

Models for within-proposition representation tested by cued recall

KARL F. WENDER and ULRICH GLOWALLA
Universität Braunschweig, 3300 Braunschweig, Germany

A model is presented to account for the data from incremental cuing experiments that have been carried out to identify the representation of propositions in memory. In such experiments subjects first learn a list of sentences and are afterward cued for recall with words from the learned sentences. The model proposed distinguishes between a memory structure and stimulus and response processes. The all-or-none tendency in the data is captured by a Gestalt-like memory code. The model is compared with the stochastic theory of Anderson and Bower and the fragmentation hypothesis of Jones.

This paper is concerned with a problem in semantic memory, that is, the form of the within-proposition representation. Many of the current theories of human memory use propositional networks as a means for representing knowledge in memory. These networks consist of nodes that correspond to concepts and links connecting the nodes that correspond to the relations between the concepts. Propositional networks have the further property that they possess subconfigurations that correspond to propositions. These subconfigurations have truth value and are structured according to syntactic rules. These syntactic rules specify what is called here the within-proposition representation.

Most of the current theories of human memory differ with respect to the within-proposition representation they assume. The present paper tries to gain further insight into this question by discussing three stochastic models for one particular experimental paradigm that has been used to identify the within-proposition representation. This paradigm is known under the name of incremental cuing.

We show in the following one example of a within-proposition representation, and then we sketch the paradigm of incremental cuing. Third, we briefly state the stochastic assumptions that have been proposed by Anderson and Bower (1973) and, subsequently, we formulate an alternative. Finally, we sketch the fragmentation hypothesis presented by Jones (1978) and compare the three approaches.

In the Anderson and Bower (1973) theory, a proposition is represented as a set of memory nodes and as a set of labeled associations between them. These associations are binary relations. The types of

This research was in part supported by Grant We 498/6 from the Deutsche Forschungsgemeinschaft to the first author. We would like to thank H. Colonius and H.-H. Schulze for many fruitful discussions. Furthermore, we thank H. Bittner and I. Bauer for their programming assistance. Requests for reprints should be sent to Ulrich Glowalla, Institut für Psychologie, Universität Braunschweig, Spielmanstrasse 19, Braunschweig 3300, Germany.

associations and the ways they combine to form a proposition are shown in Figure 1.

As can be seen in Figure 1, a proposition of four content words is represented as a rooted, labeled tree with 19 nodes, 4 of which are terminal nodes, and with 23 links connecting the nodes. In addition, the links incident with the terminal nodes are directed. It is also assumed that subjects enter (and leave) the memory network via the word nodes.

It may be noted that Anderson (1976) has changed the within-proposition structure in his theory. However, for the stochastic assumptions to be discussed, it does not matter from which structure we start; so, we are not going to describe the details here.

THE INCREMENTAL CUIING PARADIGM

In an incremental cuing experiment, subjects first learn a list of unrelated sentences. Each sentence consists, for example, of a location (L), an agent (A), a verb (V), and an object (O). In a following test phase, the experimenter selects one sentence, and from this sentence a word is presented as a cue. The subject tries

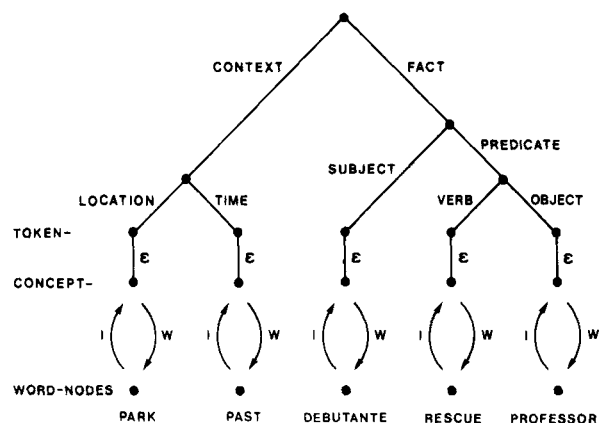


Figure 1. Representation of the sentence, "In the park the debutante rescued the professor," according to Anderson and Bower's theory (1973).

to recall the sentence. A second word is then selected and presented together with the first cue, and, again, the subject tries to recall the sentence. This is continued until all words from one sentence have been used for cuing. The same procedure is repeated with other sentences in the list. The data from such an experiment contain very detailed information. In an experiment with sentences consisting of four content words, 120 different experimental events can be distinguished depending on the sequence of cues and recalls.

