

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/225559877>

The active use of grammar in speech perception

Article in *Attention Perception & Psychophysics* · January 1966

DOI: 10.3758/BF03207817

CITATIONS

132

READS

761

3 authors, including:



Merrill F. Garrett

The University of Arizona

78 PUBLICATIONS 7,241 CITATIONS

[SEE PROFILE](#)



Thomas Bever

The University of Arizona

155 PUBLICATIONS 8,051 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Language, Cognition and Biology [View project](#)



Aspectual Interpretation [View project](#)

The active use of grammar in speech perception

MERRILL GARRETT¹, UNIVERSITY OF ILLINOIS

THOMAS BEVER² AND JERRY FODOR, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Judgments of the location of short bursts of noise in sentences were used to reveal perceptual segmentation of sentences. It was assumed that segmentation would correspond to major constituent boundaries. In order to control for correlated variables of pitch and intonation, identical acoustic material was provided with alternate constituent structures. It was found that differences in response to identical strings were predicted by the points of variation in constituent structure.

A grammar formalizes the inherent structure of a natural language. It may thus be supposed that a grammar provides an explanation of the linguistic capacities of a speaker insofar as these capacities are attributable to his ability to manipulate linguistic structures. Investigation of the relationship between formal descriptions of language and explanations of actual linguistic performance of speakers have been undertaken only recently.

Fodor & Bever (1965) and Garrett (1964) demonstrated a perceptual effect that is directly attributable to the employment of grammatical rules. It was shown that the perceptual processing of a sentence must involve at least the analysis of the sentence into grammatically defined derived constituents. That is, in perception the continuous speech input is segmented according to the constituent structure of the sentence: the immediate constituent is the unit of speech perception.

In order to provide evidence of such a perceptual segmentation, the following method was employed. In Ladefoged & Broadbent (1960), short bursts of noise ("clicks") were superimposed on tape recorded versions of test strings (sentences and digit strings). Subjects were required to recognize or reproduce the strings and locate the point of occurrence of the click in the string.

Ladefoged and Broadbent, in the initial use of this technique, observed a difference in the size of errors made for sentential and non-sentential strings. Fodor and Bever, and Garrett accounted for this difference in terms of the grammatical structure of the sentential material. A strong tendency for perceived position of clicks to migrate toward the deepest constituent boundaries was demonstrated for both binaurally and dichotically presented material.

The fact that location error in these experiments can be accounted for by the derived constituent structure is susceptible to two interpretations. It may be argued either that displacement of clicks is responsive to the recognition of some acoustic correlate of the constituent

structure or to the active imposition of a constituent analysis onto the speech signal, the latter being a process that could take place even where the constituent structure of the sentence is not acoustically marked.

That response to acoustic correlates of structure is a potential explanation of the effect is suggested by the fact that clicks are attracted to artificially introduced pauses in random strings of digits (Garrett, 1964). However, pause was controlled for in Fodor and Bever and it was demonstrated that the syntactic effect was found even in sentences where no acoustic pause marked the position of the deep constituent break. It is not, however, possible to rule out the effect of other acoustic correlates of constituent structure, as for example, the effects of intonation, rhythm, stress, etc.

Since this issue is of crucial importance for the interpretation of earlier results, an experiment which controls for all immediate acoustic features which might mark the boundaries of constituents was performed. Ambiguous strings whose constituent structures in a sentence depend on previous context were used as this control.

Method

Stimulus materials.

Six pairs of sentences were constructed for which some string of lexical items was common to each member of a pair. For example:

- A. (In her hope of marrying) (Anna was surely impractical)
- B. (Your hope of marrying Anna) (was surely impractical)

The pairs of sentences were recorded on one track of stereophonic recording tape; two taped copies of each pair were made. The common portions of each pair were made acoustically identical by splicing a recorded version of a portion of one member of the pair to the opposite member of the pair. For example in the above sentence, the portion ...*hope of marrying Anna was surely impractical* taken from the recording of A was spliced to the portion *Your ...* taken from B. When this spliced version of B is paired with a copy of the original recording of A, there are two sentences in which the acoustic material for the latter portion is identical, but for which the constituent boundaries (marked by the parentheses above) are different. In sentence A, for instance, there is a deeper boundary before the word *Anna* than in sentence B. On the second track of each recording of a stimulus string a capacitor discharge click was recorded of intensity equal to the

loudest vowel sound and duration of about 20 msec. Clicks were placed in the middle of the word or words around which the constituent boundaries were manipulated (in *Anna*, above), and in the first syllable following the second deep break (in *was*, above). The result was 24 stimulus strings (six pairs by two members by two click positions). The pairs of sentences are listed below with the two click positions for each string noted by the symbol \checkmark above the sentence. Acoustically identical portions are italicized.

