Working Memory and Comprehension of Spoken Sentences: Investigations of Children with Reading Disorder*

Stephen Crain,† Donald Shankweiler,† Paul Macaruso,† and Eva Bar-Shalom†

1 INTRODUCTION

Our goal is to investigate the role of the verbal working memory system in sentence comprehension, by presenting a model of working memory in sufficient detail to allow specific predictions to be made and tested. In testing this account, we draw on experimental methods that have recently been used in research on language development. These methods are designed to control the various sources of potential difficulty in the standard laboratory tasks used to assess children's grammatical knowledge and their use of this knowledge in sentence comprehension. We illustrate how our proposals about working memory, together with the recent innovations in method, allow us to infer that abnormal limitations in phonological processing, and not absence of grammatical knowledge, is at the root of the difficulties in spoken sentence understanding that are apparent in children with reading disability.

Since reading problems are most transparent at the beginning stages of learning to read, we focus our attention there, by investigating the linguistic abilities of poor readers in the early school years.

By “poor readers” we mean children who show a marked disparity between their measured level of reading skill and the level of performance that might be expected in view of their intelligence and opportunity for instruction. Our research compares performance by these children with age-matched controls—children who are proceeding at the expected rate in the acquisition of reading skills (for discussion of the issues regarding subtypes of reading disability and choice of control groups, see Shankweiler, Crain, Brady, & Macaruso, in press).

Much of the research on poor readers finds the source of their problems in the language domain, not in the area of visual perception or in general analytic ability. We shall take this for granted (for reviews, see Perfetti, 1985; Shankweiler & Liberman, 1972; Vellutino, 1979). Within the language domain, many sources of evidence converge on the conclusion that poor readers’ problems reflect deficiencies in phonological processing (see Liberman & Shankweiler, 1985, and Stanovich, 1982, for reviews). However, there is one finding which raises the possibility that their limitations extend beyond phonological processing to syntactic processing as well: the discovery that poor readers characteristically fail to comprehend complex spoken sentences accurately under some circumstances. This finding has led researchers to the hypothesis that these children have not mastered all of the complex syntactic properties of the adult grammatical system (Byrne, 1981; Fletcher, Satz, & Scholes, 1981; Stein, Cairns, & Zurif, 1984). We have called this the structural lag hypothesis (SLH).

---

* Portions of this research were supported by a Program Project Grant to Haskins Laboratories from the National Institute of Child Health and Human Development (HD-01994). We wish to thank the second-grade students and teachers, reading instructors, and administrators at the Coventry, CT, elementary schools. We also thank Suzanne Smith for her help with Experiment III, Henry Hamburger for extensive discussion of the issues raised in this paper, and Brian Butterworth, Myrna Schwartz and Tim Shallice for their comments on an earlier draft.
The SLH provides a coherent account of some factors that may make reading hard to learn and which may distinguish good and poor readers. This hypothesis attributes poor readers’ difficulties in spoken language comprehension to their level of attainment in the acquisition of syntax. According to the SLH, language is acquired in stages, beginning with simple syntactic structures and culminating only when all of the complex structures have been mastered. To explain why language acquisition conforms to a developmental schedule, the SLH endorses the idea that syntactic structures are ordered in inherent complexity. The late emergence of a structure in the course of language development is taken as an indicator of its relative complexity as compared to structures that appear earlier.

It is clear that the SLH deserves serious consideration. Reflecting some common assumptions about language acquisition and linguistic complexity, this hypothesis makes the following prediction about the language related difficulties of poor readers: the linguistic structures that beginning readers and unsuccessful older readers will find most difficult are just those that appear last in the course of language acquisition. Thus, the SLH would point to findings of language acquisition studies that suggest that some syntactic structures emerge later than others in language development, and to studies showing the late mastery of these structures by poor readers. Though the SLH gives a plausible account of some of the difficulties encountered by poor readers, it has a major limitation. It gives no way to tie together poor readers’ problems at the level of the sentence with their problems at the level of the word. Specifically, the postulated syntactic deficit of poor readers is independent of their deficit in processing phonological information. This means that the SLH abandons the possibility of achieving a unitary explanation of the whole symptom picture of reading disability.

In our research we have sought support for an alternative hypothesis, which we call the processing limitation hypothesis (PLH). In contrast to the SLH, this hypothesis attempts to tie together all of the symptoms of the poor reader, viewing them as derived from inefficient processing of phonological structures. Several problems can be securely tied to a deficiency in phonological processing, including the difficulties that poor readers have in word segmentation, object naming and verbal working memory. Consider first their well-attested problems in bringing phonologic segments to consciousness. It has been shown in several language communities that on analytic tests requiring conscious manipulations of the phonemic structure of spoken words, poor readers are less proficient than children who are more successful in learning to read (Bradley & Bryant, 1983; Cossu, Shankweiler, Liberman, Tola, & Katz, 1988; Lundberg, Olofsson, & Wall, 1980; Morais, Cluytens, & Alegría, 1984). Another problem that has claimed a good deal of attention is their impaired performance on tests of object naming (Denckla & Rudel, 1976; Jansky & de Hirsch, 1972; Wolf, 1981). Analysis of the errors reveals that the mistakes are often based on phonological confusions rather than on semantic confusions (Katz, 1985). This suggests that this problem, too, is a manifestation of underlying phonological impairment.

This same line of reasoning also applies to verbal working memory. Because the verbal working memory system depends on the ability to gain access to phonological structure and use it to (briefly) maintain linguistic information, we might expect people who have phonological difficulties to show various limitations on tests of ordered recall (Baddeley, 1986; Conrad, 1964, 1972; Liberman, Mattingly, & Turvey, 1972). For poor readers, as in other language-impaired populations, there is ample evidence in the literature testifying to deficiencies in short-term retention of verbal materials. Differences in recall have been obtained with a variety of verbal materials, including words and spoken sentences, but they are not typically found with materials that cannot be coded linguistically (see Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977; Wagner & Torgesen, 1987). Moreover, there is direct evidence from memory experiments that poor readers in the beginning grades are less affected by phonetic similarity (rhyme) than age-matched good readers. This is another indication of their failure to fully exploit phonologic structure in working memory (Mann, Liberman, & Shankweiler, 1980; Olson, Davidson, Kliegl, & Davies, 1984; Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979).

In addition to these symptoms, we noted earlier that poor readers are sometimes unable to comprehend spoken sentences as well as comparable good readers. The central aim of this paper is to explain how the difficulties of poor readers in understanding spoken sentences may be derived from deficient phonologic processing. On the face of it, these difficulties might seem to require another kind of of explanation. But suffice it to say here that the findings of our recent
research, including the results of the experiments presented in Section 4, have persuaded us that the source of their spoken language comprehension failures is also tied to an underlying deficiency in phonological processing, as proposed by the PLH, and is not the result of a lag in syntactic development, as predicted by the SLH.

Given these sharply contrasting hypotheses about poor readers' problems in sentence comprehension, we now turn to the kinds of evidence that can decide between them. One source of evidence may be obtained by examining the pattern of errors good and poor readers make in response to sentences of different types. If poor readers suffer from a limitation in processing, it makes sense that the pattern of errors on different structures should be similar for both groups, with the poor readers showing a decrement of roughly the same magnitude on each sentence type. The prediction that the error pattern of poor readers should parallel that of good readers serves as the foundation for one of the experiments reported in Section 3.

Another research strategy which has proven useful in distinguishing between the PLH and SLH is to examine the performance of good and poor readers on laboratory tasks which differ in how severely they tax the resources of working memory. Marked improvement in performance in the face of reduction in memory load is anticipated by the PLH but not by the SLH. In the absence of requisite structures, poor readers should fail in comprehension even when memory load is minimal. On the other hand, if a processing limitation is the source of the problem, even the most unskilled reader should prove competent with highly complex linguistic constructions in spoken language, within the constraints imposed by their limitations in processing capacity. This prediction too is tested in the experiments we report below. Before we give details of the experiments, it will be useful to describe our view of the working memory system and its role in language processing.

