

Cluster analysis of word-order judgment latencies for German active and passive sentences

ALAN KENNEDY and MARLYN MILNE
University of Dundee, Dundee DD1 4HN, Angus, Scotland

A group of nine German-speaking Ss judged pairs of words as "forward" or "reversed" with regard to a previously learned sentence. Response latencies for all "forward" judgments were subjected to hierarchical cluster analysis. The resulting tree diagrams reflect some of the differences marked in the active and passive constructions in German. The results were consistent with the hypothesis that Ss store sentences as hierarchically organized subgroups.

A number of investigators have applied the technique of hierarchical cluster analysis (HCA) to data derived from sentence processing. Levelt (1970), using a measure of intelligibility in noise and applying Johnson's HCA technique, showed that Ss appear to "chunk" segments of a sentence into units corresponding roughly with major syntactic constituents, although minor constituents were not marked. Kennedy & Wilkes (1971) showed that when response latencies of judgments regarding the order of pairs of words taken from previously learned sentences are subjected to HCA, clear hierarchical structures are obtained. The groupings, identified from pausing records, related to Ss' subjective organization of the sentences during learning. Both of the above experiments provide a demonstration that Ss store and process memorized sentences in hierarchically organized subunits.

In the present study, German active and passive constructions are contrasted. In the passive form used here, certain words closely associated logically are widely separated in their surface structure; in particular, the verb and its auxiliary are separated (e.g., *Ihre arme Katze wurde von diesen kleinen Jungen geschlagen*—literally, *their poor cat was by those small boys hit*). The word order for this form is fixed, and thus the construction allows the possibility of assessing the functional importance of segments which are "interrupted" in the surface structure by whole sequences of other elements. To some extent, this passive construction parallels English sentences containing discontinuous constituents, e.g., *John phoned the girl up*. The German active form, apart from complications introduced as a consequence of selectional relations of government and agreement, parallels normal English active sentences. By contrasting the two types in a task

involving judgments of word order, it is possible to assess to what extent "functional" units relate to left-right processing of the word string and to what extent they relate to perceived

syntactic relationships within the sentence.

Decision latencies of judgments of word order were subjected to Johnson's HCA technique in an attempt to assess the organizational strategies adopted by Ss both when learning and when recalling the two types of sentence construction.

MATERIALS

Six sentences were used, consisting of three active forms and their passive equivalents. The sentences were: (1) *Diese kleinen Jungen schlugen ihre arme Katze*; (2) *Sein alter Vater empfing jenes schöne Geschenk*; (3) *Meine hübsche Tante kaufte einige neue Kleider*; (4) *Ihre arme Katze wurde von diesen kleinen Jungen geschlagen*; (5) *Jenes schöne Geschenk wurde von seinem alten Vater empfangen*; (6) *Einige neue Kleider wurden von meiner hübschen Tante gekauft*. The sentences were prepared

