652

Grammaticality Judgments and Reading Skill in Grade 2

Anne E. Fowler

Reprint

ANNALS of DYSLEXIA

VOLUME XXXVIII 1988

THE REPRINT SERIES
The Orton Dyslexia Society
724 York Road
Baltimore, MD 21204

Grammaticality Judgments and Reading Skill in Grade 2

Anne E. Fowler

Haskins Laboratories
New Haven, Connecticut

Do the language problems associated with reading disability extend to sentence-level knowledge? If so, how should these problems be characterized? In this study, second graders were presented with two different oral syntactic tasks, equivalent in grammatical complexity, but varying in the actual task demands placed upon the reader. In the Judgment *task, subjects were asked to distinguish between grammatical and ungrammatical sentences; in the* Correction *task, subjects were asked to remedy violations of grammaticality. Performance on these sentence-level tasks was, in turn, examined in relation to reading scores, memory span, and metaphonological skill. The Judgment task revealed a wide range of ability among second-graders, and scores were systematically affected by the syntactic nature of the violation. However, performance on Judgment was not associated with reading ability or metaphonological skill, and scores were minimally affected by short-term memory factors. Mean scores on the Correction task were comparable to those achieved in the Judgment task and were also systematically affected by syntactic factors. However, performance on Correction was strongly associated with reading ability and metaphonological skill, and was greatly affected by short-term memory factors. These results suggest that reading disability does not reflect problems in basic syntactic knowledge. Instead, sentence-level problems that have been observed in less-skilled readers may be caused by the nonlinguistic processing demands of the syntactic measures typically employed.*

This research was supported in full by an NIH postdoctoral fellowship (#1-F32-HD-06543). This research benefited greatly from the insights of colleagues at Haskins Laboratories: Isabelle Liberman, Donald Shankweiler, Stephen Crain, Leonard Katz, and Virginia Mann. I am especially grateful to I. Liberman and D. Shankweiler for their careful comments on an earlier draft of this paper. Many thanks are extended to Virginia Fowler, who assisted with the creation of the tape; to Kraig Dobensky, who constructed the apparatus; and to Robin Story, who aided in data collection. Finally, and most importantly, I thank the children and staff of St. Francis and St. Bernadette Schools who were so generous with their time, patience, and enthusiasm.

Annals of Dyslexia, Vol. 38, 1988.
Copyright © 1988 by The Orton Dyslexia Society
ISSN 0474-7534

Background

It is now clear that children who have difficulties learning to read often fail to perform well on some other verbal tasks that do not involve reading (for reviews, see Liberman and Shankweiler 1985; Mann 1984; Shankweiler and Crain 1986b; Stanovich 1982; Vellutino 1979). Poor readers are especially apt to perform at a disadvantage on spoken language tasks when the task involves explicit and conscious manipulation of the phonological elements of language. Recent data suggest that problems with reading and phonological analysis may reflect still more basic, more broadly pervasive phonological deficits evident in aspects of speech perception and production (Brady, Shankweiler, and Mann 1983); word-naming (Denckla and Rudel 1976; Katz 1986); or verbal short-term memory (Mann, Liberman, and Shankweiler 1980; Torgesen and Houck 1980; see Brady 1986 and Jorm 1983 for reviews). Such basic language problems of poor readers are subtle and may not show up in everyday communication. Typically, sophisticated and challenging measures are required to reveal them. For example, Brady, Shankweiler, and Mann (1983) found that although the performance of poor readers under normal listening conditions was equivalent to that of good readers when the task required them to identify words heard on tape, their performance was hindered by the addition of white noise much more than was the performance of the good readers. Clearly, this was not a function of impaired hearing or of difficulties with understanding the task, since the poor readers were as good as, or even better than, the good readers in identifying nonspeech environmental sounds (motors, mooing, etc.) when they, too, were masked in noise.

In most of the studies mentioned above, the word and its segments were the focus of investigation. In this paper, the focus is on whether poor readers also have basic underlying language difficulties at the sentence level. In fact, just this possibility has been suggested by a number of studies conducted over the last 15 years. When compared to good readers of similar age and grade level, poor readers are reported to *produce* sentences of lesser complexity in free speech samples (de Hirsh, Jansky, and Langford 1966; Fry, Johnson, and Muehl 1970; Vogel 1974) and to be less able to *reproduce* sentences that are long and grammatically complex (Mann, Liberman, and Shankweiler 1980; Mann, Shankweiler, and Smith 1984; Perfetti and Goldman 1976; Ryan and Ledger 1984; Weinstein and Rabinovich 1971; Wiig and Roach 1975). Several other studies find that poor readers make more errors than good readers in tasks requiring the *comprehension* of complex sentence structures that normally develop late in childhood, even when these are presented orally (Byrne 1981; Fletcher, Satz, and Scholes 1981; Mann, Shankweiler, and Smith 1984; Smith, Mann, and Shankweiler 1986; Stein, Cairns, and Zurif 1984; Whitehouse 1983). Finally, a large literature finding poor readers to be deficient in tasks requiring syntactic *awareness* has sometimes been interpreted as an

indication that poor readers may lack certain syntactic structures in their internal grammars. For example, poor readers typically perform poorly on grammatical cloze tasks, in which they must select the appropriate word or morpheme to make incomplete sentences grammatical (Fletcher, Satz, and Scholes 1981; Rubin and Liberman 1983; Wiig, Semel, and Crouse 1973; Vogel 1974). They have a more difficult time than better readers in appreciating puns and jokes turning on ambiguous sentence elements (Hirsh-Pasek, Gleitman, and Gleitman 1978). And they are often found to lag behind good readers when asked to detect, correct, or explain semantically and/or syntactically anomalous sentences (Bohannon, Warren-Leubecker, and Hepler 1984; Bowey 1986; Flood and Menyuk 1983; Forrest-Pressley 1983; Scholl and Ryan 1980; see Ryan and Ledger 1984 for an extensive review).

The foregoing results are certainly suggestive of some problem at the level of syntax, but it is not evident exactly what the problem is. Getting to the root of the problem is made difficult by two confounding variables in the experimental methodology. The first has to do with the highly unnatural demands of the particular tasks employed. As noted above, it is well established that children who have—or are destined to have—difficulties learning to read will also perform poorly on a variety of tasks that tap explicit awareness of the phonological structure of language. The strong link between reading ability and phonological analysis skills is consistent with the theoretically motivated view that poor readers are lacking in metalinguistic awareness, the ability to treat language structures as an object of conscious reflection (Gleitman and Rozin 1977; Liberman 1971; Mattingly 1972). Because the limitations in metalinguistic awareness of poor readers may also extend to levels beyond the phonology, then it is not clear whether failure on sentence-level tasks derives from fundamental deficits in acquiring and processing language structures, or whether such failure is a direct reflection of an inability to access and manipulate available structures in playing metalinguistic games.