2 ORGANIZATION OF THE LANGUAGE APPARATUS

Our conception of the language apparatus shares much common ground with the modularity proposal advanced by Fodor (1983). It grows out of a biological perspective on language that has long guided research on speech at Haskins Laboratories. According to this viewpoint, the language faculty functions autonomously in the sense that it is supported by special brain structures and operates according to principles that are specific to it and not shared by other cognitive systems. One source of evidence for this conception of modularity comes from studies of speech perception (Mattingly & Liberman, 1988). Another source is from the study of aphasia and related disorders where there is evidence that a circumscribed lesion in the left hemisphere may selectively perturb certain aspects of language performance, leaving other linguistic and nonlinguistic abilities relatively intact (Linebarger, Schwartz, & Saffran, 1983; Marin, Saffran, & Schwartz, 1976; Shankweiler, Crain, Gorrell, & Tuller, 1989). There is also evidence that ability to process language may be preserved in the face of massive losses to other systems, as in cases of "isolation aphasia" (e.g., Whitaker, 1976).

Another source of evidence for modularity comes from the study of language development, where it has been found that complex linguistic principles emerge in young children at a characteristic pace that is independent of the emergence of other cognitive systems or principles (e.g., Hamburger & Crain, 1984). Also important are research findings demonstrating children's early mastery of linguistic principles that go beyond the data provided by the environment (e.g., Crain & McKee, 1985; Crain & Nakayama, 1987; Crain, Thornton, & Murasugi, 1987). Taken together, all of these findings sustain the notion that language is a biologically coherent system, as the modularity proposal maintains.

An extension of the modularity proposal supposes that the language faculty itself is composed of several autonomous subcomponents (or submodules). This componential view of sentence production and comprehension postulates several structures and processors. Roughly, each structure is a stored system of rules and principles corresponding to a level of linguistic representation: phonology, syntax and semantics. In addition to the independent levels of structural representation, the language apparatus contains special processors, including the phonological, syntactic and semantic parsers. Each parser is a special-purpose device for rule access and ambiguity resolution corresponding to a specific level of representation. Each parser operates on principles and rules in assigning constituent structure to linguistic input. Because the parsers operate on constituent structure, and not on sequences of words themselves, we can understand sentences of great length, but can retain only relatively short lists of unrelated material.
Two further architectural features of the language apparatus are essential to our explanation of the difficulties poor readers have in sentence understanding. We assume, first, that the various submodules are arranged in a hierarchical fashion, with a unidirectional and vertical ("bottom up") flow of information such that a lower level passes results to higher levels but not the reverse. It is also critical to our view that transactions between the parsers take place "on-line," with the results of low level analyses being quickly discarded, to make room for subsequent input (for related discussion, see Carpenter & Just, 1988).

**How working memory functions in the language processing system.** In keeping with the modularity hypothesis, we conceive of verbal working memory as a domain-specific system that subserves the language apparatus. The primary function of verbal working memory is to facilitate the extraction of a meaning representation corresponding to the linguistic input. Assuming that the extended modularity hypothesis is correct, this involves the interaction of several structures and processors. As we conceive of it, verbal working memory is an active processing system in which the analysis of verbal material by these structures and processors takes place during language processing.

In common with other contemporary approaches, we assume that there are two components to the working memory system (Baddeley, 1986; Baddeley & Hitch, 1974; Carpenter & Just, 1988; Daneman & Carpenter, 1980; Perfetti & Lesgold, 1977). First, there is a storage buffer where rehearsal and initial (phonological) analysis of phonetically coded information takes place. This buffer has the properties commonly attributed to short-term memory. It can hold information only briefly, perhaps only for 1-2 sec, in the order of arrival, unless the material is maintained by continuous rehearsal. The limits on capacity of the buffer mean that information must be rapidly encoded in a more durable form if it is to be retained for subsequent higher level analysis. Our conception of the storage buffer bears obvious similarities to other discussions in the literature.

What is new in our conception of the verbal working memory system concerns its other component. We view this as a control mechanism whose primary task is to relay the results of lower-level analyses of linguistic input upward through the language apparatus. Its regulatory duties begin at the lowest level by bringing phonetic (or orthographic) input into contact with phonological rules, for word level analysis. Phonologically analyzed information must be rapidly transferred out of the storage buffer and shunted to the syntactic processor, at the same time freeing the storage area to accept the next chunk of phonetic material. By synchronizing information flow with input, the control mechanism is able to push results upward through the system rapidly enough to promote on-line extraction of meaning (Crain & Steedman, 1985; Marslen-Wilson & Tyler, 1980; Wingfield & Butterworth, 1984).

In processing spoken language, on-line parsing explains how individuals with drastically-curtailed working memory capacity—capable of holding only two or three items of unstructured material—are sometimes able to comprehend sentences of considerable length and complexity (Martin, 1985; 1990; Saffran, 1985). Previous research has found it paradoxical that aphasic patients with a severely restricted phonological short-term store are sometimes capable of understanding at a level far exceeding what would be expected on the basis of their span limitations. This result is fully consistent with our model of working memory.

In reading, on-line processing of syntactic and semantic representations necessarily depends on prior orthographic and phonological processing. Until the reader is proficient in decoding from print, we would expect that reading is more demanding than speech of working memory resources. Sometimes it is assumed that print confers an advantage because the reader can look back. It is important to appreciate, however, that only the skilled reader can exploit the opportunity to reexamine sentences in text which were not successfully parsed on first reading. In the unskilled reader, the working memory system is usually preoccupied with orthographic decoding.

**3 IDENTIFYING THE SOURCE OF READING DISABILITY**

We are now in a position to show how the architectural arrangement of the language faculty can be exploited to provide an explanation of the sentence comprehension difficulties of poor readers. A modular view of the language apparatus raises the possibility that a single component may be the source of the entire symptom complex that characterizes reading disability. It is clear that failures in sentence comprehension could arise, in principle, from a deficit (or deficits) at any level that ultimately feeds into the semantic component. It is also conceivable, however, that the entire symp-
tom complex of poor readers, including their difficulties in spoken language comprehension, implicates the phonological component. Let us explain how.

Recall that the submodules of the language faculty act in strict sequence ("bottom up") to assign a partial structural analysis, which can then be passed on to higher levels. To keep information flowing smoothly, the control mechanism must avoid unnecessary computation that would delay the rapid extraction of meaning. This means that, in ordinary circumstances, the working memory buffer need not store many segments of unanalyzed linguistic material. But suppose that the phonological analysis of material in the buffer is impeded for some reason. Given the architectural features of the language apparatus we have proposed, this would also have the effect of curtailing the operation of higher level analyses of verbal material. In short, the functions of an otherwise intact system would be depressed.

This is exactly what happens in cases of reading disability, in our view. Since poor readers are deficient in setting up and organizing phonological structures, sentence comprehension is compromised because inefficient phonological analysis creates a "bottleneck" that constrains information flow to higher levels of language processing. Although the remaining components of the language apparatus may be completely intact, their operation will be hobbled by poor readers' limitations in phonological processing. In effect, a lower-level deficit in phonologic processing masquerades as a deficit at higher levels. At this point, however, we cannot rule out the possibility that the comprehension problems of poor readers are caused by a deficiency in some other component of the language apparatus (e.g., in syntactic parsing). But since poor readers' comprehension problems follow automatically from their well-attested limitations in phonological processing, it becomes unnecessary to postulate additional impairments within the language system. Moreover, we will provide evidence of the acquisition of complex syntax for both good and poor readers, as anticipated by the modularity hypothesis (see also Shankweiler & Crain, 1986).

It is important to underscore another expectation of our model, that poor readers should display successful comprehension on sentences that are not especially taxing of phonological resources. This distinguishes our view from other proposals about the relation between working memory and sentence comprehension (e.g., Baddeley, Vallar, & Wilson, 1987; Vallar, Basso, & Bottini, 1990). As long as the control mechanism of working memory is intact, even persons with abnormal limitations in phonological short-term storage capacity should be able to understand sentences of considerable complexity, if they do not impose excessive demands on phonological resources. Since the control mechanism of working memory plays such a prominent role in explaining why impaired comprehension should appear on specific sentences and not others, it will be worthwhile to describe it in more detail.

