DOCUMENT RESUME

ED 037 463 | 24 | TE 001 773

AUTHOR	Ammon, Paul R.
TITLE	Aspects of Speech Comprehension by Children and Adults: Two Experimental Studies of the Comprehension of Sentences. Final Report.
INSTITUTION	California Univ., Berkeley.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-7-I-114
PUB DATE	Sep 69
GRANT	OEG-9-8-070114-0029-(010)
NOTE	77p.
EDRS PRICE	EDRS Price MF-$0.50 HC-$3.95
DESCRIPTORS	Cognitive Measurement, *Cognitive Processes, Cognitive Tests, College Students, Elementary School Students, Expectation, *Listening Comprehension, *Memory, *Oral Communication, Reaction Time, *Sentence Structure, Time Factors (Learning)

ABSTRACT

Fifth-grade and college students in two experiments listened to tape-recorded sentences and completed tests of immediate recall. The first experiment investigated the effect of a listener's expectations upon his understanding of a sentence. Sometimes the "most probable" answer was not congruent with the "correct" answer, suggesting that the listener's expectation of what was probable influenced the manner in which he interpreted the experimental sentence. The second experiment investigated the effects of varying the delay interval from .02 to 2.0 seconds between the sentence's presentation and the memory test. Subjects either responded to questions or repeated, from memory, the word following a specified word in a sentence. Results indicated that delay caused the listener to reconstruct the wording of a sentence from a deeper interpretation of it. The results of both experiments suggested that a listener formulates hypotheses about the underlying meaning of a sentence through an initial analysis of the sentence and his own expectations of what is likely to be said, and then tests the hypotheses against syntactic cues within the sentence. Because of the effect of his own expectations and the rapid loss of essential information about the sentence, the listener may misinterpret a completely plausible sentence. (LH)

Final Report

Project No. 7-I-114

Contract No. OEG-9-8-070114-0029-(010)

ASPECTS OF SPEECH COMPREHENSION BY CHILDREN AND ADULTS

Two Experimental Studies of the Comprehension of Sentences

Paul R. Ammon

University of California

Berkeley, California

September 1969

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

ACKNOWLEDGEMENTS

The following individuals and schools were most helpful in arranging for the testing of subjects: Mrs. Helen Bacon, Mrs. Elizabeth Lott, and Dr. John Matlin; Pleasant Valley Elementary School, Novato, Calif.; Marin Elementary School, Albany, Calif.; and Whittier Elementary School, Berkeley, Calif.

The data were collected and analyzed with the able assistance of Keith Alward, Kay Byers, Stuart Horance, and Bernard Ostrowski. Dr. Joel Levin was consulted regarding some of the statistical analyses. Joan Mokray typed the final report.

The author is grateful to all concerned for their patience and diligence.

TABLE OF CONTENTS

SUMMARY

Two experiments were performed in which fifth grade children and college students listened to tape-recorded sentences and responded to tests of immediate memory. Systematic variations in the syntactic structure and the content of the sentences were found to produce significant differences in the time it took for the subjects to respond and in the frequency of correct responses. From these measures, it was possible to make inferences about the nature of the cognitive processes involved in sentence comprehension.

The first experiment investigated the effect of a listener's expectations on his understanding of a sentence. The subjects answered questions designed to see whether or not they understood "who did what to whom" in the sentences which they heard. A question was presented one half second after each sentence, and the subjects responded orally. Sometimes the most probable answer to a question was not congruent with the correct answer, as specified by the grammatical structure of the sentence. Under these conditions, the subjects in both age groups took longer to respond and gave fewer correct responses. Apparently their expectations as tc what was probable influenced the manner in which they interpreted the experimental sentences.

The second experiment investigated the effects of varying the delay interval between the presentation of a sentence and the test of immediate memory. Intervals of 0.2 seconds and 2.0 seconds were compared. Some subjects answered questions, as in the first experiment. Other subjects responded to "probe" words; that is, one word from each sentence was repeated after the

delay interval, and the subjects responded by saying the word which had followed this probe word in the sentence. The responses to probe words and to questions provide evidence concerning the ways in which a subject organizes and recodes the information in sentences. The delay interval was manipulated in order to show changes in organization and recoding during the comprehension process. It was found that subjects generally took longer to respond when there was a longer delay interval, especially if the probe word or the response was a pronoun. These results suggest that, even after a brief delay, the listener must reconstruct the exact wording of a sentence from a deeper interpretation, and that some parts of the sentence are more difficult to reconstruct than others.

The results of both experiments are consistent with an "active" model of comprehension processes. According to the model, the listener generates hypotheses about the underlying meaning of a sentence, based on a preliminary analysis of the sentence plus the listener's expectations concerning what is likely to be said. The hypotheses are then tested against other information, including syntactic cues in the sentence itself. It appears that the listener does not retain all of this additional information very long after hearing the sentence. As a result, the listener is sometimes slow to reject erroneous hypotheses, and he may even misinterpret a perfectly plausible sentence.

The same general model of comprehension seems appropriate for both of the age groups studied, but some of the effects observed in the experiments varied according to the age and sex of the subjects, indicating the need for consideration of individual differences in the comprehension process.

INTRODUCTION AND BACKGROUND TO THE RESEARCH

The two experiments reported here grew out of an earlier study of sentence comprehension (Ammon, 1968). In that study, children and adults listened to sentences and responded to a question or probe word immediately after each sentence. Response latencies were found to vary as a function of syntactic structure. It was concluded that further experimentation with the same methodology might contribute to the development of a theoretical model of the cognitive processes by which people understand sentences. Before the present experiments can be described, it will be necessary to review some conditions and findings of the original study.

Two stimulus items from the original study are shown in (1) and (2).

(1) A noisy squirrel scolded the playful children; he climbed the shady tree. (half-second pause) Who climbed the tree?

(2) The honest lawyers praised the old judge; he sent the nice letters. (pause) Who sent the letters?

The subjects (Ss) were told to answer the questions with one word as quickly as possible, and that, to be scored correct, a response had to be the same word used in the sentence, e.g., squirrel in (1). The sentences in (1) and (2) differ with respect to a variable called "separation". In (1) climbed the tree is separated from its underlying subject, squirrel, by an intervening verb phrase. This k nd of separation does not occur in (2), where judge is the subject in question.

Another way of varying separation is shown in the next two sentences.

(3) The noisy squirrel who scolded the playful children climbed the shady tree.

(4) The honest lawyers praised the old judge who sent the nice letters.

In (3) and (4) the degree of separation is determined by the location of a relative pronoun, who, rather than by the gender and/or number of a personal pronoun. The contrast between sentences containing personal vs. relative pronouns is referred to as "form".

One more variable, "confusability", is illustrated by the difference between sentences (3) and (5).

(5) The noisy squirrel who climbed the shady tree scolded the playful children.

The grammatical relationships underlying both sentences are the same, but (5) seems easier to understand, probably because tree is unlikely to be confused with the underlying subject of scolded the children. In (3), children could easily be interpreted as the subject of climbed the tree.

All three of these variables affected the latencies of correct responses. That is to say, responses were slower when the sentence had separation, relative form, or confusability. The implications of these findings will be discussed in detail below. It suffices to say at this point that response latency appears to be a sensitive indicator of the kinds of information used by listeners in processing sentences.

The five sentence-types described thus far were presented not only with questions, but also with "probe" words. That is, in another part

of the experiment, each sentence was followed by the repetition of one word from within the sentence, such as playful in (1). The Ss responded with the word which came right after the probe, e.g., children. A phrase-structure analysis of the sentences tended to predict the pattern of response latencies across probe positions within each sentence-type. Responses generally came faster within phrases than across phrase boundaries. This finding supports the hypothesis that phrases function as coding units in the processing of a sentence (see also Fodor and Bever, 1965; Johnson, 1965). Where phrase structure failed to predict the probe latencies, the effects of recency and of uncertainty about the location of who seemed to account for the deviations, possibly indicating some constraints of short-term memory on the perception of phrase structure. Sentence-types 1 through 4 above were used in the present experiments, providing an opportunity to replicate the original findings with respect to separation, form, and probe position. Two additional variables were introduced in an attempt to answer questions which had been raised by the previous results.

EXPERIMENT I

The principal variable in this experiment is called "congruence". Sentences (6) and (7) differ with respect to congruence.

(6) The friendly neighbors who helped the sick man made the good dinner.

(7) The brave firemen who saved the little kitten drank the warm milk.

In both sentences, the first noun is the underlying subject of the

final verb phrase, as indicated by the location of who. Without this grammatical cue, neighbors would still seem to be the more probable subject of made the dinner. In (7), however, the kitten seems more likely to have drunk the milk, even though the sentence says that the firemen drank the milk. The more probable interpretation and the correct interpretation, as pecified by syntax, are congruent in (6), but not in (7).

Why is congruence an interesting variable? The effect of "confusability" has already been mentioned. With a non-confusable sentence, the grammatical impossiblity of an alternate predication facilitates comprehension. Slobin (1966) reported a similar finding for "reversible" and non-reversible sentences. Ammon (1969) has suggested that the listener uses his knowledge of English selectional rules in the very process of understanding sentences; an interpretation which violates selectional restrictions -- the tree scolded the children in (5) -- will be rejected in favor of an interpretation which does not violate these restrictions. It remains to be seen whether the listener makes similar use of his expectations about what is likely to be said, when selectional restrictions permit two possible interpretations but one of them seems more probable than the other. If congruent sentences are easier to understand, this might indicate that the listener's expectations are brought to bear on the immediate process of comprehension, steering the listener toward the more probable interpretation.

Sentences (6) and (7) both have the same grammatical structure as sentence (3). In fact, sentence-types 1, 2, and 4 can also be written in a congruent or an incongruent fashion. In other words, congruence can

vary independently of separation and form, so that one can assess the interaction of congruence with these other two variables. It is possible, for instance, that the listener's expectations come into play only when another factor, like separation, increases the difficulty of interpreting the sentence.

Method

Design. The experiment was planned as a 2x2x2x2x2 factorial design, with repeated measures on the first three factors. The five factors were congruence, separation, form, sex, and age (adults vs. children). Control variables were stimulus list and presentation order (see description of materials and procedure below). The dependent variable was to be the latency of correct responses. It became necessary to depart from the planned design in some respects, as indicated in the results section.

Subjects. The Ss were 32 fifth graders and 32 college students, with each age group divided equally as to sex. The fifth grade sample was drawn from two Elementary schools serving predominantly middle-class populations in the San Francisco Bay area. The college Ss were enrolled in psychology and education courses at the University of California, Berkeley.

Stimulus materials. Each S listened to 96 sentences and responded to a question after each sentence. Half of these sentences, the experimental items, were followed by a question concerning the subject of the final verb phrase (see sentence-type 1 in the background section). In 24 of these items, noun 1 was the more probable response, while noun 2 was more probable in the other 24. These probabilities were determined by obtaining pre-experimental

judgments from comparable Ss (see Appendix A for details). Sentence-types 1 through 4 were represented equally in each set of 24 items. Overall, the combinations of congruence, separation, and form (2x2x2) yielded eight kinds of experimental items, with 6 instances of each.

The remaining 48 items were included as filler items to prevent Ss from developing a set to answer one type of question. The filler sentences were not calibrated for congruence. Half of these sentences were followed by a question about the first adjective (e.g., Who was noisy?), and half by a question about the second adjective (Who was playful?). The four sentence-types were equally represented.

Actually, there were four parallel lists of 96 items, because each sentence could be rewritten as any of the four sentence-types. The use of four lists permits partial control of the effects of sentence content while structure is varied. Each S heard only one list. Within a list, the various kinds of items appeared in quasi-random order (see Appendix B for a complete list of items and information about its construction).

All stimulus items were tape-recorded with normal sentence intonation by a male reader. The sentences and questions were read at a rate of about 150 words per minute. The word ready was said about 1 sec before each sentence and there was a delay of about 0.5 sec between the sentence and the question. A new item was begun every 15 secs, so that S had about 7 secs in which to answer each question.

Procedure. The Ss were tested individually in two sessions at least one day apart. On the first day, S heard pre-recorded instructions

and practice items (see Appendix C), and then responded to 48 of the test items. The instructions were to respond as quickly as possible with the one word from the sentence which answered the question. On the second day, the practice items were reviewed and the remaining test items were presented. Half of the Ss heard items 1-48 on day 1 and the other half heard these items on day 2. The stimulus items were presented on a Wollensak Model 1500 tape recorder, while a similar machine recorded both the stimulus items and S's responses.