This experimental paradigm has been used in several studies by Anderson and Bower (1973) and also by Colonius, Glowalla, Schulze, and Wender (Note 1). More recently, however, Anderson (1976, p. 164f) has argued against the use of recall data to identify within-proposition structures. Anderson believes that when subjects learn a list of unrelated sentences, they elaborate these sentences and, thereby, introduce extra connections between the concepts. These extra connections, then, would change the recall probabilities. Although this might well be the case, it remains to be demonstrated. The evidence up to now consists mainly of informal reports from subjects. In any case, if a model without additional links fits the data from a recall experiment, one would assume that a more elaborate model with additional links could do the same. Furthermore, if extra connections are set up, this might also affect latency measures in reaction time experiments, at least if speed-accuracy tradeoff functions are considered, since these functions are derived from the percentage of correct responses. Thus, if elaborations pose a problem, this would affect other experimental paradigms as well.

THE STOCHASTIC MODEL OF ANDERSON AND BOWER

The Anderson and Bower (1973) model distinguishes between a working memory and a long-term store. When hearing a sentence, a parser builds up a structure like that in Figure 1 in working memory. The content of the working memory is subject to time-dependent decay. During the time the tree resides in the working memory, the links and nodes are duplicated into the long-term store with some probability, namely:

$$p(t) = 1 - e^{-t/a}. \quad (1)$$

This is the probability that one particular link has been coded up to time t . The parameter a is assumed to be constant for all links within one proposition. But a varies across sentences with a density,

$$f(a) = \frac{1}{ba^2} e^{-1/ba}. \quad (2)$$

The parameter b corresponds to the overall mean rate of encoding links. The result is that $p(t)$ follows a beta distribution. Hence, one obtains the probability of recoding one particular configuration of k links from a tree of n links within a period of time t . It is as follows:

$$Q(k,n) = \frac{1}{ntb + 1} \prod_{i=1}^k \frac{tbi}{(n-i)tb + 1}. \quad (3)$$

This is the equation that governs the memory structure, in which the effective parameter is tb . Next, one computes the probability that a complete path exists from secondary node i to secondary node j . This probability is obtained by summing the probabilities for all those configurations that contain a complete path $w(i,j)$ from i to j . The model so far describes the memory structure. To complete the picture, stimulus effectiveness and response availability processes have to be added. It is assumed that different classes of words may be differently effective as cues. Therefore, for each class of words, a stimulus effectiveness parameter s_i is assumed. The s_i are assumed to be independent from each other and from the memory structure. Furthermore, when a subject tries to recall a sentence, the words from different word classes may differ in availability. Therefore, response availability parameters r_i are introduced. The r_i are assumed to have the distribution function,

$$F(r_i) = 1 - (1 - r_i)^{1/\alpha_i}. \quad (4)$$

However, they are not thought to be independent. Rather, it is assumed that the r_i come from identical percentiles of their respective distributions, if a subject actually tries to recall a particular sentence. From these assumptions the probabilities for each possible configuration of word links, that is, the word availability parameters, can be obtained.

To summarize, the model is characterized by three independent sets of assumptions: (1) a memory structure governed by the encoding parameter tb , (2) a set of independent stimulus effectiveness parameters, and (3) a set of covarying response availability parameters.

THE MULTICOMPONENT MODEL

Three things have motivated us to formulate an alternative model. The first is more technical in nature. In the Anderson and Bower (1973) model, it follows that the encoding of each of the links in the memory structure is not stochastically independent. Furthermore, Anderson and Bower introduce a special kind of dependency between response processes that makes

it somewhat clumsy to determine the probability for a given event. This has motivated us to look for a model in which stochastic independence is preserved as far as possible in order to make the model conceptually more parsimonious.

The second point that we question involves, as confirmed in several studies, the strong all-or-none tendency exhibited by data in incremental cuing experiments. That is, the probability of recalling the complete sentence is large compared with the probabilities of recalling its parts. The relative frequency of events in which the complete sentence is recalled by far exceeds the product of the relative frequencies for recalling the separate constituents. In the Anderson and Bower (1973) model, this tendency is captured mainly by the dependency between the response processes. In fact, it was this all-or-none tendency that led Anderson and Bower to introduce the response dependency. But we feel that this should not be the place to bury the issue. In our view, the all-or-none tendency should be looked upon as a property of the memory structure and not of the interfacing. Furthermore, the way in which the Anderson and Bower model accounts for the all-or-none effect has as its consequence that the model also predicts a substantial all-or-none tendency following the second cue in those cases in which the first cue has been ineffective. This is not found in the data. In the model to be presented below, an all-or-none effect to the first cue does not necessarily imply a similar effect for the second cue.