The second potential source of confounding concerns the growing evidence that poor readers have an impoverished verbal short-term memory. Whether the material to be remembered is presented in oral or written form, poor readers are at a disadvantage as compared to good readers when tested for recall of sentences, connected text, or lists of digits, words, or other phonetically recodable material (for reviews, see Brady 1986 and Jorm 1983). Although all measures of short-term memory do not consistently show a high relationship to reading disability (Perfetti and Goldman 1976; Torgesen and Houck 1980), differences between good and poor readers in memory tasks are especially pronounced at younger ages and when language material is the basis for comparison (Jorm 1983).

The importance of short-term memory problems in sentence-level tasks is potentially great; almost any response requires the child to hold the entire sentence in memory, derive some meaning from it, and typically to perform some more complex function. In this connection, Shankweiler and Crain (1986b; see also Crain and Shankweiler 1988) have pro-

posed that special limitations in short-term memory, which may in turn reflect more basic phonological problems, may have repercussions throughout the language system, masquerading as a variety of apparently distinct deficits. Limitations in short-term memory may prove particularly deleterious in studies of sentence repetition or comprehension, where increased "syntactic" difficulty is typically achieved, not simply by introducing more complex structures, but by simultaneously increasing the length of the sentences and/or the processing requirements involved. Tasks which require children to hold a sentence in mind, while they select from among a number of deliberately distracting props or pictures, may place additional demands on short-term memory. Indeed, it is not easy to devise a language task challenging to grade-school children which does not require them to hold many items in mind or to manipulate them in some way that does not stress short-term memory. The view that short-term memory problems may mask basic language competence has received some support from recent studies conducted by investigators at Haskins Laboratories (Crain and Shankweiler 1988; Shankweiler and Crain 1986a; Shankweiler and Crain 1986b; Smith 1987). By reducing extraneous processing requirements and presenting more plausible contexts, they have found that poor readers possess more comprehension skill than heretofore recognized, even with the very sentence structures found to distinguish reading groups in previous research.

To summarize, studies of sentence-level skill of poor readers typically confound syntactic complexity with demands on short-term memory and/or metalinguistic insight. Therefore, it is possible that poor readers' apparently basic and specific syntactic deficit might be better understood in terms of poor general processing skills. Thus, despite their poor showing in studies of syntactic skill, the possibility remains that given more time, fewer items, or a more straightforward task, poor readers may turn out to possess knowledge that standard tasks fail to reveal—even as syntactic complexity increases. A lack of reading level differences on well-designed syntax tasks would better delimit the poor reader's underlying problem either to the phonological domain, where some evidence for basic problems exists, and/or to general processing limitations, without also implicating a syntactic deficit. If, on the contrary, unambiguous evidence for syntactic deficits were obtained, that would suggest that reading difficulties may stem from basic language problems, rather than from the specific and somewhat unnatural demands of reading itself.

Purpose

In the present research, a measure of syntactic competence was sought that would minimize metalinguistic and short-term demands, yet at the same time allow for presentation of syntactically complex materials. To this end, a grammaticality judgment task was selected which would

assess sensitivity to violations of grammaticality as a function of differences in syntactic structure and of varying degrees of stress on short-term memory. The procedure adopted was designed to engage the interest of second graders and was presented in a simple and straightforward fashion, with the hope that the obtained variance would arise from the syntactic demands of the task, instead of from confounding factors.

This task was chosen, in part, because of the implications of some recent work on adult asphasics who have sentence processing deficits as a result of stroke. Such individuals have a drastically reduced short-term memory on all standard measures, as well as impairments both in production and comprehension of language, showing special difficulty with function words such as auxiliaries, or prepositions. Despite these limitations, "agrammatic" aphasics can make grammaticality judgments. That is, they can listen to sentences recorded on tape and indicate which ones have grammatical errors (such as "the girl was angry her brother") at a level far above chance (Linebarger, Schwartz, and Saffran 1983; see also Crain, Shankweiler, and Tuller 1984). The explanation proposed to account for these findings is that knowledge of language structures is intact in agrammatic aphasia, but is inaccessible for ordinary language functions as a result of processing difficulties.

Of course, making grammaticality judgments is not a part of ordinary communication, and, in principle, success is not guaranteed by basic knowledge in syntax. Indeed, the procedure has standardly been used as a measure of *meta*-linguistic knowledge in young children (Bohannon, Warren-Leubecker, and Hepler 1984; Flood and Menyuk 1983; Gleitman, Gleitman, and Shipley 1972; Ryan and Ledger 1984). However, developmental research suggests that the task is within the grasp of many 7-year-olds, if certain precautionary measures are observed (e.g. Gleitman, Gleitman, and Shipley 1972; Scholl and Ryan 1980). The study presented here was designed to remedy two factors that have confounded interpretation of earlier studies using sentence judgment. First, many studies that have found substantial difficulties in the early school years allow for confusion between semantic anomaly and syntactic ill-formedness (e.g. Gleitman, Gleitman, and Shipley 1972; Hakes 1980). Children who are presented with some sentences that are semantically anomalous or downright nonsensical (e.g. "the string chased the kitten" or "the chased string kitten the") are apt to misinterpret the task; they will look for semantic sense and hence fail to detect those items which are merely ungrammatical ("the kittens chases the string"). To avoid such confusion, the sentences in the present study were designed to be semantically transparent and plausible, thus focusing attention on grammatical violations. That is, sentences were excluded that violate real-world semantic sense (e.g. "The sleeping rock was in the middle of the road").

In some earlier studies of children's ability to make sentence judgments, part of the subjects' task has been to figure out the nature of the game. They are required to infer that what distinguishes one set of sen-

tences from another is the parameter of well-formedness. Indeed, good and poor readers appear to differ in their ability to make just this inference (Bohannon, Warren-Leubecker, and Hepler 1984). In the present study, children were provided with explicit instructions highlighting the distinction between grammatical and ungrammatical constructions; these were supplemented with training and feedback. To minimize processing demands, the task itself required only pointing to indicate whether the sentence was "okay" or whether the speaker "made a mistake." There was no need to manipulate materials, provide an explanation, or even to construct a verbal response.

Second graders were selected to be subjects in this study for several reasons. First, although it is possible at this age to make a fair assessment of reading skill, syntactic skill should not yet have been deeply affected by experience with reading. Second, by this age, acquisition of basic syntactic structures should be largely complete; failure on a task would signify a syntactic problem of some interest. And third, as discussed above, developmental research suggests that the task of making grammaticality judgments can be readily grasped at this time. And yet, because they are also far from ceiling performance at this age, testing 7-year-olds maximizes the opportunity to observe individual differences.

The *Judgment* task was selected to provide a measure of sentence knowledge relatively free from demands on metalinguistic skill or short-term memory. Using this procedure, it was possible to manipulate various types of syntactic violation systematically as well as to vary degrees of stress on short-term memory; these manipulations are elaborated upon in the Method section below. In addition, in a *Correction* task, subjects were asked to correct violations of grammaticality, using the same sentences and the same manipulations as in the Judgment task. Although the syntactic knowledge brought to bear in the two tasks should be the same, it was expected that the Correction task would place greater demands on both metalinguistic skill and short-term memory. Thus, differences in performance between the tasks should be attributable to these extra-syntactic processing demands.

