive. In contrast, metamemory is insensitive when it fails to reflect the differences

shown in memory, either by showing no reliable pattern or by showing a divergent pattern from the one obtained in memory. By examining the overall picture of memory-metamemory comparisons, it becomes possible to assess the informational basis behind metamemory estimates. The present results suggest an "echoing" in metamemory occurring reliably to the strongest recall patterns of task instruction and idea unit level. Prediction/postdiction estimations interacted with these two strongest recall variables in such a way that postdictions more clearly reflected obtained memory results than did predictions. Organization and passage topic traced similar patterns of results in metamemory as in recall, but failed to reach significance. Hence, the weaker recall variables of organization, passage, and higher-order interactions were less likely to be monitored reliably in estimations. Rather than showing insensitivity due to distortion in estimations, metamemory reflected insensitivity due to lack of monitoring of weaker recall variables. The echo effect depicted only the strongest memory variables as being monitored reliably in metamemory. The greater the strength in obtained recall patterns, the greater the liklihood it was accessed in metamemory. In all cases, however, the strengths of the obtained metamemory effects were diminished compared to those of memory.

Three possible reasons could account for this echoing. First, metamemory might echo only salient memory effects, where "salience" is derived from its strength in memory, as manipulated experimentally. Whether this strength is due to greater attention, a greater number of experimentally-induced

associations, or greater activity in memory processing is not important for this particular explanation. Rather, if certain memory effects are particularly salient in any specified experiment, their strength (as measured by omega^2) is enhanced. The outcome of memory salience is what is monitored in metamemory. Greater salience allows the memory effects to be monitored in metamemory. Hence, the greater the salience in memory patterns, the greater the liklihood it is monitored reliably in metamemory. This explanation implies that if any cognitive variable's salience is strengthened or diminished, as measured by omega^2, then concomitant monitoring differences would be seen in metamemory. If text organization, for example, were manipulated to have greater memory salience, then significant effects would be obtained in metamemory measures for text.

A second possibility explains metamemory as differentially sensitive to task demand information. Task demands tap into subjects' control processing, orienting them towards specific strategies. Both depth of processing and idea unit level constitute task demands, promoting active processing of text. Both require subjects to think about the text in specific ways. In contrast, organization and passage content do not require subjects to actively process text. Although both are concerned with defining the memory representation, neither indicates which aspect of text information is important for the memory representation. Hence, metamemory, being only sensitive to task demand activity, did not monitor organization or content. If organization or content had entered into task demands, then such information would have been monitored. For example,

if subjects were instructed to reorder scrambled passages, or compare the different organizations or content of passages, then metamemory would have reflected these variables reliably. According to this interpretation, metamemory locks into task demands and stores them for future use. Estimations can only tap into the task demands. If metamemory, in the form of estimations, echoes task-oriented processing alone, then estimations provde another index of depth, since depth of processing itself is a task orientation. Hence, the criticism of circular explanations of depth (Baddeley, 1976) could be partially circumvented by appealing to a metamemory index.

A third explanation as to how memory and metamemory might operate posits a single representation in memory which is accessed through different retrieval measures. Alternative response decisions tap into the stored representation with varying degrees of specificity. Instead of metamemory filing away a separate representation based on memory salience, or based on task demand activity, and instead of metamemory as being described as a distinct and separate system, it is viewed as another aspect of cognitive processing. One memory representation is formed and different retrieval processes access this representation. Estimates of recall only retrieve the general characteristics of the memory representation. According to this interpretation, estimates of recall constitute a weaker test of the memory representation, just as a recall test is typically less sensitive than a recognition test in tapping memory. If the echo effect resulted from diminished sensitivity

of estimation measures, then other selected measures of metamemory would not necessarily echo their concomitant memory measures. This view would invoke explanations of how different retrieval processes activate a single memory representation.

These three approaches can be reconciled only through further research designed to uncover the processing mechanisms involved in memory monitoring. The approach illustrated here represents a general paradigm in which to test various theoretical implications of metamemory. In this paradigm, memory and metamemory processes were compared by manipulating existing cognitive theories. Levels of processing and cognitive effort were selected as examples of theories relevant to memory-metamemory comparisons. In addition, idea unit level was manipulated since it was particularly appropriate to prose. Not only was the generalizeability of levels of processing and idea unit l. extended to include the metamemory system, but in turn, all manipulated variables pointed to a descriptive account of memory and metamemory processing referred to as the "echo effect." Further research should uncover the stability of this effect as well as its underlying causes.

