

Semantic focus and sentence comprehension*

ANNE CUTLER and JERRY A. FODOR

Massachusetts Institute of Technology

Abstract

Reaction time to detect a phoneme target in a sentence was found to be faster when the word in which the target occurred formed part of the semantic focus of the sentence. Focus was determined by asking a question before the sentence; that part of the sentence which comprised the answer to the sentence was assumed to be focussed. This procedure made it possible to vary position of focus within the sentence while holding all acoustic aspects of the sentence itself constant. It is argued that sentence understanding is facilitated by rapid identification of focussed information. Since focussed words are usually accented, it is further argued that the active search for accented words demonstrated in previous research should be interpreted as a search for semantic focus.

The meanings of individual words and the combinatorial information provided by syntax do not always constitute the entire meaning of a sentence. The relationship of a sentence to its context is indicated, for example, by the focus structure of the sentence, where the focus is that information which is new and unrecoverable from preceding discourse. Speakers commonly indicate focus by assigning primary sentence accent to focussed words. Thus the new information in (1) is that it was the man on the corner, not some other man, who wore the blue hat, whereas in (2) the new information is that the hat was blue and not some other color:

- (1) The man on the CORNER was wearing the blue hat.
- (2) The man on the corner was wearing the BLUE hat.

Although accenting a word is the most usual way of expressing focus, various syntactic devices can produce the same effect, for instance the topicalisation of “blue” in (3):

*This research was supported by N. I. H. grant number 5-ROI-HD05168-05 to Jerry A. Fodor and Merrill F. Garrett. The authors are grateful for comments from Dianne Bradley, David Fay, Merrill Garrett, David Swinney, Virginia Valian and Edward Walker. Our thanks also to Dumont Billings, who ran the experiment, and to Edward Walker for providing the ALICE statistical analysis program. Reprint requests should be addressed to the first author at the Laboratory of Experimental Psychology, University of Sussex, Brighton, BN1 9OG, England.

- (3) Blue, the hat was, that the man on the corner was wearing.

Linguists have devoted considerable attention to the manifestations of focus (see e.g., Halliday (1967), Jackendoff (1972)). Psycholinguistic studies of comprehension have not, so far, explicitly addressed the question of how hearers identify the focus of a sentence. Nevertheless, some indirect evidence can be adduced; focus is very often expressed as accent, and the processing of a sentence's suprasegmental structure (accent and intonation contour) has been extensively investigated. The results of a series of phoneme-monitoring studies of accent appear to offer a starting-point for investigating the comprehension of focus.

In the phoneme-monitoring task, subjects are asked to comprehend sentences and at the same time to listen for the occurrence within them of a word beginning with a specified target sound. Reaction times to the target sound in this task are faster if the target begins a word which bears accent (sentence stress) than if it begins a non-accented word (Shields, McHugh and Martin, 1974; Cutler and Foss, 1977). Since accented words are acoustically clearer in several ways (Lehiste, 1970), it might be argued that this result reflects merely the greater ease of identifying acoustically clearer words. Cutler (1976a), however, concluded that at least part of the reaction time advantage of accented targets was due to an active search for accented words on the part of the listener. In Cutler's experiment each sentence was produced in two suprasegmental versions, one in which the target-bearing word was accented and one in which it was not. The target-bearing word was then spliced out of each version and replaced by identical copies of the same word taken from a third, fairly monotone rendition of the same sentence. Thus the target-bearing word was acoustically identical in the two versions of each sentence to be compared, but the suprasegmental context in which it occurred differed greatly from one version to another: in one version the intonation of the early part of the sentence was consistent with accent falling on the target-bearing word, in the other version it was consistent with accent falling elsewhere. Reaction times were significantly faster to targets which occurred in positions which should have borne accent than to targets which occurred where no accent was predicted. Since the acoustic clarity of the target-bearing word did not differ for the two versions, the reaction time difference in this experiment must reflect the suprasegmental differences in the part of the sentence which preceded the target. In other words, cues in the intonation contour must have enabled listeners to direct their attention to a part of the sentence where accent was about to occur.

The function of sentence accent is, above all, to indicate new information. It is clear that locating accented words as quickly as possible could, therefore,

be a useful strategy in sentence comprehension. Once the accented words are identified, the hearer knows which part of the sentence is new information, i.e., which part is most essential for grasping the speaker's message. Thus the reaction time advantage of accented words in phoneme-monitoring seemed to us to be due at least in part to the fact that accent expresses focus. To test this explanation, we designed the present experiment. If the reason that accent facilitates phoneme-monitoring reaction times is because it indicates focus, then, we predicted, a similar reaction time advantage for focussed words should be obtainable when focus is indicated by non-suprasegmental cues.