Treatment of data. The tape recordings made during the experimental sessions were later played through an oscillographic recorder (Brush Recorder Mark 220) at a chart speed of 25 mm/sec. The latency of correct responses was measured in millimeters from the onset of the question to the onset of the response. (The onset of a question is sometimes easier to pinpoint than its offset.) To be scored correct, a response had to be an exact reproduction of the word in the stimulus sentence that answered the question. The raw latencies (in millimeters) were converted to reciprocals, multiplied by 1000, and rounded to the nearest whole number. Then the mean of these scores was computed for each S under each condition of the experiment. These individual cell means will be referred to as transformed latency (TL) scores.

Results and Discussion

Latency of correct responses. Fourteen of the fifth grade Ss produced no correct responses in one or more cells of the design. These Ss were dropped from the first analysis, along with the corresponding college Ss matched by sex, list, and presentation order. As a result, the total N was reduced from 64 to 36 and the experimental variables were not completely

balanced for the control variables of list and order.

Mean TL scores for the reduced samples are presented in Table 1. It is important to remember that a high score indicates a quicker response. Because the analysis of variance showed no significant effects associated with sex, the data for male and female Ss have been collapsed in Table 1. The other four main effects were all significant. That is, higher scores occurred with congruent sentences, $F(1,32) = 19.89$, $p < .001$ -- with unseparated sentences, $F(1,32) = 17.84$, $p < .001$ -- with the personal pronoun form, $F(1,32) = 4.82$, $p < .05$ -- and with college Ss, $F(1,32) = 4.99$, $p < .05$. There were no significant interactions.

The significant effects of separation, form, and age confirmed the reuults of the original study (Ammon, 1968). The relative difficulty of incongruent sentences was demonstrated for the first time. Congruence did not interact with age or the other stimulus variables, but this finding was not replicated in further analyses.

To see whether the results of the first analysis would hold up with all stimulus variables counterbalanced, a second analysis of variance was performed on the complete college sample (see Table 2). There was a significant main effect of sex, $F(1,30) = 7.28$, $p < .05$. The means for females and males were 23.0 and 20.3 respectively. However, since none of the interactions with sex approached significance, the means in Table 2 are for combined sex groups. As in the first analysis, congruence and separation produced significant effects, $F(1,30) = 24.48$, $p < .001$ and 9.21, $p < .01$. The main effect of form was not significant in this analysis, $F < 1$. On the other hand, there were significant interactions of congruence by separation, $F(1,30) = 4.59$, $p < .05$, and of

TABLE I

Mean TL Scores for Correct Responses to Question 1 as a Function of Age, Congruence, Separation, and Form.

		Fifth Grade		College	
Congruence	Form	Sep	$\overline{\text{Sep}}$	Sep	$\overline{\text{Sep}}$
Con	Per	19.8	20.7	21.9	22.8
	Rel	18.9	20.7	21.7	22.7
$\overline{\text{Con}}$	Per	17.8	19.4	20.0	20.9
	Rel	16.1	18.8	20.1	19.9

TABLE 2

Mean TL Scores for College Ss' Correct Responses to Question 1 as a Function of Congruence, Separation, and Form

	Congruent		Congruent	
Form	Sep	$\overline{\text{Sep}}$	Sep	$\overline{\text{Sep}}$
Per	22.4	23.3	20.2	21.2
Rel	21.6	23.3	20.9	20.5

congruence by separation by form, $F(1,30) = 4.17$, $p < .05$. Both of these interactions seem to result from unusually high scores for sentence-type 3 (separated relative) in the incongruent condition. The high mean in that cell reverses the expected effects of separation and form. It was hoped that an analysis of the number of correct responses would facilitate further interpretation of the latency results.

Number of correct responses. Both age groups made enough errors to permit another analysis of variance using the number of correct responses in each cell as the dependent variable. The means appear in Table 3. These means are collapsed across sex groups, due to the absence of any significant effects associated with sex. As with the latency data, there were significant effects of congruence and separation, in the expected direction. That is, more correct responses occurred with congruent and with unseparated sentences -- $F(1,60) = 156.53$ and 52.48 respectively, $p < .001$. There was a significant interaction between congruence and separation -- $F(1,60) = 38.50$, $p < .001$ -- but the nature of this interaction was quite different from that obtained in the latency analysis. Here, the effect of separation was greater for the incongruent sentences. The scores were especially low for incongruent sentence-type 3, the same condition in which unusually high TL scores were obtained for correct responses.

It is possible that each cell contains a small number of items which are easy enough to be answered quickly and correctly by most Ss under any conditions in the experiment. Items at a higher level of difficulty may be answered correctly under most conditions, with latency varying as a function of those conditions; but under the most difficult conditions (incongruent

TABLE 3

Mean Number of Correct Responses to Question 1 as a Function of Age, Congruence, Separation, and Form

		Fifth Grade		College	
Congruence	Form	Sep	$\overline{\text{Sep}}$	Sep	$\overline{\text{Sep}}$
Con	Per	4.03	4.91	5.78	5.56
	Rel	4.06	4.72	5.34	5.66
$\overline{\text{Con}}$	Per	1.84	3.38	4.31	4.72
	Rel	1.53	3.44	3.72	5.00

separated relative), these items are answered incorrectly, leaving only the very easy items to contribute to the latency score.

The form effect was not significant, as in the latency analysis for the complete college sample, but there was a significant interaction between form and separation in the present analysis. The lowest scores occurred with type 3 sentences (separated relatives) and the highest scores with type 4 (unseparated relatives). A similar interaction was predicted but not observed in the original study (Ammon, 1968).

The overall mean number of correct responses was higher for college students than for fifth graders, $F(1,60) = 51.65$, $p < .001$. Furthermore, age interacted significantly with congruence and with separation -- $F(1,60) = 9.19$ and 11.70, $p < .01$. These interactions reflect the fact that congruence and separation produced larger differences with children than with adults. Because of the interactions with age, it seemed advisable to do separate analyses within age groups. The error terms for these analyses were estimated for each group separately, because of significantly greater variance in the fifth grade group. To summarize the results of these separate analyses briefly, the main effects of congruence and separation were significant for both age groups, as well as the interaction of these two variables. The separation-by-form interaction was found only with the college Ss.

The effects of congruence and separation suggest that two kinds of perceptual bias operate in the comprehension of the experimental sentences. Apparently the listener tends to perceive as the subject of the final verb phrase the noun which is more probable and/or the noun which is closer to the verb phrase. When the actual structure of the sentence is not consistent

with these biases, the listener must overcome his biases in order to comprehend accurately. Thus he either takes longer to respond or he makes an error.

Errors. The foregoing explanation of the data leads to the prediction that a large percentage of the errors should be ones in which S responds with the alternative noun in the stimulus sentence, especially with incongruent and/or separated sentences. "Alternative" errors did, in fact, account for 62% of all the errors in the experiment. (The other 38% consisted of miscellaneous error-types, such as no responses, changes from singular to plural or vice versa, synonyms, and other intrusions.)

Table 4 shows the mean number of alternative (A) errors as a function of the sentence variables and age. It is obvious from this table that A errors were much more frequent for incongruent sentences. In fact, the greater frequency of total errors with incongruency is due entirely to A errors; the number of miscellaneous errors was equal in the congruent and the incongruent conditions. Table 4 also shows that, with one exception, A errors were more frequent for separated sentences, other things being equal. The means for congruent sentences are so low that the differences among those means are probably unreliable. But the means for incongruent sentences show quite clearly the higher frequency of A errors with separation. It is not true, however, that A errors alone account for the effect of separation on total correct responses; the number of miscellaneous errors also increases with separation. Apparently, then, the separation effect reflects more than a bias toward the closer noun, although the A errors do seem to show such a bias, especially in the incongruent condition.

TABLE 4

Mean Number of 'A' Errors on Question 1 as a Function of Age, Congruence, Separation, and Form

		Fifth Grade		College	
Congruence	Form	Sep	$\overline{\text{Sep}}$	Sep	$\overline{\text{Sep}}$
Con	Per	.41	.25	.09	.25
	Rel	.81	.19	.38	.09
$\overline{\text{Con}}$	Per	3.03	1.59	1.38	1.00
	Rel	3.31	1.72	1.78	.94

Latency of errors. When S makes an A error in response to an incongruent sentence, has he actually perceived it as a congruent sentence? Obviously the response itself seems to indicate that such is the case. But is the process by which S arrives at the response the same process involved in the correct perception of a congruent sentence? If so, the latency of A errors ought to be about the same as the latency of correct responses to those sentence-types for which the A errors would be correct.

Table 5 shows the mean TL scores for A errors made to each sentence-type in the incongruent condition, along with the mean TL for correct responses in the same cell and the mean TL for correct responses in the corresponding congurent condition, i.e., the sentence-type for which the A errors would be correct. (The N is different for each triad of means because the means are based only on those Ss who produced scorable responses in all three categories in each cell.) In every case the mean for A errors is considerably lower than the mean for correct responses in the congruent condition where the A errors would be correct. Moreover, in all cases but two, the mean TL of A errors is considerably lower than the mean for correct responses to the same incongruent sentence-types.

These results suggest that, in making A errors, the Ss did not process incongruent sentences in the same way as congruent sentences. They did not simply leap to an erroneous conclusion about the subject of the final verb phrase. Their slowness in producing A errors may indicate a cognitive conflict between their expectations and the grammatical cues in the sentence. When the conflict was resolved in favor of the expectation, an A error occurred. What is surprising is the frequency with which the

TABLE 5

Mean TL Scores for 'A' Errors in the Non-Congruent Condition

Compared with Mean TL Scores for Correct Responses

		Fifth Grade			College		
Separation	Form	Correct	A Error	A Correct	Correct	A Error	A Correct
Sep	Per	17.7	18.2 (N=27)	20.6[a]	19.4	18.1 (N=23)	22.9
	Rel	15.9	18.0 (N=22)	20.8[b]	20.0	16.6 (N=23)	23.0
$\overline{\text{Sep}}$	Per	19.0	16.9 (N=27)	19.5[c]	21.6	18.6 (N=18)	22.5
	Rel	19.3	16.9 (N=25)	19.0[d]	19.3	17.4 (N=15)	21.2

a = Con, $\overline{\text{Sep}}$, Per, Correct

b = Con, $\overline{\text{Sep}}$, Rel, Correct

c = Con, Sep, Per, Correct

d = Con, Sep, Per, Correct

grammatical cues lost out in the conflict, even though the Ss knew they had to listen carefully. Perhaps the listener considers the correct interpretation of an incongruent sentence but he just can't believe his ears. The information which is needed to reassure the listener that he can believe his ears must be lost quite rapidly from immediate memory.

The only part of Table 5 showing evidence that Ss simply ignored the grammatical cues is in the fifth grade data for separated sentences. The fact that A errors occurred somewhat faster than correct responses might indicate that sometimes the younger Ss did leap to erroneous conclusions. It should be noted that these are the cells in which both biases -- noun probability and noun position -- are pitted against the grammatical cues. The interactions of age with congruence and with separation reported earlier suggest that children are less able than adults to suspend their heuristic biases or, alternatively, that they are less able to retain the grammatical information. Of course these are not mutually exclusive interpretations.

Conclusions

Experiment I has produced evidence that two conditions inhibit the accurate perception of a subject-predicate relationship in a sentence. The relationship is more difficult to perceive (1) if it does not seem as probable to the listener as an alternate relationship, and (2) if the subject and predicate are separated by another phrase containing a noun which might be mistakenly interpreted as the subject. The evidence concerning a third condition -- personal vs. relative pronouns -- is ambiguous and will not be discussed further right now.

It seems reasonable to interpret these results in terms of the listener's expectations, which he applies actively in processing a sentence. That is, the listener apparently is set to perceive subject-predicate relationships between consecutive noun phrases and verb phrases. He also seems to draw upon his expectations concerning what is likely to be said and, in doing this, he uses semantic information of a sort which goes beyond the selectional rules in recent theoretical analyses of linguistic competence (Katz and Fodor, 1963; Chomsky, 1965). When these expectations are not congruent with the grammatical cues actually present on the "surface" of the sentence, the resulting conflict slows down the comprehension process and may lead to a misinterpretation of the sentence's underlying meaning.

The emerging picture of the comprehension process shows that the various levels of sentence analysis interact in very complex ways. A model of sentence comprehension must utlimately show how the listener's expectations are employed relative to each other and to the grammatical cues in the sentence. But a lot more data are needed to fill in such details of the model. The structure and content of the stimuli should be varied in many other ways. The task can also be varied, with stimuli held constant. (For example, it would have been interesting to look for effects of congruence on performance of the probe task with the sentences used in Experiment I, but it was not possible to do that experiment during the present project.) Individual differences in processing remain largely unexplored. Finally, the temporal aspects of the stimuli can be manipulated, as in Experiment II.

EXPERIMENT II

Studies like Experiment I proceed from the assumption that response latency and errors in a test of immediate memory will reflect the way a listener organizes and recodes the information in sentences. Presumably the processes of organizing and recoding take time, just as the stimulus itself unfolds over time. Thus the listener has a different perception of the sentence at different points in time. It follows, then, that latencies and errors should vary as a function of the time at which the listener's perception of the sentence is tested. Experiment II was a first attempt at varying the delay between the end of the stimulus sentence and the test of immediate memory.