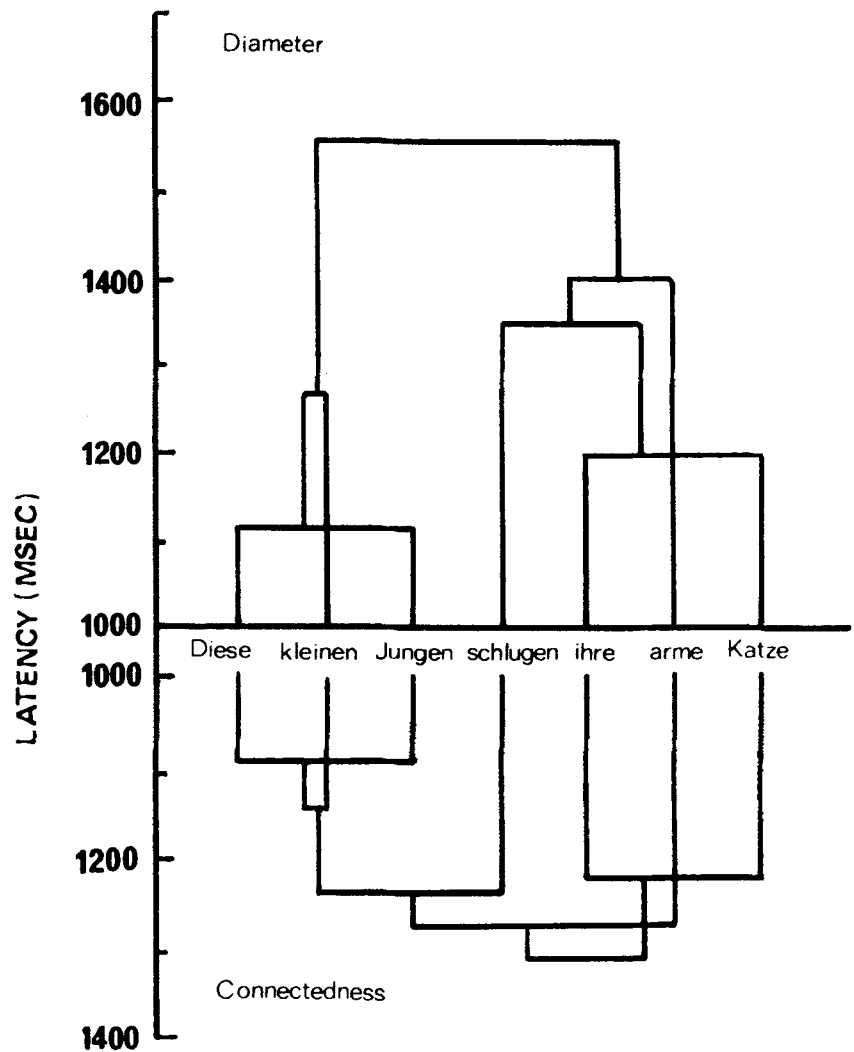


Fig. 1. Hierarchical cluster analysis of active sentence type (example sentence) derived from forward word-order comparison latencies. Diameter and connectedness methods of analysis are shown separately.

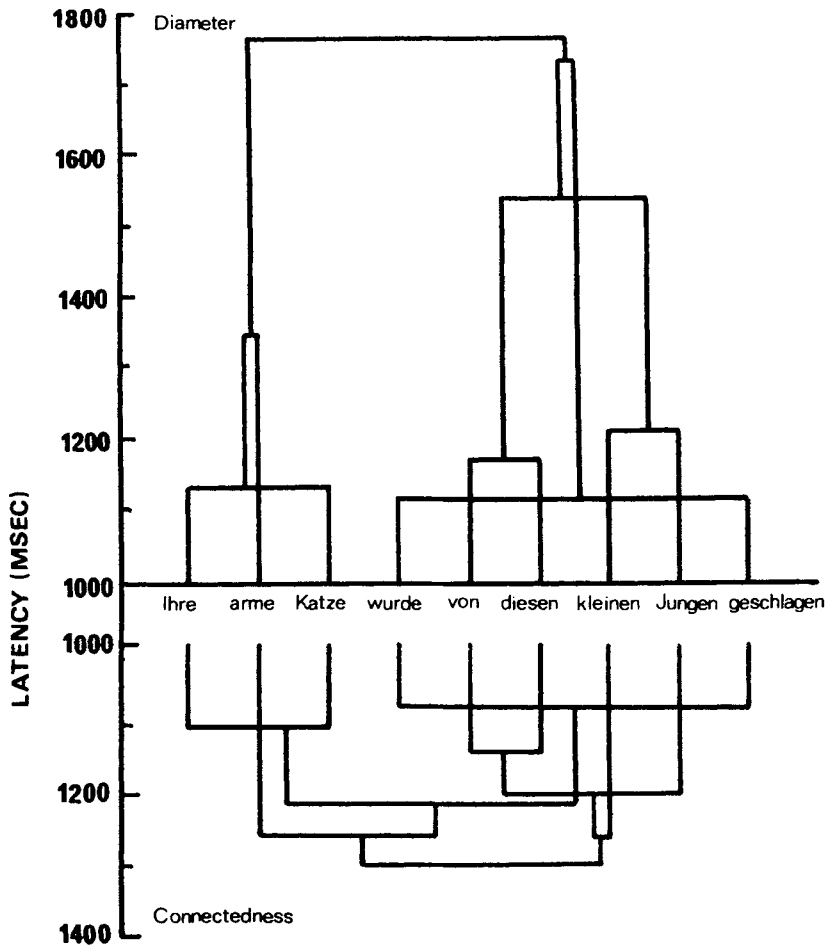


Fig. 2. Hierarchical cluster analysis of passive sentence type (example sentence) derived from forward word-order comparison latencies. Diameter and connectedness methods of analysis are shown separately.

in uppercase typescript on tachistoscope display cards. For each, a further set of cards was drawn up, displaying all possible pairs of words allowing both "forward" and "reversed" order with respect to the original sentence.

PROCEDURE

Each S learned one active and one passive construction, with the restriction that no S learned both forms of a particular sentence. Half the Ss learned the sentences in the order of active-passive and half in the reverse order. The learning phase consisted of S reading the sentence from a card and then recalling it with the card concealed until a criterion of three consecutive correct recalls was achieved. Immediately following this, all pairs of words from the sentence, both "forward" and "reversed," were presented in random order, each pair occurring twice. Word pairs were presented for 2 sec in a tachistoscope, and S responded by pressing a right-hand button to indicate "forward" order and a left-hand

button to indicate "reverse." Learning and responding to the second sentence took place on the following day. Testing for active sentences lasted approximately 30 min and for passives, approximately 45 min.

SUBJECTS

The Ss were 10 volunteers whose native language was German. All had spoken English for at least 10 years.

RESULTS

One S was lost as a result of a failure to report for the second session. The analysis was based on correct "forward" responses, following the procedure adopted by Kennedy & Wilkes (1971). Mean error rates were 5.5% for actives and 7.8% for passives.

