The interaction of veracity and syntax in the processing of sentences [1]

JACQUES MEHLER, [2] *MASSACHUSETTS INSTITUTE OF TECHNOLOGY*
PETER CAREY, *HARVARD UNIVERSITY*

The reaction times of 80 Ss in judging sentences true or false with respect to pictures were analyzed, and it was noted (a) that true sentences containing an expected surface structure required less time than false sentences of the same structure, (b) that latencies to true sentences containing an unexpected surface structure were longer than latencies to the same sentences when their structure was identical to that of previous sentences, and (c) that transitive-verb constructions appeared easier to judge than predicate nominatives of the same length.

McMahon (1963) and Wason (1959) found that it takes a subject less time to indicate that a true sentence is true than to indicate that a false sentence is false. Mehler and Carey (1967) demonstrated that for sentences in noise an unexpected change in the surface structure of a sentence causes great difficulties in perceiving that sentence. The present experiment employed lists of sentences similar to those of Mehler and Carey (1967) and required Ss to determine whether each sentence was true or false with respect to a picture. The experiment provides information on the question of whether the time to process the veracity of sentences interacts with the time needed to process the surface structure of sentences. Although McMahon found that latencies to indicate correctly whether a sentence is true or false increase with certain grammatical transformations (e.g., passivization, negation), these findings do not determine whether there is an interaction between the computation of truth and the processing of syntax. Such an interaction is part of the general problem of the relation of syntax and semantics in the understanding of sentences.

Procedure

Eighty undergraduate and graduate students at the Massachusetts Institute of Technology served as paid Ss in the experiment. All were native speakers of English. Each S was assigned to one of four experimental groups, and every S was run individually by the same E. All Ss received 11 sentences through earphones, and each sentence was preceded by a specially prepared picture. S's task was to flip a switch to indicate whether each sentence was true or false with respect to the picture. For half of the Ss in each group of 20, an upward deflection signified "true," whereas for the other half, the same judgment was indicated by a downward deflection of the switch. An auditory signal at the end of each recorded sentence activated a timer, which ran until S's response broke the circuit. The timer therefore measured the latency of the response by which Ss judged the veracity of the sentence. Ss were instructed to respond as quickly as possible without making errors; all groups received practice on warm-up sentences.

For two groups of 20 Ss each, the first 10 sentences contained a predicate nominative construction, as in (A), which was their 10th sentence. The 11th and final sentence for these Ss was (B), which contained a transitive-verb construction that differed from the surface structure of the preceding sentences.

(A) They are performing monkeys.

(B) They are bombarding cities.

For one group, (A) was true and (B) false; for the other group, the sentences had the opposite truth values. The remaining two groups first received 10 transitive-verb sentences as in (B), which was their 10th sentence. The 11th sentence for these Ss was (A), which contained the predicate nominative construction. For one group, (B) was true and (A) false; for the other group, the truth values were reversed. The experimental design therefore allowed the collection of latency data on test sentences (A) and (B) when each was (a) true and compatible with the surface structure of previous sentences, (b) true and incompatible with the previous structure, (c) false and compatible, and (d) false and incompatible.

For all groups the pictures preceded the sentences, and S had as much time to study the picture as he wished before E turned on the tape recorder that presented the sentence. The relation between pictures and sentences was always such that S could correctly determine the truth or falsity of the sentence only after hearing the last word. For example, when *They are performing monkeys* was true, the sketch depicted circus monkeys, and circus horses when it was false. The sequence of appropriate true and false responses was randomized, with the restriction that two lists contained six true sentences and five false sentences, whereas the remaining two lists contained five true sentences and six false sentences.