The compiling analogy. Pursuing an analogy with the compiling of programming languages, we view the control mechanism of working memory as a control structure whose function is to carry out a series of translations, each being a translation from a relatively high level language (the source language) to a more detailed language (the target or object language). This concept is familiar in computer science, where high level languages like Pascal or Lisp are compiled into lower-level languages such as assembly language or machine language. But the notion of compiling is quite general, and has proven useful in modeling human language processing as well.

Cognitive compiling occurs in natural language processing experiments in which a subject is asked to act out the interpretation of a sentence using toys and figures provided in the experimental workspace. Here, the source language, e.g., English, must be translated into a more detailed language that underlies the overt actions the subject makes in response to the input. We will refer to the mental language that serves as the target language for observable physical actions as the language of plans. In our view, several interesting properties of the control component of working memory can be illuminated by considering the translation between input sentences and the plans that they evoke (see Hamburger & Crain, 1984, 1987, for further discussion and for empirical data). In the paragraphs that follow, we focus on the difficulties that may arise for the executive component of working memory in the process of translating from language input to plans. We consider first situations that are amenable to simple translation between source and target language (Hamburger and Crain, 1984). Then we will look at particular linguistic forms which deviate from the best-case scenario, thereby exacting a toll from the resources of working memory.

In the simplest case, (i) each well-formed fragment of target language code is associated with a
single constituent of source language code, (ii) the fragments of target language code can be concatenated to form the correct representation of the input, (iii) the fragments can be combined in the same order they are accessed, and (iv) each fragment is processed immediately after it is formed, permitting the source code to be discarded. These conditions form a straightforward process of sequential look-up-and-concatenation. Rarely, however, are all the conditions met in ordinary language. And when they are not, the computations involved in reaching the target code (e.g., the semantic interpretation or plan associated with a linguistic expression) could stretch the resources of verbal working memory. It will not be possible to spell out each condition in detail, but it may be helpful to make a few remarks, about each, focussing on the linguistic constructions that appear in the experiments reported in Section 4.

(i) The first condition is an isomorphism between any two levels of representation. Correspondence of this kind is maintained between syntactic and semantic constituency in Montague Grammar in order to provide a systematic account of the assignment of semantic values to linguistic expressions. The foundation of this account is the principle of compositionality, which states that the meaning of a linguistic expression is determined from the meanings of its constituent expressions (and their mode of combination). Despite the appeal of a straightforward relationship between the syntax and semantics, linguists working in the generative framework have argued that syntax and semantics are largely autonomous. In our terms, this is liable to add to the complexity of translating between syntactic and semantic structural representations.5

(ii) Whether or not the first condition is met, it seems reasonable to suppose that the simplest way to combine nodes of the target language is by concatenation. Unfortunately, it is clearly not possible to concatenate meanings even in parsing simple natural language phrases like “expensive socks” or “second bear.” Since expensive socks are not expensive, it would be a mistake to evaluate this phrase on a word-by-word basis, e.g., by forming a semantic value for “expensive” (say, the set of expensive things), and then combining this with the semantic value of the following word, “socks.” Similarly, the second bear is not necessarily in second position in an ordered array. On occasion, concatenation of word meanings is possible, for example with NPs that contain absolute adjectives, like “green,” “fuzzy,” “Albanian,” and so on, where the denotation of the adjective is not dependent on the linguistic context (e.g., naked Albanian wrestler). But since no unique semantic value can be given to relative adjectives, (e.g., expensive) or to ordinals, the human sentence processing mechanism must hold off interpreting these prenominal modifiers until the head noun has been received.

Translations that require the parser to splice together dissociated pieces of code at some level also violate the simple process of look-up-and-concatenation (see Hamburger & Crain, 1984, 1987). An example of this source of distress for working memory is “second striped ball.” An analysis of the logical structure of the plan corresponding to this phrase shows it to consist of a nested loop structure in which fragments of plan associated with “striped ball” are inserted into the piece of code associated with the ordinal “second.” Breaking apart the code needed to increment a counter is required in order to test objects (for stripedness and ballhood), to ensure that the counter is advanced only as stripped balls are located. This process is referred to as compiling discontinuity by Hamburger and Crain.6

Empirical support for the claim that phrases like this present difficulties for young children comes from several acquisition studies which find that children often choose object (b) from an array like the following in response to a request such as “point to the second striped ball” (Matthei, 1981; Roeper, 1972). That is, children incorrectly select the object that is second and striped and a ball, instead of the second of the striped balls (d).

Hamburger and Crain (1984) suggest that this error may be the spurious result of premature interpretation of “the second...” as applying to the entire set of objects in the array. They found that children were dissuaded from this concatenative response if they were asked to handle the subsets of objects before these were placed in the array. This presumably inhibits premature execution, since it is unclear in this circumstance which (sub)set of objects the ordinal “second” modifies.

(iii) There is another locus of difficulty in translating from a source language form to target language code: condition (iii). This condition requires the order of concatenation of plans to mirror the linguistic input. Let us call any violation of this condition a sequencing problem. In addition to
compiling discontinuity, a sequencing problem arises in the example of "second striped ball." As we saw, the locus of the difficulty with this phrase is not with either the source code or the object code, but only in their relationship. This suggests a possible alternative to the merging of code in the formation of a plan. The alternative would be to hold onto the code associated with the ordinal "second" until after the remaining elements have been combined (establishing the data structure for "striped ball."). But setting aside a constituent to await the preparation of other elements with which it is to associate is assumed to be costly of memory resources. In terms of the model of working memory we are considering, this would constitute a violation of condition (iii) and, as a consequence, also condition (iv). 7

Relative clauses. A second example of the difficulties a sequencing problem may pose for comprehension is from a study on the acquisition of restrictive relative clauses. This study (Hamburger & Crain, 1982) discovered that many children who performed the correct actions associated with sentences like (1) often failed, nevertheless, to act out these events in the same way as adults.

(1) The cat scratched the dog that jumped through the hoop.

Most 3-year-olds and many 4-year-olds acted out this sentence by making the cat scratch the dog first, and then making the dog jump through the hoop. Older children and normal adults act out these events in the opposite order, the relative clause before the main clause. Intuitively, acting out the second mentioned clause first seems conceptually more correct since "the dog that jumped through the hoop" is what the cat scratched.

It is reasonable to suppose that this kind of conflict between the order of mention and conceptual order (and most appropriate order of execution) stresses working memory because both clauses must be available long enough to enable the hearer to formulate the plan which represents the conceptual order. Presumably, this difference between the responses of children and adults reflects the more severe limitations in children's working memory in coping with sentences that pose sequencing problems.

The response we have characterized as conceptually correct (with the relative clause acted out first) requires the formation of a two-slot template, and a specification of the particular sequence in which the actions are to be carried out. Since on the simple look-up-and-concatenate scenario, processing occurs on line (i.e., in a left-to-right word-by-word basis), it seems to us that the difficulty presented by the conflict between order-of-mention and conceptual order occurs because the information in both clauses must be held in memory long enough to put the first-mentioned action into the second slot. If memory is overloaded, a subject may adopt a default procedure of acting out clauses in their order of mention—that is, according to the simple translation routine of look-up-and-concatenate.

To explain this phenomenon we draw upon another analogy to translation among programming languages. Here we appeal to the distinction between compiling, which completes the translation before starting to execute, and interpreting, which interleaves translation and execution. We can use this distinction in explaining children's conceptually incorrect responses to sentences like (1). Since children are unable to hold information long enough in working memory to compile a conceptually correct plan, it makes sense to suppose that they opt instead to interpret in cases like (1). Consistent with this supposition is the observation that children often begin to act while the sentence is still being uttered.

Temporal terms. A third example of the sequencing problem arises with sentences containing the temporal terms before and after. These terms explicitly dictate the conceptual order of events, and they too may present a sequencing problem by introducing conflicts between conceptual order and order-of-mention. This is illustrated by sentence (2).

(2) Jabba flew the X-Wing fighter after Han Solo sped away in the Millennium Falcon.

