NOTE

SPOKEN SENTENCE COMPREHENSION BY GOOD AND POOR READERS: A STUDY WITH THE TOKEN TEST

Suzanne T. Smith[1], Virginia A. Mann[2] and Donald Shankweiler[1]

([1]Haskins Laboratories and University of Connecticut, and [2]Haskins Laboratories and Bryn Mawr College)

INTRODUCTION

Previous research has indicated that disabled readers in early elementary grades make inefficient use of working memory for linguistic materials (Liberman, Shankweiler, Liberman, Fowler and Fischer, 1977; Stanovich, 1982; Jorm, 1983; Mann, 1984; Liberman and Shankweiler, 1985). Because sentence processing requires temporary retention of successive words during parsing, impaired comprehension might be expected in cases of verbal short-term memory deficit, particularly when materials are syntactically complex or lacking in contextual support (Mann, Liberman and Shankweiler, 1980). We have investigated this possibility by administering the Token Test (De Renzi and Faglioni, 1978) to good and poor readers as a measure of comprehension, and a test of immediate recall of word strings as a measure of verbal working memory.

The Token Test requires full processing of lexical content and syntax for successful performance, and places progressively increasing demands upon short-term memory. Poor performance on the Token Test could be due to a number of factors, including poor retention of information in the items or inadequate understanding of the grammar. We predicted that, in light of their known short-term memory deficit, poor readers would make more errors than good readers on Part 5, in which two tokens, each modified by two adjectives, are indicated. We also predicted that their memory problems could lead to an increased number of requests for commands to be repeated. If syntax is selectively impaired in poor readers, they would be expected to make more errors on Part 6 in which complex syntactic constructions are introduced. An analysis based upon the work of Noll and Randolph (1978) provides a further test of this hypothesis. Errors on Part 6 are classified as involving either grammatical terms or token attributes.

MATERIALS AND METHODS

Subjects

The subjects were 35 third-grade children from a suburban public school system. All were native speakers of English with no known speach or hearing deficiencies, and all had an intelligence quotient of 90 or above as measured by

the Peabody Picture Vocabulary Test (Dunn, 1965). Reader groups were formed based on teachers' evaluations of reading ability and scores on the reading subtest of the Iowa Test of Basic Skills (Hieronymus and Lindquist, 1978) which had been administered four months earlier. Eighteen students comprised the good reader group (three boys and fifteen girls), and seventeen students comprised the poor reader group (nine boys and eight girls). The mean Iowa grade-equivalent score for good readers was 4.59 (range = 4.1 to 5.2) and 2.3 for poor readers (range = 1.7 to 2.6). The groups did not differ significantly in IQ (good readers, 109.3; poor readers, 107.7), nor in age (good readers, 110.5 months; poor readers, 107.4 months). Reading ability was not significantly correlated with IQ.

Tests and procedure

Subjects were tested individually in one session during which the Peabody Picture Vocabulary Test, the test of word-string recall, and the Token Test were administered in that order.

Word Strings

Eight strings of six monosyllabic words distributed over a range of frequencies (Carroll, Davies and Richman, 1971) were compiled by random selection with the restriction that there be no rhymes or highly-associated words within a string. The words were recorded on audiotape by a male speaker of English who produced each with neutral intonation at a rate of approximately one per second. The onset of each string was signalled by the word "ready". The subjects' task was to repeat each word string in the order given.

Token Test

The shortened version of the Token Test (De Renzi and Faglioni, 1978) was employed with one modification: a blue token was substituted for the yellow token and test sentences were modified accordingly. The test was administered following the standard instructions with the exception that items were repeated on all sections of the test if the subject so requested.

RESULTS

Recall of Word Strings

Recall was scored in two ways: (a) order-free scoring, giving credit for correct items in whatever order they occurred; (b) order-strict scoring, crediting only words repeated in proper sequence as correct. The mean order-free recall score (max. = 48) for good readers was 28.2 (S.D. 2.91) and 25.3 (S.D. 4.37) for poor readers ($t = 2.22$, d.f. = 33, $p = .03$). The mean order-strict recall score for good readers was 21.5 (S.D. 4.31) and 17.5 (S.D. 4.69) ($t = 2.56$, d.f. = 33, $p = .015$). Both word-recall scores were significantly, though not strongly, correlated with reading ability (free-recall, $r = .33$, d.f. = 33, $p = .05$; ordered-recall, $r = .37$, d.f. = 33, $p = .03$). Neither recall score was significantly correlated with age or IQ.

TABLE I

Mean Percent Correct Responses on Each Section of the Token Test

Section	Reader Group	
	Good (N=18)	Poor (N=17)
1	100.00	100.00
2	100.00	98.53
3	100.00	100.00
4	98.61	98.53
5	97.22	86.76
6	81.20	76.47

Token Test Performance

Poor readers requested significantly more repetitions of the Token Test commands than good readers — an average of 1.41 vs. .16, respectively, $t = 2.81$, d.f. $= 33$, $p = .009$. Requests occurred primarily on Sections 5 and 6, and an analysis of variance based on these sections revealed a significant reader group effect ($F = 6.01$; d.f. $= 1, 33$; $p = .02$), but no differential effect attributable to test section.