In turn, scores on these metasyntactic measures were related to *reading skill*, to a verbal *IQ* measure, and to two abilities previously found to correlate with reading: *sentence recall* and *metaphonological skill*. Thus it was possible to compare the performances of this sample of children with findings of other studies, as well as to calculate the contribution of these abilities to performance on sentence-level tasks. By use of partial correlation procedures, it was also possible to arrive at a clearer understanding of the relationship between reading skill and syntactic knowledge uncontaminated by these other factors.

The expectations were as follows:

I) That second graders would be able to detect and correct some syntactic violations at a level above chance, but that the type of violation would have an effect on performance.

II) That reading and error correction would both require greater metalinguistic awareness and more processing skill than a basic measure of syntactic skill like making judgments of grammaticality. Thus, there should be a higher correlation between reading and Correction than between reading and Judgment.

III) Metaphonological scores were expected to pattern like reading scores, correlating with performance on the Correction task, but not with judgment performance.

IV) In view of the findings of Linebarger, Schwartz, and Saffran (1983) on aphasics, it was anticipated that individual differences in short-term memory would bear lightly or not at all on judgment scores. Similarly, experimental manipulations varying short-term memory requirements were expected to have little effect. Short-term memory factors were hypothesized to have a larger effect on correction.

Method

Subjects

The sample included 36 second graders (18 boys and 18 girls) attending two parochial elementary schools in predominantly white working-class and middle-class neighborhoods in New Haven. The only children excluded from the study were non-native speakers of the language and children who were repeating the second grade. All other children who responded to our request for permission were screened on the Peabody Picture Vocabulary Test (PPVT—Dunn, 1959). All those children obtaining an IQ score of 85 or higher were included in the study. IQ scores of the subjects included ranged from 86 to 144, with a mean of 108.6 (s.d. 15.2). The mean chronological age was 7;6 years (range 6;11 to 8;2). All subjects were receiving some phonics-based instruction.

Procedure

Children were tested in six sessions, spaced over the fall and spring semesters. The reading tests and the PPVT were administered on an individual basis in two 15-minute sessions near the end of the fall semester, following standard procedures. The experimental tasks measuring syntactic judgment, sentence repetition, and sentence correction were administered during the spring semester according to the procedures outlined below. The Judgment task was presented by earphones in two 20-minute sessions given on separate days in the first phase of testing. Separate tapes were made for each of the test sessions. All the sentence types were represented equally on each tape. The order of presentation was counterbalanced across subjects. The sentence recall and correction tasks were presented together on a third day, taking a total of 20 minutes. In the fourth and final session, children received the phonological segmentation

test; this took approximately 12 minutes and followed the procedures outlined in Rosner and Simon (1971). Children were tested separately in a quiet room.

Reading Measures

The *Decoding Skills Test* (DST—Richardson, DiBenedetto, and Adler (1982) provided the measure of reading ability used here. This measure was selected on the basis of the considerable evidence that decoding skill accounts for a large portion of the variance in other reading measures including both comprehension and sight reading vocabulary (see Stanovich 1982 for a review). The DST is a three-part criterion-referenced test of word recognition and decoding skill shown to be valid and reliable at or below the fifth-grade level. The Basal component measures the recognition of representative words from ten standardized basal reader series (thus familiar words) and yields grade norms from Pre-Primer to 5-2 (translatable into grade levels of 1.3 to 5.8), as well as a raw score. The second subtest consists of 60 Real words, systematically representing regular orthographic patterns and carefully selected to be in the speaking (but not necessarily reading) vocabularies of children at 11 different reading levels from Pre-Primer to 5-2. The third part includes 60 orthographically regular Nonsense words, derived from the real words by substitution of one or two segments (e.g. *jit* from *hit* or *bemonthrate* from *demonstrate*). Because the ability to identify individual pseudowords has proven to best distinguish good from poor readers (e.g. Perfetti and Hogaboam 1975), scores on the Nonsense section served as the independent measure of reading skill relied upon here; this measure is free of the confounding factors of reading experience and oral language vocabulary.

Phonological Segmentation

The *Auditory Analysis Test* (AAT—Rosner and Simon 1971) was used to assess metaphonological skill. This task requires subjects to repeat a word spoken by the experimenter, and then to isolate and remove a specified syllable or phonological segment, e.g. "can you say 'smile' without the 's'?" The AAT includes two practice items and 40 test items presented in order of difficulty, beginning with "Can you say birthday without the day?" and ending with "Can you say offering without the er?". The units to be eliminated include initial, final, and medial syllables; and initial, final, and medial consonants occurring alone or in clusters. The score is based on the number correct out of 40; if the subject makes five consecutive errors, testing is stopped and all subsequent items are treated as wrong. This test has been normed and validated by Rosner and Simon (1971) for grades K through 5.

Short-term Memory

A measure of short-term memory was taken to enable us to partial out individual differences in memory, in order to arrive at a less confounded measure of syntactic skills. The measure selected was a *Sentence Repetition*

Task in which subjects were asked to repeat verbatim five tokens of the grammatical sentences used in the study. Each sentence was presented twice, following which the time to respond was not restricted. Following Mann, Liberman, and Shankweiler (1980), error scores were derived by subtracting one point for each word deleted, transformed, or permuted. Maximum score correct on this task was 45.

Grammaticality Judgment Test

As discussed above, individual sentences for the *Judgment Test* were designed to be sematically transparent and plausible in an attempt to insure that the observed variance would arise from the syntactic demands of the task alone. Accordingly, grammatical errors were selected that would not prevent the child from arriving at the correct meaning of the sentence.

A total of 200 sentences were presented to each child, sampling evenly across five classes of grammatical violation and two memory stress conditions, as outlined below. These sentences included 100 pairs of sentences, with each pair including a grammatical and an ungrammatical variant. All sentences were of the same length (eight or nine words); within each sentence category, complexity was kept constant by use of a set number of nouns, verbs, and clauses. The sentences were recorded by a skilled female speaker under the guidance of the author.

Following Linebarger, Schwartz, and Saffran (1983), ungrammatical sentences were read with the intonation contour appropriate to a well-formed sentence derivable by substitution of one or two words. Sentences were presented in blocks of 20 separated by a 15-second break; each block sampling evenly across all conditions. Five blocks (100 sentences) were presented on each of two different days, with each day sampling evenly across all sentence types. Four seconds separated each sentence, but subjects were encouraged to work at their own pace, with the experimenter stopping the tape whenever the subject had not yet made a decision. Each sentence was presented once and no repetitions were allowed. Each day's recording lasted 15 minutes; actual testing took approximately 20 minutes a day. Each child's response was recorded on paper and scored as correct or incorrect.

