Sentence Verification and Language Comprehension of Aphasic Persons

Robert H. Brookshire and Linda E. Nicholas
Veterans Administration Medical Center, Minneapolis, Minnesota

Since the late 1960's, a large number of studies of sentence comprehension by non-brain-damaged adults have been carried out, and a considerable amount is known about the process of sentence comprehension, at least for normal listeners and readers.

Many investigators of sentence comprehension have made use of some form of the sentence verification procedure. In this procedure, an individual is asked to judge whether sentences are true or false, either on the basis of the individual's general knowledge (using sentences such as, "Winter is colder than summer.") or with reference to a picture presented along with the sentence (using sentences such as, "The boy is pushing the girl."). Sentence verification procedures have been used extensively in examining sentence comprehension by non-brain-damaged listeners (Slobin, 1966; Chase and Clark, 1972; Carpenter and Just, 1975).

Sentence verification procedures were first used in attempts to verify various predictions generated by Chomsky's generative-transformational model of language (Chomsky, 1957). The time needed to verify the truth or falsity of a sentence was seen as an indirect measure of the number of transformations required to deduce the truth of the sentence. The results of sentence verification studies subsequently were used to generate and support processing models of language comprehension, in which comprehension was seen as a linear series of steps or stages. The time needed to verify the truth or falsity of a sentence was seen as a measure of the number (and kind) of steps necessary to deduce the truth of the sentence. These processing models include Chase and Clark's (1972) response change model, Trabasso's (1972) optional recoding model, and Carpenter and Just's (1975) constituent comparison model. Other investigators have used sentence verification experiments to assess the role of cognitive and experimental factors in sentence comprehension (Just and Clark, 1973; Clark, 1977). These latter investigations have demonstrated that extralinguistic and experiential information are important sources of information about sentence meaning in many experimental and real-life situations. For example, Slobin (1966) among others, has demonstrated that experientially (but not grammatically) anomalous sentences created by reversing sensical sentences, such as

The girl is watering the flowers.
to
The flowers are watering the girl.

are identified as false faster than similarly false, but not anomalous, sentences generated by reversing sentences such as

The girl pushes the boy.
to
The boy pushes the girl.

Carpenter and Just (1975) have suggested that when non-brain-damaged subjects are given sentences of the form

The girl pushes the boy.

along with pictures which either represent the sentence, or differ from it in subject, verb, or object, verification is carried out by means of a "find and compare" strategy. In this strategy, sentence and picture elements are sequentially compared until a mismatch is detected, in which case the subject responds "false," or until the sentence is completed and no mismatch is detected, in which case the subject responds "true." According to Carpenter and Just, such "find and compare" strategies permit verification of sentence truth value without the need for generating an internal representation of the overall meaning of the sentence.

Two studies of sentence verification by aphasic subjects have appeared in the literature. Both resembled previous studies with non-brain-damaged persons, and both delivered simple affirmative or negative sentences to which subjects responded "true" or "false." West, Gelfer, and Rosen (1976) placed aphasic subjects in a sentence verification task in which subjects were asked whether spoken affirmative sentences of the form

This is a shoe.

were true or false, based on pictures which accompanied the sentences. They found that aphasic subjects' reaction times to true sentences were faster than their reaction times to false sentences.

Just, Davis, and Carpenter (1977) placed aphasic and non-brain-damaged subjects in a verification task in which subjects were presented with the spoken or printed affirmations and negations, "red," "not red," "blue," and "not blue," along with colored dots which either matched the affirmation or negation or mismatched it. Subjects' reaction times and error rates were measured. In general, their results suggested qualitative similarity between aphasic and non-brain-damaged subjects' performance.

These two studies of sentence verification suggest that the verification technique may be appropriate for aphasic subjects, and that aphasic subjects' performance seems to be quantitatively inferior to, but qualitatively similar to performance of non-brain-damaged subjects. However, they do little more than scratch the surface, in terms of aphasic subjects' performance in sentence verification tasks.