A sequencing problem arises in (2) because the order in which events are mentioned is opposite to the conceptual order. Again, research in language acquisition has found that young children frequently interpret these sentences in an order-of-mention fashion (Clark, 1970; Johnson, 1975). As with relative clause sentences, it is likely that this response reflects an inability to hold both clauses in memory long enough to formulate a plan for acting them out in the correct conceptual order. Once again, children's failure to segregate translation and execution explains their incorrect, default decision to adopt the simple look-up-and-concatenate translation.

In this case, however, an alternative account of children's difficulties has been proposed. It has been argued that a structural explanation, and
not a processing explanation, is called for. Proponents of the structurally-based explanation (Amidon & Carey, 1972) point out that the same children who failed to act out sentences like (2) correctly emitted a high rate of correct responses to sentences similar in meaning, but with simpler syntactic structure, as in (3).

(3) Push the motorcycle last; push the helicopter first.

There is direct evidence that processing factors, and not lack of syntactic competence, are responsible for children's errors in comprehending sentences with temporal terms. The evidence is this: once processing demands are reduced, most 4- and 5-year-old children usually give the correct response to test sentences like (4) and (5).

(4) Push the helicopter after you push the motorcycle.

(5) Before you push the motorcycle, push the helicopter.

To minimize processing load, one must take cognizance of a presupposition on the use of temporal terms. The presupposition associated with sentences (4) and (5) is that the hearer intends to push a motorcycle. These sentences are felicitous only if the subject has already indicated an intent to play with a motorcycle prior to receiving the test sentence. To satisfy this presupposition, one simply has to ask the child in advance to select one of the toys to play with before each trial. When young children were given this contextual support, they displayed unprecedented success in comprehending sentences with the temporal terms "before" and "after" (Crain, 1982; Gorrell, Crain, & Fodor, 1986). The same finding was also obtained in a recent study of mentally retarded adults (W. Crain, 1986).

We should also mention a superficial linguistic property that forestalls premature execution, and thereby eases the burden on working memory, in sentences like (5). This is the presence of a temporal term in the initial clause, which indicates that a two-slot template is required. Notice that in the corresponding sentence with "after," the temporal conjunction appears in the second clause. The account of memory difficulties we have proposed would therefore lead us to expect this type of sentence to be harder, especially if it contains "after." This prediction is confirmed in the experiments reported in section 4 below.

In our discussion of the control component of working memory, we have assigned to it as few combinatorial duties as possible. This makes it essentially a simple-minded traffic controller for symbolic representations that are being composed within the submodules of the language apparatus. It is also apparent that the structure building operations that take place within these modules are frequently at odds with the efficient management of information flow. Much is gained, however, by having them incorporated into the language faculty, since they supply the generative capacity for producing and understanding (an uncountable number of) novel sentences.

Garden path effects. Corresponding to each intermediate level of representation is a processing mechanism, or parser. The task of each parser is to assign structure to the incoming code as it is being transmitted from the next-lower parser. This analysis phase of the compiling process was aptly referred to by Miller (1956) as chunking.

The syntactic parsing mechanism is probably the best understood of the parsers. This mechanism consists of a number of routines for accessing syntactic rules and principles and resolving ambiguities that arise when more than one analysis is compatible with the current input. We assume that access to rules during on-line processing uses hard-wired portions of the language apparatus—almost reflex like in character—that are sparing of processing capacity in most cases. However, natural languages permit massive local ambiguity and, despite the flexibility that this allows, it surely incurs some cost to memory resources.

In fact, there is considerable evidence that local ambiguities are quickly resolved, perhaps within one or two words after they arise. One parsing tendency that seems to have evolved to meet the twin exigencies of ambiguity and working memory limitations is called Right Association (see Frazier, 1978; Kimball, 1973.) Right Association explains why listeners or readers connect an incoming phrase as low as possible in the phrase marker that has been assigned to the preceding material. This 'strategy' reflects the functional architecture of the language apparatus, which has many computations to perform and little space for their compilation and execution. As a result, strategies like Right Association dictate that incoming material is integrated into the most readily available (i.e., local) node in the phrase marker under construction. So, for example, Right Association dictates that the adverb "yesterday" will be attached to the lower of the two VPs in the ambiguous sentence (6) and will therefore be interpreted as related to the last mentioned event.
(6) Bush said he apologized to the UAW, yesterday.

In keeping with Right Association, there is a strong tendency for people to interpret (6) to mean that Bush apologized yesterday, and not that he uttered a sentence to that effect yesterday.

It is reasonable to suppose that memory limitations promote rapid on-line integration of material into a structural representation. Although parsing strategies may enable the parser to circumvent the limitations of working memory, they sometimes introduce problems of their own, because the decision dictated by a strategy may turn out to be incorrect in the light of subsequent input. In this case, the perceiver is led down a garden path by the parser. The existence of ‘garden path’ effects (illustrated in (7)) shows that for some sentences even full knowledge of the grammar is not powerful enough to overcome the liability of a tightly constrained working memory.

(7) Bush said that he will apologize to the UAW, yesterday.

Recovery from garden paths is possible only within the limits of working memory, because this determines whether the grammatically correct attachment site is still available. Since sentences that tax working memory heavily have been found to present problems for poor readers, they should be less able than good readers to recover from incorrect analyses prompted by parsing strategies like Right Association. Therefore, they should be even more susceptible than good readers to garden path effects. Experiment III (reported below) tests this prediction by asking good and poor readers to respond to several types of garden path sentences.

An examination of how the test sentences were constructed may help to clarify the logic of this experiment. Suppose you are looking at a picture in which a girl (Mary) is using a crayon to draw a picture of a monkey who is drinking milk through a straw. The corresponding sentence is given in (8). What is the unspecified NP in this situation? Both “a crayon” and “a straw” are grammatically well-formed, but the analysis favored by Right Association has “with NP” modifying “drinking milk” rather than modifying “drawing a picture,” so the general preference is to cash out the NP as “a straw.”

(8) Mary is drawing a picture of a monkey that is drinking milk with NP

This parsing preference is still present if the NP in (8) is extracted by Wh- Movement, as in (9).

(9) What is Mary drawing a picture of a monkey that is drinking milk with?

The preposition “with” again coheres strongly with the relative clause, rather than with the main clause. The result is that one is tempted to make an ungrammatical analysis of (9) in which “what” has been extracted from the relative clause, violating a putative universal constraint on extraction called Subjacency. Research in language acquisition, using a picture verification task, found that many children succumb to this temptation, in an apparent violation of Subjacency, responding to (9) by saying “a straw,” rather than giving the grammatically correct response, “a crayon” (Crain & Fodor, 1985; Otsu, 1981).

This incorrect response clearly bears on the choice of the two hypotheses we are considering about the source(s) of reading disability. Since Subjacency is part of Universal Grammar, the PLH would maintain that it should be adhered to by good and poor readers alike from the earliest stages of language development. On the other hand, the processing limitations of poor readers would lead us to expect them to make more apparent violations of the Subjacency constraint. It is then incumbent on the PLH to show that the relatively poor performance of poor readers on sentences like (9) is due to parsing pressures (viz. the effect of Right Association) rather than to ignorance of universal constraints on syntax.

There are two critical ingredients in determining whether responses which violate the Subjacency constraint reflect a processing limitation or, instead, arise from a structural deficit. As noted earlier, if poor readers suffer from a processing limitation, this should be revealed in the pattern of errors across sentence types for both reader groups: poor readers should show a decrement in performance across sentence types, but there should be no group-by-sentence-type interaction. This pattern emerges from comparison of the responses of the reader groups in Experiment III. The final ingredient is a demonstration of the grammatical competence of poor readers with the construction under investigation. This is the objective of Experiment IV.