To engage their interest, the children were presented with a large five-point scale (30" long × 7" high), with two large disks at either end (6" diameter), schematically representing a "happy face" and a "sad face." Three smaller disks (2" diameter) in between had eyes only, but were neutral with regard to expression (see Figure 1). Subjects listened to the sen-

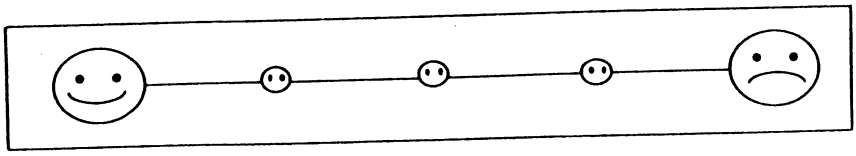

Figure 1. Apparatus used in grammaticality judgment task.

tences over soft earphones and pointed to the happy face when the sentence was "said the right way" and to the sad face when the speaker "made a mistake." The intermediate faces were presented as alternatives for when the child was unsure (didn't know) or when it was "just a little bit not okay." The children responded well to this apparatus and appeared to enjoy the task.

Each experimental session was preceded by a training session, with explicit instruction and feedback. Training was presented orally by the experimenter and involved short sentences that violated rules not actually tested in the study, e.g. "was you eating?" vs. "were you eating?" Demonstrating with pointing gestures, the experimenter highlighted the contrast between sentences which were "said the right way" (e.g. "I eat cookies") and bad sentences in which the speaker "made a mistake" (e.g. "Me eat cookies") and continued in this way until each child seemed to understand the task (up to ten pairs).

Types of grammatical rule violation. Five different classes of grammatical rule violation were selected, based in part upon some of the errors studied by Linebarger, Schwartz, and Saffran (1983), but revised to suit the present purpose. The categories represent different kinds of syntactic knowledge including rules specific to English and rules found universally, as well as rules specific to individual lexical items and those general to a grammatical category. Some constructions violated basic constraints on movement; others concerned the fine points of surface agreement. The grammatical constructions were selected to vary sufficiently in difficulty, both within and across sentence types, that they would capture individual differences in syntactic skill. An example of each of the five error types is presented below.

AUXILIARY AGREEMENT: VIOLATIONS OF AGREEMENT AND COPYING WITHOUT DELETION (*AGR*)

(1) *Are* you **find/finding* the math homework too difficult?
(2) *Was* your sister's new boyfriend *(*was) invited* to dinner?

TAG QUESTIONS: VIOLATIONS OF AGREEMENT ACROSS CLAUSES (*TAG*)

(3) That gorgeous red maple *tree* is tall, isn't **he/it?*
(4) Cindy *was* very sick last Thursday, **hasn't/wasn't* she?

ILLEGAL AUXILIARY MOVEMENT (*AUX*)

(5) *Is* the man who is lying on the floor ___ sleeping?
Is the man who ___ lying on the floor *is* sleeping?
**Is* the man who is lying on the floor *is* sleeping?
(6) *Did* the large tree branches that fell *break* anything?
**Did* the large tree branches that *fall broke* anything?

INCOMPLETE WH-MOVEMENT (*WH*)

(7) Which flowers will he buy for his mother?
**Which* will he buy *flowers* for his mother?

(8) Which records are you and Bill giving to Louise?
Which are you and Bill giving *records* to Louise?

STRICT SUBCATEGORIZATION (*SUB*)

(9) Grandpa *wants* you **would/to* go to the store now.
(10) They *hoped* the family next door **to/would* leave soon.

Types of memory stress. Although the success of agrammatic aphasics in making grammaticality judgments suggests that memory factors do not figure centrally in grammaticality judgments, this task was designed to test the role of short-term memory explicitly. Four of the five sentence categories under study (AGR, TAG, WH, SUB) were represented with a set of sentences designed to stress short-term memory and a set of sentences where short-term memory stress would be at a minimum. This manipulation was internal to the sentence, systematically varying the number of words that the child would have to store verbatim to detect a particular violation. The total number of words per sentence was kept constant across all sentence tokens (eight or nine words). To stress short-term memory, four or more intervening elements were placed between the two critically important (underlined) items in the sentences, as in items (2), (4), (8), and (10). In those sentences designed to minimize the burden on the short-term memory, the fewest possible intervening elements (0 to 3) were placed between critically dependent items, as in items (1), (3), (7), and (9). Because the sentences represented in items (5) and (6) were, perforce, locally ungrammatical, this manipulation did not apply there.

Error Correction Test

On a separate day, subjects were asked to make the appropriate *Corrections* on 50 of the ungrammatical sentences used in the Judgment Test. These sentences were randomly selected so as to sample evenly across the same experimental manipulations. In this case, the subjects were told they would hear sentences with a mistake and were asked to say it "the right way." The sentences were recorded on tape, as in the Judgment task. In this case, however, each sentence was presented twice, with an 8-second interval after each repetition: Although further repetitions were not allowed, subjects were again given as much time to respond as they desired. Sentences were presented in blocks of ten, followed by a 15-second break. Individual blocks were balanced across sentence type and memory condition. The taped sequence lasted 13 minutes altogether.

Again, the taped experimental session was preceded by an orally presented training session, with explicit instructions and feedback. The training sentences were very short and involved violations not tested for in the experimental component. For example, when presented with "*Is you eating?," subjects were asked to respond with "Are you eating?". Subjects appeared to understand this task readily. Sessions, including training and testing, took approximately 20 minutes. Subjects' attempts

at correction were recorded on tape and simultaneously transcribed online by the experimenter, to be checked later against the recorded version.

Results

Reading Measures

The Decoding Skills Test proved to be a sensitive test of reading ability in the second grade, revealing a wide range of skill without any subject approaching the minimum or maximum score. The mean score correct on Basal Vocabulary was 66.3 out of 110, with Instructional levels ranging from Primer (translatable to a grade level of 1.6 on normed reading tests) to 5-1 (grade level 5.3). The modal and mean grade levels obtained were 2.3 and 2.6, consistent with the actual time of testing. The mean score correct on Part 2, testing orthographically regular Real words, was 38.7 out of 60 (range 10 to 60); on Nonsense words the mean score correct was 29.5 out of 60 (range 0 to 57). The Phonic Transfer Index (the ratio score of real words to nonsense words), which measures the ability to generalize decoding patterns to novel stimuli, ranged from .50 for the poorest readers up to near perfect (.92) for the most advanced readers.

Consistent with previous findings (Perfetti and Hogaboam 1975; Richardson, Di Benedetto, and Adler 1982), performance on these various reading measures correlated highly with each other ($r(36) > .90$, $p < .001$). On the other hand, the correlation between scores on Nonsense decoding and the PPVT failed to reach significance ($r(36) = .31$, $p > .05$); suggesting that the decoding measure, as designed, is relatively independent of vocabulary knowledge.