Results and Discussion

In order to correct for the characteristic positive skewness of reaction time data, all latencies (X) were transformed to $\log_{10}X$. This transformation ren-

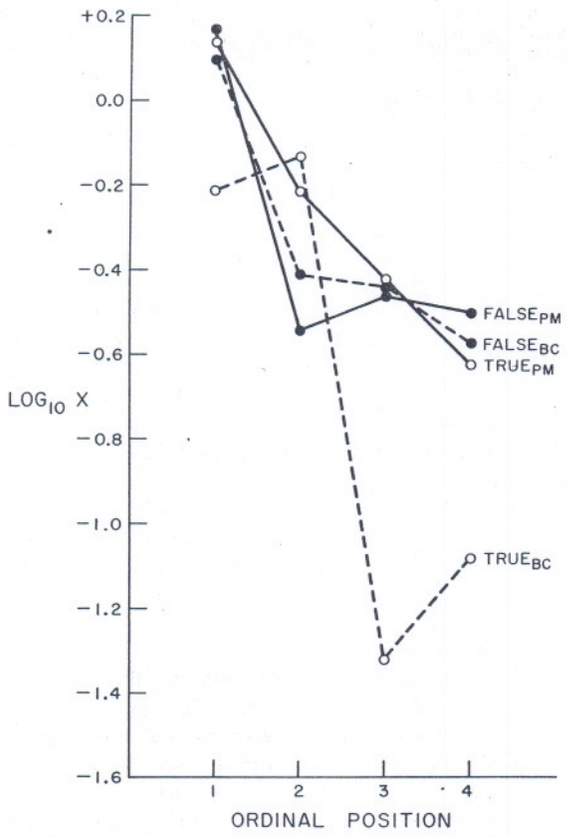


Fig. 1. The practice effect on the first four true and the first four false sentences of the two syntactic structures. Mean log latencies.

dered the latency distributions effectively normal and allowed a series of t tests to be carried out on the transformed data.

Figure 1 shows[3] that for both sentence types and both truth values a practice effect obtained, in which latencies to successive sentences of the same structure and truth value declined quite regularly. The figure also reveals that mean latencies for true sentences of the type *They are bombarding cities* (BC) were generally much shorter than those for true sentences of the type *They are performing monkeys* (PM). For false sentences the difference between sentence types disappeared. It seems likely that the difference between the latencies for true sentences of the two syntactic structures is a genuine syntactic effect, in which transitive-verb constructions are easier to process than predicate nominative constructions of the same length. This interpretation is supported by the fact that for the two test sentences the average latency for *They are bombarding cities* was significantly shorter than for *They are performing monkeys* ($t = 2.01$, $p = .05$, two-tailed, $df^4 = 147$). For the test sentences, neither the interaction of sentence type and position, nor the interaction of sentence type and truth value was significant.

The fact that transitive-verb constructions were processed more rapidly than predicate nominatives is interesting in view of the fact that a computation of depth according to Yngve (5) reveals that the predicate nominative is the more complicated structure. The Yngve count for *They are bombarding cities* is 4/4,

whereas the count for *They are performing monkeys* is only 3/4.

According to Martin and Roberts (1966), a low Yngve number should be related to relative ease of processing. Our results suggest that the Yngve count is not a reliable predictor of experimental difficulty.[5]

Figure 2 reveals that when the test sentences were true and compatible in structure (10th position), the latency was much shorter than when the same true sentences were incompatible in structure (11th position). This difference is highly significant ($t = 3.94$, $p < .001$, two-tailed, $df = 72$). When the test sentences were false, however, the difference between 10th and 11th positions was not significant ($t = 0.546$, $p > .10$, two-tailed, $df = 75$). The effect observed by Mehler and Carey (1967), in which a sentence with unexpected surface structure was misperceived, would seem to

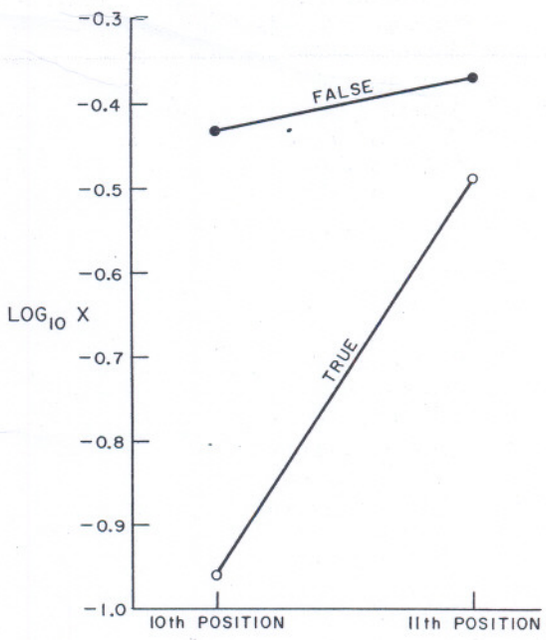


Fig. 2. The interaction of truth value and position. Mean log latencies.

have its analogue in increased processing time only when the sentence being processed is true. The effect of unexpected syntax seems to disappear when the sentence is false.