4 APPLYING THE WORKING MEMORY MODEL TO IDENTIFY THE CAUSES OF SENTENCE COMPREHENSION FAILURES IN POOR READERS

In this section we elaborate on the specific problems that should be incurred by poor readers, given our model of language processing. Four experiments are reported here. These experiments were designed to test between the two competing
hypotheses (sketched in Section 1) about the source of impaired comprehension of spoken sentences by poor readers. Specifically, we ask whether the sentence processing difficulties are due to a syntactic deficit, as claimed by the SLH, or alternatively, whether they reflect a limitation in processing involving working memory, as claimed by the PLH. To explore both possibilities, we selected good and poor readers in the second grade. Reader groups were established on the basis of combined word and non-word scores on the Decoding Skills Test (DST) of Richardson and DiBenedetto (1986). To ensure that the difficulties experienced by the poor reader group could not be attributed to a general deficiency in cognitive function, the reader groups were equated on intelligence as well as on chronological age. (For discussion of the general efficacy of this experimental design, see Shankweiler, Crain, Brady, & Macaruso, in press.)

Temporal terms (Experiments I and II)

In the first two experiments, we were interested to discover how variations in processing load affect the performance of poor readers relative to good readers. In the last section, we saw that sentences which contain temporal terms are of particular interest in deciding between the competing hypotheses because (i) temporal terms have been found to emerge late in the course of normal language development, and (ii) the source of late mastery has been attributed to syntactic complexity, as the SLH would suggest, as well as to their demands on memory resources, as the PLH would have it. In order to test between these hypotheses, Experiments I and II used a figure manipulation paradigm, with input sentences containing adverbial clauses with the temporal terms “before” and “after.” This task engages children in a game in which they are asked to move toys as dictated by orally presented sentences. The set of objects available in the experimental workspace was the same in both experiments; it comprised nine objects (cars, trucks, horses) of different colors and sizes.

The purpose of the first experiment was to establish a baseline of linguistic competence by good and poor readers with sentences containing temporal terms. In the second experiment, we sought to manipulate processing demands in two ways. First, an additional modifier was added to one of the noun phrases in half of the test sentences. This maneuver increased the possibility that subjects would make errors in selecting the objects to be moved on each trial. A second change involved presenting the test sentences in contexts that satisfied the presupposition associated with the use of the temporal term. We hypothesized that poor readers would show appreciable performance gains when processing demands were minimized through the satisfaction of this presupposition. It should be kept in mind that if the poor reader group displayed a sufficiently high level of correct performance in any condition, this would argue against the hypothesis that the relevant syntactic structures are missing from their grammars. But, in addition, an increase in successful comprehension in felicitous contexts would lend credence to a processing explanation of their performance failures in less than optimal contexts.

Each experiment was carried out with a different set of 14 good and 14 poor readers. The mean combined reading scores (on the DST) for the good and poor readers were 92.9 and 23.7 out of 120, respectively (Experiment I), and 97.2 and 37.9 (Experiment II). The IQ of subjects was calculated on the basis of their performance on the Peabody Picture Vocabulary Test—Revised (Dunn & Dunn, 1981). Performance on this test was used to ensure that both groups were in a similar IQ range, and that the differences between good and poor readers could not be attributed to different levels of vocabulary knowledge. The mean Peabody score for good and poor readers was 110.6 and 105.0, respectively (Experiment I), and 115.4 and 109.0 (Experiment II).

Experiment I

The purpose of Experiment I was to assess the level of linguistic competence for both reader groups with sentences containing temporal terms. This experiment employed simple NPs and, like many previous studies in the acquisition literature, provided no contextual support. In half of the twelve test sentences the order-of-mention of events corresponded to the conceptual order of events, as in (10). In the other half, the order in which events were mentioned was opposite to the conceptual order, as in (11).

(10) Push the red car before you push the largest horse.
(11) Push the smallest horse after you push the blue car.

First of all, we found that poor readers made significantly more errors than good readers, $F(1,27) = 4.92, p < .04$. However, the overall performance of both groups was high, with the poor reader group performing well above chance (87.5% correct). This indicates that the poor readers were not lacking the necessary competence to successfully interpret temporal term sentences even
when they contain inessential prenominal modifiers. The near-ceiling performance of the good reader group (96% correct) meant that subsequent analyses of their error patterns would not be revealing, so the remainder of our analyses focuses on the pattern of errors by the poor readers.

In particular, we were interested in determining whether or not the sentences we expected to be most demanding of memory resources do indeed cause special problems for poor readers. These sentences are the ones which present a conflict between the conceptual order and the order-of-mention and contain the temporal term after, as in (11). Poor readers' 21.4% errors on these sentences reflects their highest error percentage for any sentence type. In fact, it is a significantly higher error rate than for before sentences (7.1% errors) of the same type, $F(1,13) = 4.50, p = .05$. This confirms our expectation that sentences like (11) would be the most difficult for poor readers given their inherent memory limitations.

In Experiment II, we asked whether a high proportion of correct responses is still characteristic of poor readers in contexts that are even more demanding of working memory resources. If not, the combined data would lend support to the hypothesis that poor readers suffer from a limitation in processing. This difference across tasks would defy explanation on the hypothesis that they suffer from a developmental lag in the acquisition of complex syntax.

Experiment II

The purpose of this experiment was to test the effects of varying memory demands on good and poor readers. According to the account of the working memory system presented earlier, poor readers should be highly sensitive to alterations in processing load which give rise to problems in cognitive compiling. We sought first to exacerbate the processing load beyond the level imposed in Experiment I by including an additional prenominal modifier in half of the test sentences. As exemplified in (12) and (13), these sentences contained the ordinal term 'second,' which introduces discontinuity in related statements in the plan that one must compile in order to respond accurately to the noun phrase in which the ordinal appears. We will refer to sentences with NPs of this sort as complex NPs.

A second change in design was introduced in order to increase the ease of processing. We took advantage of a pragmatic property that is often associated with sentences containing subordinate clauses, namely, their presuppositional content, and we exploited this property to reduce the burden imposed on memory by satisfying the presupposition associated with test sentences of this type, as discussed earlier. In the revised procedure, children are asked, before each test sentence is presented, to identify one object they want to play with in the next part of the game. The experimenter subsequently incorporates this information into the subordinate clause introduced by the temporal term. For instance, sentence (12) would have been presented only after a subject had selected the blue car. This will be referred to as the felicity condition. In the other, no felicity condition, the presupposition inherent in the use of temporal terms was not satisfied; sentences were presented in the "null context," as in Experiment I. In the null context, unmet presuppositions must be "accommodated" into the listener's mental model of the discourse setting (Lewis, 1979). In order to compensate for unmet presuppositions, the subject must revise his/her current mental model by averring that the presupposition was met. Updating one's knowledge representation in this way is known to be costly of processing resources (see Crain & Steedman, 1985, and references therein). In light of these considerations, the PLH anticipates a high rate of successful comprehension by both reader groups in the felicity condition, but it predicts that poor readers' performance will suffer in contexts that are more taxing of working memory, as in the no felicity condition.

The stimuli in Experiment II consisted of 16 sentences with temporal terms before and after. In contrast to Experiment I, only four sentences were presented in which the order-of-mention of events was the same as their conceptual order, as in (12). In the remaining twelve sentences, order of mention was opposite to the conceptual order, as in (13). All children encountered the test sentences in both contexts, i.e., in the felicity and no felicity conditions. This required two testing sessions for each child, with half of the children receiving contextual support in the first session, and half in the second session.

Overall analyses of the results reveal main effects of reader group ($F(1,26) = 14.16, p < .001$), felicity ($F(1,26) = 6.50, p < .02$), and NP complexity ($F(1,26) = 6.13, p < .02$). In addition, there is a marginally significant NP complexity × reader...
group interaction ($F(1,26) = 3.92, p=.06$) and a trend toward a felicity × reader group interaction ($F(1,26) = 2.89, p=.10$). The main effect of reader group tells us that poor readers performed less well than good readers. However, the main effect of felicity indicates that satisfying the felicity conditions, i.e., reducing the processing demands created by conflicts in sequencing, produced a significant reduction in errors for both groups. The marginally significant felicity × reader group interaction suggests that the satisfaction of presuppositions increased performance for poor readers to a greater extent than for good readers. As displayed in Figure 1, there is a greater disparity between their performance for no felicity than for felicity. This lends credence to the hypothesis that, without contextual support, poor readers' limitations in working memory are exacerbated.