Therefore, we designed two experiments in which we set out to examine several questions about aphasic persons' sentence comprehension using sentence verification procedures. The experimental questions addressed were: (1) Does the performance of aphasic subjects to true and false sentences parallel that of non-brain-damaged subjects? (3) Does the structural complexity of sentences affect aphasic subjects' response accuracy or reaction time in a sentence verification task? (3) Do aphasic and nonaphasic subjects match pictures and sentences as holistic units, or do they decide sentences' truth or falsity on the basis of sequential comparisons between sentence parts and pictures? (4) How strong is the relationship between aphasic subjects' performance on a traditional test of "auditory comprehension (such as the Token Test; DeRenzi and Vignolo, 1962) and their performance in a sentence verification task?

In the first study, we evaluated the performance of ten aphasic and ten age-matched non-brain-damaged subjects when pictures were presented simultaneously with spoken sentences and subjects' errors and reaction times to the sentences were tabulated.

Ten black and white line drawings depicting transitive actions were selected (Figure 1). Seven sentences were developed for each of these pictures. These pictures and sentences were used to develop 70 sentence

Figure 1. Example of picture used in experiment.

and slide combinations. Sentences were either active or passive. The activity depicted in each slide and the sentence which accompanied it either matched or mismatched. For active sentences the mismatch could occur on either the subject, the verb, or the object. A mismatch could also occur on active sentences if the subject and object were reversed. For passive sentences, a mismatch could only occur if the subject and object were reversed; no mismatches of subject, verb, or object occurred in passive sentences. Examples of the sentences used in this experiment are:

Active sentences	
Match:	The woman is washing the clothes.
Subject Mismatch:	The man is washing the clothes.
Verb Mismatch:	The woman is ironing the clothes.
Object Mismatch:	The woman is washing the dishes.
Subject-Object Reversal:	The clothes are washing the woman.
Passive Sentences	
Match:	The clothes are being washed by the woman.
Subject-Object Reversal:	The woman is being washed by the clothes.

Subjects responded true or false by operating pushbuttons. Their reaction times were recorded electronically and errors were tabulated by the experimenter. Each subject was also given the Token Test prior to participation in the sentence verification task.

The impressive characteristic of aphasic subjects' performance in this sentence verification task was its qualitative similarity to the performance of non-brain-damaged subjects. Aphasic subjects as a group made more errors in sentence verification than non-brain-damaged subjects (although most of these errors were committed by two subjects), but the pattern of their errors according to sentence type and truth value was remarkably similar to that of non-brain-damaged subjects. A similar situation existed with regard to reaction times. Reaction times of aphasic subjects were generally longer than reaction times of non-brain-damaged subjects, but the reaction times of the two groups were affected similarly by location of mismatch and type of sentence (Figures 2 and 3). These findings support previous conclusions that aphasic subjects' performance on tasks such as the one used here is qualitatively similar to, but quantitatively inferior to the performance of non-brain-damaged subjects of similar age and educational history (Warren et al., 1977).

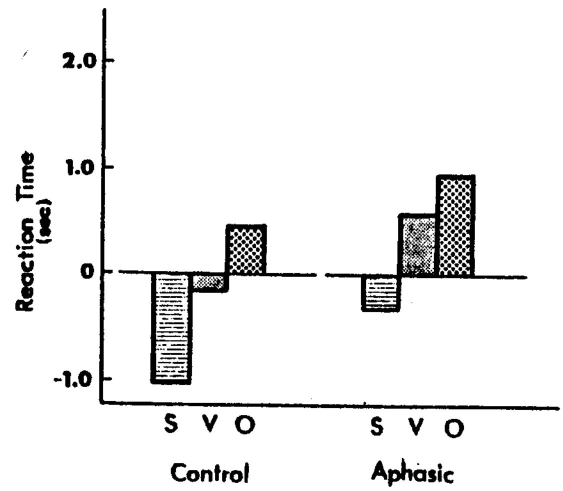

Figure 2. Reaction times of aphasic and control subjects to mismatches on subjects, verbs, and objects in sentences.