The third reason for a revision of the model comes from data reported by Foss and Harwood (1975) and also by Wender (1979). These authors compared recall of an object when the subject or verb of a sentence was given as a single cue with recall to both subject and verb given as a double cue. The interesting result was that recall to the double cue was higher than what the Anderson and Bower (1973) model in principle can predict. From the Anderson and Bower model, it follows that the probability of recall following the double cue can never exceed the sum of the probabilities of recall to the single cues that was, in fact, found in the data. Foss and Harwood propose one possibility to account for this result. The following model contains a similar feature.

The stochastic assumptions we have tried are as follows. We keep the distinction among stimulus effectiveness, memory structure, and response availability. With respect to the memory structure, we start with a somewhat more general case; but by putting constraints on the parameters, the model structure proposed by Anderson and Bower (1973), for example, can be obtained. The memory structure we assume is shown in Figure 2. This is basically the same tree as in the Anderson and Bower model except that it has been homeomorphically reduced. The links in the tree are assumed to be of an all-or-none character. The

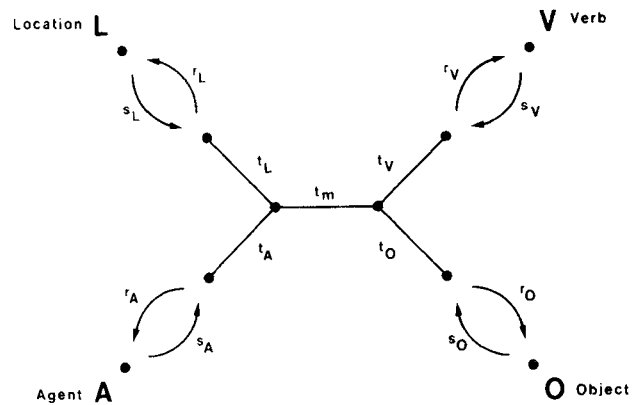


Figure 2. Homeomorphically reduced within-proposition representation of the memory structure in Figure 1.

parameters t_L , t_A , and so on, represent the probability that the respective link exists in memory at the time of recall. The s_i and r_i correspond to the stimulus effectiveness and response availability. All these parameters have to be estimated from the data. The links in the tree are assumed to be independent and, therefore, the probability for one particular path through the net is obtained as the product of the probabilities for those links contained in that path. This is also true for the response availability and stimulus effectiveness parameters.

Because of the tendency toward total recall in the data, the model outlined so far would certainly not be able to account for the data. Hence, we introduce the following additional assumptions. We assume that under certain conditions a proposition is represented in memory not as a tree, but rather, as one unstructured whole. That is, if one part of the sentence is remembered, all other parts will come to mind immediately. To put it another way, there is some probability g_{LAVO} that a sentence will be coded so perfectly in memory that the probabilities for all links in the propositional tree become unity. This will be called the unit code.

We generalize the idea of the unit code in the following way. In the most general model, we assume that any subconfiguration connecting two or more concept nodes may be stored in the unit code with some probability. If a concept node is part of a subconfiguration, this node has to be connected to at least one other concept node. The resulting possible subconfigurations are shown in Figure 3. The last of the configurations, labeled g , corresponds to the case in which no subconfiguration in the unit code is present.

The subconfigurations are assumed to be mutually exclusive events and their probabilities sum to unity. Therefore, the probability for any of the experimental events can be obtained by adding the probabilities for every subconfiguration multiplied by corresponding

Subconfiguration	Parameter
	g_{LAVO}
	g_{LAV}
	g_{LAO}
	g_{LVO}
	g_{AVO}
	g_{LA-VO}
	g_{LA}
	g_{LV}
	g_{VO}
	g_{LO}
	g_{AV}
	g_{AO}
	g

Figure 3. Subconfigurations of the multicomponent model. The solid lines represent the part in the unit code.