Phonological Segmentation

Performance on the Auditory Analysis Test (AAT) ranged from 10 correct to 37 correct (maximum = 40), with a mean correct score of 24.9 (s.d. 8.4), consistent with Rosner and Simon's (1971) second grade norms. The correlation between performance on this task and Nonsense reading score was .65 ($p < .001$). The phonological analysis skills in this sample account for 40 percent of the variance in the Nonsense reading scores; this closely matches the 39 percent figure Rosner and Simon reported as accounting for second grade performance on a different reading test. The correlation in this sample remained moderately high even after partialling out differences in IQ ($r(36) = .60$, $p < .05$).

Short-term Memory

The results suggest that the sentence length and sentence constructions selected in the Sentence Repetition Test do not overly tax the memory span of second graders, as assessed by their ability simply to repeat grammatical versions verbatim. In the repetition of five nine-word sentences, the number of words omitted, changed, or transmuted ranged in

Table I
Performance on Syntactic Measures

	Judgment (max = 200)	Correction (max = 50)
Mean % correct	75.0	76.0
Standard deviation	12.7	15.0
Range	47.5 to 93.0	46.0 to 100.0
Sentence-type effect	$p < .001$	$p < .001$
Memory-stress effect	$p > .01$	$p < .001$

number from 0 to 7; the mean score correct was 43.3 (maximum = 45). Individual differences in memory (as assessed by the number of words correctly recalled) accounted for 10 percent of the variability in reading, consistent with the figure of 10 to 12 percent reported in other studies (Mann and Liberman 1984). The correlation between memory span and reading scores just failed to reach significance ($r = -.31$, $p = .06$); however, children making three or more memory errors had significantly lower Nonsense reading scores than did those with two or fewer errors (F [2,33] = 4.29, $p < .05$). Performance on the repetition task correlated significantly with scores on the phonological deletion task ($r = .48$, $p < .01$) and with the PPVT IQ measure ($r = .46$, $p < .01$). When IQ was partialled out, the correlation between short-term memory and the metaphonological task remained significant ($r = .35$, $p < .05$), but the correlation with reading was very small ($r = .20$, $p > .05$).

Grammaticality Judgment Test

Second grade children scored well above chance, on average, in detecting ungrammatical sentences (see Table I). Whereas a chance performance was 50 percent, overall mean percentage correct was 75.0 (s.d. 12.7). Mean scores for individuals ranged from 47.5 to 93 percent correct, with only three out of 36 subject failing to score well above chance on at least two of the ten sentence categories.[1] As predicted, performance

[1]Following the procedures of Linebarger, Schwartz, and Saffran (1983), performance on the judgment task was also analyzed in terms of signal detection analysis, using the statistic A', a nonparametric index of sensitivity which takes into account the tendency to false alarms (rejection of a well-formed sentence) in calculating the ability to reject ill-formed sentences. This measure, derived by the formula given below, can be interpreted quite naturally as an expected score of percent correct on a good/bad forced choice procedure. Chance on this measure is 0.50, indicating an equal tendency to reject both well-formed and ill-formed sentences.

$$A' = 0.5 + (y - x)(1 + y - x)/4y(1 - x)$$

where x = proportion of false alarms and y = proportion of hits.

The primary difference emerging from the A' measure, compared to a simple measure of proportion correct, was a greater range of scores, because it exaggerates performance at both ends. The mean A' score was .80 with a range of 0.36 to 0.96 and standard deviation of 0.15. Correlation results using this analysis techniques corroborated closely with those based upon overall percentage correct. The decision not to rely upon the A'

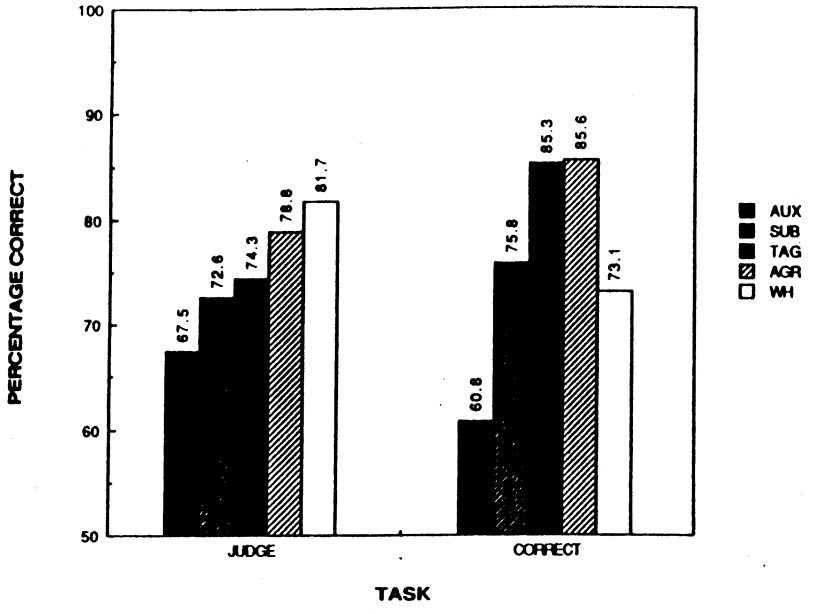

Figure 2. Order of difficulty on sentence categories as a function of task.

on this task was significantly affected by sentence type (F [4, 140] = 22.92, $p < .0001$), with the order of difficulty as presented in Figure 2.

Overall percentage correct was neither consistently nor significantly affected by the desired response (accept, reject), (F[1,35] = 0.00, $p = .96$). Similarly, overall Judgment performance was not consistently affected by the memory stress manipulation described above (F[1,35] = 0.01, $p = .94$). Memory stress did, however, interact with sentence type (F[3,105] = 5.02, $p < .01$). Posthoc least significant difference (lsd) comparisons (Winer 1971, pp. 199–201) indicated that violations were more apt to be noticed when there were fewer intervening items only in the case of Wh-movement.

As can be seen in Table II, performance on the judgment task correlated significantly with IQ, with metaphonological skill, and with short-term memory. Judgment skill did not, however, correlate significantly with reading scores. When, as is the standard procedure, the contribution of IQ was partialled out of these correlations, the correlation between Judgment and memory span remained significant. However, with the contribution of IQ removed, Judgment no longer correlated significantly with either reading or the associated metaphonological task. Similarly,

---

measure is based on the fact that in those cases where children seemed to be operating at chance (e.g. 45 percent correct overall), the A' score assigned was heavily influenced by number of false alarms. Given a preponderance of false alarms relative to hits, scores ranged as low as −0.375 which we felt skewed the data unnecessarily.