Comprehension of each Token Test item was scored as correct or incorrect. Table I displays the mean percent correct for each reader group on each section of the test. Analysis of variance indicated that good readers made more correct responses than poor readers ($F = 4.79$; d.f. $= 1, 33$; $p = .03$). There was a significant main effect of test section ($F = 42.60$; d.f. $= 5, 165$; $p < .001$) and a significant reader group $\times$ section interaction ($F = 2.64$; d.f. $= 5, 165$; $p = .025$). The scores were significantly correlated with performance on the Iowa Test ($r = -.40$, d.f. $= 33$, $p = .02$), but not with word-string recall, age or IQ. Subsequent analyses revealed that both reader groups had attained nearly perfect scores on Sections 1-4, with no significant differences. They differed significantly on Section 5 ($t = 2.05$, d.f. $= 33$, $p = .04$), with poor readers performing less well than good readers. On Section 6, although poor readers averaged fewer correct responses than good readers, the difference is not statistically significant.

The pattern of errors on Section 6 (grammatical term vs. token attribute errors) revealed that both reader groups made proportionally more errors on grammatical terms than on token attributes (85.4% vs. 14.6% for good readers; 71.7% vs. 28.3% for poor readers). Significant differences between reader groups obtained, however, only in the incidence of token attribute errors: .55 errors for good readers vs. 1.76 for poor readers ($t = 2.42$, d.f. $= 33$, $p = .02$).

DISCUSSION

The finding that poor readers exhibited significantly more errors than good readers in recall of word strings is consistent with other reports in the research literature that indicate that poor readers are less efficient than good readers in

working memory for linguistic materials. In keeping with the expectation that inefficient working memory may lead to impaired sentence comprehension, it was found that poor readers made significantly more errors than good readers on the Token Test. The finding of inferior performance by poor readers on certain Token Test items is in keeping with results that have been reported on other groups of children with language impairment (Tallal, 1975; Lapointe, 1976; Whitehouse, 1983). Since reading ability was not significantly correlated with IQ, it is unlikely that performance on the Token Test reflected general cognitive ability; instead, the performance of poor readers reflects problems that are language-related.

We consider two hypotheses in accounting for poor readers' difficulty on the Token Test. On one hypothesis, syntactic competence is deficient in poor readers. That is, poor readers' lower scores could be attributed to failure to master the grammar of English at a level commensurate with that of good readers as, for example, Byrne (1981) has suggested. An alternative hypothesis would suppose that poor readers incur comprehension losses on linguistic constructions that stress their inefficient processing capacities. In that event, poor readers' lower scores could be a result of their known difficulty with holding verbal material in short-term memory (Mann, Shankweiler and Smith, 1984; Shankweiler, Smith and Mann, 1984). A consideration of the pattern of performance across test items may help us choose the better hypothesis.

We first note that, while items on Part 6 were the most difficult for both groups, good readers did not surpass poor readers. This fails to support the hypothesis that poor readers have a specific syntactic deficit. It is likely that some of the difficulty associated with these items may stem from procedural factors rather than from a specific syntactic deficit. For example, approximately one fourth of the subjects made an error on Item 26, "Touch the black circle and the red square," whereas only a few subjects made errors on similar items that appeared in Part 4. This finding cannot easily be accounted for either by a syntactic deficit or impaired linguistic memory. Instead, it suggests that items in Part 6 are confounded by procedural factors that mask the expression of syntactic ability.

The only section of the test on which good readers surpassed poor readers was Part 5, in which items require manipulation of two tokens, each modified by two adjectives. Importantly, these items have the same grammatical structure as those in the previous two test sections on which good and poor readers did not differ. The fact that performance on these items alone distinguished good and poor readers may plausibly reflect the amount of information that must be held during processing such items, a possibility which lends support to the verbal short-term memory hypothesis.

The further finding that reader groups were significantly differentiated by token attribute errors in the error analysis on Part 6 could be taken as additional support for the working memory hypothesis. We note, however, that this analysis fails to take into account the full range of memory demands associated with sentence comprehension; retention errors can reflect syntactic misanalysis of constituent structure as well as incorrect recall of individual words. Failure to retain information adequately at either level may result in impaired comprehension.

Possibly, a more direct indication that comprehension problems in the poor readers reflected short-term memory limitations is the fact that the poor readers requested more repetitions than the good readers. This finding, together with the

significant difference in performance between reader groups on Part 5, and the poor readers' greater incidence of errors of retention of token attributes on Part 6 suggests that their difficulty on the Token Test was associated with difficulty in retaining verbal material in short-term memory, even though performance on the word-string measure of short-term retention was not significantly correlated with Token Test performance.