The main design problem with such an experiment lies in the very premise on which it is based, namely, that sentence focus and accent coincide; since accent affects phoneme-monitoring reaction times, to demonstrate an analogous effect for focus we must remove the confounding with accent, i.e., keep the acoustic representation of the sentence constant across changes in focus. However, as we have mentioned, accent is not the only method of focussing a word in a sentence.¹ In this experiment focus was manipulated by preceding the sentence with a question, the answer to which was the word to be focussed. Thus when (4) is preceded by (5) it can be said to have a focus structure identical to that of (1), whereas preceding it by (6) will produce the focus structure of (2):

- (4) The man on the corner was wearing the blue hat.
- (5) Which man was wearing the hat?
- (6) What hat was the man wearing?

In the present experiment, a set of sentences similar to (4) was presented to subjects. Half the subjects heard the sentence preceded by (5), half heard it preceded by (6). The experimental tapes were spliced together from separate recordings of sentences and questions, so that the monitored sentence (4) was actually acoustically identical in each presentation. Because acoustic cues in the sentence could favor one monitoring location over another, each sentence contained two possible phoneme targets; within each of the two question groups, half the subjects listened for one target, half for the other. Thus

¹ These methods include topicalisation, clefting ("It is the man who is wearing the hat") and pseudo-clefting ("The one who is wearing the hat is the man"), all of which are less suitable than the method actually adopted since they involve variations of the surface structure of the sentences to be compared.

Allen and O'Shaughnessy (forthcoming) report that various devices used to indicate sentence focus—clefting, pseudo-clefting, topicalisation, preposed question—produce reliable and essentially similar effects on the suprasegmental contour; fundamental frequency accent occurs in each case on the element which is focussed. (In the present experiment these explicit cues were of course not present since the questions were produced separately from the sentences).

in (4) half the subjects listened for a word beginning with /k/, half listened for a word beginning with /b/.

If phoneme-monitoring reaction time were indeed to prove sensitive to focus variation in the manner suggested by the results from the accent experiments, then it was expected that subjects monitoring in (4) for the /k/ sound on the word "corner" would respond faster if (4) was preceded by (5) than if it was preceded by (6), while subjects monitoring in the same sentence for the /b/ sound on "blue" would produce faster reaction times if (4) was preceded by (6) than if it was preceded by (5).

Thus we predicted an interaction between the two variables of Focus Position and Target Position, such that reaction times to earlier targets would be faster when the preceding question focussed upon the earlier part of the sentence, whereas reaction times to later targets would be faster when focus was upon the later part of the sentence.

On the basis of other findings, a main effect of Target Position was also predicted. It is a reliable finding in phoneme-monitoring experiments that targets near the end of a sentence produce faster reaction times than targets near the beginning of a sentence (Foss, 1969; Shields, McHugh and Martin, 1974). It has been argued (Cutler, 1976b) that this effect reflects the construction and testing of semantic hypotheses during sentence comprehension, with a greater probability of correct hypotheses being constructed as the sentence progresses towards its end.

Methods

Materials

Thirty-two experimental sentences were constructed and recorded along with thirty-two distractor sentences. Also recorded were two questions for each experimental sentence, one question directed to information in the earlier part, one to information in the later part of each sentence. For each experimental sentence two possible target specifications were determined, with one target-bearing word occurring in each of the two parts of the sentence to which the questions were directed (as in (4) above). Thus each experimental sentence occurred in four versions, with all possible combinations of target and question. Each distractor sentence was also preceded by a target specification and a question. The experimental sentences with their questions are listed in Appendix I.

Care was taken in recording the experimental sentences to keep the intonation contours as neutral as possible; in particular, no potential target phrase was assigned emphatic accent.

Since the targets in the experimental sentences always occurred in modifiers (attributive adjectives or prepositional phrases), the distractor sentences were designed to prevent this regularity becoming apparent. Targets in distractor sentences began verbs, head nouns of noun phrases, adjectives in predicate position and prepositions. Twelve distractor sentences did not contain an instance of the specified target.

Four materials sets were constructed, each containing all distractor sentences and a version of each experimental sentence. No experimental sentence occurred in the same version on more than one tape, and the two variables of Target Position (early versus late) and Focus Position (early versus late) were counterbalanced across sets. By means of tape-splicing the need for multiple recording of the sentences was avoided; the experimental sentences were acoustically identical in each set although preceded by different target-question combinations.