In the original study (Ammon, 1968) and in Experiment I, the delay after the sentence was held more or less constant at about half a second. For the present study, two delay intervals were chosen so that one was shorter and the other longer than half a second. It seemed desirable to make the short delay as close to zero as possible without interfering with the sentence itself. The long delay interval was selected to provide a contrast which would not drastically alter the nature of the task to be performed. On the basis of subjective judgment, delay intervals of 0.2 and 2.0 seconds were selected.

Actually, two experiments were performed on delay -- one with questions like those used in Experiment I, and one with the probe words described in the background section. The research question was, "does a difference in delay interval have a measurable effect on responses to questions or probes?" A main effect of delay, by itself, would not be very

interesting, because it would not show changes over time in the listener's way of organizing the sentence. Thus a more precise statement of the research question is, "does the amount of delay interact with other stimulus variables which have been manipulated systematically?" Without an explicit theory, it would be arbitrary to predict specific interactions, so no such predictions were made for this exploratory study. Experiment II also permits further replication of previous findings.

Method: Experiment II-A

Design. The study assessing the effects of delay interval on question responses comprised a 2x2x2x2x2 factorial design, with repeated measures on the first two factors. The five factors were separation, form, delay (0.2 sec vs. 2.0 sec), age (children vs. adults), and sex. All of these factors were counterbalanced for two presentation orders. The dependent variables were the latency of correct responses and the percentage of correct responses.

Subjects. The Ss were 16 fifth graders and 16 college students, drawn from the populations described under Experiment I above. Half of the Ss in each age group received the 0.2 sec delay and half received the 2.0 sec delay. Each subgroup contained an equal number of males and females.

Stimulus materials. Each S listened to the 96 sentences from Experiment I and responded to a question after each sentence. As in Experiment I, there were four sentence-types and three question-types, making twelve different types of item altogether. (The three types of questions referred respectively to the noun which was the subject of the final verb phrase, to the noun modified by the first adjective, and to the noun modified

by the second adjective.) The design called for eight replications of each item-type with different content. Although congruence was not intended to be an experimental variable in the present study, the results of Experiment I suggested that the eight items within each cell should be balanced for congruence. Thus each cell was to contain two congruent items, two incongruent items, and four items for which the congruence was unknown (the "filler" sentences from Experiment I). Appendix D contains the complete list of stimulus items. a few items were misplaced, so that not all cells in the design had the same number of items, as Appendix D shows. Unlike Experiment I, all Ss in this study heard the same list of items.

All stimulus items were tape-recorded in a manner similar to Experiment I. Two tapes were made -- one containing a delay of approximately 0.2 sec after each sentence, the other containing 2.0 sec delays after each sentence. The tapes were essentially the same in all other respects.

Procedure. The procedure was identical to the procedure used in Experiment I.

Treatment of data. Transformed latency (TL) scores were derived in the same manner as in Experiment I. Because the number of items was not equal in all cells, it was necessary to use percent correct, rather than number correct, as the second dependent variable. Only the responses to question-type 1 were scored for the present analyses.

Results and Discussion

Latency of correct responses. The mean TL scores for correct responses to question 1 are presented in Table 6. An analysis of variance showed three significant effects associated with the delay interval ($p < .05$, unless otherwise noted). Subjects receiving the shorter delay interval (0.2 sec) had significantly higher scores -- $F(1,24) = 4.55$. There was a significant interaction of delay with form, such that sentences with the relative form had a higher mean than sentences with the personal pronoun form in the short delay condition, but there was no difference between forms in the long delay condition -- $F(1,24) = 5.24$. The three-way interaction of delay, form, snd sex was also significant -- $F(1,24) = 4.83$. For female Ss, the superiority of the relative form was evident in both delay conditions, but for the males this difference was present only in the short delay condition and was actually reversed with the long delay.

Two other effects were significant. There was a main effect of form, such that higher scores occurred with sentences of the relative form -- $F(1,24) = 5.24$. And the college Ss had significantly higher scores than the fifth graders -- $F(1,24) = 9.91$, $p < .01$. This kind of age difference has appeared in all previous analyses and needs no further comment here.

The results concerning delay are difficult to interpret. The size of the delay interval was found to interact with form, but this interaction is attributable entirely to the sex difference which showed up as a three-way interaction. Furthermore, although the variable of form seems to be implicated, there is some reason to doubt that form is actually the effective factor. In past experiments, the significant main effects of form

TABLE 6

Mean TL Scores for Correct Responses to Question 1

as a Function of Delay, Age, Sex, Separation, and Form,

			Fifth Grade		College	
Sex	Delay	Form	Sep	$\overline{Sep}$	Sep	$\overline{Sep}$
F	0.2	Per	17.8	18.8	22.0	22.0
		Rel	18.8	20.3	21.5	23.5
	2.0	Per	17.3	17.3	20.8	20.8
		Rel	20.3	18.5	19.8	20.8
M	0.2	Per	19.8	19.8	21.5	22.8
		Rel	22.0	22.3	23.0	25.5
	2.0	Per	18.0	17.3	22.8	20.8
		Rel	16.3	18.0	20.8	20.5

have always been in the opposite direction, with faster responses occurring after the personal pronoun form. Only the male Ss in the long delay condition followed this pattern in the present experiment. It is surprising, too, that the separation effect, which has always been significant in past analyses, was not significant in this one. These unusual findings with regard to form and separation suggest that the use of only one stimulus list with all Ss may have precluded adequate control over other stimulus variables affecting latency. Thus, if the delay interval did interact jointly with sex and with properties of the stimulus, it is not clear what the stimulus variable(s) might have been.

Once again, the frequency of correct responses was analyzed in order to clarify the meaning of the latency data.

Percentage of correct responses. Table 7 shows the mean percentage of correct responses in each condition of Experiment II-A. There were no significant effects associated with delay. However, there were some results which are pertinent to the preceding analyses. A higher percentage of correct responses occurred with unseparated sentences than with separated sentences -- $F(1,24) = 19.66$, $p < .001$. Form interacted significantly with separation -- $F(1,24) = 4.33$. In this interaction, the superiority of unseparated sentences was greater for the relative form, the lowest and the highest means occurring with sentence-types 3 and 4 respectively.

Both the separation effect and the interaction of form with separation are consistent with previous findings (see Experiment I). It seems likely, then, that the apparent effect of form on TL scores was, in fact, due to other, uncontrolled stimulus variables affecting only latency.

TABLE 7

Mean Percentage of Correct Responses to Question 1 as a Function of Delay, Age, Sex, Separation, and Form

			Fifth Grade		College	
Sex	Delay	Form	Sep	$\overline{\text{Sep}}$	Sep	$\overline{\text{Sep}}$
F	0.2	Per	53	47	83	88
		Rel	43	66	78	90
	2.0	Per	50	66	86	84
		Rel	39	72	89	84
M	0.2	Per	67	68	75	84
		Rel	57	71	64	84
	2.0	Per	72	66	70	94
		Rel	53	63	79	90

It is interesting, too, that there was no main effect of delay on percentage of correct responses. The delay effect on latency may, therefore, simply indicate a "pacing" phenomenon. In other words, when the experimenter takes his time in presenting a question, the subjects take their time in responding, but this does not affect the accuracy of their responses.

As usual, adults had more correct responses than children -- $F(1,24)=36.76$, $p < .001$. There were also interactions of age with sex, and of age with sex and separation -- $F(1,24) = 4.27$ and 7.74, respectively.

Delay in relation to congruence. Although congruence was not included as an experimental variable, its presence as a control variable permitted the analysis of delay in relation to congruence with reasonable control over the other stimulus variables which were manipulated. For each S, a TL score and a percentage correct were computed for all congruent items and for all incongruent items. The means for these variables are shown in Tables 3 and 9.

In addition to the usual age effect, analysis of the latency data showed main effects of delay and congruence similar to those found in previous analyses -- $F(1,24) = 5.76$, $p < .05$, and 44.61, $p < .001$. Among the female Ss, congruence had a greater effect in the long delay condition. For males, the congruence effect was greater wtth a short delay. This interaction will be interpreted in a later discussion.

As for the percentage of correct responses, the only significant effects were those of congruence and of age -- $F(1,24) = 72.47$ and 14.88, respectively, $p < .001$.

TABLE 8

Mean TL Scores for Correct Responses to Question 1 as a Function of Delay, Congruence, Age, and Sex

		Fifth Grade		College	
Sex	Delay	Con	$\overline{\text{Con}}$	Con	$\overline{\text{Con}}$
F	0.2	20.8	19.0	24.3	21.8
	2.0	21.3	13.3	22.8	16.5
M	0.2	22.3	17.8	24.0	18.8
	2.0	18.0	16.3	21.5	19.3

TABLE 9

Mean Percentage of Correct Responses to Question 1

as a Function of Delay, Congruence, Age, and Sex

		Fifth Grade		College	
Sex	Delay	Con	$\overline{\text{Con}}$	Con	$\overline{\text{Con}}$
F	0.2	61	32	95	64
	2.0	84	28	89	68
M	0.2	86	28	89	50
	2.0	84	46	89	57

Method: Experiment II-B

Design. The study relating delay interval to performance of the probe task proceeded from a 4x6x2x2x2 factorial design, with repeated measures on the first two factors. The five factors were sentence-type, position of the probe word, delay, age, and sex. The data from the younger age group (fifth grade) contained too many errors to be analyzed, so age was not a variable in the data to be reported here. The overall design was broken down for three further analyses, as explained in the discussion of the results. All independent variables were counterbalanced for two presentation orders. The dependent variable was the latency of correct responses.

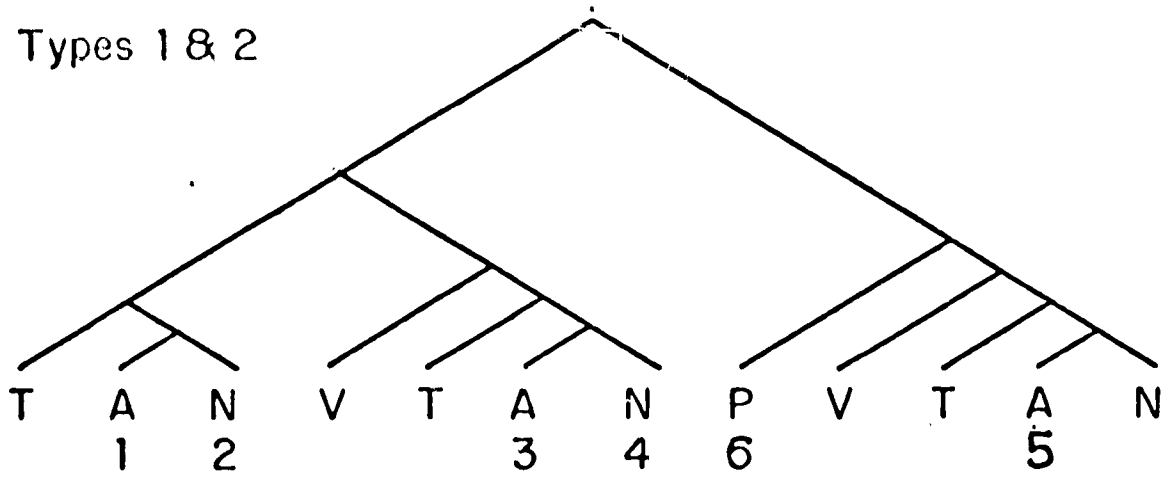
Subjects. The Ss were 16 students enrolled in education and psychology courses at the University of California, Berkeley. Half of the Ss received the 0.2 sec delay and half received the 2.0 sec delay. Each subgroup contained four males and four females.

Stimulus materials. Each S listened to the 96 sentences used in Experiment II-A, but responded to a probe word rather than a question after each sentence. The plan was to select the probe words from six positions in each of the four sentence-types. These positions are indicated by their numbers in Figure 1. A completely balanced set of materials would contain four instances of each type-by-probe combination. Once again, however, there were some slight departures from a completely balanced set (see Appendix D).