The fact that poor readers perform at a success rate of 82.4% when the felicity conditions were satisfied, even when half of the test sentences contained complex NPs, calls into question the claim of the SLH that poor readers lag in their mastery of complex syntactic structures.

Averaged over the felicity and no felicity conditions, the main effect of NP complexity tells us that complex NPs evoking significantly more errors than simple NPs. However, the marginally significant NP complexity × reader group interaction (see Figure 2) indicates that poor readers were more adversely affected by changes in NP complexity than good readers. The special difficulties that the poor readers displayed with the sentences containing complex NPs presumably reflect the fact that these sentences are more taxing on working memory resources.

*Figure 1. Percentage of incorrect responses to temporal term sentences (Experiment II).*

*Figure 2. Percentage of incorrect responses to simple and complex NP sentences (Experiment II).*
As predicted, the worst case for poor readers was the after sentences that pose a conflict between order-of-mention and conceptual order and no contextual support in the form of satisfied presuppositions. Good readers, on the other hand, were not handicapped by the memory demands this situation places on the subject.

A glance at the type of errors made by poor readers in the most difficult situation indicates almost as many responses in which they identified the wrong object (19%) as responses in which they acted out the clauses in the wrong order (25%). In a few instances both error types occurred in processing the same sentence. However, in the corresponding situation in Experiment I (i.e., in response to after sentences when order-of-mention conflicts with conceptual order), poor readers made relatively few wrong object responses (7%) compared to wrong order responses (14%). Most likely, this difference in the relative frequency of error types is due to the manipulation of NP complexity across the experiments. Poor readers were more likely to choose a wrong object in sentences containing complex NPs in Experiment II, since choosing an object denoted by a complex NP requires overcoming the additional problems of compiling discontinuity.

Taken together, the findings of Experiments I and II indicate that as processing demands are increased, poor readers’ performance on an object manipulation task involving temporal terms sentences is eroded much more than good readers’ performance. Decreasing the processing demands, either by satisfying the felicity conditions or by using less complex NPs, elevates performance by poor readers, such that group differences diminish, and all subjects perform at a high level of success. Results showing improved performance by poor readers when processing demands are lowered cannot be explained on the hypothesis that poor readers are lacking the relevant syntactic structures. Only an account in terms of processing limitations can explain these findings. We hypothesize, based on the results of Experiments I and II, that if we conducted an additional experiment which presented only simple NPs (as in Experiment I), but with the felicity conditions satisfied, we would find near-perfect performance by both groups.

Garden path effects (Experiments III and IV).

In Experiment III we examined the responses of good and poor readers to the kind of garden path sentences discussed in section 2. As we noted, an incorrect response to these sentences can be explained in two ways, corresponding to the two
hypotheses we have been considering about the source of poor readers' problems in sentence comprehension. First, errors could reflect the absence in a subject's grammar of a structural constraint on extraction, Subjacency. Alternatively, they could be the result of an inability to overcome the effect of the parsing strategy Right Association, presumably due to limited working memory resources.

Three types of garden-path sentences were presented, in order to vary the processing load placed on working memory. As indicated earlier, varying sentence types puts us in a position to examine the error pattern across sentence types to help distinguish between the PLH and the SLH. The three types of syntactic constructions are illustrated in (14)-(16). There were four sentences of each type (adapted from Crain & Fodor, 1985).

(14) Relative Clause: Who is Bill pushing the cat that is singing to?
(15) Prepositional Phrase (deep): What is Jennifer drawing a picture of a boy with?
(16) Prepositional Phrase (distant): Who is Susan handing over the big heart-shaped card to?

Sentence (15) is labeled deep because the origin of the 'extracted' Wh-phrase is a prepositional phrase that is embedded in an NP which is itself embedded in an NP. This contrasts with the distant case (16), in which there is only one level of embedding. Although the sentences are matched for length, we anticipated that distant PP sentences would be easier to process than either the deep PP sentences or the relative clause sentences. The depth of syntactic embedding in both relative clause and deep PP sentences means that they deviate more than the distant PP sentences from the simple look-up-and-concatenation translation process presented in the last section.

A set of 44 good readers and 46 poor readers (which includes all of subjects of Experiment I and II) participated in this study. The mean combined word and non-word reading scores on the DST for the good and poor readers were 96.3 and 37.5, respectively. On each trial subjects were asked to listen carefully to a tape-recorded set of sentences which described a scene depicted in a large cartoon drawing placed in front of them. Immediately following the description, they were asked to respond to a question about some aspect of the drawing. As an example, the context sentences in (17) preceded the test question (14).

(17) Bill's father is waiting for Bill to bring him the cat. The cat loves to sing and has made up a song for his toy mouse.

The grammatically correct response to this question is "his father." The response "the mouse" is incorrect, since it represents an apparent violation of Subjacency. The PLH would argue that an examination of the pattern of errors across sentence types for each group may provide evidence that these errors are not in fact violations of Subjacency at all, but are the result of the processing burdening these sentence impose on working memory. It is difficult to say exactly what the SLH would predict about the pattern of responses by good and poor readers for any of the sentence types presented in this experiment. It seems reasonable to suppose, however, that the SLH might anticipate that the reader groups would display different profiles, since these sentences are exceedingly complex. To reiterate, the PLH predicts that both groups will manifest a similar pattern of errors across sentences of varying syntactic types, with poor readers penalized to a greater degree than good readers on sentences that are costly of working memory resources (e.g., the PP 'deep' and relative clause sentences).

This is exactly what was found. Analysis of the percentage of incorrect responses made by both groups reveals main effects of sentence type \((F(2,176) = 21.53, p<.001)\) and reader group \((F(1,88) = 8.95, p<.004)\), but no interaction of sentence type and reader group. Figure 4 shows that the effect of sentence type is due mainly to the higher percentage of errors for relative clause (30.5% errors) and PP deep (32.4% errors) sentences than for PP distant (15.8% errors) sentences, as anticipated by the model of working memory we presented. The reader group difference reflects the fact that good readers (21.4% errors) performed significantly better than poor readers (31.0% errors).

The absence of interaction means that poor readers show a general decrement in performance, but responded in the same way as good readers to the three different types of garden path sentences. As may be seen in Figure 4, there are no constructions which provide disproportionately greater difficulty for poor readers. This invites the inference that errors for both reader groups should receive the same interpretation. Since good readers exhibit a high proportion of correct responses on these sentences, it is reasonable to conclude that their errors are ones of performance and do not reflect an underlying deficiency in
syntactic knowledge. The significant reader group difference across sentence types might, at first glance, suggest that poor readers are lacking this knowledge, were it not for the absence of a group by sentence type interaction.13

Experiment IV

Following the proposal presented in the previous discussion, this experiment used the relative clause as a further testing ground between the hypotheses. To this end, we sought to establish that the reader group differences in response to relative clause sentences in Experiment III do not reflect lack of grammatical competence on the part of poor readers, by showing that poor readers are able to successfully comprehend relative clauses in contexts that ease memory load. Ideally, we would like to provide evidence that the apparent violations of Subjacency by poor readers disappear when the demands on memory are reduced. However, reducing processing load for sentences which could induce Subjacency violations is nearly impossible. Therefore we chose an alternative approach: to assess the ability of poor readers to comprehend the most complex substructure contained in the sentences which evoked the greatest number of errors, i.e., the ones containing relative clauses. Thus, this experiment was designed to assess the competence of poor readers in interpreting relative clauses with memory demands held to a minimum. Positive results in this case would lend further support in favor of the joint claim that poor readers possess the universal constraint of Subjacency but failed to perform as well as the good readers in Experiment III because of their deficient processing capabilities.14

The relative clause has been a focus of research in normal language acquisition as well as in the literature on reading disability. Relative clauses are found to evoke difficulties in interpretation for preschool children (Tavakolian, 1981), and for older children who are poor readers (Byrne, 1981; Stein et al., 1984). Early research in both areas led some researchers to the conclusion that children's poor performance was due to a lack of syntactic knowledge. However, Mann et al. (1984) tested good and poor readers' comprehension of relative clauses using an act-out task and found that, although good readers performed significantly better than poor readers overall, both groups were affected in the same way by the difficulty of the type of relative clause. This familiar error pattern of good and poor readers is further evidence that they differ in processing capabilities, rather than in structural competence, as we have seen.