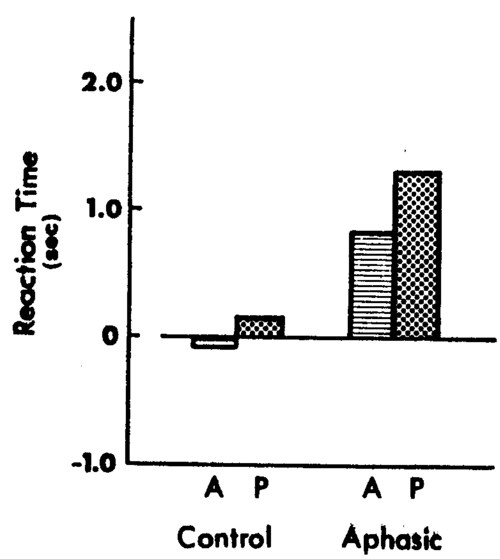

Figure 3. Reaction times of aphasic and control subjects to active and passive sentences.

We were somewhat surprised by the ability of most of our aphasic subjects to verify the accuracy of spoken sentences, and by their skill at detecting mismatches on lexical items. Only two of ten subjects committed enough errors on the sentence verification task to allow one to discriminate them from non-brain-damaged subjects on the basis of their sentence verification performance. This was in spite of the fact that nine of the ten aphasic subjects had Token Test error scores which clearly indicated deficient "auditory comprehension," and the tenth had marginal Token Test performance. A Pearson correlation coefficient calculated between subjects' Token Test scores and their sentence verification scores yielded $r=.67$, suggesting a moderate relationship between Token Test and sentence verification performance. At this point, it appears that the Token Test may give a more pessimistic estimate of auditory language comprehension than sentence verification tests.

In our next study, we were interested in the ways in which availability of pictures to subjects during sentence verification might affect their performance. Our interest came from two sources. Carpenter and Just (1975) feel that when pictures are presented along with transitive action sentences in sentence verification tasks, non-brain-damaged subjects serially compare sentence and picture elements until a mismatch is detected, when they respond "false," or until the sentence is completed without a mismatch, when they respond "true." According to Carpenter and Just, it is unnecessary that subjects comprehend the meaning of the entire sentence to carry out this strategy.

In addition Glucksberg and Danks (1975) have suggested that the temporal arrangement of sentence and picture presentation in verification tasks is important, and is likely to influence the results obtained. For example, when non-brain-damaged subjects are shown a transitive action picture, they have a strong tendency to encode it as an active sentence. If they are first shown such a picture, and then presented with the sentence, their reaction times to active voice sentences which match the picture will be shorter than their reaction times to passive sentences which match the picture. This occurs, according to Glucksberg and Danks, because subjects have coded the picture in active voice, and must either recode the picture into passive voice, or recode the sentence into active voice, in order to decide whether or not sentence and picture match.

Consequently, we set up a study to investigate the effects of manipulation of picture presentation relative to sentence presentation upon subjects' performance. Ten aphasic (five fluent and five disfluent) and ten age-matched non-brain-damaged subjects were placed in a sentence verification task in which pictures were presented either simultaneous with, before, or after the sentences. Sentence and slide combinations were like those used in the first study.

In Simultaneous Condition, picture and sentence occurred simultaneously with a five-second interval between trials. In Picture First Condition, the picture was presented for ten seconds with the sentence beginning when the picture went off. Again, there was a five-second interval between trials. In Sentence First Condition, the sentence was presented first and the picture was presented when the sentence ended, with a five-second interval between trials.