Subjects

Seventy-nine members of the MIT community, recruited by advertisement on campus, took part in the experiment. Each subject was paid \$2 for participating. Twenty subjects heard each of three materials sets, nineteen subjects heard the fourth materials set.

Procedure

Subjects were tested in groups of 2 or 3 at a time. They were instructed first to pay careful attention and comprehend the sentences as they would be tested on them later, and second to listen within the sentences for the occurrence of the target sound for that sentence, and to press a button as soon as they heard a word beginning with that sound. The target sound, which could be /b/, /d/ or /k/, was specified immediately prior to the question which preceded the sentence. The subjects were informed that the target-bearing word would occur only in the sentence, never in the question. No question contained a word beginning with the target sound for the sentence following it.

The instructions and materials were presented binaurally over Telephonics TDH-49 headphones. A signal on the tape, synchronised with the onset of the target sound but inaudible to the subjects, started a separate timer for each subject which was stopped by the subject's action in pressing the button.

After hearing the sentences the subjects were given a post-test which covered all 32 experimental sentences. The subjects were required to make the correct choice among four alternatives replacing one word in the sentence. In half of these judgments the subjects were providing words which had constituted the answer to the question they had heard, in half they were pro-

viding words from the other part of the sentence; in half of the cases, the words provided had been target-bearing words, in half they had not.

Results

1. Reaction Times

The data from three subjects who had missed a large number of targets were omitted from this analysis. Seven per cent of the remaining data was lost for varying reasons: of a total of 2432 (76×32) target presentations, 26 were not recorded because of equipment failure or experimenter error; in 57 cases, the subject failed to respond; 16 very long reaction times were excluded as possible reprocessing ("double take"); and 73 very short reaction times were excluded as anticipations. No more than 5 responses were lost for any one subject.

A mean reaction time for each condition was computed for each of the remaining 76 subjects; the overall means for each condition are shown in Table 1. A mean reaction time for each condition was also computed for each sentence. Separate analyses of variance were performed on these two sets of means; the combined result of the two analyses allowed computation of the *min F'* statistic (Clark, 1973).

Both main effects and the interaction between them reached significance. Reaction times to targets in later positions were significantly faster than reaction times to targets in earlier positions (*min F'* (1,50) = 20.86, $p < 0.001$). Reaction times were faster when the preceding question focussed upon the earlier part of the sentence than when it focussed upon the later part of the sentence (*min F'* (1,95) = 8.45, $p < 0.01$). The predicted interaction was also found: reaction times to the earlier target were faster in the early than in the late focus condition, whereas reaction times to the late target were faster in the late than in the early focus condition (*min F'* (1,84) = 19.36, $p < 0.001$).

Table 1. *Mean RTs (msec) to presence of target phoneme*

| | | Target position | | \bar{X} |
|----------------|-----------|-----------------|------|-----------|
| | | early | late | |
| Focus position | early | 376 | 360 | 368 |
| | late | 445 | 344 | 394 |
| | \bar{X} | 410 | 352 | |

The difference between focussed and non-focussed targets in the late target condition, which was small, was examined separately and found to be significant ($t(31) = 1.7, p < 0.05$ (one-tailed)).

The additional variable of Materials Sets in the analysis by Subjects was not significant, and did not interact with the Target Position variable or with the Target Position by Focus Position interaction. It did, however, interact weakly with the Focus Position variable ($F(3,72) = 2.91, p < 0.05$); although all four Materials Sets showed a Focus Position effect in the same direction, one set showed a stronger effect than the others. It is doubtful whether much importance should be attributed to this finding.

2. Post-test

Overall performance on this test maintained a high level: mean percent correct across all 79 subjects was 70% (chance performance would be 25%). When the word to be chosen had been the target-bearing word the mean number of errors across subjects was 4.66, when it had not borne the target the mean number of errors was 4.99, a non-significant difference ($t(78) = 0.288, p > 0.1$). However, as had been expected, words which had been focussed were retained better, so that significantly fewer errors were made when the word to be chosen had been the focussed word than when it had not been focussed (4.23 versus 5.41; $t(77) = 2.368, p < 0.025$ (one-tailed)).

Discussion

As predicted, reaction times to the earlier target in a sentence were faster when the preceding question had focussed on the earlier part of the sentence, reaction times to the later target were faster when the focus was upon the later part of the sentence. Furthermore, as predicted, targets near the end of the sentence were responded to faster than earlier targets.