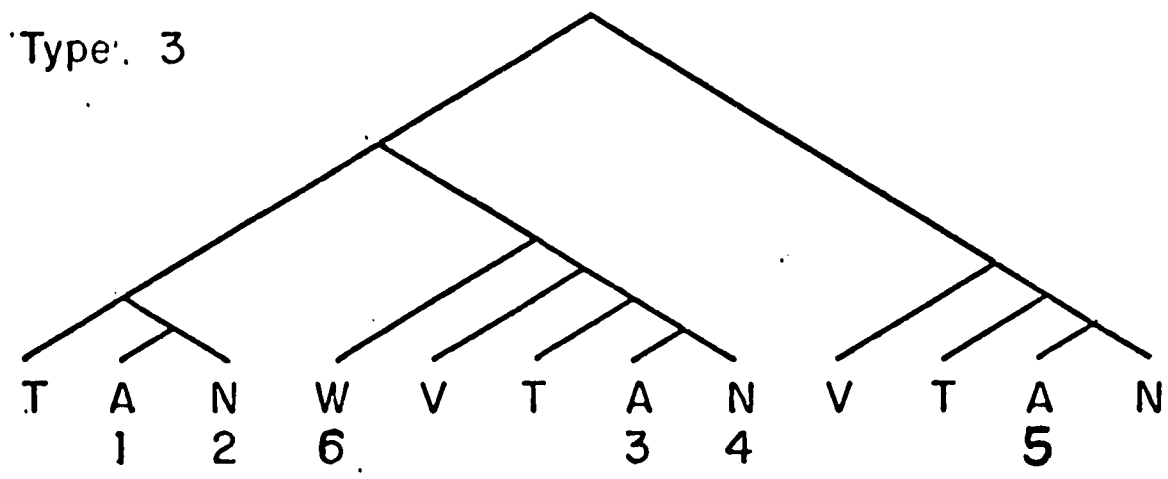
The stimulus list was tape-recorded as in Experiment II-A, except for the substitution of probe words for questions.

Procedure. The procedure was identical to the procedure used in Experiments I and II-A.

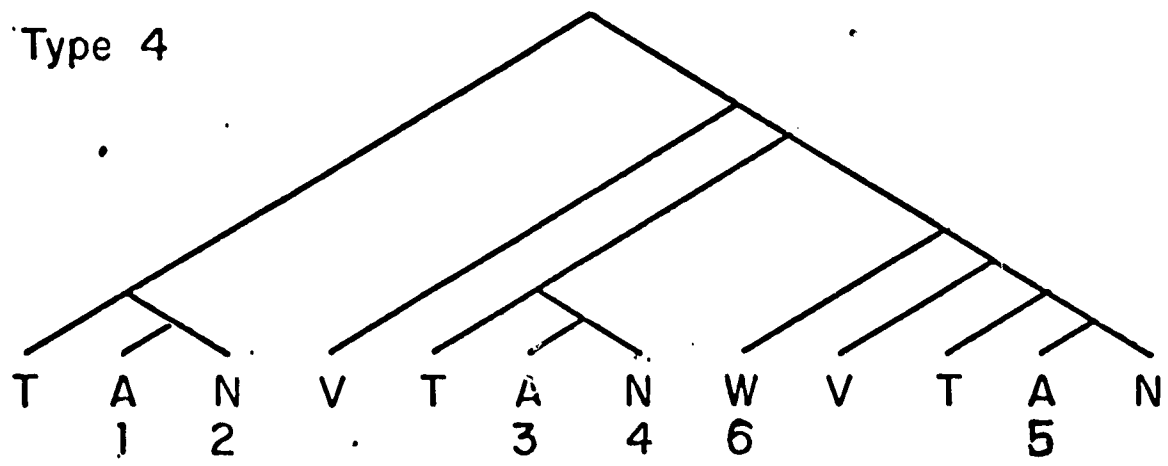
Types 1 & 2



Type 3



Type 4



T = the, a, some

A = adjective

N = noun

V = verb

P = personal pronoun

W = who

1—6 = probe positions

FIG. 1. Tree diagrams of phrase structure for each sentence-type, showing probe positions.

Treatment of data. Transformed latency (TL) scores were derived in the same way as in Experiments I and II-A.

Results and Discussion

All sentence-types. The mean TL scores for each condition of the experiment are shown in Table 10. In the first overall analysis, sentence-type was treated as a four-level factor. There were two significant effects associated with delay. One was the interaction of delay with probe position -- $F(5,60) = 2.37$, $p < .05$. The profiles in Figure 2 suggest that the interaction reflects relatively slow responses to probe 6 when the delay interval was increased from 0.2 sec to 2.0 sec. The sixth probe position in each sentence-type was occupied by a pronoun. Since pronouns are not represented in the semantic interpretation of a sentence, information about their location may be lost rather quickly from immediate memory, once the listener has used this information for determining the basic grammatical relations in the sentence. Thus, when a pronoun is presented as a probe word after a longer delay, it is relatively difficult for S to retrieve the appropriate response.

A second factor causing the interaction of delay and probe position might have been recency. Subjects can sometimes give a quick, echoic response to a probe word, without having done much analysis of the sentence. The probability of such an echoic response seems likely to be greatest for words near the end of the sentence because of their recency, especially in the short delay condition. If the delay interval is increased, however, the advantage of recent words over others ought to be reduced, resulting in an interaction of delay interval and probe position. In all sentence-types except type 3, the pronoun was in the next to the last probe position. The

TABLE 10

Mean TL Scores of College Ss in the Probe Task

Sex	Delay	Type	Probe Position 1	2	3	4	5	6
F	0.2	1	37.8	32.5	33.0	31.8	38.3	40.8
		2	34.5	32.3	34.3	32.0	45.0	46.0
		3	40.5	31.5	36.0	30.0	38.5	36.5
		4	34.3	28.5	43.0	32.5	40.3	41.8
	2.0	1	41.0	29.8	34.8	28.5	38.5	39.8
		2	45.3	33.8	36.0	33.5	38.0	39.0
		3	44.3	29.0	35.3	27.0	42.5	28.5
		4	37.8	27.0	38.5	30.0	42.0	37.3
M	0.2	1	41.8	27.8	41.0	28.0	41.8	49.8
		2	43.8	31.8	40.0	27.0	44.8	46.8
		3	44.8	29.0	44.0	26.3	45.3	41.5
		4	43.3	31.8	41.0	29.0	46.8	47.5
	2.0	1	35.3	22.8	34.8	24.0	33.8	36.8
		2	32.0	24.5	33.8	29.5	37.0	36.0
		3	34.5	23.5	32.5	24.0	34.3	28.3
		4	29.5	24.8	27.5	25.3	34.3	31.0

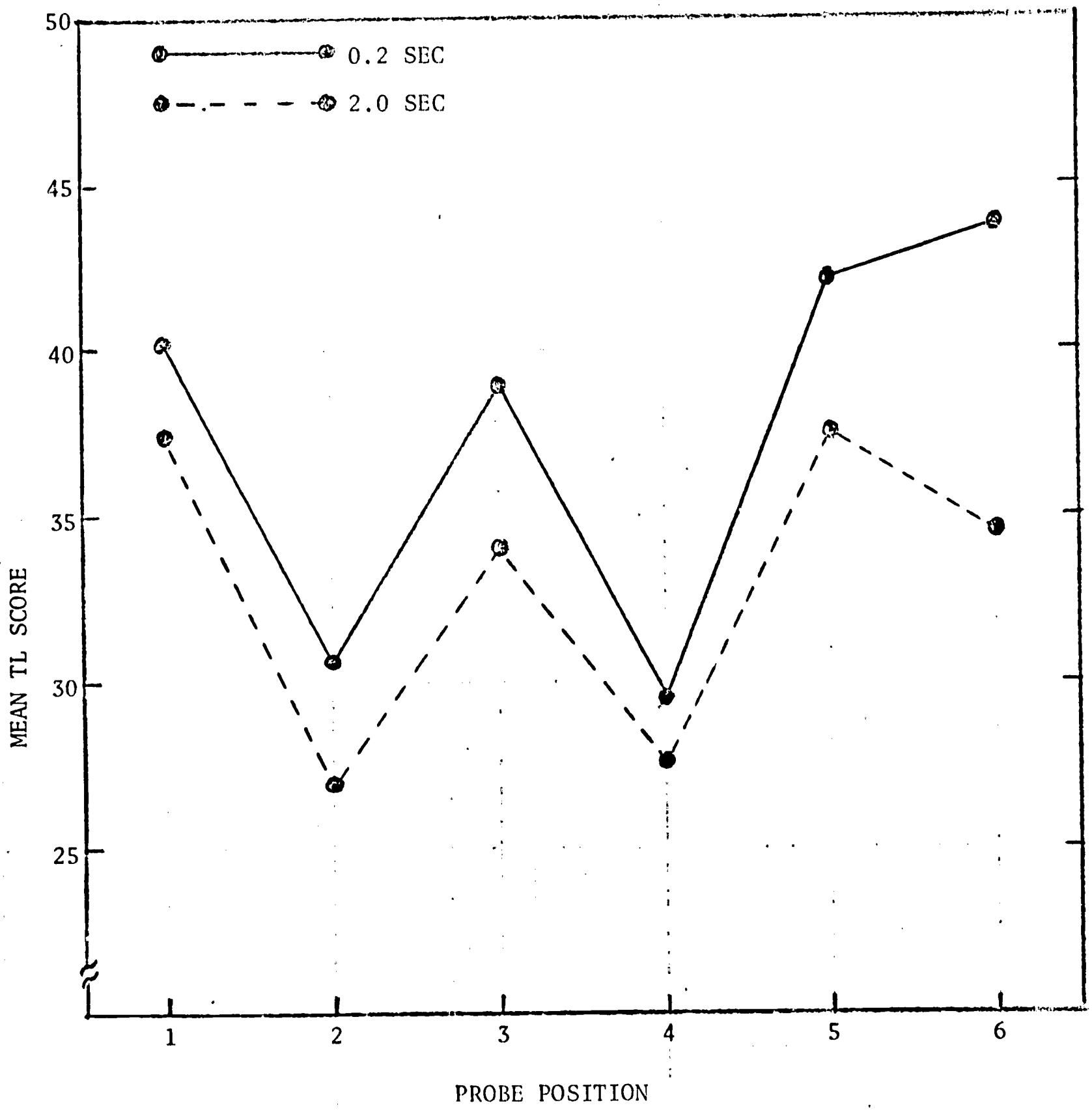


FIG. 2. Mean transformed latency score as a function of probe position, all sentence-types combined.

pronouns (probe 6) would therefore have shown more of the recency effect than words in earlier probe positions. This recency hypothesis is supported by the fact that probe position 5 -- always the last in the sentence -- showed the second largest difference between the long and short delay groups.

The foregoing interpretations seem plausible, but they must be regarded as tentative at best, in light of two additional facts. First of all, the Scheffé method of post hoc comparisons showed no significant simple contrasts when the delay differences at various probe positions were compared. The contrast which came closest to significance was that between the largest difference (probe 6) and the two smallest differences (probes 1 and 4) -- a difference of 7.02 where a difference of 7.15 was needed for significance. Secondly, there was a significant three-way interaction of delay, probe, and sex -- $F(5,60) = 2.40$. Inspection of Figure 3 indicates that, in the short delay condition, the females had a much flatter profile than the males, but, with the long delay, the profiles of males and females were roughly parallel. The three-way interaction may account, in part, for the two-way interaction of delay and probe. It is interesting, however, that the large difference between delay conditions at probe position 6 was still evident within both sex groups.