REFERENCES

Baddeley, A. D. The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139-152.

Baker, L. A. Comprehension monitoring: Identifying and coping with text confusions. (Tech. Rep. #145), Urbana, Ill.: Center for the study of reading. University of Illinois, September, 1979.

Bransford, J. D. & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. (Tech. Rep. # 47). Urbana, Illinois: Center for the study of reading. University of Illinois, June,

Brown, A. L. & DeLoache, Skills, plans, and self-regulation. (Tech. Rep. #48). Urbana, Illinois: Center for the study of reading. University of Illinois, July, 1977.

Brown, A. L. & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. Child development, 1977, 48, 1-8.

Brown, R. & McNeill, D. The "tipof the tonge " phenomenon. Journal of verbal learning & verbal behavior, 1966, 5, 325-337.

Cavanaugh, J. C. & Borkowski, J. G. Searching for metamemory-memory connections: A developmental study. Developmental psychology, 1980, 441-453.

Craik, F. I. M. & Lockhart, R. S. Levels of processing: a framework for memory research. Journal of verbal learning and verbal behavior, 1972, 11, 671-684.

Craik, F. I. M. & Tulving, E. Depth of processing and the retention of words in episodic memory. Journal of experimental psychology: General, 1975, 104, 268-294.

Dooling, J. & Christiaansen, R. Levels of encoding and retention of prose. In G. H. Bower (Ed.), The psychology of learning and motivation, (Vol. 2), N. Y.: Academic, 1977.

Flavell, J. H. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American psychologist, 1979, 24, 906-911.

Flavell, H. G. Cognitive monitoring. In W. P. Dickson (Ed.), Children's oral communication skills, N. Y. : Academic Press, 1981.

Hart, J. T. Memory and the memory monitoring process. Journal of verbal learning and verbal behavior, 1967, 6, 685-691.

Hart, J. T. Memory and the feeling of knowing experience. Journal of educational psychology, 1965, 56, 208-216.

Hays, W. L. Statistics for the social sciences. New York: Holt, Rinehart & Winston, 1973.

Johnson, R. E. Recall of prose as a function of the structural importance of the linguistic units. Journal of verbal learning and verbal behavior, 1970, 9, 12-20.

Keniston, A. H. & Flavell, J. H. A developmental study of intelligent retrieval. Child development, 1979, 50, 1144-1152.

Kintsch, W., Kozminsky, E., Stroby, W. G., McKoon, G., & Keenan, J. M. Comprehension and recall of text as a function of content variables. _Journal of verbal learning and verbal behavior_, 1975, _14_, 196-214.

Markman, E. J. Comprehension monitoring. In W. P. Dickson (Ed.), _Children's oral communication skills_, N. Y.: Academic Press, 1981.

Mistler-Lachman, J. L. Depth of comprehension and sentence memory. _Journal of verbal learning and verbal behavior_, 1974, _13_, 98-106.

Mistler-Lachman, J. L. Queer sentences, ambiguity, and levels of processing. _Memory and cognition_, 1975, _3_, 395-400.

Nelson, T. O. Repetition and depth of processing. _Journal of verbal learning and verbal behavior_, 1977, _16_, 151-171.

Paris, S. G. Metacognitive development: Children's regulation of problem solving skills. Paper presented at the Midwestern Psychological Association, Chicago, May, 1978.

Paris, S. G. & Carter, A. Y. Semantic and constructive aspects of sentence memory in children. _Developmental psychology_, 1973, _9_, 109-113.

Schallert, D. E. Improving memory for prose: The relationship between depth of processing and context. _Journal of verbal learning and verbal behavior_, 1976, _15_, 621-632.

Stein, N. L. & Nezworski, T. The effects of organization and instructional set on story memory. _Discourse processes_, 1978, _1_, 177-193.

Time Magazine, Video Venom, 1980, 116, 81.

Time Magazine. Pioneers of the supergene. 1980, 116, 57.

Time Magazine, Riding the wind. 1980, 116, 78.

Tyler, S. W., Hertel, P. T., McCallum, M. C. & Ellis, H. C. Cognitive effort and memory. Journal of experimental psychology, 1973, 80, 352-373.

Wellman, H. M. Tip of the tongue and feeling of knowing experiences: A developmental study of memory monitoring. Child development, 1977, 48, 13-21.