While these results are partially supportive of the memory hypothesis, they are, to be sure, not conclusive. First, in light of the lack of a significant correlation between word-string recall and the Token Test, we suggest that Token Test performance may recruit memory abilities beyond those reflected in a test of rote recall. (See, for example, Daneman and Carpenter, 1980, for indications that tests of rote recall are not as strongly associated with reading ability as some other measures of working memory.) Second, although the Token Test gives a rough measure of comprehension performance that has proven very useful in the clinical assessment of language disorders, it is not the most appropriate instrument to analyze the basis of comprehension impairment. Future studies aimed at understanding comprehension deficits in poor readers could usefully employ a wider variety of syntactic constructions in test sentences that incorporate controls for sources of processing difficulties.

Abstract

Good and poor readers in the third grade (age nine years) were examined on (a) a test of immediate memory for word strings and (b) on a version of the Token Test (De Renzi and Faglioni, 1978) to assess sentence comprehension. The poor readers made more errors than the good readers in recall of word strings and on some Token Test items. Those Token Test items that impose the greatest burden on short-term memory were the most sensitive to reader group differences; syntactic complexity alone did not distinguish the groups. The findings support other indications that poor readers make less effective use of working memory in processing spoken sentences than good readers; they do not indicate a syntactic deficit on the part of the poor readers.

Acknowledgements. We are grateful to the East Hartford, Connecticut, public school authorities for granting permission for this study. We express appreciation to the principals, the reading consultants, the third grade teachers and students at the Barnes, Center, Mayberry, Silver Lane and Stevens Schools for their cooperation.

This research was supported by NICHD Grant HD-01994 to Haskins Laboratories.

References

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

Carroll, J.B., Davies, P., and Richman, B. *The American Heritage Word Frequency Book*. Boston: Houghton Mifflin, 1971.

Daneman, M., and Carpenter, P.A. Individual differences in working memory. *Journal of Verbal Learning and Verbal Behavior, 19*: 450-466, 1980.

De Renzi, E., and Faglioni, P. Normative data and screening power of a shortened

version of the Token Test. *Cortex, 14*: 41-49, 1978.

DUNN, L.M. *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service, 1965.

HIERONYMUS, A.N., and LINDQUIST, E.F. *Iowa Test of Basic Skills*. Boston: Houghton Mifflin, 1978.

JORM, A.R. Specific reading retardation and working memory: A review. *British Journal of Psychology, 74*: 311-342, 1983.

LAPOINTE, C.M. Token Test performances by learning disabled and achieving adolescents. *British Journal of Disorders of Communication, 11*: 121-133, 1976.

LIBERMAN, I.Y., and SHANKWEILER, D. Phonology and the problems of learning to read and write. *Remedial and Special Education 6*, 8-17, 1985.

LIBERMAN, I.Y., SHANKWEILER, D., LIBERMAN, A.M., FOWLER, C.A., and FISCHER, F.W. Phonetic segmentation and recording in the beginning reader. In A.S. Reber and D.L. Scarborough (Eds.), *Toward a Psychology of Reading: The Proceedings of the CUNY Conferences*. Hillsdale, NJ: Erlbaum, 1977, pp. 207-225.

MANN, V.A. Reading skill and language skill. *Developmental Review 4*: 1-15, 1984.

MANN, V.A., LIBERMAN, I.Y., and SHANKWEILER, D. Children's memory for sentences and word strings in relation to reading ability. *Memory and Cognition, 8*: 329-335, 1980.

MANN, V.A., SHANKWEILER, D., and SMITH, S.T. The association between comprehension of spoken sentences and early reading ability: The role of phonetic representation. *Journal of Child Language, 11*: 627-643, 1984.

NOLL, J.D., and RANDOLPH, S.R. Auditory semantic, syntactic, and retention errors made by aphasic subjects on the Token Test. *Journal of Communication Disorders, 11*: 543-553, 1978.

SHANKWEILER, D., SMITH, S.T., and MANN, V.A. Repetition and comprehension of spoken sentences by reading-disabled children. *Brain and Language, 23*: 241-257, 1984.

STANOVICH, K.E. Individual differences in the cognitive processes of reading: II. Text-level processes. *Journal of Learning Disabilities, 15*: 549-554, 1982.

TALLAL, P. Perceptual and linguistic factors in the language impairment of developmental dysphasics: An experimental investigation with the Token Test. *Cortex, 11*: 196-205, 1975.

WHITEHOUSE, C.C. Token Test performance by dyslexic adolescents. *Brain and Language, 18*: 224-235, 1983.

Suzanne T. Smith, Virginia A. Mann, Donald Shankweiler, Haskins Laboratories, 270 Crown Street, New Haven, CT 06511, U.S.A.