link, stimulus effectiveness, and response availability probabilities. For example, the probability that a subject recalls the whole sentence when cued with the location is given by

$$\begin{aligned}
 P(L \rightarrow AVO) = & s_{L^t A^t V^t O} (g_{LAVO} + g_{LAV}^t O \\
 & + g_{LAO}^t V + g_{LVO}^t A + g_{AVO}^t L \\
 & + g_{LA-VO}^t m + g_{LA}^t m^t V^t O \\
 & + g_{LV}^t A^t O + g_{LO}^t A^t V + g_{AV}^t L^t O \\
 & + g_{AO}^t L^t V + g_{VO}^t m^t L^t A \\
 & + g_{L^t A^t m^t V^t O}). \tag{5}
 \end{aligned}$$

On intuitive grounds, not all configurations in Figure 3 are equally likely to be stored in the unit code. When considering pairs of concept nodes, for example, it seems more plausible that agent and verb form one unit, than, say, verb and location. However, such more intuitive models are easily obtained from the general one by setting one or more of the subconfiguration parameters equal to zero. In fact, as it turned out, a very simple version of the model fits the data quite well. Moreover, this approach has the advantage that several versions of the multicomponent model are nested in more general ones, and, therefore, χ^2 difference statistics can be applied in these cases.

Before presenting applications of the multicomponent model, we will sketch one further model, which is proposed by Jones (1978). His fragmentation hypothesis states, in essence, that every fragment of a sentence, that is, any subset of content words, has a certain probability of being recalled from memory. Access to a fragment depends on correspondence of one or more of its components to the word or words given as retrieval cues. The fragment types are the parameters of the model from which the probabilities for different recall patterns are directly obtained. Jones' hypothesis appears to be a special case of our model obtained by setting the parameters for all links in the memory structure equal to zero and the parameters for stimulus effectiveness and response availability equal to one. The various fragment types, then, resemble the subconfigurations of our multicomponent model.

APPLICATION TO DATA

Anderson and Bower (1973) report four experiments with the incremental cuing paradigm. In Experiment 1 sentences consisting of location (L), agent (A), verb (V), and object (O) were used. An example is, "In the park the hippie touched the debutante." The data from this experiment have been reported in detail in their Table 10.5 (Anderson & Bower, 1973, p. 303f). Hence, we could use these data to test our model. The

parameters were estimated by an iterative search routine (Chandler, 1969) that minimized a χ^2 criterion comparing observed and predicted frequencies for the 82 experimental events. We applied several versions of our multicomponent model. The most general version includes one parameter for each subconfiguration. Furthermore, the probabilities for all single links, including stimulus effectiveness and response availability, are all different. In a sense, this general model is not satisfactory, because it contains too many free parameters. The general model was applied to the data in order to provide a baseline against which more special models can be compared. The special models are obtained from the general one by setting some of the parameters equal to zero or one, or by forcing several parameters to be equal.

As the most parsimonious model for the data of Anderson and Bower's (1973) Experiment 1, we propose a model with the following specifications. From the parameters for the various subconfigurations (compare Figure 3), only g_{LAVO} and g are estimated from the data, and all other parameters for subconfigurations are set equal to zero. The single-link parameters (i.e., the t s) are allowed to vary free. Finally, all input and output parameters are required to be equal. Hence, this model has only seven free parameters. The parameter estimates are given in Table 1.

In Table 2 we compare the observed frequencies with the predictions made by our model, and also with the predictions made by Anderson and Bower's (1973) model. The 82 recall patterns may be divided into five sections. In Lines 1-28 only the first cue led to some recall; the later cues were unsuccessful. Lines 29 and 30 contain those events in which the first cue evoked one word, and the second cue an additional word. In Lines 31-66 the first cue was not successful, but the second cue led to some recall. Lines 67-78 give those events in which the first two cues led to no recall but the third cue did. In each line the two different orders of first and second cue have been combined. In Lines 79-82 all three cues were unsuccessful. Each line contains the six possible sequences of cues.

It is interesting to note that in our preferred model only the unit code for the complete sentence has a probability other than zero. Furthermore, all stimulus effectiveness and response availability parameters have identical numerical values, which makes it easy to compute the expected probabilities. For example, the probability that L as a cue leads to the recall of A, V, and O is given by

$$P(L \rightarrow AVO) = f^4 (g_{LAVO} + (1 - g_{LAVO}) t_L t_A t_m t_V t_O), \tag{6}$$

where f stands for the effectiveness as well as the availability parameters.