The results of this study generally supported our earlier findings that performance of aphasic subjects qualitatively resembles that of non-brain-damaged subjects, but that aphasic subjects make more errors and respond more slowly. In Simultaneous Condition, which was essentially a replication of our previous study, we again found progressively longer reaction times to mismatches on subject, verb, or object (Figure 4). (We have no good explanation for the unexpectedly long reaction times of aphasic subjects to verb mismatches in this condition.) Picture First Condition generated a similar progression of reaction times to subject,

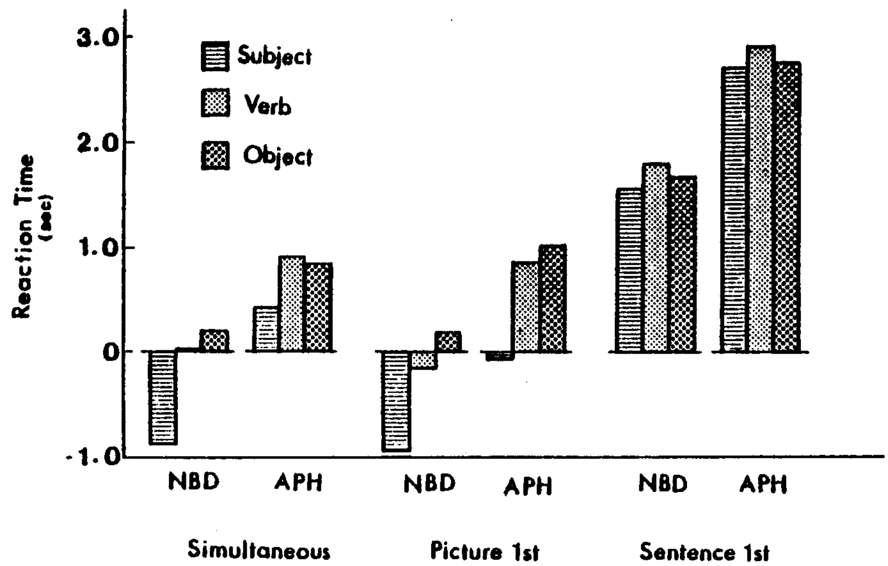

Figure 4. Reaction times to mismatches on subjects, verbs, and objects for non-brain-damaged and aphasic subjects in Simultaneous, Picture First, and Sentence First conditions.

verb, and object mismatches, suggesting that the processes involved in comprehending the sentences presented in the two conditions are probably similar. Sentence First Condition, on the other hand, generated reaction time performances for both groups that were clearly different from their performances in the other two conditions. Reaction times in Sentence First Condition were longer. However, this is simply the result of our procedures; in Sentence First Condition, timing began with sentence onset, and subjects had to wait to see the following picture before responding. In the other two conditions, they did not have to wait in the same fashion. On the other hand, the obliteration of progressively longer reaction times to mismatches on subjects, verb, and object in Sentence First Condition was not an artifact of procedures. Reaction times to verbs were slightly, but consistently, longer than reaction times to subjects or objects. These results suggest, first, that our subjects were no longer carrying out serial comparisons of sentence and picture elements in the order subject-verb-object, and second, that picturable elements (subject, object) could be verified more quickly than non-picturable elements (verb).

There were rather striking disparities in performance between fluent aphasic subjects and disfluent and non-brain-damaged subjects when reaction times to active and passive sentences within the three conditions were compared (Figure 5). Disfluent subjects resembled non-brain-damaged subjects in reaction times (and in number of errors) to active and passive sentences. Disfluent and non-brain-damaged subjects verified passive sentences about as quickly as they verified active sentences. On the other hand, fluent aphasic subjects took significantly longer to verify passive sentences than active sentences in Simultaneous and Picture First Condition. Sentence First Condition obliterated this effect. Fluent aphasic subjects also consistently made more errors to all sentences in all conditions than either disfluent or non-brain-damaged subjects did.