1. (a) In order to catch his train *George drove furiously to the station.*
(b) The reporters assigned to *George drove furiously to the station.*
2. (a) In her hope of marrying *Anna was surely impractical.*
(b) Your hope of marrying *Anna was surely impractical.*
3. (a) Because it was a most important city *Hambourg was leveled by the war.*
(b) Only the metropolitan district of *Hambourg was leveled by the war.*
4. (a) During prohibition because many were afraid to give open support *drinking liquor was made illegal.*
(b) During prohibition although a majority of people did support *drinking liquor was made illegal.*
5. (a) As a direct result of their new invention's influence the *company was given an award.*
(b) The retiring chairman whose methods still greatly influence the *company was given an award.*
6. (a) No matter how well trained these new pilots are flying planes *can be dangerous.*
(b) Living near an airport where the new pilots are flying planes *can be dangerous.*

These stimulus strings were sorted into four groups of six such that each group contained one member from each of the original six pairs. They were then placed on tapes with a variety of other sentences designed to prevent a bias for any particular click position. These "padding" sentences were, with minor variations, the same in all four groups. The six stimulus sentences for each group appeared in the same order and in the same serial position in each group (i.e., across groups, any two successive stimulus strings were separated by approximately the same number and type of padding strings).

Subjects.

The subjects were 72 high school students, all in their 3rd and 4th year. All were volunteers who were paid a small fee for their participation.

Procedure.

The subjects were run in small groups over a period of three days. The recorded material was presented dichotically (click in one ear, sentences in the other) over Jenson K-4 shielded headphones. Ss were required to write down the entire sentence and to indicate (by

TABLE 1 Frequencies of response at points of variation in constituent boundaries for six pairs of stimulus strings

| | |
|--|--|
| (1) "George" a 9 0 b 1 11 p = .00005 | (4) "liquor" a 6 0 b 0 5 p = .0022 |
| (2) "Anna" a 13 2 b 5 4 p = .1130 | (5) "the co." a 7 2 b 1 12 p = .0015 |
| (3) "Hamb." a 11 0 b 5 11 p = .0003 | (6) "planes" a 7 3 b 3 5 p = .1850 |

making a slash mark through their written version) the point in the sentence where they believed the click to have occurred. Ss were given ample time to do this and were told to indicate their confidence by noting those judgments which they felt to be guesses on their part.

Results

All the responses for each of the 12 sentences were combined (i.e., for both click positions) since only the comparison between members of a pair is relevant here. As an illustration of the distributions obtained, Fig. 1 shows one of the six pairs with all responses tabulated, the positions of clicks and the points of comparison indicated. Differences in the number of responses (as in Fig. 1) for each member of an acoustically matched pair arise from scoring. Responses were discarded where S failed to correctly record that portion of the sentence in which he located the click (i.e., his response was undeterminate with respect to the correct version of the sentence).

The frequency of response in each of the comparison categories (pts. of variation in the constituent structure) is given in Table 1 for the six pairs of sentences. Fisher's exact test was used to determine the probability of obtaining the observed frequencies (or those more extreme).

Discussion

This study was intended to determine whether earlier results should be interpreted as reflections of the assignment of constituent structure during the processing of sentences or were rather effects of correlated acoustic variables (such as pause and intonation) which tend to mark constituent boundaries. The results clearly demonstrate that the latter hypothesis is not adequate. Exactly the same acoustic signal was responded to differently in every case and the differences were uniformly as predicted by the intended variation in the constituent structure.

Table 1 shows the tests for significance of these differences between the members of each pair at the points where constituent structure was varied. Four of the six pairs are significantly different at beyond the .01 level. The two remaining sentence pairs, although not significantly different ($p = .113$ and $.185$), show concentrations of response in the predicted positions.