The input data for HCA consisted of the off-diagonal submatrices comprising mean response latencies of "forward" decisions to all pairs of words (i.e., 1 vs 2; 1 vs 3; etc.).

Figure 1 shows the obtained tree diagrams for active sentences for the "connectedness" and "diameter" methods of solution. Theoretically, the connectedness ("minimum")

method may not be applied if the "distance" values (here latencies) are asymmetrical, which is obviously the case in the present study. The degree of agreement between the two methods of solution, however, can be used as a measure of the extent to which the data may be described by a genuinely *hierarchical* structure (Levelt, 1970). Figure 2 shows the obtained tree structure for passive sentences. Following the procedure adopted by Levelt, we made an attempt to assess the "goodness of fit" of the diameter tree solutions to the input data. A complete listing of all order relationships specified in the tree diagram was tested against the latencies given in the input matrix [e.g., the inequality $d(1,3) > d(1,2)$ in the tree diagram for active sentences was compared against the pair of latencies 1 vs 2 and 1 vs 3]. The *stress* of an HCA solution is given as the number of order relations violated by the data as a proportion of the total number of order relations (inequalities) specified by the HCA. For active sentences, there were 28 violations in a total of 140 specified order relations (20%). For passives, there were 108 violations in a total of 443 (24%). Relationships involving the verb or, in the case of the passive, the auxiliary plus verb accounted for 80% of all violations in actives and 71% in passives.

DISCUSSION

Taken together, these results show a reasonable degree of regularity, which is encouraging since the latency measure used is a more remote source of data than judgments of grammaticality or perceived intelligibility, where S's attention is necessarily directed to a left-right processing order for sentences.

The differences in latency of judgments shown here can be related to earlier studies on probe retrieval or probe location, where two types of search processes were discussed (Kennedy & Wilkes, 1968, 1969). When order information is *not* required to make a decision (e.g., in the case where S simply judges presence or absence of a stimulus), the latency differences appear to reflect the semantic importance of stimulus words. When order information must be preserved, as when S gives as response some other word, then latencies reflect structural features of the sentence. In the present task, although a "yes-no" decision is used, it is obvious that order information is retained, and the results tend to show not only the salient structural groups but also the manner in which smaller elements are themselves incorporated into larger superordinate structures. It appears that decisions regarding the

order of any particular word pair must call for the reconstitution of at least those parts of the tree which contain both elements in the pair.

A more detailed examination of the present results shows three further features. First, the first and third items of a sentence often occur as an initial cluster (e.g., *diese Jungen*), with a qualifying adjective only assimilated later (*diese kleinen Jungen*); this result possibly stems from the inflected forms of German adjectives where agreement in case, number, and gender of the qualified noun is demanded and the possibility for a particular adjective of agreement with both nouns arises, which may delay adjective-noun pair decisions. Second, in the case of the passive constructions employed, it is obvious that the verb and its auxiliary, although widely separated in the surface structure, are processed as a single unit, often appearing as the first cluster identified.

Finally, the major discrepancy between the two methods of solution

appears in the placing of the verb. It is not clear from the data of the present study (and the results of the stress measure confirm this) whether the verb should properly be shown as assimilated into the subject phrase or, as most grammars would maintain, into the predicate. This difficulty in characterizing the role of the verb in "subjective phrase structure" has been pointed out by Martin (1970). When the experimental task directly involves judgments of grammatical relationships, the verb is usually shown as part of the "predicate" in an HCA tree diagram (Levelt, 1970). However, in the present task, there is reason to suppose that the response latencies reflect strategies adopted by S when organizing the sentences during learning, and in many cases, such grouping strategies place a break *after* the main verb (Martin, Kolodziej, & Genay, 1971). If the present data can be viewed as providing an index of organizational and/or retrieval strategy, then they point to a degree

of uncertainty in the sample of bilingual Ss as to where the primary division in the sentence should be placed.

REFERENCES

- KENNEDY, R. A., & WILKES, A. L. Response-times at different positions within a sentence. *Quarterly Journal of Experimental Psychology*, 1968, 20, 390-394.
- KENNEDY, R. A., & WILKES, A. L. An analysis of storage and retrieval processes in memorizing simple sentences. *Journal of Experimental Psychology*, 1969, 80, 396-398.
- KENNEDY, R. A., & WILKES, A. L. Functional structure in sentences: A performance analysis. *Quarterly Journal of Experimental Psychology*, 1971, 23, 214-224.
- LEVELT, W. J. M. Hierarchical chunking in sentence processing. *Perception & Psychophysics*, 1970, 8, 99-102.
- MARTIN, E. Towards an analysis of subjective phrase structure. *Psychological Bulletin*, 1970, 74, 153-166.
- MARTIN, J. E., KOLODZIEJ, B., & GENAY, J. Segmentation of sentences into phonological phrases as a function of constituent length. *Journal of Verbal Learning & Verbal Behavior*, 1971, 10, 226-233.