Reaction times were also faster when the preceding question had focussed upon the earlier part of the sentence than when it focussed upon the later part. This result presumably reflects completion of the focus-identifying (question-answering) task, freeing greater processing capacity for other tasks. Subjects in phoneme-monitoring experiments are in any case performing a divided attention task, listening for the target and comprehending the sentence simultaneously. In this experiment the presence of the preceding question provided yet another task, namely determining the answer to the question, despite the fact that the subjects were not required to give the answer after hearing the sentence. Extensive work in recent years on the division of attention has shown that simultaneous performance of two competing tasks can lead to a decrement in performance on one or another or both in com-

parison with the level of performance on the same tasks when performed in isolation. In the present case it is likely that simultaneous performance of the question-answering task would interfere with performance of the monitoring task, and the reaction time data lend support to this assumption. (Note that it is unlikely that the question-answering task, or for that matter the monitoring task, would interfere with comprehension. Highly overlearned tasks appear to be resistant to interference from competing tasks (Moray, 1969), and sentence comprehension is probably the most highly overlearned task that subjects in a laboratory experiment are ever asked to perform).

Our chief prediction, however, concerned the interaction between target position and position of focus. We have shown that reaction time to a phoneme target on a particular word is faster when the word is focussed than when it is not. We would like to argue that this result indicates that the previous finding of a reaction time advantage of accented words reflects at least partly the role of accent as an indicator of focus, or new information. Similarly, the finding by Cutler (1976a) that listeners actively search for words which will be accented can be interpreted in the light of the present experiment as a search for sentence focus.

The most impressive aspect of the operation of the sentence comprehension device is its speed. For a sentence to be understood, the part which above all needs to be apprehended must be the focus—the part which is totally new. Any strategy which allowed the comprehension device to find the sentence focus rapidly would presumably facilitate understanding. Thus we suggest that our results indicate that particular attention is directed to the processing of focussed words when a sentence is being understood. Once the focussed segment has been located (usually by tracking the intonation contour to locate the main accent), processing of it begins immediately, even though in some cases processing of preceding segments is not complete. Thus phoneme targets on focussed words can be identified more rapidly than targets on non-focussed words.

We should note that a possible alternative explanation would stress not the prior processing of focussed words, but rather their relative availability for controlling responses. On this account, the order of perceptual analysis of material in the sentence would be independent of focus, but once the focussed segment has been identified it becomes the item most accessible for report. The present data are compatible with this interpretation, but it seems to us relatively implausible. Phoneme-monitoring reaction times appear to be predominantly affected by on-line processes, above all lexical access (Cutler and Norris, 1979). This latter account would in any case still bolster our principal conclusion, that the identification of semantic focus is an essential component of sentence comprehension.

References

- Allen, J. and O'Shaughnessy, D. (forthcoming) The effect of focus shifting transformations on fundamental frequency contours.
- Clark, H. H. (1973) The language-as-fixed effect fallacy: A critique of language statistics in psychological research. *J. verb. Learn. verb. Behav.*, 12, 335–359.
- Cutler, A. (1976a) Phoneme-monitoring reaction times as a function of preceding intonation contour. *Percep. Psychophys.*, 20, 55–60.
- Cutler, A. (1976b) Beyond parsing and lexical look-up. In R. J. Wales and E. C. T. Walker (eds.), *New Approaches to Language Mechanisms*. Amsterdam, North-Holland.
- Cutler, A. and Foss, D. J. (1977) On the role of sentence stress in sentence processing. *Lang. Speech*, 20, 1–10.
- Cutler, A. and Norris, D. (1979) Monitoring sentence comprehension. In W. E. Cooper and E. C. T. Walker (eds.), *Sentence Processing: Psycholinguistic Essays Presented to Merrill Garrett*. Hillsdale, N.J., Erlbaum.
- Foss, D. J. (1969) Decision processes during sentence comprehension: effects of lexical item difficulty and position upon decision times. *J. verb. Learn. verb. Behav.*, 8, 457–462.
- Halliday, M. A. K. (1967) Notes on transitivity and theme in English. Part II. *J. Ling.*, 3, 199–244.
- Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge, Mass., MIT Press.
- Lehiste, I. (1970) *Suprasegmentals*. Cambridge, Mass., MIT Press.
- Moray, N. (1969) *Attention: Selective Processes in Vision and Hearing*. London, Hutchinson.
- Shields, J. L., McHugh, A. and Martin, J. G. (1974) Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *J. exper. Psychol.*, 102, 250–255.