The other significant effects in the overall analysis were sentence-type: $F(3,36) = 3.33$ -- probe position: $F(5,60) = 38.50$, $p < .001$ -- and the interaction of sentence-type and probe: $F(15,180) = 3.08$, $p < .001$. In order to facilitate further interpretation of these results, the data were broken down for separate analyses involving each of the phrase structures illustrated by tree diagrams in Figure 1. While these analyses revealed

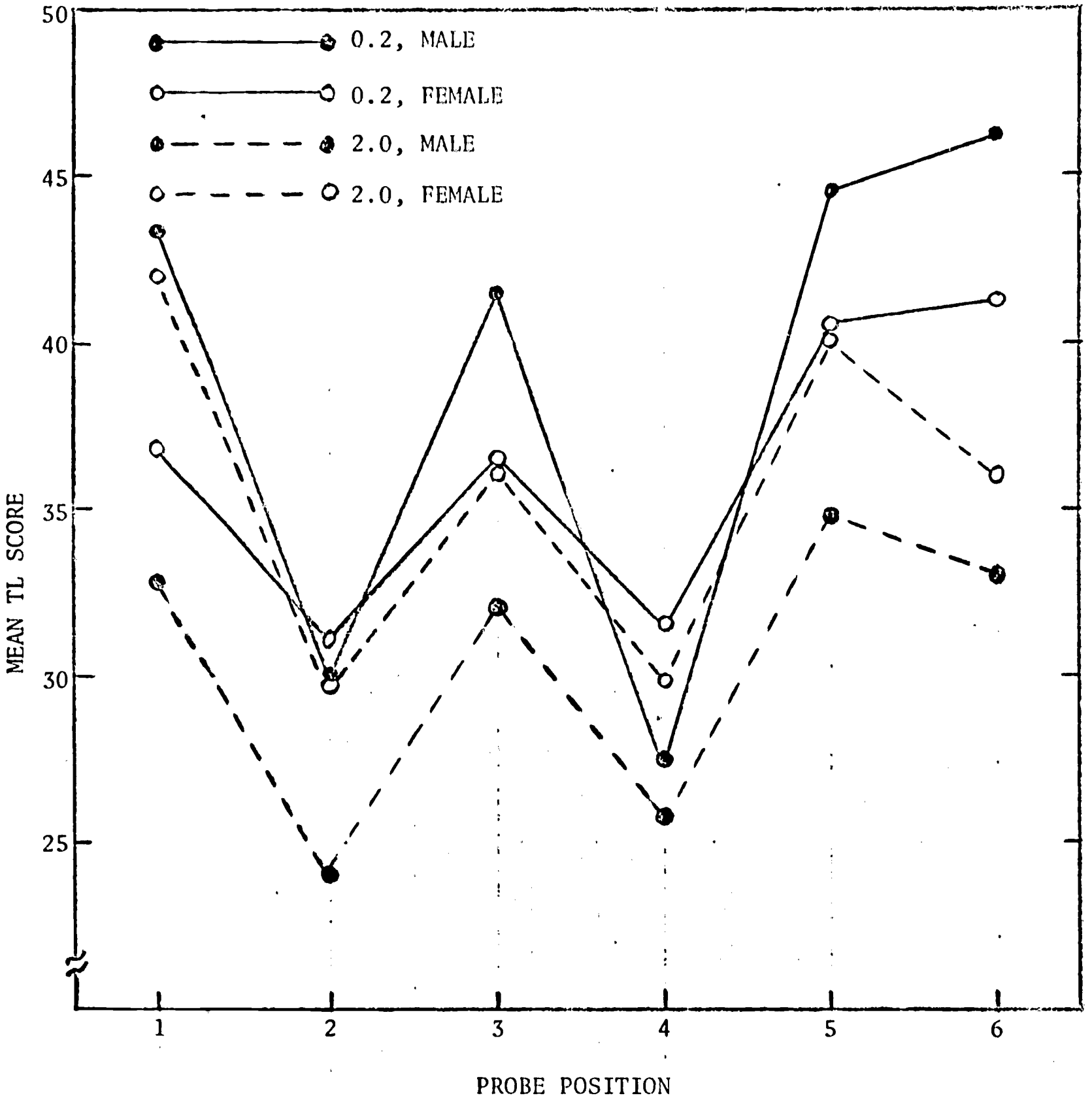


FIG. 3. Mean transformed latency score as a function of probe position, all sentence-types combined, showing interaction of delay interval, sex, and probe position.

no significant effects of delay, they are nonetheless worth reporting for the sake of comparison with previous findings.

Sentence-types 1 and 2. The first two sentence-types have the same phrase structure, but differ in terms of separation, which affects only deep structure in this case. The means for types 1 and 2 are plotted in Figure 4 as a function of delay and probe position. The analysis of variance showed a significant main effect of separation, with higher means occurring in the unseparated condition, type 2 -- $F(1,12) = 5.14$. The same effect was found in the original study (Ammon, 1968).

There was a main effect of probe position -- $F(5,60) = 30.27$, $p < .001$. Pair-wise contrasts by the Scheffé method showed that the overall means for probes 1, 3, 5, and 6 were significantly higher than the means for probes 2 and 4 (the critical value was 4.8). All of these differences are predicted by a processing model based on phrase structure and explained by Ammon (1968). According to the model, TL scores should decrease as an inverse function of the number of "nodes" (or intersections) encountered when one traces a path from the probe word to the response word in the tree diagrams in Figure 1. The Scheffé contrasts also showed the mean for probe 6 to be significantly greater than the mean for probe 3. In this case, the phrase-structure model predicts a difference in the opposite direction.

The effect of probe position interacted significantly with sex -- $F(5,60) = 2.54$. Once again, this interaction seems to be a matter of the female Ss producing a flatter profile of means across probe positions (see Figure 5). Scheffé contrasts showed a significant difference between the sex differences at probes 2 and 4 vs. probes 3 and 6 (the critical value

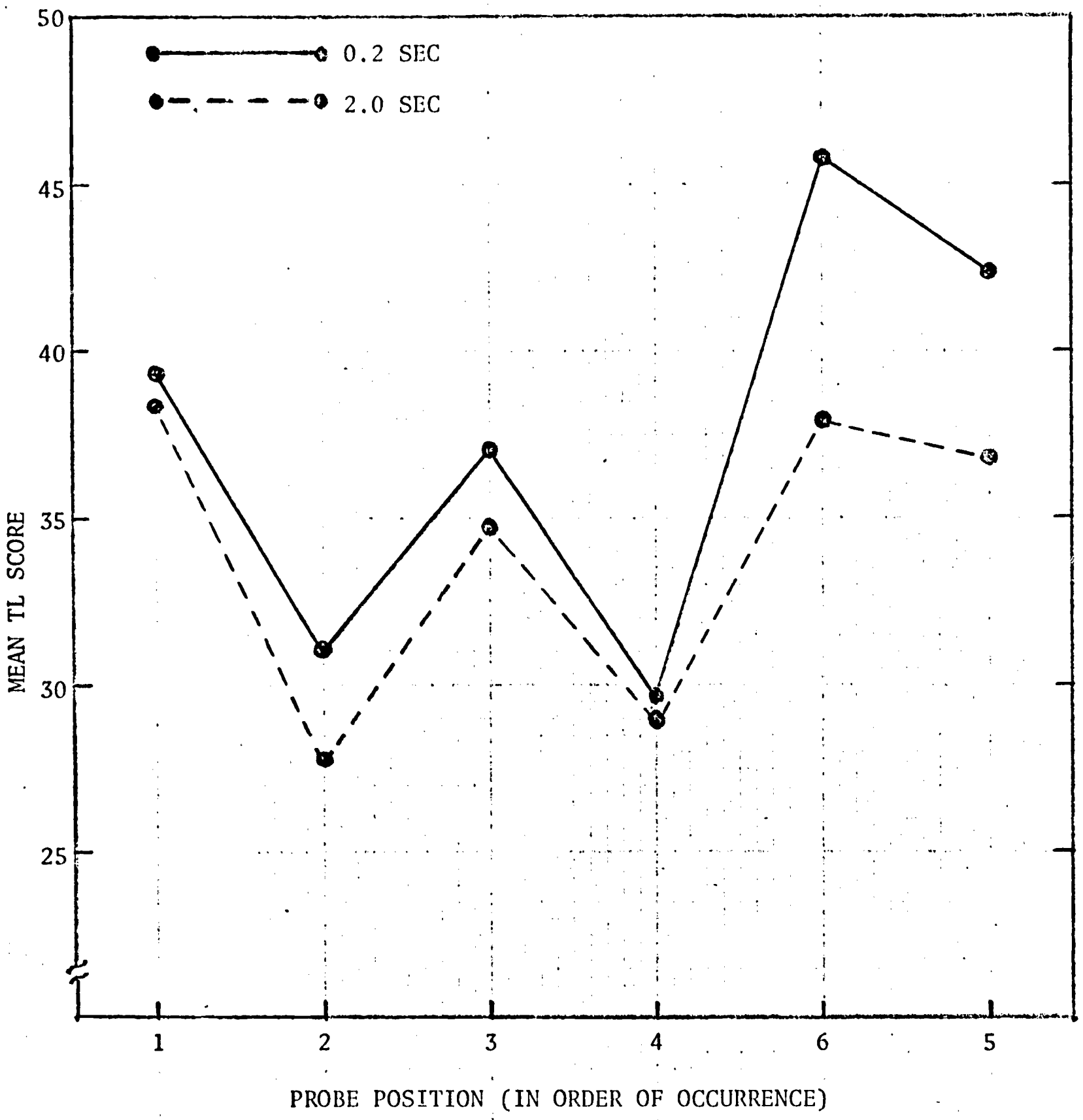


FIG. 4. Mean transformed latency score as a function of probe position, sentence-types 1 and 2 combined.

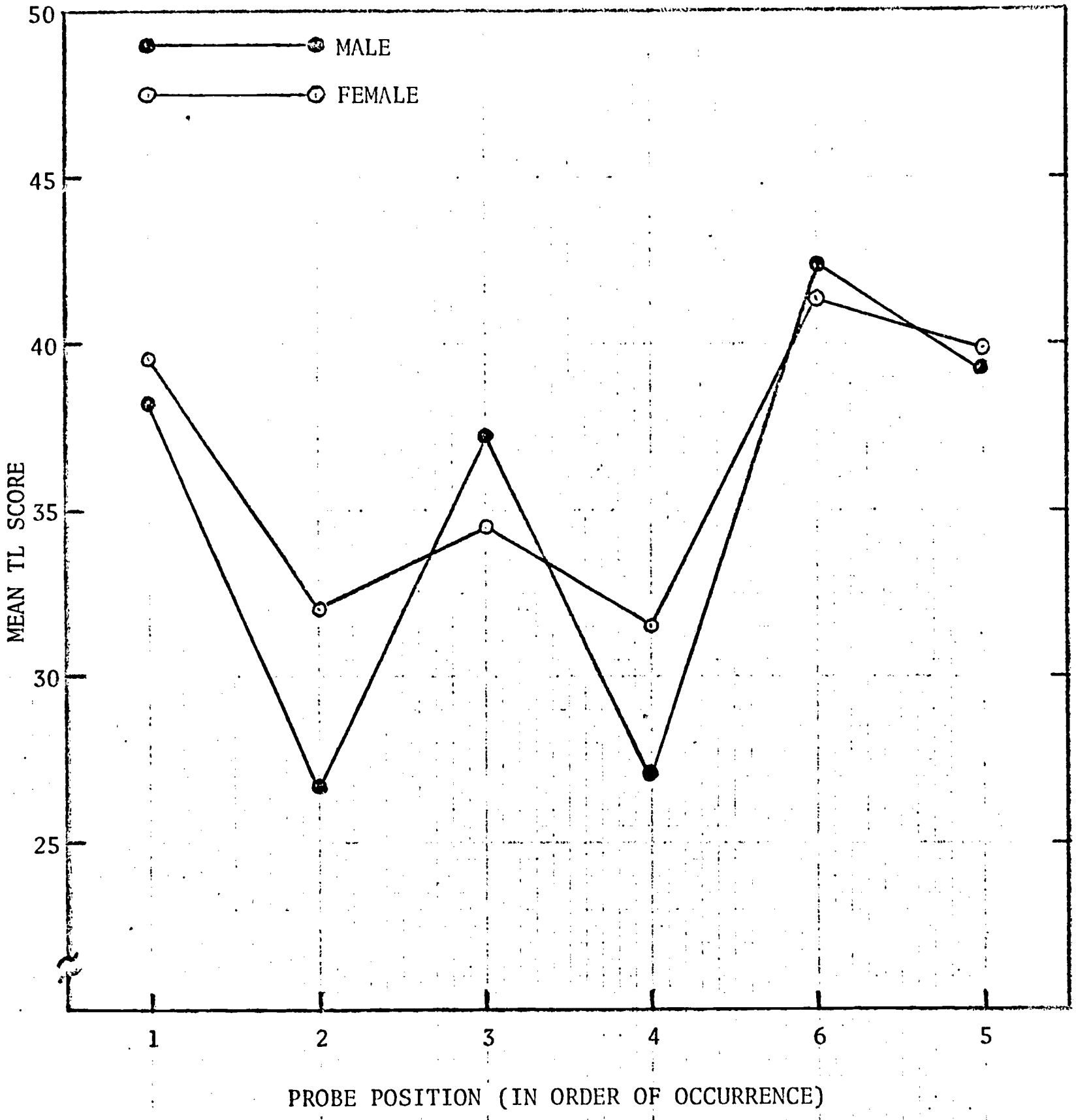


FIG. 5. Mean transformed latency score as a function of probe position, type types 1 and 2 combined, showing interaction of sex and probe position.

equalled 6.7). In other words, the two largest differences in favor of the females were significantly different from the two differences in favor of the males.

The mean for females at probe 3 seems unusually low, perhaps reflecting a simple serial position effect, with performance higher at the ends of the sentence than in the middle. This low mean may account, in part, for the unexpected significant difference between probes 3 and 6 reported above. However, the mean for probe 6 was also higher than expected. Probes 1, 3, and 5 all involve a one-node transition, whereas probe 6 has two nodes between stimulus and response. Thus probe 6 should have had the lowest mean of the four, but, in fact, it had the highest. To some extent, this may be an artifact of measuring response latency from the onset of the probe word. Such a measurement includes the length of the probe word itself. Since probe 6 was always a pronoun, and all pronouns were monosyllabic, the average length of the probe word in position 6 was probably shorter than in the other positions. The mean TL score for 6 would, therefore, have been increased in relation to the others.