Table II
Correlations Between Syntactic Tasks and Measures of IQ, Short-Term Memory, Reading & Phonological Segmentation
(Factors in parentheses have been partialled out)

	Judgment	Judge Bad	Judge Good	Correction
PPVT-R IQ	0.47**	0.53***	0.24	0.60***
Short-Term Memory	0.46**	0.54***	0.30	0.67***
DST Reading	0.23	0.01	0.38*	0.48**
Reading (IQ)	0.10	−0.23	0.33*	0.38*
Reading (STM)	0.09	−0.24	0.31	0.39*
AAT	0.38*	0.25	0.38*	0.53***
AAT (IQ)	0.22	0.02	0.30	0.37*
AAT (STM)	0.18	−0.02	0.26	0.37*

*$p<.05$
**$p<.01$
***$p<.001$

when the contribution of memory differences was removed from the correlation, Judgment skill again correlated nonsignificantly with both reading and phonological deletion. Indeed, without the contribution of memory span, even the correlation between IQ and Judgment scores failed to reach significance.

Correlations between reading and Judgment skill were nonsignificant for each of the five error types tested; and for both memory stress conditions. The only measure of judgment skill to correlate at all significantly with reading was the ability to *accept* a grammatical sentence as good (JUDGE GOOD on Table II); even then the significance of this correlation failed to survive the partialling out of short-term memory differences. This skill, however, failed to correlate with either memory span or IQ. In sharp contrast, the ability to *reject* ungrammatical sentences (JUDGE BAD on Table II), correlated with reading not at all, but did correlate with both memory span and IQ. Indeed, the two abilities to accept or reject sentences appropriately were themselves correlated only at a low level ($r = .30$, $p < .10$).

Correction Test

The analysis applied to the Correction Test distinguished only between responses in which the focal error was appropriately corrected (with or without nonsubstantive omissions or changes) and those responses in which errors were left intact or incorrectly modified. Following these criteria, subjects achieved a mean Correct score of 76 percent (range 46 to 100, s.d. 15; see Table II).

Accurate performance on the correction task was significantly affected by sentence type (F [4,140] = 13.38, $p < .0001$), with the order of difficulty as presented in Figure 2. This order of difficulty was consistent

with that observed in the Judgment task with one notable exception: incomplete Wh-movement was the easiest of all violations to detect, but proved to be one of the most difficult to remedy properly.

In contrast to its inconsequential role in the Judgment task, the memory stress manipulation had a highly significant effect on Correction task scores (F [1,35] = 13.25, $p < .001$); it tended to be more difficult to accurately reconstruct sentences with errors at a distance than those in which the errors are more local. The interaction between sentence type and memory stress was also significant (F [3,105] = 7.97, $p < .0001$); posthoc lsd analyses suggest that the memory stress factor was in effect in the predicted direction for three out of the four sentence types for which it was tested (AGR (0.92 vs. 0.79); WH (0.78 vs. 0.68) and TAG (0.90 vs. 0.81)); in all cases, these differences were significant at the .01 level.

As presented in Table II, scores on the Correction task correlated significantly, at the .01 level, with IQ, short-term memory, phonological segmentation, and with reading. When shared variance with IQ was removed, the correlations with each of the other measures remained significant, at the .05 level. These correlations also remained significant after partialling out memory span differences.

An internal look at how performance on individual subscores contributes to the overall significant correlation revealed that reading correlated significantly ($p < .05$) with three of the sentence types: AGR, AUX, and TAG. Both AGR and AUX remained significant correlated with reading after memory span or IQ was partialled out. Whereas items selected to stress memory correlated significantly with reading skill (r = .39, $p <$ .05), performance on the low stress condition only approached significance at the .05 level (r = .32).

Comparing Correction and Judgment

The similarities in mean score correct across the two syntactic measures are highlighted in Table I. In addition, correlations between accurate performance on the Judgment task and on the Correction task were high (r = .69 overall; r = .59 with IQ partialled out; and r = .55 with STM partialled out; $p < .01$ in all cases), reflecting a shared variance of 48 percent.

Discussion

This study examined the ability of second-grade children to detect and to remedy violations of grammaticality. These abilities were, in turn, examined in relation to other areas of verbal skill (reading, short-term memory, vocabulary, and phonological analysis) long known to vary in second graders. Thirty-three of the 36 subjects performed above chance on the Judgment task and all children were able to correct at least half of the deviant sentences. A significant role of grammatical knowledge was indicated in the large effect of sentence type on performance as well as in

the high proportion of shared variance across tasks. Although the performance measures were nearly identical for Judgment and Correction, as measured by percentage correct or standard deviation, the two tasks varied considerably in terms of how strongly they correlated with other abilities. Performance on the Judgment task, selected as a means of assessing grammatical knowledge without stressing processing capacity, did *not* correlate with differences in reading skill. Rather, skill on this task was much more highly associated with IQ than with either the metaphonological task or reading. In sharp contrast, the Correction task, which requires not only conscious knowledge of grammatical structure, but the ability to hold lengthy strings in memory and manipulate items to formulate new correct sentences, correlated robustly with both reading ability and phonological analysis.

The second-grade children under investigation were comparable in most respects to other groups of children participating in reading studies. This was true not only with respect to the absolute scores achieved on standardized tests (the AAT and the DST), but was also apparent in the correlations among these and measures of IQ and short-term memory. For example, in this study, consistent with previous studies, IQ and short-term memory each contributed approximately 10 percent to the variance in reading. Similarly, as in other studies, metaphonological analysis skill, assessed here on the AAT, was much more closely associated with success in reading, contributing close to 40 percent of the variance. This study also revealed a strong relationship between short-term memory and skill in phonological analysis, a relationship anticipated in earlier studies of reading (Brady 1986; Mann and Liberman 1984; Shankweiler and Crain 1986b; see Blachman 1986 for supporting results).

The results of this study agree well with the findings of previous investigators who have explored the relationship between reading skill and the ability to correct violations of grammaticality (Bowey 1986; Flood and Menyuk 1983; Forrest-Pressley 1983; Ryan and Ledger 1984). These investigators, using diverse procedures and examining a wide age range, have consistently reported a strong correlation between reading ability and correction skill when differences in age were controlled for. This study goes beyond most in demonstrating that the relationship between reading and correction skills is also independent of individual differences in vocabulary knowledge or short-term memory.

A review of the literature suggests that the relationship between reading ability and the isolated ability to make judgments of grammaticality is less well understood. Bohannon, Warren-Leubecker, and Hepler (1984), for example, report that children who showed evidence of a high level of "word order awareness" when tested at 5 to 7 years of age achieved significantly higher reading scores a year later than those children who had obtained "low awareness" scores. On the other hand, although Vogel (1974) found that large differences separated normal from dyslexic second graders on seven out of nine sentence-level measures

(e.g. sentence repetition or sentence cloze tasks), she failed to find a difference between the groups in a grammaticality judgment task. The contrast between these two studies is consistent with the hypothesis presented above that good and poor readers should differ only on those tasks which are highly demanding of memory resources or metalinguistic skill: Bohannon's guessing game may have placed quite different demands on the child than Vogel's measure of syntactic awareness.