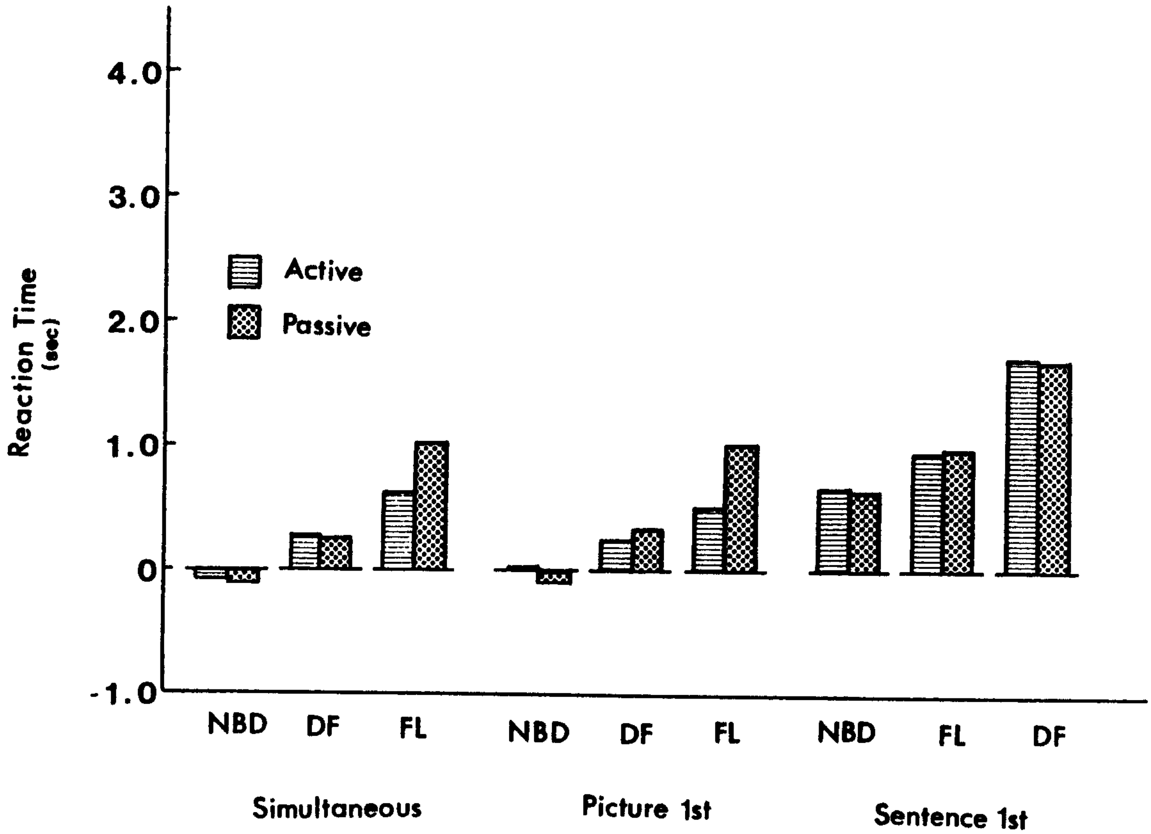

Figure 5. Reaction times of fluent and disfluent aphasic and non-brain-damaged subjects to active and passive sentences in Simultaneous, Picture First, and Sentence First conditions.

We had expected that when pictures were presented before sentences, passive sentences would take longer to verify than active sentences, because, according to Glucksberg and Danks, pictures are usually coded in active voice. This did not occur. Passive sentences did not take appreciably longer to verify in Picture First Condition than in Simultaneous Condition for any group of subjects.

Finally, we found that presenting pictures and sentences nonsimultaneously increased the relationship between Token Test and sentence verification performance. The Pearson Product-Moment Correlation Coefficient between subjects' Token Test performance in Simultaneous Condition was .69; between Token Test and Picture First Condition was .75; and between Token Test and Sentence First Condition was .79.