Sentence-type 3. A second phrase structure is represented by type-3 sentences. The means for type 3 are shown in Figure 6. In the analysis of variance, only the main effect of probe position was significant -- $F(5,60) = 19.86$, $p < .001$. Pair-wise Scheffé contrasts (with a critical value of 6.5) showed that the means for probes 1, 3, and 5 were significantly greater than for probes 2 and 4, that 1 was greater than 6, and that 6 was greater than 4. All of these differences were predicted by the phrase-structure model, but some of the predicted differences were not significant.

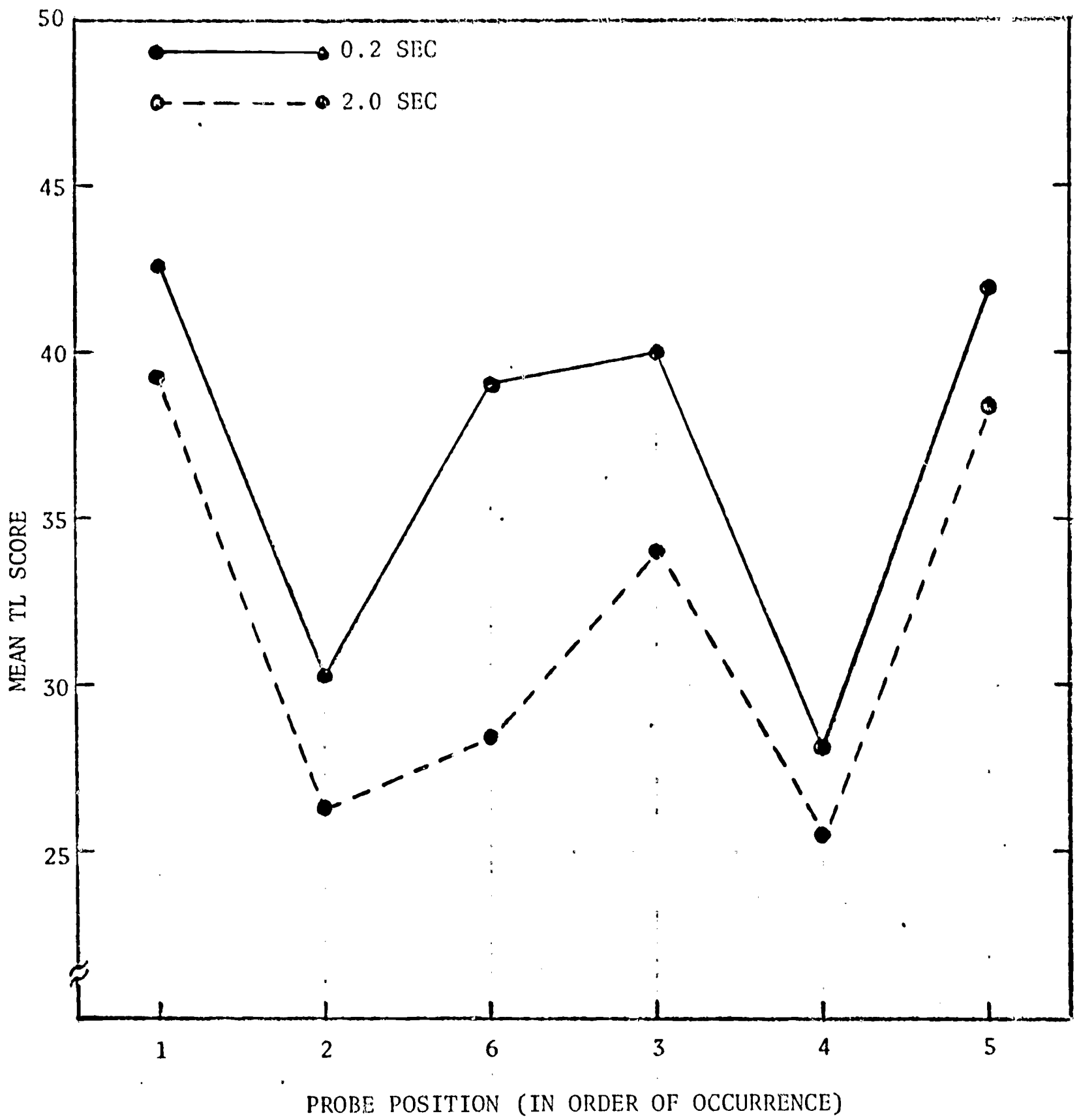


FIG. 6. Mean transformed latency score as a function of probe position, sentence-type 3.

Sentence-type 4. Means for the last phrase structure, type 4, are presented in Figure 7. Again, only the main effect of probe was significant -- $F(5,60) = 14.27$, $p < .001$. Pair-wise Scheffé contrasts (with a critical value of 6.8) showed that the means for probes 1, 3, 5, and 6 were significantly greater than the means for probes 2 and 4. All of these differences were consistent with the phrase-structure model, but some predicted differences were not significant.

An Interpretation of Experiment II

The main research question in Experiment II concerned interaction between the delay interval and other stimulus variables. Three such interactions were found. In separate analyses of the latency data, delay interacted with form, with congruence, and with probe position. In each case, however, the variable of sex was also involved in the interaction, thereby complicating any interpretation of delay effects in terms of perceptual processes. A highly speculative interpretation is presented in the next few paragraphs.

To begin with the interaction of delay and probe position, this is one case in which the two-way interaction may nave been relatively independent of the sex difference. That is, the large difference between delay groups at probe 6 was apparent within both sexes (see Figure 3). The three-way interaction between delay, probe, and sex seems to derive from an especially jagged profile for male Ss in the short delay condition (see Figure 3). Unfortunately, the independence of the two- and three-way interactions cannot be demonstrated statistically with the present data, because of the borderline significance of both interactions. In any case,

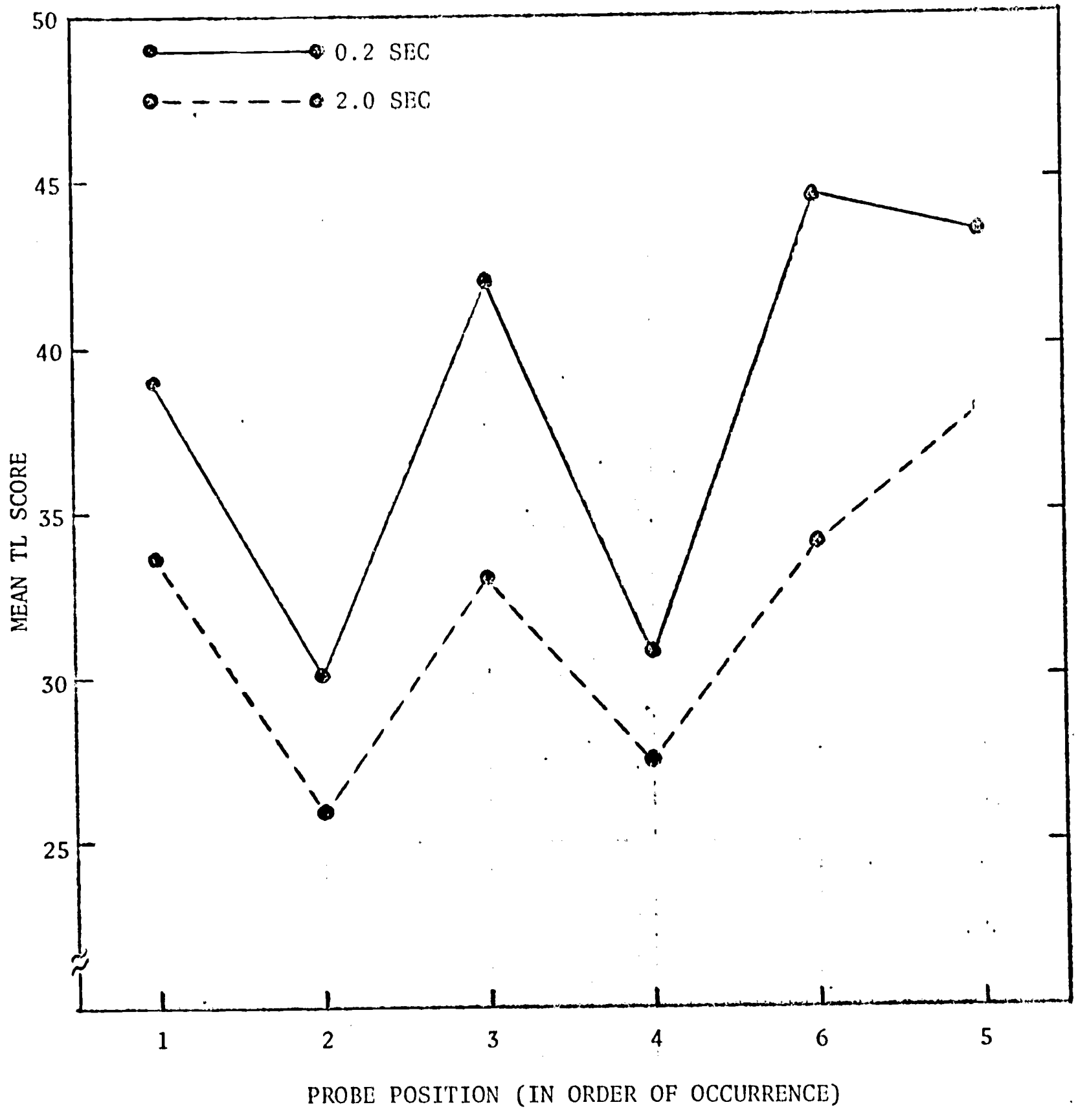


FIG. 7. Mean transformed latency score as a function of probe position, sentence-type 4.

the proposed source of the two-way interaction at probe 6 seems quite reliable across the various sentence-types (see Figures 4,6, and 7).

The three-way interaction of delay, probe, and sex suggests that, in the course of processing a sentence, males are initially more affected than females by a stimulus variable which produces a main effect. In this case, the stimulus variable is probe position, which appears to reflect the listener's analysis of a sentence into phrase-like constituents. Perhaps males are more inclined than females to employ this analytic strategy early in the processing of a sentence. Thus males would initially perceive the sentence as a few discrete "chunks", and would respond rather slowly across phrase boundaries but very quickly within phrases. Females, on the other hand, may be more inclined to respond on the basis of a global perception of the sentence as a whole, tending to smooth out the profile of latency scores across probe positions. There is, in fact, some evidence that males tend to be more analytic perceivers than females in other sorts of perceptual tasks (e.g., Witkin, et al., 1954). The sex difference would be less apparent in later stages of sentence processing because females must ultimately do some analysis in order to interpret the sentence, and both groups would be able to reconstruct the whole surface of the sentence from its underlying parts after analysis.

The suggestion that subjects responded on the basis of a reconstructed sentence after a two-second delay is supported by the observed main effects of delay. The slower responses of the long delay group may have reflected the extra time taken for reconstruction. The short delay Ss were not always significantly faster, but the difference in their favor

was remarkably consistent in the probe task (see Figures 4, 6, and 7). Similar differences occurred in the latency data for questions. Only the probe task required the subject to have a verbatim representation of the sentence at the time of responding, but even when answering a question, the subject may have attempted to check his response against a reconstruction of the sentence. The reconstruction hypothesis is an interesting alternative to the "pacing" explanation proposed in an earlier discussion of the delay effect.

Another interaction involving delay was the one between delay, congruence, and sex in Experiment II-A. Here the males showed a greater congruence effect with a short delay, while the females were more affected by congruence in the long delay condition. Once again, this difference may reflect analytic vs. global perceptual strategies. That is, the females may have initially responded to incongruent sentences on the basis of a global impression which was accurate enough to yield fairly rapid correct responses some of the time. Meanwhile, the males were already in the throes of the conflict produced by incongruence. The females experienced more of this conflict later, when they got around to analyzing the sentence further.

The third interaction pertinent to the research question involved delay, form, and sex in Experiment II-A. In this case, the male subjects in the long delay condition were different from the other three groups. This interaction does not seem to fit very well with the speculation built up around the other two interactions. The true identity of the form variable was called into question earlier anyway. Consequently, no attempt will be made to explain this portion of the data.

Conclusions

Conclusions concerning the effects of the delay interval must be very tentative, due to problems in the sampling of subjects and stimulus items. The number of subjects was rather small for detecting interactions of delay with other variables, especially in Experiment II-B, where the fifth grade data were not useable. Repeated measures on the delay variable would have provided more statistical power than independent groups in the analyses of delay effects. The problem with the stimulus material was that the use of a single set of items for all Ss increased the probability of spurious results. The latency data in Experiment II-A may have suffered from this problem. (Aside from this one instance, however, there was a striking degree of consistency between the present data and those collected in previous studies.) Despite the problems of sampling, Experiment II did provide some interesting leads for future study with more adequately designed experiments.