Ryan and Ledger (1979) relied on procedures very similar to those employed here and obtained comparable results: in both studies, the tendency of first and second graders to accept grammatical sentences correlated with reading skill, but the tendency to reject grammatical deviations did not. Why better readers should be generally more accepting of sentences than poor readers is not clear either to them or to us. However, given that neither overall scores nor weighted scores revealed any superior discrimination power on the part of better readers (see footnote 1), it does not seem appropriate to attribute greater grammatical competence to the better readers on the basis of a systematic response bias.

In a grammaticality judgment study similar in purpose to the one reported here, Kean (1984) asked six self-declared dyslexic adults to make judgments of grammaticality and ambiguity on 75 sentences reflecting 16 aspects of English sentence structure. Kean found that the overall mean percentage correct of the dyslexic subjects was within the range of the mean scores of three normal control groups (slightly above 10 and 11-year-old children; slightly below both college sophomores and randomly selected adults) and well above the mean score of a group of adults who had suffered neurological disorders very early in development. However, Kean reported an anomalous pattern of interpretation among the dyslexics on two of the sentence types involving the assignment of referential relations between nouns and pronouns on the basis of syntactic structure (e.g. "Lila loves Lila" and "John promised Mary to do the dishes and she did them.") On these sentence types, the dyslexic group patterned like the neurologically impaired group and unlike the normal controls. On the basis of these results, Kean suggests that dyslexics suffer a systematic, if subtle, linguistic impairment. She argues that such difficulties cannot be attributed to a memory limitation or a limitation in metalinguistic capacity because these same subjects can and do handle other, longer, sentences quite competently. In spite of the different criteria in selecting poor readers, the two studies are in agreement in finding dyslexics are not at a disadvantage on the sentence structures common to both studies, such as tag questions. This suggests that the technique employed is promising.

No previous study on grammatical awareness in children has systematically manipulated memory demands. Because agrammatic aphasics with severely impaired memory spans can make judgments of grammaticality (Linebarger, Schwartz, and Saffran 1983), it was hypothesized that short-term memory would contribute minimally to performance on the Judgment task. Contrary to these expectations, in the second grade

population under study here, there was a significant correlation between Judgment skill and the ability to repeat a grammatical sentence verbatim. Nonetheless, the relationship between short-term memory and Judgment is less robust than that between short-term memory and Correction skill. This is indicated in the strength of association as measured in correlation coefficients as well as in the lack of effect of a memory stress manipulation on Judgment performance. The same manipulation had a large effect on performance on the Correction task.

Disparity in the correlational patterns which indicate to what extent the Judgment and Correction tasks are associated with each of several other factors allows us to consider which of the factors pertaining to sentence-level performance are intrinsically related to individual differences in reading. Because performance on the Judgment task does not correlate with reading performance, it is apparent that the representation of grammatical structures, and the awareness of the distinction between sentences that are and are not grammatical, are not deficient in poor readers. On the other hand, those features which distinguish Correction from Judgment do vary with reading skill. The high correlation with short-term memory and the large effect of the memory stress manipulation suggest that what makes Correction difficult, at least in part, is the increased demand it makes on short-term memory. That short-term memory fails to account for all the variance in Correction performance may be explained by the need to rework the sentence explicitly. It may be that the Correction task requires more "metalinguistic" knowledge than the straightforward Judgment task; this interpretation is corroborated by the high correlation between phonological analysis skill and scores on the Correction task. In sum, the pattern of results observed here suggests that poor readers typically do not lack access to grammatical knowledge; rather they lack the ability to act on sentences in ways which require explicit accounting and/or which heavily stress short-term memory. Alternatively, they lack the flexibility to carry out re-analysis.

The results of this study are much in accord with the hypothesis of Liberman (1971), Mattingly (1972) or Gleitman and Rozin (1977) that it is not basic lexical and syntactic knowledge that is lacking in poor readers, but metalinguistic ability, the most unnatural skill of extracting and manipulating phonemic elements, or perhaps a broader underlying deficit in phonological processing (Liberman and Shankweiler 1985). These results, too, are consistent with the hypothesis set forth by Shankweiler and Crain (1986b) that basic syntactic structures are intact in the typical poor reader, but that performance will appear less than optimal when short-term memory is heavily taxed.

These findings have some clear implications for the diagnosis and treatment of reading disability. First, they suggest that the syntactic knowledge of good and poor readers alike is well in place by seven years of age; this is consistent with the indications of basic research on the acquisition of syntax in preschoolers (see deVilliers and deVilliers 1978 or

Gleitman and Wanner 1984 for an overview). Consequently, the ability to make grammaticality judgments, or to perform even more straightforward tests of syntax is not expected to be a good predictor of reading skill. Problems in syntax over and beyond those that can be attributed to the artificial and memory-demanding nature of the task point to wider problems in language and should not be construed as reading disability. Finally, these results suggest that children who are experiencing difficulties in learning to read, and in understanding what they read, are not limited by problems with syntax, per se. Instead, as stressed by Perfetti (1985), their problems even in poor comprehension can usually be attributed to lower-level processes including short-term memory, decoding skill, and underlying phonological abilities.

References

Blachman, B. 1986. Language factors in early reading. Paper delivered at the National Meeting of the Orton Dyslexia Society, Philadelphia, Pa.

Bohannon, J. N., Warren-Leubecker, A., and Hepler, N. 1984. Word-order awareness and early reading. *Child Development* 55:1541–1548.

Bowey, J. A. 1986. Syntactic awareness and verbal performance from pre-school to fifth grade. *Journal of Psycholinguistic Research* 15:285–308.

Brady, S. A. 1986. Short-term memory, processing and reading ability. *Annals of Dyslexia* 36:138–153.

Brady, S. A., Shankweiler, D., and Mann, V. A. 1983. Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology* 35:345–367.

Byrne, B. 1981. Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics* 2:201–212.

Crain, S. and Shankweiler, D. 1988. Reading acquisition and language acquisition. *In* A. Davidson, G. Green, and G. Herman (eds.). *Critical Approaches to Readability: Theoretical bases of linguistic complexity.* Hillsdale, NJ: LEA (In Press).

Crain, S., Shankweiler, D., and Tuller, B. 1984. Preservation of sensitivity to closed-class items in agrammatism. Paper read at the Academy of Aphasia, Los Angeles, CA.

de Hirsh, K., Jansky, J., and Langford, W. 1966. *Predicting Reading Failure.* New York: Harper and Row.

Denckla, M. B. and Rudel, R. G. 1976. Naming of object-drawings by dyslexic and other learning disabled children. *Brain and Language* 3:1–15.

deVilliers, J. G. and de Villiers, P. A. 1978. *Language Acquisition.* Cambridge, Mass.: Harvard University Press.

Dunn, L. M. 1959. *Peabody Picture Vocabulary Test.* Circle Pines, MN: American Guidance Service.

Fletcher, J. M., Satz, P., and Scholes, R. 1981. Developmental changes in the linguistic performance correlates of reading achievement. *Brain and Language* 13:78–90.