Table 1
Parameter Estimates of the Multicomponent Model for Anderson and Bower's Experiment 1

g_{LAVO}	=	.23	t_V	=	.31
t_L	=	.44	t_O	=	.47
t_A	=	.39	f	=	.83
t_m	=	.78			
χ^2	=	67.1	df	=	74

Note—The parameter f stands for all input and output processes.

We propose this model because of the following reasons. The estimation procedure resulted in a χ^2 of 67.1 with 74 degrees of freedom, which is quite acceptable. If we take a less restrictive model and estimate all input and output parameters separately, a χ^2 (67) of 56.6 results. As judged by the χ^2 difference, this gain in fit is not statistically significant. Even the fit of the most general version, which has 25 free parameters, is not significantly better. The general version reached a χ^2 (56) of 52.0.

Our preferred model (reported in Table 1) was better than a model in which the input and output parameters were varied and the parameters for single links, the t s, were kept equal. Such a model obtained a χ^2 (71) of 111.9, which is not acceptable. Also, more restrictive models did not fit the data.

One particular version of the more restrictive models, in which t_m is set equal to one, is of interest. Researchers have argued (e.g., Doshier, 1976; Thorndyke & Bower, 1974) that the resulting within-proposition structure corresponds to the theory of Norman, Rumelhart, and the LNR Research Group (1975). However, this additional constraint increased the χ^2 considerably to 88.2 ($df = 75$). In light of this analysis, the structure underlying our model seems preferable.

Anderson and Bower (1973) obtained for their model a χ^2 (72) of 79.4. Although this is acceptable, our model looks somewhat superior. The main argument in favor of our model is not so much the goodness of fit (which is a necessary condition) as the property that our model places the all-or-none tendency into the memory structure and not into the response process.

Jones (1978) applied his model to the same data of Anderson and Bower's (1973) Experiment 1 and obtained a fit equivalent to a χ^2 (68) of 58.9, which is acceptable. The predictions of the fragmentation hypothesis are shown in the rightmost column of Table 2. A direct test between Jones' and our models by comparing χ^2 statistics is not valid, since neither model is a special case of the other one. Nevertheless, we argue in favor of our multicomponent model, since it contains a memory structure comparable to current theories of semantic memory, and it keeps the distinction between memory structure, input, and output processes. Moreover, our model seems to be more parsimonious, requiring only 7 compared to 13 free parameters.