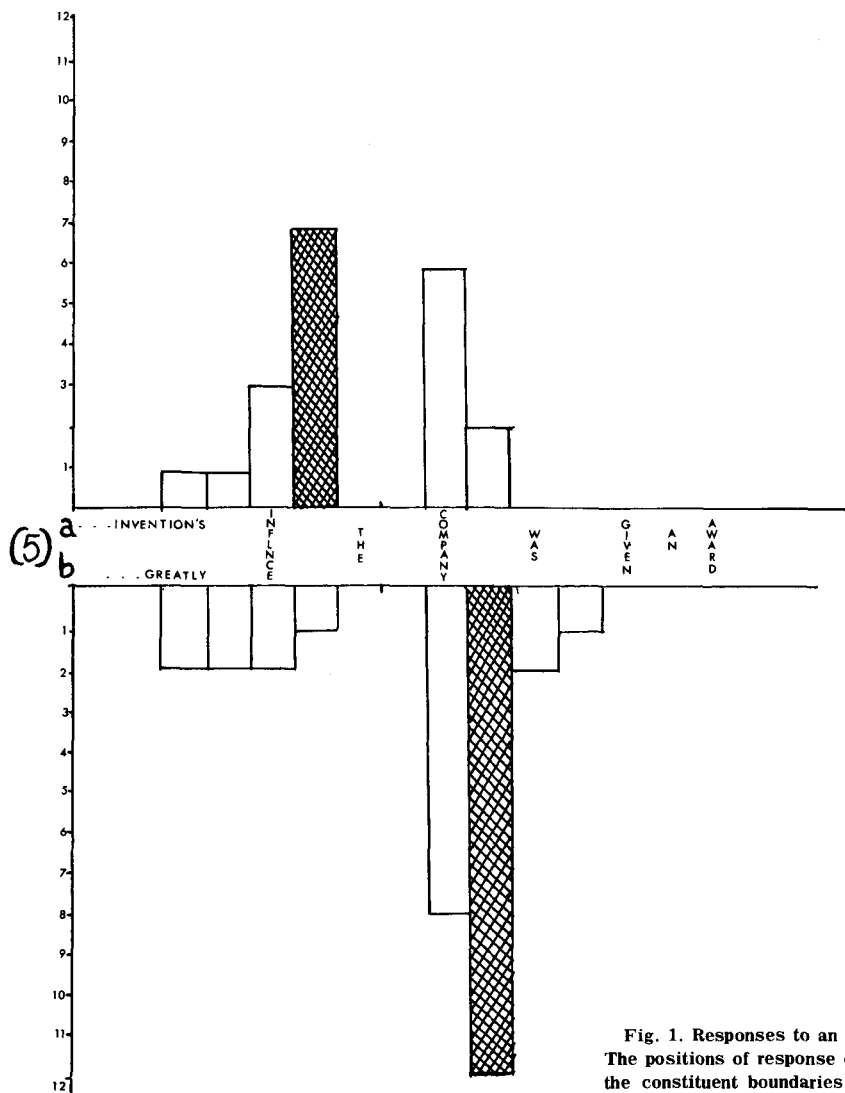


Fig. 1. Responses to an acoustically matched pair of sentences. The positions of response concentration predicted by variations in the constituent boundaries are cross-hatched areas in the graph.

If one considers only which of the two comparison categories shows the larger number of responses, the pattern conforms to variation in constituent boundaries for eleven of the twelve instances in Table 1 ($p = .003$ for this pattern). A possible explanation for the reduced differences between the members of pair No. 6 lies in a failure to provide a unique constituent structure for one of the members. After the experiment it was realized that both members of pair No. 6 may be given readings which assign the same constituent boundaries. Six (b) may be given the reading:

(Living near an airport) (where the new pilots are)
(flying planes can be dangerous)

rather than the intended:

(Living near an airport) (where the new pilots are)
(flying planes) (can be dangerous)

The possibility of this additional reading may account for the reduced differences between 6(a) and 6(b) at the predicted points.

The primary significance of these results is the support they provide for a view of sentence decoding as an active process in which the listener provides the structural analysis of the sentence rather than responding passively to some acoustic cues which mark the structure. It is clear that the differential response of Ss in this experiment cannot be accounted for with any "passive" theory of speech perception. The results here provide added evidence of the operation of an active process of perception at the level of the grammar's syntactic component.

References

- Fodor, J., & Bever, T., The psychological reality of linguistic segments. *J. verbal Learn. verbal Behav.*, in press.
- Garrett, M., Structure and sequence in judgments of auditory events. Institute of communications research, University of Illinois, 1964.
- Ladefoged, P., & Broadbent, D., Perception of Sequence in Auditory Events. *Quart. J. exp. Psychol.*, 1960, 12, 162-170.

Notes

1. Now at Massachusetts Institute of Technology.
2. Now at Harvard University.