Further support for the view that poor readers' difficulties with relative clauses reflect performance factors comes from an additional study of
good and poor readers' comprehension of relative clauses by Smith, Macaruso, Shankweiler, and Crain (1989). Adapting several experimental innovations from the literature on language acquisition, Smith found that poor readers made few errors when the pragmatic presuppositions on the use of relative clauses were satisfied (Hamburger & Crain, 1982). Smith et al. compared their findings with those of the Mann et al. study, in which the same subject selection criteria were used, but in which the presuppositions of relative clauses were unsatisfied. Taken together, the data from these studies revealed that the changes in methodology had the effect of eliminating reader group differences, with the result that both good and poor readers performed at a high level of success.

The present study employed an act-out task which incorporated some of the methodological innovations used by Smith et al. and by Hamburger and Crain, in order to assess the grammatical competence of a subset of the good and poor readers who participated in Experiment III. The same set of 14 good and 14 poor readers from Experiment I participated in the present experiment. Three types of relative clause constructions were used.15 Examples are provided in (18)-(20).

(18) SO The lion that the bear bit jumped over the fence.
(19) OO The boy touched the girl who the ice cream fell on.
(20) OS The lady hugged the man who picked up the suitcase.

In order to reduce the processing burden imposed by sentences such as (18)-(20), we satisfied one of the presuppositions of relative clauses. Specifically, we incorporated two objects in the experimental workspace corresponding to the head noun of the relative clause. For example, in (19), there were two figurines from which the subject could choose to act out the sentence. By including the extra figurine, we satisfied the requirement of a restrictive relative clause to restrict, in the example, the set of girls to the one on which the ice cream fell.

As in the Smith et al. study, we found no significant group differences and a low error rate for all subjects (good readers: 8.7% errors; poor readers: 11.1% errors). In addition, the pattern of errors across sentence types was virtually identical for both groups. This is shown in Figure 5. The main finding was that poor readers acted out relative clause sentences with an 89% success rate when the appropriate presuppositions were met. In fact, for two of the sentence types, poor readers made only 7% errors. This provides support for the contention that a syntactic deficit was not the cause of the inferior performance of poor readers in Experiment III. The results of Experiment IV strongly suggest that poor readers are able to comprehend various relative clause constructions. These results invite the inference that the poor readers' higher error rate in the context provided in the earlier study by Mann et al. was a consequence of their abnormal limitations in working memory.

Figure 5. Percentage of incorrect responses to relative clause sentences (Experiment IV).
5 CONCLUSION

The purpose of the present research was to develop an approach to language comprehension problems that is sufficiently detailed to allow specific predictions and powerful enough to identify causes. The first half of the paper lays the groundwork. It presents the major assumptions about the language apparatus that underlie our research. Working within the framework of a modular view of language, we gave an account of verbal working memory that emphasizes its control functions more than its storage functions. We proposed that the task of the control component of working memory, in extracting meaning from linguistic forms, is to regulate the flow of linguistic material through the system of parsers, from lower- to higher-levels of representation. The inherent limitation in buffer capacity entails that higher-level processing must be executed within very short stretches of text or discourse. To deal with this problem, the control component of working memory must rapidly transfer partial results of linguistic analyses between levels of structural representation, i.e., the phonology, the lexicon, the syntax, and the semantics.

A further goal was to show how this model could be successfully applied to get to the root of difficulties that children who are poor readers often display in processing spoken sentences. Alternative hypotheses about the causes of poor readers’ sentence comprehension problems were posed, and the conceptual machinery for testing between these hypotheses was introduced. The hypotheses make different predictions because they locate the source of reading difficulties in different components of the language apparatus. Roughly, the views turn on the distinction between structure and process. On the SLH, poor readers suffer from a structural deficit, i.e., a deficit in the stored mental representation of principles of syntax, in addition to their (unrelated) deficiencies in phonological processing, verbal working memory, and so on. On the PLH, each of the deficits of poor readers is a reflection of their limitations in processing phonological information.

To test between these hypotheses, we reviewed the results of four interlocking experiments. A pattern of findings emerged that indicates that the necessary syntactic structures were in place, and that the source of the poor readers’ difficulties in comprehension of spoken sentences stemmed from inefficiencies in on-line processing of sentences that for one reason or another stressed working memory. Thus, inefficiency of verbal working memory and not failure to acquire critical language structures was identified as the factor responsible for the comprehension difficulties. We argued that, ultimately, the problems of poor readers originate at the phonological level, and that difficulties that might appear to reflect a syntactic deficiency are, in reality, manifestations of a special limitation in accessing and processing phonological structures.

REFERENCES


CriticIU approtlChes

children.

Oxford University


of short-term

P., Crain, S.,

langUllge acquisition

so-called


Cohen,

&


&

Cambridge, MA: MIT Press.

Liberman, A M. (in press). Specialized


Psychologiall, comp"tatimull, and theoretical

11, 101-136.

of

&

of

compassion of syntax by


Claremont Graduate School.


working memory and reading. Journal of Verbal Learning and

Behavior, 19, 450-466.


dyslexic and other learning disabled children. Brain and


de Villiers, J. G., Tager-Flusberg, H. B. T. Hakuta, K., & Cohen,


of Psycholinguistic Research, 8, 499-518.


changes in the linguistic performance correlates of reading


Press.


strategies. Unpublished doctoral dissertation, University of

Connecticut.

Goldsmith, S. (1980). The psycholinguistic bases of reading disability:


City University of New York.

Greer, P., Crain, S., & Fodor, J. D. (1986). Contextual information

and temporal terms. Paper presented at the Boston University

Conference on Language Development, Boston, MA.


Kuczaj, II (Ed), Language development, Volume 1: Syntax and


Johnson, M. L. (1975). The meaning of before and after for pre

school children. Journal of Experimental Child Psychology, 19, 88-

99.


reading disability: Evidence from an object-naming task.

Cognition, 22, 225-257.


Cognition, 2, 15-47.


Philosophical Logic, 8, 339-359.


Language codes and memory codes. In A. W. Melton & E.

Martin (Eds.), Coding processes and human memory. Washington,

DC: Winston and Sons.


problem of learning to read and write. Remedial and Special

Education, 6, 8-17.


the beginning reader. In A. S. Reber & D. L. Scarborough

(Eds.), Toward a psychology of reading: The Proceedings of the

CUNY Conferences. Hillsdale, NJ: Lawrence Erlbaum

Associates.


Sensitivity to grammatical structure in so-called agrammatic


aphasia. Unpublished doctoral dissertation, University of

Connecticut, Storrs.

Lundberg, I., Oflofson, A., & Wall, S. (1980). Reading and spelling

skills in the first school years predicted from phonemic

awareness skills in kindergarten. Scandinavian Journal of

Psychology, 21, 159-173.


memory for sentences and word strings in relation to reading

ability. Memory and Cognition, 8, 329-335.


between comprehension of spoken sentences and early reading

ability: The role of phonetic representation. Journal of Child

Language, 11, 627-643.


Dissociations of language in aphasia: Implications for normal

functions. Annals of the New York Academy of Sciences, 280, 866-

884.

Marlens-Wilson, W. D., & Tyler, L. K. (1980). The temporal

structure of spoken language understanding: The perception of


and sentence comprehension deficits in agrammatic and

conduction aphasics. Paper presented at the Annual Meeting of

the Academy of Aphasia, Pittsburg, PA.

Martin, R. C. (1990). Neuropsychological evidence on the role of

short-term memory in sentence processing. In G. Vallar & T.

Shallice (Eds.), Neuropsychological impairments of short-term


containing reciprocals. In S. L. Tavakolian (Ed.), Language

acquisition and linguistic theory. Cambridge, MA: MIT Press.


perceiving systems for speech and other biologically significant

sounds. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.),

Functions of the auditory system. New York: Wiley.