Yussen, S. R. & Paquette, N. S. Developmental changes in predicting recognition memory for semantically related and unrelated sentences. Developmental psychology, 1978, 13, 107-113.

Table 1

Mean Proportion Recalled as a Function of

Task Instruction and Text Organization

Text Organization	Task Instruction: Fluency	Concreteness	Category-Relatedness	Mean
Matched	.106	.172	.164	.147
Scrambled	.106	.166	.186	.153
Narrative	.143	.231	.213	.196
Mean	.118	.190	.188	

Table 2

Mean Proportion Estimated Recall as a Function of

Task, Organization, and Prediction/postdiction Testing

	Task Instruction			
Text Organization	Fluency	Concreteness	Category Relatedness	Mean
Matched	.190	.246	.246	.227
Scrambled	.195	.260	.236	.230
Narrative	.237	.277	.278	.264
Mean	.208	.261	.253	
Prediction	.218	.236	.266	.240
Postdiction	.197	.285	.240	.240

Table 3

Summary Table of

Comparisons and Omega^2 for

Recall and Estimation Measures

Effect	Recall		Recall Estimation	
Task:	Con=Cat>Flu	(11.5)*	Con=Cat>Flu	(2.1)*
Idea Unit:	$L_1 > L_2 > L_3$	(6.8)*	$L_1 > L_2 > L_3$	(4.7)*
Organization:	Nar>Scr=Mat	(4.7)*	Absent	(.8)
Task X Idea Unit:	Cat: $L_1 = L_2 > L_3$	(1.4)*	Cat: $L_1 > L_2 > L_3$	(.1)
	Con: $L_1 > L_2 > L_3$		Con: $L_1 > L_2 > L_3$	
	Flu: $L_1 > L_2 > L_3$		Flu: $L_1 > L_2 > L_3$	
Pre/post X Task:	Absent	(0)	Pre: Cat=Con=Flu	(.9)*
			Post: Con>Cat=Flu	
Pre/post X IU:	Absent	(0)	Pre: $L_1 > L_2 = L_3$	(.1)*
			Post: $L_1 > L_2 > L_3$	

Con = concreteness
Cat = category-relatedness
Flu = fluency
Nar = narrative
Scr = scrambled
Mat = matched
L = level (L_1=High; L_2= Medium; L_3= Low)
Pre/post = prediction/postdiction
IU = idea unit

*Effect was significant at $p < .05$.

Figure 1. Proportion recalled as a function task instruction and idea unit level

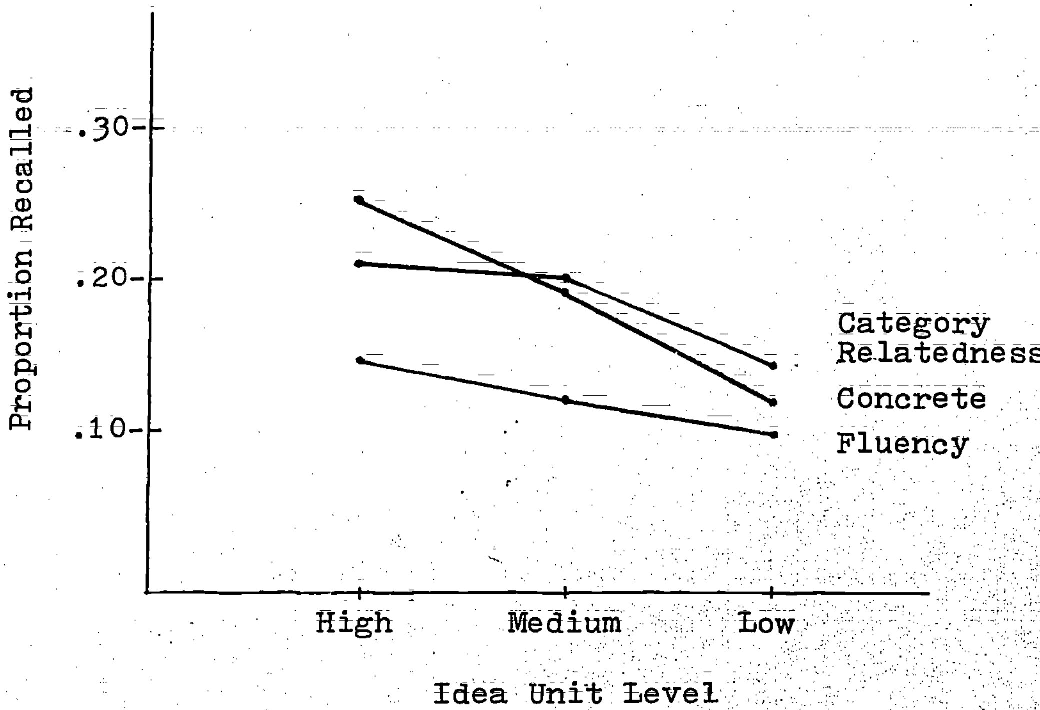


Figure 2. Proportion estimated recall as a function of pre/post estimations and idea unit level