Table 2
Observed and Expected Frequencies of Anderson and Bower's Experiment 1

Event	Observed	Expected		
		Multicomponent	HAM	Fragmentation
1. L→AVO	93	87.8	89.0	88.3
2. L→AV \bar{O}	17	24.4	21.2	20.2
3. L→A $\bar{V}O$	33	31.7	39.2	35.3
4. L→A $\bar{V}\bar{O}$	41	40.0	35.1	36.8
5. L→ $\bar{A}VO$	33	27.5	32.6	28.3
6. L→ $\bar{A}\bar{V}O$	23	20.3	23.4	21.3
7. L→AVO	37	35.7	39.4	34.0
8. A→LVO	87	87.8	88.8	88.3
9. A→LV \bar{O}	20	24.4	23.0	20.2
10. A→LV \bar{O}	36	31.7	42.4	35.3
11. A→LV \bar{O}	43	40.0	39.1	36.8
12. A→LV \bar{O}	36	25.5	26.3	25.0
13. A→LV \bar{O}	19	17.1	22.0	17.3
14. A→LV \bar{O}	36	29.7	35.5	29.2
15. V→LAO	78	87.8	74.2	88.3
16. V→LA \bar{O}	18	24.4	19.8	20.2
17. V→L $\bar{A}O$	30	27.5	30.3	28.3
18. V→L $\bar{A}\bar{O}$	24	20.3	24.0	21.3
19. V→L $\bar{A}O$	21	25.5	22.4	25.0
20. V→L $\bar{A}\bar{O}$	16	17.1	19.2	17.3
21. V→L $\bar{A}O$	30	33.8	34.9	30.8
22. O→LAV	95	87.8	87.8	88.3
23. O→LAV \bar{V}	38	31.7	39.6	35.3
24. O→L $\bar{A}V$	30	27.5	32.9	28.3
25. O→L $\bar{A}\bar{V}$	41	35.7	40.7	34.0
26. O→L $\bar{A}V$	22	25.5	24.7	25.0
27. O→L $\bar{A}\bar{V}$	29	29.7	31.7	29.2
28. O→L $\bar{A}\bar{V}$	31	33.8	37.7	30.8
29. L→A $\bar{V}\bar{O}$ or A→LV \bar{O}	5	3.0	3.7	5.0
30. V→L $\bar{A}\bar{O}$ or O→L $\bar{A}V$	5	3.0	3.7	5.0
31. L→:A→VO	6	8.5	9.7	8.3
32. L→:A→V \bar{O}	6	5.7	4.9	5.8
33. L→:A→V \bar{O}	11	9.9	7.7	9.7
34. L→:V→AO	4	8.5	8.1	8.3
35. L→:V→A \bar{O}	5	5.7	4.3	5.8
36. L→:V→A \bar{O}	11	11.3	7.7	10.3
37. L→:O→AV	11	8.5	9.5	8.3
38. L→:O→AV \bar{V}	8	9.9	7.3	9.7
39. L→:O→AV	11	11.3	8.4	10.3
40. A→:L→VO	6	9.2	11.9	9.4
41. A→:L→V \bar{O}	4	6.8	4.8	7.1
42. A→:L→V \bar{O}	4	11.9	8.2	11.3
43. A→:V→LO	5	9.2	10.2	9.4
44. A→:V→L \bar{O}	7	6.8	4.7	7.1
45. A→:V→L \bar{O}	13	11.3	7.7	10.3
46. A→:O→LV	9	9.2	11.8	9.4
47. A→:O→LV \bar{V}	10	11.9	8.4	11.3
48. A→:O→LV	13	11.3	8.4	10.3
49. V→:L→AO	12	10.6	18.6	11.8
50. V→:L→A \bar{O}	14	13.3	8.8	12.3
51. V→:L→AO	11	11.9	10.9	11.3
52. V→:A→LO	8	10.6	19.1	11.8
53. V→:A→L \bar{O}	13	13.3	9.9	12.3
54. V→:A→L \bar{O}	7	9.9	10.0	9.7
55. V→:O→LA	14	10.6	18.0	11.8
56. V→:O→L \bar{A}	14	11.9	10.9	11.3
57. V→:O→L \bar{A}	11	9.9	9.1	9.7
58. O→:L→AV	13	8.1	8.7	6.7
59. O→:L→AV \bar{V}	8	13.3	6.9	12.3
60. O→:L→AV	4	6.8	4.9	7.1
61. O→:A→LV	5	8.1	8.8	6.7
62. O→:A→LV \bar{V}	10	13.3	7.6	12.3
63. O→:A→LV	3	5.7	5.0	5.8

Table 2 Continued

Event	Observed	Expected		
		Multicomponent	HAM	Fragmentation
64. O→:V→LA	8	8.1	7.4	6.7
65. O→:V→LĀ	9	6.8	4.7	7.1
66. O→:V→LĀ	4	5.7	4.3	5.8
67. LA→:V→O	7	11.3	7.0	10.3
68. LA→:O→V	7	11.3	7.9	10.3
69. LV→:A→O	7	9.9	8.7	9.7
70. LV→:O→A	8	9.9	8.3	9.7
71. LO→:A→V	11	5.7	4.6	5.8
72. LO→:V→A	5	5.7	3.9	5.8
73. AV→:L→O	9	11.9	9.2	11.3
74. AV→:O→L	10	11.9	9.2	11.3
75. AO→:L→V	7	6.8	4.4	7.1
76. AO→:V→L	7	6.8	3.9	7.1
77. VO→:L→A	7	13.3	7.9	12.3
78. VO→:A→L	11	13.3	8.4	12.3
79. AVO→:	386	363.0	366.2	369.1
80. LVO→:	368	363.0	370.2	369.1
81. LAO→:	358	363.0	363.8	369.1
82. LAV→:	365	363.0	367.2	369.1
Total	2952	2952.0	2952.0	2952.0

To further test our model, we took data from Anderson and Bower's (1973) Experiment 4.¹ A comparison with these data is of special interest, since both Anderson and Bower and Jones (1978) report that they had difficulties in fitting the data of this experiment. The sentences used in Experiment 4 were of the form, "In the location (L) the object (O) during the time (T) was related (R) by the subject (S)." For example, "In the park the professor during the night was rescued by the debutante." The problems Anderson and Bower and Jones had with Experiment 4 arose with certain response patterns following the first cue. The fragmentation hypothesis, for example, predicts that L prompts T equally as often as T prompts L. This is not found in the data. We took the 80 experimental events from Experiment 4 that contain the responses to first cues to see whether our model can handle these difficulties.