Miller, G. A. (1956). The magical number seven, plus or minus

two: Some limitations of our capacity for processing

information. Psychological Review, 60, 81-97.


abilities of dyslexics and normal readers. Perceptual and Motor

Skills, 58, 221-222.


Development of phonetic memory in disabled and normal


toddlers: Toward a theory of syntactic development. Unpublished

doctoral dissertation, MIT, Cambridge, MA.


Press.


and sources of individual differences. In M. A. Just & P. A.

Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale,

NJ: Lawrence Erlbaum Associates.

Richardson, E., & DiBenedetto, B. (1986). Decoding skills test. Parkton,

MD: York Press.

Roeppeer, T. W. (1972). Approaches to a theory of language acquisition

with examples from German children. Unpublished doctoral

dissertation, Harvard University, Cambridge, MA.


FOOTNOTES


1Also University of Connecticut, Storrs.

2For instance, poor readers have been found to be less accurate than their age-matched controls in understanding the infinitival complements of object-control adjectives like “easy” (Byrne, 1981), which Chomsky (1969) found problematic for children as old as nine. Relative clauses also pose difficulties for poor readers in some contexts (Byrne, 1981; Goldsmith, 1980; Mann, Shankweiler, & Smith, 1984; Stein et al., 1984; but also see Smith, 1987); the relative clause is a grammatical structure which many researchers believe to develop late (Brown, 1973; de Villiers, Tager-Flusberg, Hakuta and Cohen, 1979; Sheldon, 1974; Tavakolian, 1981; but see Hamburger and Crain, 1982). Poor readers have also been shown to encounter difficulty with some dative constructions (Fletcher, et al., 1981) that have also been found to evoke comprehension errors in young children (Scholes, 1978).

3Further discussion of the implications of this hypothesis concerning the causes of reading disorder is presented in Crain and Shankweiler (1987), Shankweiler and Crain (1986), Shankweiler et al. (in press).

4This distinguishes our conception of verbal working memory from proposals by Baddeley (1986) and Carpenter and Just (1988). These researchers see working memory as a general purpose device which plays a central role not only in language, but also in reasoning, problem solving and in other forms of complex thinking.

5As we conceive of it, a plan is a mental representation used to guide action. The formation of a cognitive plan is an important aspect of any comprehension task involving the manipulation of objects, since the complexity of a plan is a potential source of difficulty in sentence comprehension. Therefore, when children perform poorly in language comprehension tasks that involve planning, it is important to consider the possibility that formulating the relevant plan is the locus of the problem, rather than other purely linguistic aspects of the task, such as their imperfect knowledge of linguistic rules. Where our objective is the assessment of the extent of children’s linguistic knowledge (i.e., competence), we must develop our understanding of the nature and relative complexity of plans (whereby that competence is demonstrated), and we must also devise experimental paradigms in which the impact of plan complexity on linguistic processing is minimized (see Section 4 below). In addition to plan complexity, the compiling analogy allows us to entertain several other types of difficulty that a sentence can present to a subject whose task it is to plan and execute an appropriate response.

6Although the issue is too thorny to take up here, it may be useful to mention just one of many apparent counterexamples to the isomorphism between syntax and semantics. The example involves the phrase “three consecutive rainy days.” The preferred, but noncompositional interpretation of this phrase means “three consecutive days on which it rained.” The compositional interpretation requires a composite set of ordered rainy days, perhaps taken from a set of weeks, each containing a rainy day or two. So if it had rained on three consecutive Thursdays, Donald might have forgotten his umbrella on three consecutive rainy days. Since the compositional interpretation requires a more complex set of presuppositions, the default is to interpret this phrase in a way that violates compositionality. As we noted, it is reasonable to suppose that this could add to processing difficulty.

7There are other circumstances in which noncontiguous elements must be combined, in violation of condition (ii). This problem arises whenever there are discontinuous dependencies in the source language. In English, for example, the problem of inverse compiling discontinuity occurs in verb-particle constructions such as look up, and in cases of verb subcategorization (e.g., the verb put takes both an NP and a PP complement). Clearly, these properties of the linguistic input...
are seen to impose demands on memory, by postponing the closure of syntactic categories.

7It is worth noting that changes in word order could circumvent the sequencing problem, by permitting a nearly optimal translation between source and target language. Consider a language in which nominal modifiers appear after the noun, as in "...bear brown second." In this language, fragments of the target plan would be composed in the order in which they were mentioned in the phrase. However, it should be pointed out that in many situations, the construction of plans which satisfy condition (ii) may still require a commitment of memory resources. Specifically, the memory system must retain a data structure corresponding to the semantic value of each element of the phrase. For example, interpreting the noun "bear" as part of the phrase "bear brown second" requires the formation of a data structure delimiting the set of bears in the discourse setting. This data structure must somehow be stored, awaiting the formation of the semantic value for the remainder of the phrase. This burden on memory could be eliminated in certain contexts, such as situations in which the objects denoted by the noun are aligned in a way that allows one to restrict attention to that set alone. This would occur, for instance, if all of the bears were in the same field of view. A simple shift of gaze would allow the parser to restrict attention to the remainder of the phrase.

8At present, we cannot say whether unmet presuppositions thwart cognitive compiling and thereby result in premature execution or whether, instead, the satisfaction of presuppositions facilitates compile-mode behavior. In any event, it is clear that children's mistaken responses in situations that flout presuppositions exonerates the syntactic component and implicates limitations in working memory capacity.

9An example may be helpful in illustrating this point. It is proposed in Hamburger and Crain (1987) that certain principles of compositional semantics restrict the set of word-order possibilities in natural language. In determining semantic constituency relations, the following principle is proposed:

(P) An operator must be composed with appropriate operands.

To apply this principle, let us consider two categories of modifiers, selectors and restrictors. Modifiers like "striped" are restrictors, since they perform the semantic operation of acting on a set to produce a set as output. A selector also acts upon a set, but it returns an element as output. For example, in the phrase "second striped ball" the value of applying "second" to the set designated by "striped ball" is an individual ball. Principle (P) allows a restrictor to apply before a selector, but not after one, since the output produced by a selector is not appropriate input to a restrictor. This explains the illicit character of "striped second ball," even though this alternative phrasal order might be preferred from the standpoint of the working memory system.

10It is important to note that even the test sentences used in Experiment I are not the least complex case of the relevant construction, since the noun phrases in both clauses contained an inessential prenominal modifier. We chose to introduce this slight additional complexity in order to avoid inducing ceiling performance by subjects so as to make reader group differences more pronounced.

11In evaluating the results of this experiment, two points should be kept in mind. First, in the temporal terms experiments, there were two sources of errors: ordering errors and object selection errors. For example, in (11), an ordering error would consist of pushing the smallest horse first, whereas an object selection error might result in picking up the middle-sized horse. Unless otherwise indicated, errors of both types were combined in the analyses. Second, all of the ANOVAs we report treated subjects, and not items, as a random factor. Therefore, the findings cannot be generalized to items outside of these experiments.

12It should be noted that the obvious control sentences—those sentences which express agreement between order-of-mention and conceptual order—are not adequate for our purposes.

13It should be noted that a processing limitation account would be compatible with one kind of interaction, in which reader group differences increase to a greater extent on sentences that are particularly costly of working memory resources, such as relative clauses and 'deep' PPs. There is a hint of this type of interaction in Figure 4, and this interaction has been obtained with other populations who exhibit more severe limitations in memory capacity (e.g., Lukatela, 1989).

14It is worth noting that even normal adults respond incorrectly to garden path sentences in apparent violation of Subjacency when demands on working memory are raised substantially (Crain and Fodor, 1985). This supports our conclusion that the problems poor readers have with these same sentences are also due to processing factors.

15The two-letter code indicates how these sentence types differ. The first letter represents the noun phrase in the matrix sentence (Subject or Object) that bears the relative clause; the second letter represents the 'empty' noun phrase position in the relative clause. For example, an OS relative like (20) is one in which the Object of the main clause is modified by a relative clause with a superficially null Subject.