The results of our studies appear to justify several conclusions about sentence comprehension of aphasic subjects. First, they suggest that aphasic people go about the task of making sense out of spoken sentences in much the same way that non-brain-damaged people do, at least for active sentences. Our aphasic subjects' reaction times to mismatches on subject, verb, or object in active sentences generally paralleled the reaction times of non-brain-damaged subjects. These results are consistent with other studies that have suggested that many differences in auditory processing performance between aphasic and non-brain-damaged persons are quantitative, rather than qualitative. Our subjects' reaction times to subject, verb, and object mismatches were consistent with Carpenter and Just's predictions of what will occur if subjects carry out a series of find-and-compare operations, in which they check sentence elements against picture elements in the order in which sentence elements are heard.

It seems appropriate, then, that we structure sentence verification tasks to preclude the use of find-and-compare strategies which permit subjects to respond correctly without deducing the meaning of the entire sentence. If such find-and-compare operations are permitted, the task becomes essentially a spoken-word-to-picture matching task, instead of a test of sentence comprehension.

These results should make us generally cautious about the extent to which such find-and-compare strategies may be used in generally available tests of auditory language comprehension. Because the strategy is such an easy one, it seems likely that aphasic people will make use of it whenever they can. We need to be cautious then, about making assumptions regarding comprehension in any situation in which we have visual stimuli such as pictures present, creating the opportunity for find-and-compare strategies.

Our results suggest that one way to force subjects to deal with the meaning of the entire sentence is to stagger the presentation of pictures and sentences in time, presenting pictures and sentences separately, rather than simultaneously. Whether picture or sentence comes first would seem to be unimportant. However, our results suggest that it does make a difference. When pictures were presented first, reaction times resembled those obtained when picture and sentence occurred together, and suggested that subjects were carrying out find-and-compare strategies even when pictures were taken away before sentences were presented. However, when sentences were presented first, differential reaction times to subject, verb, and object mismatches disappeared, suggesting that find-and-compare was not being used. At this time, we suspect that when pictures are presented first, subjects are able to retain a visually-encoded image of the picture, and then carry out a find-and-compare operation between the sentence and this image. Confirmation of this suspicion awaits further experimentation.

Not all of our results support the idea that there are quantitative, but not qualitative differences between aphasic and non-brain-damaged persons. Our non-brain-damaged and disfluent aphasic subjects took about as long to verify active sentences as they did to verify passive sentences. Fluent aphasic subjects, on the other hand, took consistently longer to verify passive sentences than active sentences, except in Sentence First Condition. It looks as though fluent aphasic subjects may be relying primarily on a find-and-compare strategy whenever pictures are available, either before or during sentence presentation. Consequently, they get into trouble when such front-to-back comparisons are not appropriate, as when passive sentences are presented.

One of the reasons for our interest in sentence verification was our dissatisfaction with currently available tests of auditory language comprehension; especially the Token Test. We suspected that sentence verification might provide better estimates of daily-life language comprehension than the Token Test does. When we compared sentence verification scores with Token Test scores, group correlation coefficients generally approximated .70. This is a significant value; however, the relationship between Token Test and sentence verification scores accounts for less than half of the total variability. The meaningfulness of this correlation is further threatened by the existence of a number of subjects who performed almost normally on sentence verification, but were grossly impaired on the Token Test. We are generally more impressed by the discrepancies between Token Test and sentence verification performance than by the similarities between them. At this point, we cannot be certain which measure is the better estimator of daily-life auditory language comprehension abilities. However, the reports of our aphasic subjects, together with our observations of their daily-life comprehension behavior, lead us to believe that in a significant number of cases the Token Test gives an unnecessarily pessimistic estimate of daily-life language comprehension abilities. We believe that sentence verification may give an estimate that is closer to the mark, but confirmation of this belief awaits continued investigation.

REFERENCES

Carpenter, P.A. and Just, M.A., Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 82, 45-73, 1975.

Chase, W.G. and Clark, H.H., Mental operations in the comparison of sentences and pictures. In L. Gregg (Ed.), Cognition in Learning and Memory. New York: Wiley, 1972.

Chomsky, N., Syntactic Structures. The Hauge: Mouton, 1957.