Figure 2 also shows that for test sentences in 10th position, true responses required much less time than false responses ($t = 4.31$, $p < .001$, 2-tailed, $df = 75$). For 11th position, the difference is in the same direction, but it is not significant ($t = 0.93$, $p > .10$, 2-tailed, $df = 72$). The effect observed by McMahon (1963) and Wason (1959), in which true sentences have shorter decision times than false sentences, would therefore appear to obtain only when sentence structure does not violate S's expectation.

When a sentence was false and contained an unexpected surface structure, then, response time was not significantly longer than when only one of these disrupting factors was present. This fact suggests that the time necessary to process falsity *or* unexpected structure is already so long that the other process can simultaneously occur without significantly increasing response time. This suggestion of simultaneous, parallel processing is interesting in terms of the following conceptualization. If surface structure is incompatible with S's expectation, the processing must return to the node at which the incongruity first because apparent. If, as seems likely, Ss expect sentences to be true, the presentation of a false sentence may cause S to retrace his analysis of the sentence, much as he retraces unexpected surface structure. In the sentences employed in the experiment, the last word determined both veracity and compatibility of surface structure. When a sentence was both false and unexpected, therefore, recomputation for veracity and structure critically involved only the last word, and the two sorts of analyses co-occurred. Had separate phrases been critical for veracity and for structure, however, parallel processing might not have occurred. Such might have been the case, for example, in the sentence *They are bombarding cities on the coast* where *cities* determines compatibility of surface structure, and *coast* determines veracity. Additional experiments are necessary to validate this prediction.

References

Martin, E., & Roberts, K. H. Grammatical factors in sentence retention. *J. verbal Learn. verbal Behav.*, 1966, 5, 211-218.

McMahon, L. Grammatical analysis as part of understanding a sentence. Unpublished doctoral dissertation, Harvard University, 1963.

Mehler, J., & Carey, P. The role of surface and base structure in the perception of sentences. *J. verbal Learn. verbal Behav.*, 1967, 6, 335-338.

Wason, P. C. The processing of positive and negative information. *Quart. J. exp. Psychol.*, 1959, 11, 92-107.

Yngve, V. H. The depth hypothesis. *Structure of language in its mathematical aspect.* 130-138. In R. Jakobson (Ed.), *Amer. Math. Soc.*, Providence, R. I., 1961.

Notes

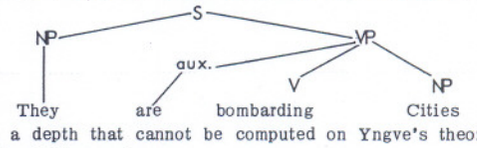
1. The research reported here was supported by Air Force Contract # AF 19 (628) - 5705 to Professor Jerry A. Fodor and NIH Grant 5-T01-HD00111 to Professor Morris Halle.

2. Present address for J. Mehler: Laboratoire de Psychologie, 17 Rue Richer, Paris 9, France.

3. Fig. 1 depicts only the first four true and the first four false sentences of each syntactic type, since in some cases the fifth true or false sentence was (1) or (2), and these latter data are treated elsewhere in the paper as part of the analysis of "test sentences".

4. The number of degrees of freedom for the 80 Ss' responses to sentences (1) and (2) is 147, and not 152, because 5 of the 160 responses were incorrect and were therefore discarded.

5. Although many linguists would even today disagree about the structure to be assigned to the sentence *they are bombarding cities*, our point does not suffer since the only alternative is



which has a depth that cannot be computed on Yngve's theory which requires binary nodes.

(Accepted for publication October 15, 1967.)