APPENDIX I

Sentences used in the experiment

1. The man on the corner was wearing the blue hat. Which man was wearing the hat? Which hat was the man wearing?
2. The reporter with the daily newspaper was responsible for the candid story. Which reporter was responsible for the story? Which story was the reporter responsible for?
3. The opening of the concert was spoiled by the director's outburst. Which opening was spoiled by the outburst? Whose outburst spoiled the opening?
4. The value of the bonds was altered with the devalued currency. Which value was altered with the currency? Which currency altered the value?
5. The house with the carport must belong to the doctor's widow. Which house must belong to the widow? Which widow must the house belong to?
6. The woman with the bag went into the dentist's office. Which woman went into the office? Which office was it that the woman went into?
7. The checking of the ballots was interrupted when the computer failed. What checking was interrupted when the failure happened? Which failure interrupted the checking?
8. The author of the bestseller refused to go to the Congressman's party. Which author refused to go to the party? Which party did the author refuse to go to?
9. The motion to dismiss was proposed by the Californian senator. Which motion was proposed by the senator? Which senator proposed the motion?

10. The manager of the dairy will check on his bank account. Which manager will check on his account? Which account will the manager check on?
11. The mother of two daughters wrote to the boarding school. Which mother was it that wrote to the school? Which school was it that the mother wrote to?
12. The interviewer from the digest questioned the book's publisher. Which interviewer was it that questioned the publisher? Which publisher was it that the interviewer questioned?
13. The association of consumers objected to the new delivery procedure. Which association was it that objected to the new procedure? Which new procedure was it that the association objected to?
14. The price of bread was raised at the Council session. Which price was raised at the session? Which session raised the price?
15. The group of dancers negotiated with the Broadway agent. Which group negotiated with the agent? Which agent was it that the group negotiated with?
16. The owner of the department store fired the bicycle salesman. Which owner fired the salesman? Which salesman was it that the owner fired?
17. The voice of the caller was hard to hear on the defective telephone. Which voice was hard to hear on the telephone? Which telephone was the voice hard to hear on?
18. The road to the Cape was washed out in the December rains. Which road was washed out in the rains? Which rains washed out the road?
19. The suspension of the boy was protested at the college meeting. Whose suspension was protested at the meeting? Which meeting protested the suspension?
20. The flavor of the coffee was ruined by the dirty water. Which flavor was ruined by the water? Which water ruined the flavor?
21. The remains of the camp were found by the deer hunter. Which remains were found by the hunter? Which hunter found the remains?
22. The actions of the committee focussed on the dangerous situation. Which actions focussed on the situation? Which situation was it that the actions focussed on?
23. The attitudes of the businessman aroused his colleagues' anger. Whose attitudes aroused anger? Whose anger did the attitudes arouse?
24. The watcher on the balcony saw the driver's escape. Which watcher was it that saw the escape? Whose escape was it that the watcher saw?
25. The regulations of the commission set new billboard standards. Which regulations set new standards? Which new standards did the regulations set?
26. The residents of the district were annoyed at the building plans. Which residents were annoyed at the plans? Which plans were the residents annoyed at?
27. The program about Britain interested the common viewers. Which program interested the viewers? Which viewers did the program interest?
28. The fear of death inspired the courageous fighters. Which fear inspired the fighters? Which fighters were inspired by the fear?
29. The tourists from Denmark photographed the Bicentennial parade. Which tourists photographed the parade? Which parade was it that the tourists photographed?
30. The personnel officer of the company interviewed the baseball player. Which personnel officer interviewed the player? Which player did the personnel officer interview?

31. The rising price of boxes worried the candy manufacturer. Which rising price worried the manufacturer? Which manufacturer was worried at the rising price?
32. The janitor at the ballpark joined the custodians' union. Which janitor joined the union? Which union did the janitor join?

Résumé

On a trouvé que le temps de réaction pour détecter un phonème cible est plus rapide quand cette cible se trouve dans le mot qui fait partie du focus sémantique de la phrase. Ce focus est déterminé par la pose d'une question avant l'énoncé de la phrase: on suppose que la part de la phrase qui contient la réponse à cette question est le focus. Cette procédure permet de varier les positions du focus dans une phrase tout en n'en tenant constants les aspects acoustiques. On soutient que la compréhension d'une phrase est facilitée par l'identification rapide de l'information focalisée. Les mots focalisés étant le plus souvent accentués, on peut aussi soutenir que la recherche active des mots accentués que des expériences précédentes avaient mis en évidence peut être interprétée comme une recherche sémantique.