Clark, H.H., Inferences in comprehension. In D. Laberge and S.J. Samuels (Eds.), Basic Processes in Reading: Perception and Comprehension. New York: Wiley, 1977.

Clark, H.H. and Chase, W.G., On the process of comparing sentences to pictures. Cognitive Psychology, 3, 472-517, 1972.

DeRenzi, E. and Vignolo, L.A., The Token Test: A sensitive test to detect receptive disturbances in aphasics. Brain, 85, 556-678, 1962.

Glucksberg, S. and Danks, J., Experimental Psycholinguistics: An Introduction. Hillsdale, New Jersey: Lawrence Erlbaum, 1975.

Just, M.A. and Clark, H.H., Drawing inferences from presuppositions and implications of positive and negative sentences. J. Verbal Learning and Verbal Behavior, 12, 21-31, 1973.

Just, M.A., Davis, G.A. and Carpenter, P.A., A comparison of aphasic and normal adults in a sentence verification task. Cortex, 13, 402-423, 1977.

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

Trabasso, T., Mental operations in language comprehension. In R.O. Freedle and J.B. Carroll (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington, D.C.: Winston, 1972.

Warren, R.L., Hubbard, D.J. and Knox, A.L., Short-term memory scan in normal individuals and individuals with aphasia. J. Speech and Hearing Research, 20, 497-509, 1977.

West, J., Gelfer, C. and Rosen, J., Processing of true and false affirmative sentences by aphasic subjects. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1976. Minneapolis, MN: BRK Publishers, 1976.

DISCUSSION

Q: Were your response times for aphasic subjects only for correct responses?
A: Yes, we only looked at reaction times to correct responses.

Q: What was the proportion of correct responses for your aphasic subjects?
A: Error rates averaged around ten percent for each of the conditions for aphasic subjects.

Q: Were the error rates of aphasic subjects comparable to those for normal subjects?
A: No, non-brain-damaged subjects had lower error rates than aphasic subjects.

Q: Do you ever feel at all uncomfortable with interpreting these kinds of results in terms of a qualitative vs. quantitative distinction as possibly being a little simplistic? I wonder if some of the decisions that we make about data that we call qualitative might actually be quantitative.
A: Do you mean that it might be just a matter of degree?

Q: Yes, I wonder if there might be some kind of fuzziness in that distinction like the discomfort we've come to have with the competence-performance distinction.
A: It seems that what we might be seeing, though, are quantitative differences when the task is relatively easy for aphasic patients and then qualitative differences when we start pushing the aphasic patients.

Q: We as a profession seem to be down on the Token Test because it is difficult and the types of commands that make it up are different from daily life comprehension tasks. I wonder if we can't, through clinical experience, come to interpret the test results in relation to their meaning for the patient's day-to-day activities.

Q: Did you use any reversible sentences?
A: No, they were all nonreversible, transitive action sentences.

Q: You didn't try reversible sentences such as "The lion was killed by the tiger" which are much more difficult than nonversible sentences for aphasic subjects?
A: No, that would definitely have been a more difficult task and we chose not to use those kinds of sentences in this study.

Q: Would you explain to me how you get negative reaction times?
A: The electronic timer started at the beginning of the sentence. The subject was allowed to respond whenever he decided whether the sentence was true or false. The subject's response stopped the timer. If he made his decision before he had heard the entire sentence, (as frequently happened with mismatches on subjects or verbs) a negative reaction time resulted because each sentence was timed and subtracted from the subject's reaction time to that sentence.

Q: In the Sentence First Condition, how long after the sentence was over did the pictures come on?
A: Immediately.

Q: Was it controlled by machine?
A: Yes, everything was electronically controlled so that time variables were always the same, and there was never a gap between presentation of picture and slide combinations.

Q: Did you conclude that sentences should be presented first in order to test for sentence comprehension?
A: Yes, if you want a measure of sentence comprehension rather than a word-to-picture match, it appears that you should present the sentence first and then the picture, at least for active sentences.