To apply our model, we first tried the structure proposed by Anderson and Bower (1973, p. 322) in their Figure 10.10a, again, homeomorphically reduced (see Figure 4a). The most general version for five-element sentences of the multicomponent model has 44 parameters, one for each possible subconfiguration (27), memory structure link (7), stimulus effectiveness (5), and response availability (5). The application resulted in a $\chi^2(35)$ of 108.6, which is not acceptable. So, we followed Anderson and Bower's (1973, Figure 10.11) example by changing the supposed underlying memory structure. The homeomorphically reduced structure can be seen in Figure 4b. The only difference between Figures 4a and 4b is an interchanging of the terminal nodes S and O. The structure of Figure 4b is obtained by introducing passive relations as terminal nodes.

Anderson and Bower suggest accepting this structure as a HAM representation of these sentence types as well.

After incorporating the structure of Figure 4b, our model worked much better. Again, we applied several versions of the model; the most general one resulted in a $\chi^2(35)$ of 40.7, which is acceptable. From the other versions of the multicomponent model, we will discuss only two, including the one we finally accepted. This final model yielded a $\chi^2(66)$ of 79.0. The parameter

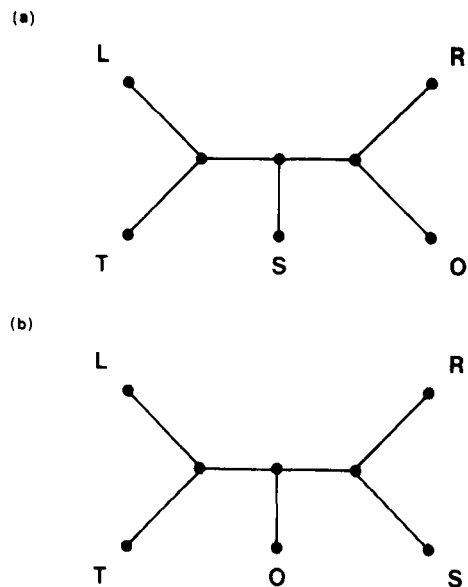


Figure 4. Homeomorphically reduced within-proposition representation of the sentences in Experiment 4 of Anderson and Bower (1973).

estimates can be seen in Table 3. The gain in fit of the most general version is statistically not significant. The model in Table 3 again is justifiably economical, requiring only 13 parameters.

Application of Anderson and Bower's (1973) HAM model to the 135 recall patterns observed in Experiment 4 resulted in a $\chi^2(123)$ of 204.1, which is significant. Jones (1978) applied his model only to overall single-cue probabilities. Single-cue probability refers to the average probability of correctly recalling words to first cues regardless of the precise pattern of recall using the data of Anderson and Bower's (1973, p. 326) Table 10.12a. This application to an even smaller portion of the data resulted in a $2\hat{I}$ statistic (Kullback, 1959) of 63.1 with 10 degrees of freedom, which is highly significant. To facilitate a comparison between the fragmentation hypothesis and the multicomponent model, we fit Jones' model to the first-cue data of Experiment 4 using our program. This resulted in a $\chi^2(52)$ of 193.8, which demonstrates once more the difficulties of the fragmentation hypothesis with respect to the data of Experiment 4.

Comparing the results of applying the existing models to the data of Experiment 4, it appears that the multicomponent model is the most successful. To interpret this result, one should recall arguments of Anderson and Bower (1973), as well as of Jones (1978), in which they allege having explained the difficulties their models had in fitting the data. Anderson and Bower, in particular, point to four events that account for much of the χ^2 , namely, the events in which L evokes recall of just O, in which O evokes just L, in which R evokes just S, and the event in which S evokes just R. The sum of the χ^2 components for these events is 50.8, which is a result of the fact that the HAM model consistently underpredicts the frequencies. Whereas 134 such events have been observed, the HAM model predicts just 76.9 cases. In comparison, the multicomponent model predicts overall 124.4 such cases, and accordingly, 7.1 results as sum of the χ^2 components. The better fit of the multicomponent model with respect to these experimental events is essentially covered by the separate parameter for the pair subconfiguration gLO (see Table 3).