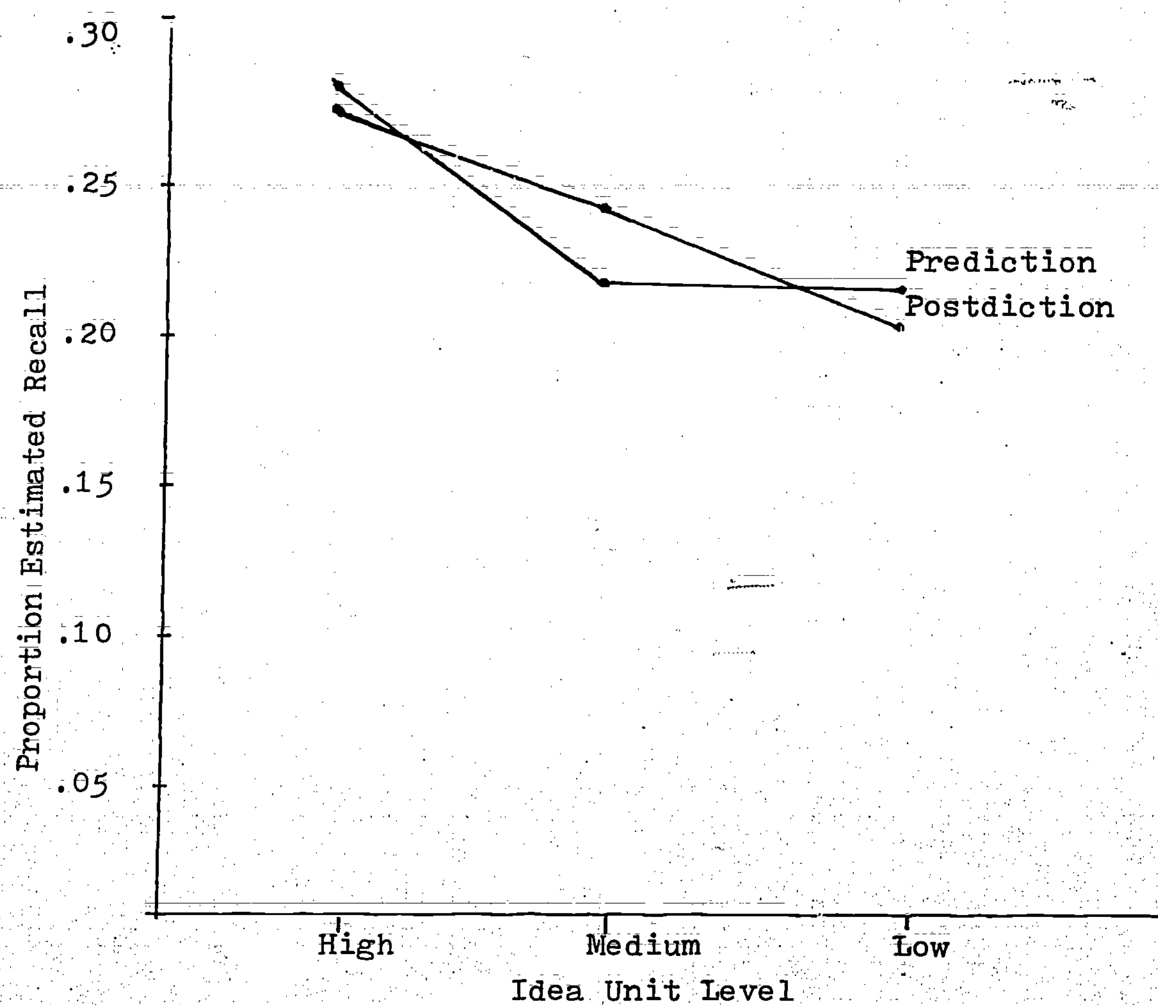


Figure 3. Signed accuracy as a function of pre/post estimation and idea unit level