The amount of delay following the stimulus sentences interacted with other variables affecting the latency of responses to questions and probe words. It may be concluded that the perception of a sentence changes during the first two seconds after the sentence is heard, and that these changes are associated with different levels of perceptual organization and recoding. Early in the process of interpretation, the sentence is represented as a string of words, or even sounds, either as a whole or as a hierarchy of phrase-like chunks. At this point, responses are relatively fast, especially for the most recent part of the sentence. The listener

has the surface characteristics of the sentence more or less "at his finger tips."

Other research (e.g., Sachs, 1967) has indicated that the surface information from sentences is forgotten quite rapidly -- after less than 100 syllables of connected discourse. This rapid forgetting is of no consequence in most communication settings because, after the preliminary analysis of a sentence, the surface information is probably used only as a double check on a deeper interpretation (see the discussion of the conflict generated by incongruent sentences in Experiment I). The deeper interpretation itself is retained for a relatively long time.

From the present data, it would appear that the forgetting of surface information has begun even after two seconds have elapsed following a sentence. At this point, the listener probably has arrived at a deep interpretation, having identified the grammatical relationships underlying the sentence. If the listener needs surface information -- as in responding to probe words, or in checking a deeper interpretation -- he may actually have to reconstruct the surface representation, at least in part, from the deeper level of analysis. The process of reconstruction takes time and increases response latencies. It takes even more time to recover those surface characteristics which do not follow more or less automatically from the deeper level, such as information regarding pronouns. These characteristics must be filled in on the basis of whatever weak traces are left over from the original surface representation.[1]

---

[1]The data from the fifth grade subjects in Experiment II-B support the inference that information regarding pronouns was relatively difficult

Because of sex differences in the effects of delay, it became necessary to speculate about the role of cognitive variables correlated with sex in the processing of sentences. The hypothesis was made that males, having a more "analytic" perceptual style than females, proceed more rapidly to deeper levels of analysis. This speculation about analytic vs. global perception merely illustrates the type of theory which is required to account for individual and group differences in sentence processing.

FINAL SUMMARY AND CONCLUSIONS

The experimental results which have been reported support the contention that understanding a sentence is an "active", hypothesis-testing process. Upon hearing a sentence, the listener seems to formulate hypotheses very quickly as to its underlying meaning. These hypotheses are based on selected cues from the surface of the sentence, on the listener's knowledge of a lexicon and a set of grammatical rules, and on the listener's expectations concerning what is likely to be said. The hypotheses are then t sted against additional inputs to the proeessing system, either from memory or from later stimulation, and one hypothesis is finally accepted. Psycholinguists have proposed similarly active models of sentence comprehension

---

[1]cont., to retrieve after the longer delay. The percentage of correct responses was calculated for the two rrobe positions involving pronouns in all sentence-types. The short delay group had 69% correct and the long delay group had only 49%, a difference of 20 percentage points. At the remaining probe positions, the two groups were practically the same -- 61% correct with the short delay and 60% with the long delay. Essentially the same results were observed within each of the four sentence-types. It was not possible to do a similar analysis of the college data because of the generally high level of performance -- about 90% correct in both of the delay groups.

before (e.g., Miller and Chomsky, 1963). What is surprising about the model suggested by the present data is the dominant role played by the listener's expectation of what is likely to be said, as opposed to his perception of the actual surface properties of the sentence.

In Experiment I, the subjects (fifth graders and college students) took longer to answer questions about sentences in which the grammatical structure was not congruent with the most probable interpretation. The increase in response latency was probably symptomatic of a conflict between hypotheses based on the listener's expectations and his memory of grammatical cues on the surface of the sentence. When the subjects made errors in responding to incongruent sentences, the majority of their errors indicated that the more probable, alternate interpretation had been accepted. Apparently the subjects' memory of grammatical cues was usually strong enough to challenge erroneous hypotheses, but it was not always strong enough to cause their rejection.

"Alternative" errors were especially frequent when the alternate interpretation was not only more probable than the correct one, but when it also involved a subject-predicate relationship between a noun phrase and a verb phrase which were not separated by another phrase. In other words, the contiguity of nouns and verbs may be one of those surface cues that the listener uses to generate hypotheses about deeper relationships in a sentence, or to test hypotheses generated in other ways.

The results from Experiment I are all the more striking when the general conditions of the experiment are taken into account. They were not inordinately long or complex in structure. Even the relatively improbable

sentences were plausible and perfectly grammatical. The subjects were instructed to listen carefully, and the great majority of them seemed highly motivated to follow instructions and perform well. Finally they were required to respond within a few seconds after hearing each sentence, which is when the sentences should have been easiest to remember. (It might be argued that the pressure to give quick responses reduced the level of performance, but there was generally an inverse relationship between the time taken to respond and the probability that the responses would be correct.)

In sum, the conditions of the experiment would lead one to expect practically perfect comprehension, regardless of how improbable some of the sentences might have been. The surface properties of the sentences ought to have dominated in the comprehension process, resulting in correct interpretations most of the time. But the results of Experiment I showed that surface cues, in the form of pronouns, were sometimes overriden by the listener's expectation of what was likely to be said. The fact that this happened even under the conditions of the experiment suggests that the listener's expectations must be terribly important in everyday communication, where there is usually much less incentive to pay close attention to the surfcce of a message. In the experimental situation, the subjects might have disregarded surface details -- because that is the way they normally process sentences -- or else they might have been unable to retain the surface information long enough to rely on it. Some results from Experiment II-B support the latter interpretation.

In Experiment II-B, the subjects performed a task which required them to retain each sentence word by word. When the delay interval following

the sentences was increased from 0.2 to 2.0 seconds, the subjects seemed to have more trouble keeping track of the pronouns than other parts of the sentences. A pronoun is one of those surface details which mark grammatical relationships but do not have to be remembered once the sentence is understood. Apparently information regarding pronouns is difficult to remember, even when the listener needs to retain it for just a few seconds. Given such a limited capacity for immediate memory, it is not surprising that some surface information plays a rather limited role in the very process of comprehension.

It follows from the foregoing discussion that the listener should have a great deal of trouble understanding sentences which contradict his expectations. Of course most everyday sentences are embedded in extended discourse or dialogue, plus a nonverbal context. As a result, the total message is highly redundant and it affords the listener several opportunities to correct his misinterpretations. Nevertheless, redundancy itself is a variable, and the listener must also have the capacity to make use of whatever redundancy there is. So it seems quite likely that the listener's expectations would lead him astray some of the time, even in a natural communication setting.

Individual and group differences in expectations may account for some of the breakdowns which occur in communication. This, of course, is not a novel idea, but the present research indicates that communication breakdowns may occur at the very immediate and elementary level of perceiving basic grammatical relations in a sentence, even under optimal listening conditions. However, the present research does not provide any direct

evidence that group differences in expectation lead to differences in comprehension, because the stimulus sentences were selected to insure common expectations among the age and sex groups involved.

There were a few noteworthy differences in performance associated with age and with sex. The fifth grade subjects generally took longer to respond and made more errors than the college students. In Experiment I, the fifth graders were more likely to be mislead by their expectations and by the contiguity of a noun phrase and a verb phrase. These results suggest that children simply have more limited capacity than adults for remembering the surface details of sentences. Sex differences appeared as interactions between sex, amount of delay following the sentences, and other stimulus variables in Experiments II-A, and II-B. To some extent, these interactions could be interpreted as evidence that males employ a more "analytic" strategy than females in perceiving sentences, but this interpretation is quite tentative.

The age and sex differences underline further the importance of considering individual differences in the process of comprehending speech. Two listeners who know essentially the same words and grammatical rules, and who share the same expectations, may nonetheless go about interpreting a sentence in different ways, due to differences in memory capacity, or in preferred strategies for using memory, expectations, and linguistic knowledge. Without systematic study of such individual differences, it may not be possible to go beyond the broad outlines of the process model presented in this report.

REFERENCES

Ammon, P.R. The perception of grammatical relations in sentences: A methodological exploration. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 869-875.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Mass.: M.I.T. Press, 1965.

Fodor, J.A., & Bever, T.G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Johnson, N.F. The psychological reality of phrase-structure rules. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 469-475.

Katz, J.J. , & Fodor, J.A. The structure of a semantic theory. Language, 1963, 39, 170-210.

Miller, G.A., & Chomsky, N. Finitary models of language users. In R.D. Luce, R. Bush, & E. Galanter (Eds.), Handbook of mathematical psychology, Vol. II. New York: Wiley, 1963.

Sachs, Jacqueline S. Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 1967, 2, 437-442.

Slobin, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.

Witkin, H.A., et. al. Personality through perception. New York: Harper, 1954.

ABSTRACT

Two experiments were performed in which fifth grade children and college students listened to tape-recorded sentences and responded to tests of immediate memory. Systematic variations in the syntactic structure and the content of the sentences were found to produce significant differences in the time it took for the subjects to respond and in the frequency of correct responses. From these measures, it was possible to make inferences about the nature of the cognitive processes involved in sentence comprehension.

The results of both experiments are consistent with an "active" model of comprehension processes. According to the model, the listener generates hypotheses about the underlying meaning of a sentence, based on a preliminary analysis of the sentence plus the listener's expectations concerning what is likely to be said. The hypotheses are then tested against other information, including syntactic cues in the sentence itself. It appears that the listener does not retain all of this additional information very long after hearing the sentence. As a result, the listener is sometimes slow to reject erroneous hypotheses, and he may even misinterpret a perfectly plausible sentence.

The same general model of comprehension seems appropriate for both of the age groups studied, but some of the effects observed in the experiments varied according to the age and sex of the subjects, indicating the need for consideration of individual differences in the comprehension process.

APPENDIXES

A. Procedure for Obtaining Pre-Experimental Judgments of Predication Probabilities

The variable of congruence could only be manipulated if one of the two possible predications in each item was clearly more probable. The relative probabilities were measured by obtaining subjective judgments from fifth graders and college students. The items were administered orally to classroom groups (about 30 Ss at each age level). Each item had the following format:

> The brave fireman saved the little kitten.
>
> Someone drank the warm milk.
>
> Who was more likely to drink the milk?

The Ss had answer sheets with three response alternatives for each item, e.g., "fireman", "kitten", or "equally likely". The instructions were simply to circle the response which seemed most appropriate.

If 75% or more of the Ss in both age groups agreed that one predication was more likely, the item was used in Experiment I as one of the experimental items.

B. List of Stimulus Items for Experiment I and a Description of Its Construction

1. A*

The grumpy elves
pinched the sleepy crow;
they leaped from the low branch.

Who leaped from the branch?

2. C

The polite actor
who smiled at the happy woman
brought a black umbrella.

Who was happy?

3. B

The plump wife
hired the husky gardener;
he planted the yellow flowers.

Who planted the flowers?

4. D

The famous conductor
instructed the graceful musicians
who composed the pretty tunes.

Who was graceful?

5. C

The tough general
who walked with the gloomy men
wore the shiny medals.

Who wore the medals?

6. A

The busy bees
stung the speedy runner;
they guarded the golden hive.

Who guarded the hive?

7. D

The quick father
saved the weak swimmers
who rested on the sandy beach.

Who was weak?

8. B*

The skinny detective
captured the cruel criminals;
they overheard the secret plot.

Who overheard the plot?

---

* Incongruent item.

9. C

The honest lawyers
who praised the old judge
sent the nice letters.

Who was honest?

10. B

The angry ducks
bit the nosey butcher;
he wanted the speckled eggs.

Who was nosey?

11. A

The crafty uncle
talked with the handsome nephews;
he buried the precious jewels.

Who was crafty?

12. D*

The fine captain
kissed the thin maiden
who returned from a difficult voyage.

Who returned from a voyage?

13. B

The stupid monkey
chased the funny soldiers;
they made some silly faces.

Who was stupid?

14. D

The strict aunt
punished the naughty twins
who took the chocolate cookies.

Who took the cookies?

15. A

The weird daughter
lifted the slimy snakes;
she heard the slow music.

Who was slimy?

16. C*

The cute girlscouts
who rooted for the powerful athlete
got the silver trophy.

Who got the trophy?

17. C

The thoughtful husband
who introduced the important guests
hung up the warm coats.

Who was thoughtful?

18. A

The tidy wife
thanked the friendly butcher;
she counted the silver coins.

Who was friendly?

19. D

The handsome grocer
advised the proud farmers
who owned an orange tractor.

Who owned a tractor?

20. B

The gay rooster
woke the cross maid;
she entered the rickety barn.

Who was cross?

21. B

The rich widow
instructed the faithful butler;
he set the large table.

Who set the table?

22. A*

The playful kindergarteners
teased the short monkey;
they wanted the ripe bananas.

Who wanted the bananas?

23. C

The evil bandit
who murdered the jealous actress
had a black mask.

Who had a mask?