The weakness of Jones' (1978) model in fitting the

Table 3
Parameter Estimates of the Multicomponent Model for
Anderson and Bower's Experiment 4

gLTORS	=	.29	g*	=	.07
gLO	=	.10	t	=	.29
s _L	=	.71	r _L	=	.95
s _T	=	.43	r _T	=	.82
s _O	=	.79	r _O	=	.90
s _R	=	.54	r _R	=	.83
s _S	=	.68	r _S	=	.89
χ^2	=	79.0	df	=	66

Note—g* stands for all pair subconfigurations, except LO; t stands for all single-link parameters.

data of Experiment 4 consists in not being able to predict that certain combinations work better as "A prompts B" than as "B prompts A." For instance, L prompts recall of T 171 times, while T prompts recall of L only 122 times. The predictions of our model are 172.9 and 120.7, respectively; the fragmentation hypothesis predicts 174.5 in both cases. These and similar data properties are captured by the different parameters for stimulus effectiveness and response availability in our multicomponent model. For instance, the estimates of s_L and s_T are .71 and .43, respectively.

The reason for the weakness in Jones' (1978) model in this respect stems from a main feature of his model, namely, the prediction of specified sets of symmetry. This is a consequence of the aspect of the fragmentation hypothesis that certain recall patterns should occur with the probability of the same fragment type. For example, P(L→TORS), P(T→LORS), P(O→LTRS), and P(S→LTO \bar{R}) should display equal values. But this was not found in the data—at least with respect to T prompts LOS. The observed frequencies are 22, 7, 20, and 19, respectively; the mean estimate of Jones' model for fragment type LTOS is, therefore, .0227, leading to a predicted frequency of 15.7 for all four experimental events. This deviation from the symmetry prediction might be seen in other cases as well. As a matter of fact, Jones (1978, p. 362f) points out that the deviation from his fragmentation hypothesis is indeed localized to the "time" word. So, Jones himself suggests the constraints of his model, namely, that asymmetric observed frequencies with respect to recall patterns involve the same fragment type.

In conclusion, we restate the main arguments in favor of the multicomponent model. The first argument is more theoretical in nature, and the two following ones refer to properties of the data gathered in cued recall experiments. Our model preserves the distinction between memory structure and input and output processes, which keeps it in line with current theories of semantic memory.

As outlined before, data in incremental cuing experiments show a strong all-or-none tendency. In our model these dependencies in subjects' responses are represented by the memory structure and not by dependencies between output processes. Finally, our model is able to predict asymmetries with respect to corresponding response patterns to different cues, observed under certain experimental conditions (viz., Experiment 4). Therefore, we propose the multicomponent model as a plausible alternative to the stochastic model of Anderson and Bower (1973) and the fragmentation hypothesis of Jones (1978).

REFERENCE NOTE

1. Colonijs, H., Glowalla, U., Schulze, H.-H., & Wender, K. F. *Stochastic models for sentence learning*. Braunschweiger Berichte aus dem Institut für Psychologie, Universität Braunschweig, 1, 1976.

REFERENCES

- ANDERSON, J. R. *Language, memory, and thought*. Hillsdale, N.J.: Erlbaum, 1976.
- ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.
- CHANDLER, J. P. STEPIT—Finds local minima of a smooth function of several parameters. *Behavioral Science*, 1969, **14**, 81-82.
- DOSHER, B. A. The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology*, 1976, **8**, 291-310.
- FOSS, D. J., & HARWOOD, D. A. Memory for sentences: Implications for human associative memory. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 1-16.
- JONES, G. V. Tests of a structural theory of the memory trace. *British Journal of Psychology*, 1978, **69**, 351-367.
- KULLBACK, S. *Information theory and statistics*. New York: Wiley, 1959.
- NORMAN, D. A., RUMELHART, D. E., & THE LNR RESEARCH GROUP. *Explorations in cognition*. San Francisco: Freeman, 1975.
- THORNDYKE, P. W., & BOWER, G. H. Storage and retrieval processes in sentence memory. *Cognitive Psychology*, 1974, **6**, 515-543.
- WENDER, K. F. Über die interne Struktur der Gedächtnisrepräsentation von Sätzen. In H. Ueckert & D. Rhenius (Eds.), *Komplexe menschliche Informationsverarbeitung*. Bern: Huber, 1979.

NOTE

1. We thank J. R. Anderson for providing us with the data for this analysis.

(Accepted for publication June 28, 1979.)