Flood, J. and Menyuk, P. 1983. The development of metalinguistic awareness and its relation to reading achievement. *Journal of Applied Developmental Psychology* 4:65–80.

Forrest-Pressley. 1983. *Cognitive and Meta-Cognitive Aspects of Reading.* New York, NY: Springer-Verlag.

Fry, M., Johnson, C., and Muehl, S. 1970. Oral language production in relation to reading achievement among select second-graders. *In* D. J. Bakker and P. Satz (eds.). *Specific Reading Disability: Advances in theory and method.* Rotterdam: Rotterdam University Press.

Gleitman, L. R., Gleitman, H., and Shipley, E. 1972. The emergence of the child as a grammarian. *Cognition* 2:137–164.

Gleitman, L. R. and Rozin, P. 1977. The structure and acquisition of reading: Relation between orthography and the structured language. *In* A. S. Reber and D. L. Scarborough (eds.). *Toward a Psychology of Reading*. Hillsdale, NJ: Erlbaum.

Gleitman, L. R. and Wanner, E. 1984. Current issues in language learning. *In* M. H. Bornstein and M. E. Lamb (eds.). *Developmental Psychology: An advanced textbook*. Hillsdale, NJ: Erlbaum.

Hakes, D. 1980. *The Development of Metalinguistic Abilities in Children*. New York, NY: Springer-Verlag.

Hirsh-Pasek, K., Gleitman, L. R., and Gleitman, H. 1978. What did the brain say to the mind? A study of the detection and report of ambiguity by young children. *In* A. Sinclair, R. J. Jarvella and W. J. M. Levelt (eds.). *The Child's Conception of Language*. New York, NY: Springer-Verlag.

Jorm, A. R. 1983. The cognitive and neurological basis of developmental dyslexia: A theoretical framework and review. *Cognition* 7:19–33.

Katz, R. B. 1986. Phonological deficiencies in children with reading disability: Evidence from an object-naming task. *Cognition* 22:225–257.

Kean, M-L. 1984. The question of linguistic anomaly in developmental dyslexia. *Annals of Dyslexia* 34:138–151.

Liberman, I. Y. 1971. Basic research in speech and lateralization of language: Some implications for reading disability. *Bulletin of the Orton Society* 21:71–87.

Liberman, I. Y. and Shankweiler, D. 1985. Phonology and the special problems of learning to read and write. *Remedial and Special Education* 6:8–17.

Linebarger, M. C., Schwartz, M. F., and Saffran, E. M. 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 13:361–392.

Mann, V. A. 1984. Reading skills and language skill. *Developmental Review* 4:1–15.

Mann, V. A. and Liberman, I. Y. 1984. Phonological awareness and verbal short-term memory: Can they presage early reading problems? *Journal of Learning Disabilities* 17:592–599.

Mann, V. A., Liberman, I. Y., and Shankweiler, D. 1980. Children's memory for sentences and wordstrings in relation to reading ability. *Memory and Cognition* 8:329–335.

Mann, V. A., Shankweiler, D., and Smith, S. T. 1984. The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language* 11:627–643.

Mattingly, I. 1972. Reading, the linguistic process, and linguistic awareness. *In* J. F. Kavanagh and I. G. Mattingly (eds.). *Language by Ear and by Eye*. Cambridge, MA: MIT Press.

Perfetti, C. A. 1985. *Reading Ability*. New York: Oxford University Press.

Perfetti, C. A. and Goldman, S. 1976. Discourse memory and reading comprehension skill. *Journal of Verbal Learning and Verbal Behavior* 14:33–42.

Perfetti, C. A. and Hogaboam, T. 1975. The relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology* 67:461–469.

Richardson, E., DiBenedetto, B. and Adler, A. 1982. Use of the Decoding Skills Test to study differences between good and poor readers. *In* K. Gadow and I. Bialer (eds.). *Advances in Learning and Behavioral Disabilities* (Vol. 1). Greenwich, CT: JAI Press.

Rosner, J. and Simon, D. P. 1971. The auditory analysis test: An initial report. *Journal of Learning Disabilities* 4:40–48.

Rubin, H. and Liberman, I. Y. 1983. Exploring the oral and written language errors made by language disabled children. *Annals of Dyslexia* 33:111–120.

Ryan, E. B. and Ledger, G. W. 1979. Grammaticality judgments, sentence repetitions and sentence corrections of children learning to read. *International Journal of Psycholinguistics* 6:23–40.

Ryan, E. B. and Ledger, G. W. 1984. Learning to attend to sentence structure: Links between

metalinguistic development and reading. *In* J. Downing and R. Valtin (eds.). *Language Awareness and Learning to Read*. New York: Springer-Verlag.

Scholl, D. M. and Ryan, E. B. 1980. Development of metalinguistic performance in the early school years. *Language and Speech* 23:199–211.

Shankweiler, D. and Crain, S. 1986a. Syntactic versus performance factors in the study of language disorders. Paper presented at the Academy of Aphasia.

Shankweiler, D. and Crain, S. 1986b. Language mechanisms and reading disorder: A modular approach. *Cognition* 24:139–168.

Smith, S. T. 1987. Syntactic comprehension in reading-disabled children. Doctoral dissertation. Storrs, CT: University of Connecticut.

Smith, S., Mann, V. A., and Shankweiler, D. 1986. Good and poor readers comprehension of spoken sentences: A study with the Token Test. *Cortex* 22:627–632.

Stanovich, K. E. 1982. Individual differences in the cognitive processes of reading: 1. Word decoding. *Journal of Learning Disabilities* 15:449–512.

Stein, C. L., Cairns, H. S., and Zurif, E. B. 1984. Sentence comprehension limitations related to syntactic deficits in reading-disabled children. *Applied Psycholinguistics* 5:305–322.

Torgesen, J. K. and Houck, D. 1980. Processing deficiencies of learning-disabled children who perform poorly on the digit span test. *Journal of Educational Psychology* 72:141–160.

Vellutino, F. R. 1979. *Dyslexia: Theory and research*. Cambridge, MA: MIT Press.

Vogel, S. A. 1974. Syntactic abilities in normal and dyslexic children. *Journal of Learning Disabilities* 7:47–53.

Weinstein, R., and Rabinovitch, M. S. 1971. Sentence structure and retention in good and poor readers. *Journal of Educational Psychology* 62:25–30.

Whitehouse, C. C. 1983. Token test performance by dyslexic adolescents. *Brain and Language* 18:224–235.

Wiig, E. H. and Roach, M. A. 1975. Immediate recall of semantically varied sentences by learning disabled readers. *Perceptual and Motor Skills* 40:119–125.

Wiig, E. H., Semel, E. M., and Crouse, M. A. B. 1973. The use of English morphology by high-risk and learning disabled children. *Journal of Learning Disabilities* 7:59–67.

Winer, B. J. 1971. *Statistical Principles in Experimental Design (2nd Ed.)*. New York: McGraw-Hill.