24. D

The selfish gardener
yelled at the dirty beggars
who picked the sweet fruit.

Who was dirty?

25. A

The blond boy
visited the sick lady;
he painted a colorful card.

Who painted a card?

26. D

The brave doctor
fixed the wounded soldiers
who hated the terrible war.

Who was brave?

27. C

The horrible ghost
who scared the busy waitress
had a white face.

Who was busy?

28. B*

The nasty pirate
kidnapped the lonely princess;
she hid the heavy treasure.

Who hid the treasure?

29. D*

The considerate librarian
helped the new sisters
who mended the torn bookcovers.

Who mended the bookcovers?

30. A

The deaf artist
petted the old cats;
he ate the stale dinner.

Who was deaf?

31. C*

The important citizens
who met with the popular senator
won in the last election.

Who won in the election?

32. B

The best swimmer
guided the tired astronauts;
they climbed up the wet ladder.

Who was best?

33. A

The polite janitors
listened to the smart teacher;
they swept up the broken glass.

Who swept up the glass?

34. B*

The jolly babysitter
tucked in the happy kids;
they told a spooky tale.

Who told a tale?

35. C

The jealous twins
who teased the best student
dropped the heavy books.

Who was jealous?

36. D

The chubby midget
saw the worried clowns
who stood in the dark shadows.

Who was worried?

37. C*

The terrible monster
who picked up the sad queen
let out a loud scream.

Who let out a scream?

38. A

The fair judge
warned the rude sailors;
he left the crowded room.

Who was rude?

39. B

The generous neighbors
cared for the sick dentist;
he prepared the evening meal.

Who was generous?

40. D

The tall officer
arrested the greedy lawyers
who thought up a tricky plan.

Who thought up a plan?

41. A

The lively horse
kicked the strong cowboys;
he jumped over the high fence.

Who was lively?

42. D

The cheerful widow
waited for the slow mailman
who whistled a merry tune.

Who was cheerful?

43. B

The tidy boss
scolded the sloppy secretary;
she typed the long letters.

Who typed the letters?

44. C

The cross waitress
who chased the gloomy dogs
went out the back door.

Who was gloomy?

45. D*

The French skiers
cheered the German ballerina
who slid down the steep mountain.

Who slid down the mountain?

46. B

The mean soldiers
kidnapped the clever scientist;
he blew up the stone house.

Who was clever?

47. C

The small girl
who yelled for the quick lifeguard
fell into the cold pool.

Who fell into the pool?

48. A*

The sly hunter
followed the fat bears;
he smelled the fresh honey.

Who smelled the honey?

49. A

The brown spider
killed the gray flies;
he crawled down the sticky web.

Who crawled down the web?

50. B

The cute dancer
fed the furry rabbits;
they wore some pink ribbons.

Who was cute?

51. C

The Italian chef
who worked for the wealthy woman
baked the tasty pastry.

Who baked the pastry?

52. D*

The bashful mermaid
watched the drunken seamen
who swam in the deep ocean.

Who swam in the ocean?

53. B

The eager rancher
raised the chubby sheep;
they liked the green grass.

Who liked the grass?

54. C

The proud uncle
who smiled at the nervous bride
got in the blue car.

Who was nervous?

55. D

The clever magician
fooled the short girl
who held a silk handkerchief.

Who was clever?

56. A

The wicked outlaws
beatup the blind hermit;
they swung a crooked stick.

Who was wicked?

57. A*

The careful doctors
examined the curious scientist;
they created a mysterious chemical.

Who created a chemical?

58. B*

The fierce lion
played with the healthy cubs;
they growled in the huge cage.

Who growled in the cage?

59. C

The thirsty customers
who paid the kind clerk
looked for some cold soda.

Who was thirsty?

60. D

The husky butler
tripped the clumsy thieves
who carried a small flashlight.

Who was clumsy?

61. D

The brave firefighter
rescued the little kittens
who drank the warm milk.

Who drank the milk?

62. B

The generous lady
greeted the lame beggar;
he brought some red roses.

Who was lame?

63. A

The noisy squirrel
scolded the playful children;
he climbed the shady tree.

Who was playful?

64. C*

The strange people
who frightened the gentle cow
ran through the great pasture.

Who ran through the pasture?

65. B

The fast sailors
harpooned the giant whale;
he wrecked the tiny boat.

Who wrecked the boat?

66. A

The sneaky brothers
spied on the crazy inventor;
they stood by the large window.

Who was crazy?

67. C

The pleasant teacher
who helped the shy students
gave the right answers.

Who was shy?

68. D

The worried sheriff
met the wise grandma
who told about the bad accident.

Who was worried?

69. A*

The angry cowboy
caught the wild Indians;
he shot the swift arrows.

Who shot the arrows?

70. C

The bored king
who welcomed the gay acrobats
enjoyed the skillful tricks.

Who enjoyed the tricks?

71. B*

The wicked witch
scared the honest prince;
he used some magic powers.

Who used some powers?

72. D

The good-looking customers
questioned the courteous saleslady
who bought the comfortable shoes.

Who was courteous?

73. D

The neat businessman
paid the efficient waitress
who worked at the Chinese restaurant.

Who worked at the restaurant?

74. B

The lazy boss
hired the poor farmers;
they did the easy work.

Who was lazy?

75. A

The nasty guard
hit the stubborn prisoners;
he started a big fight.

Who was nasty?

76. C*

The cheerful nurse
who talked with the young patient
took some pink pills.

Who took some pills?

77. B

The jolly mother
tickled the plump babies;
they shook the plastic rattles.

Who was plump?

78. D*

The famous artist
drew the beautiful models
who sold some expensive paintings.

Who sold some paintings?

79. A

The nervous passengers
called for the pretty stewardess;
they wished for some different
magazines.

Who wished for some magazines?

80. C

The frightened boys
who remembered the ugly gypsy
hid behind a dead tree.

Who was frightened?

81. C

The kind mother
who cooked for the noisy children
washed the dirty dishes.

Who washed the dishes?

82. A*

The mad boxer
hit the rude reporters;
he wrote the interesting stories.

Who wrote the stories?

83. B

The pleasant hiker
fed the shy raccoons;
they crept alongside the wooden
cabin.

Who crept alongside the cabin?

84. D

The rich pilot
married the talented dancer
who owned an expensive house.

Who was rich?

85. A

The sneaky gypsy
robbed the trusting neighbors;
he discovered the open drawer.

Who discovered the drawer?

86. C

The strong Indian
who caught the wild buffaloes
pulled on the long rope.

Who was strong?

87. B

The lazy barber
liked the bald men;
they bought the morning newspapers.

Who was lazy?

88. D

The lively campers
spoke with the helpful ranger
who knew of the nicest lake.

Who knew of the lake?

89. B*

The thoughtful hermit
greeted the hungry guests;
they lit a cozy fire.

Who lit a fire?

90. C

The young genius
who amazed the smart experts
built the new machine.

Who was smart?

91. D*

The graceful singer
performed for the friendly grandma
who appeared on a television program.

Who appeared on a program?

92. A

The grateful driver
thanked the handy mechanics;
he fixed the broken car.

Who was grateful?

93. B

The tiny squirrels
tickled the sleepy bear;
he rolled on the soft ground.

Who was sleepy?

94. C*

The silly comedian
who entertained the big audience
laughed at the funny jokes.

Who laughed at the jokes?

95. A

The huge dogs
bit the grumpy officer;
they ran through the dark woods.

Who was grumpy?

96. D

The husky king
petted the tame lions
who chewed the juicy meat.

Who was tame?

All items with a given letter (e.g., A) have essentially the same type of grammatical construction. Each block of 16 items contains 8 experimental items and 8 filler items. The 8 experimental items represent all possible combinations of the variables called congruence, separation, and form. No content words are used in more than one item within a block of 16.

C. Instructions and Practice Items for Experiment I

I will tell you what we are going to do. You will hear me say a sentence. Then I will ask you a question about the sentence. You listen to the sentence and then you answer the question.

Here is the sentence:

The nice queen thanked the foreign musicians; they went down the steep stairs.

Who went down the stairs?

The right answer is "the musicians" -- the musicians went down the stairs.

Here is another sentence with another question:

The crafty eskimos who trapped a wet seal fell on the slippery ice.

Who was crafty?

The right answer is "the eskimos" -- the eskimos were crafty.

I am going to say a lot of different sentences and I will ask you a question after each one. You will be able to give the right answer if you listen carefully to the sentence. You only have to say one word for each answer.

Now here are a few more sentences for you to practice on before we start the real test. See how fast you can give the answer to each question. I will say "ready" before each sentence.

Ready: The helpful secretary
worked for the popular president;
she read the interesting magazine.

Who was popular?

Ready: The bad alligator
who swallowed the spotted turtles
swam in the muddy river.

Who swam in the river?

(Repeat any practice item which S has answered incorrectly.)

OK, now we can start. Remember to listen carefully, to say one word for each answer, and to give your answer quickly.

D. List of Stimulus Items for Experiment II

The sentences listed in Appendix C were presented with following questions (Experiment II-A) or probe words (Experiment II-B).

1. Who leaped from the branch?

   GRUMPY

2. Who was happy?

   WHO

3. Who planted the flowers?

   PLUMP

4. Who was graceful?

   GRACEFUL

5. Who was gloomy?

   GLOOMY

6. Who guarded the hive?

   RUNNER

7. Who was quick?

   SANDY

8. Who was cruel?

   THEY

9. Who was honest?

   LAWYERS

10. Who was angry?

    SPECKLED

11. Who was handsome?

    HE

12. Who was thin?

    THIN

13. Who was funny?

    FUNNY

14. Who was strict?

    AUNT

15. Who was weird?

    DAUGHTER

16. Who was powerful?

    POWERFUL

17. Who was important?

    IMPORTANT

18. Who counted the coins?

    BUTCHER

19. Who was proud?

    WHO

20. Who entered the barn?

    GAY

21. Who set the table?

    BUTLER

22. Who was short?

    THEY

23. Who was evil?

BLACK

24. Who was dirty?

WHO

25. Who painted a card?

BLOND

26. Who was brave?

DOCTOR

27. Who was horrible?

WHITE

28. Who was nasty?

PIRATE

29. Who was considerate?

LIBRARIAN

30. Who ate the dinner?

CATS

31. Who won in the election?

SENATOR

32. Who was tired?

THEY

33. Who was polite?

BROKEN

34. Who was happy?

HAPPY

35. Who dropped the books?

JEALOUS

36. Who was chubby?

DARK

37. Who was terrible?

MONSTER

38. Who left the room?

FAIR

39. Who prepared the meal?

DENTIST

40. Who was tall?

TRICKY

41. Who was strong?

STRONG

42. Who whistled a tune?

CHEERFUL

43. Who was sloppy?

SLOPPY

44. Who went out the door?

DOGS

45. Who slid down the mountain?

BALLERINA

46. Who was clever?

CLEVER

47. Who was quick?

WHO

48. Who smelled the honey?

BEARS

49. Who was gray?

GRAY

50. Who was cute?

PINK

51. Who baked the pastry?

WOMAN

52. Who was drunken?

WHO

53. Who was eager?

RANCHER

54. Who was proud?

BLUE

55. Who held a handkerchief?

GIRL

56. Who was blind?

BLIND

57. Who was careful?

DOCTORS

58. Who growled in the cage?

CUBS

59. Who was kind?

WHO

60. Who was husky?

BUTLER

61. Who drank the milk?

BRAVE

62. Who brought some roses?

GENEROUS

63. Who was noisy?

SHADY

64. Who was gentle?

WHO

65. Who was giant?

HE

66. Who stood by the window?

SNEAKY

67. Who was pleasant?

RIGHT

68. Who told about the accident?

GRANDMA

69. Who was wild?

WILD

70. Who was bored?

KING

71. Who used some powers?

WICKED

72. Who was courteous?

COURTEOUS

73. Who worked at the restaurant?

WAITRESS

74. Who did the work?

FARMERS

75. Who was stubborn?

HE

76. Who was cheerful?

PINK

77. Who was jolly?

MOTHER

78. Who sold some paintings?

FAMOUS

79. Who was pretty?

THEY

80. Who was frightened?

BOYS

81. Who washed the dishes?

KIND

82. Who was mad?

INTERESTING

83. Who was pleasant?

WOODEN

84. Who was talented?

WHO

85. Who was sneaky?

GYPSY

86. Who was strong?

INDIAN

87. Who was lazy?

BARBER

88. Who was helpful?

HELPFUL

89. Who was thoughtful?

COZY

90. Who built the machine?

YOUNG

91. Who was graceful?

TELEVISION

92. Who was grateful?

DRIVER

93. Who was sleepy?

HE

94. Who laughed at the jokes?

SILLY

95. Who ran through the woods?

OFFICER

96. Who chewed the meat?

HUSKY