Young, R. K. A comparison of two methods of learning serial associations. *American Journal of Psychology*, 1959, 72, 554–559.
Young, R. K. The stimulus in serial verbal learning. *American Journal of Psychology*. 1961, 74, 517–528.
Young, R. K. Tests of three hypotheses about the effective stimulus in serial learning. *Journal of Experimental Psychology*, 1962, 63, 307–313.
Young, R. K. Serial learning. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968. Pp. 122–148.
Young, R. K., & Casey, M. Transfer from serial to paired associate learning. *Journal of Experimental Psychology*, 1964, 67, 594–595.
Young, R. K., Hakes, D. T., & Hicks, R. Y. Effects of list length in the Ebbinghaus derived-list paradigm. *Journal of Experimental Psychology*, 1965, 70, 338–341.
Young, R. K., Hakes, D. T., & Hicks, R. Y. Ordinal position number as a cue in serial learning. *Journal of Experimental Psychology*, 1967, 73, 427–438.
Young, R. K., Patterson, J., & Benson, W. M. Backward serial learning. *Journal of Verbal Learning and Verbal Behavior*, 1963, 1, 335–338.

FRAN: A SIMULATION MODEL OF FREE RECALL

John Robert Anderson[1]

STANFORD UNIVERSITY, STANFORD, CALIFORNIA

I. Introduction

The basic free recall experiment with which we will be concerned involves the presentation to a subject (*S*) of a list of words, one word at a time. After seeing all the words in the list, the *S* is asked to recall them in any order he chooses. The experimental paradigm derives its name from the fact that the *S* is not constrained to recall items in a particular order. The free recall paradigm has recently attracted much research interest because of evidence indicating the strong influence of various types of conceptual organization upon the *S*'s recall. This evidence is of two sorts:

1. The experimenter (*E*) may choose sets of words with certain

[1] This research was supported by a grant, MH-13950-04, from the National Institutes of Mental Health to my research advisor, Gordon Bower. I am indebted to Professor Bower for his advice and assistance in the preparation of this paper.

organizational characteristics and note how that organization is reflected in recall. Bousfield (1953) reported that words which came from the same conceptual category (i.e., were instances of the same superordinate concept) tended to be recalled together. Bousfield and Cohen (1953) found that this clustering increased over trials, and Underwood (1964) found better recall for lists organized into a small number of conceptual categories. Jenkins and Russell (1952) examined free recall in a list that contained 24 pairs of words, each pair consisting of a stimulus word and its highest frequency associate. Although such associated pairs were not presented adjacently in study, there was a high probability for the items to be recalled together. Deese (1959a) found that sets of words which tend to elicit one another as free associates are more easily recalled than sets of words which are not interassociated.

2. Even when the *E* deliberately minimizes relationships among the list words, *S*s apparently still impose their own idiosyncratic organization upon the word list. In their introspective reports subjects almost universally describe some sort of organizational strategy. Typical strategies include developing visual images in which the referents of the words interact, creating stories that involve the words as main actors, and finding some obscure property that applies to a number of words (e.g., "mother," "butter," and "cup" are all found in the kitchen). Subjects also claim to recall together the words they have grouped together during study rather than recalling words in a random order. It is as if the *S*s created a small number of multiword units out of a long list of words and recalled these units rather than the words individually. If so, it is clear why categorized and highly interassociated lists are better recalled than those which are not. Such lists are more conducive to such a "unitization" strategy.

Alongside these introspective reports, there is evidence in the recall protocols themselves to indicate that *S*s organize their recall even with "unrelated" words. By tabulating the frequency with which the same pairs of words are recalled adjacently across trials, it becomes clear that the order of recall shows much greater across-trial consistency or stereotypy than could be produced by randomly ordered recall (Bousfield, Puff, & Cohen, 1964; Tulving, 1962). Such stereotypy, occurring although the words are rerandomized on each presentation, is just what would be predicted if the *S*s were recalling the words in groups or "subjective units."

This stereotypy of recall increases across trials on a list as does the number of words recalled. This correlation between stereotypy and amount recalled has led some theorists (e.g., Mandler, 1967; Tulving, 1968) to postulate that free recall learning is a consequence of progressive organization of the list into a small number of subjective units. This is called the "unitization hypothesis." Mandler and Pearlstone (1966) reported an experiment which offers impressive support for this hypothesis. Their *S*s sorted 52 cards, each containing a word, into two to seven groups based on similarity of meaning. After sorting the cards once, the *S* was given the cards a second time in randomized order and instructed to resort them in the same way. This procedure was repeated until the *S* was able to sort the cards identically twice in a row. One group of *S*s was told to study the words as they sorted them in order to be able to recall them later. The other *S*s were not told that there would be a free recall test. In a subsequent free recall test both groups recalled equally well. From this result, Mandler and Pearlstone concluded that organization of a list of words into stable categories is a *sufficient* condition for a high level of recall. This is to be predicted from the unitization hypothesis which relates level of recall to the degree to which the list has been "chunked" into a small number of stable groups.

Despite this kind of evidence, we feel that the unitization hypothesis as it stands is inadequate as a theory of free recall. A first objection is that several experiments are now available indicating that measures of organization and of recall can vary independently (Cofer, 1967; Cofer, Bruce, & Reicher, 1966; Dallett, 1964; Puff, 1970; Rosner, 1970). Such failures of the general correlation indicate the need to make more explicit the structures and processes underlying the formation and retrieval of the "units." Once this is done, it might be possible to determine boundary conditions on the general prediction of "more organization, more recall."

A second objection to the unitization hypothesis arises from informal observations of ours when we have asked *S*s to introspect on what they were doing as they studied and recalled a list of words. Consider the representative introspections, given in the Appendix, of one *S* as she studied a list of 40 words during a free recall trial. She had already studied and recalled this list once. At first glance she and most of her cohorts we have analyzed appear to do roughly what the unitization hypothesis claims, viz.; by using imagery, categorization, storytelling, etc., they try to organize words into groups during the input trial. Also *S*s tend to recall the words in the groups they have created. However, there is a consistent sort of *instability* to the groupings which is totally alien to the spirit of the unitization hypothesis. Groups of words break into parts, which merge into new groups, which break apart again and remerge into old configurations. It would appear that *Ss* develop plans for retrieval that do not rely upon rigid subjective units. During the input trial, associations are established between individual words, but there is little effort to define stable units. A network of such interword associations constitutes a powerful retrieval system; if the *S* can recall just one word, he can recall all the words he has associated with that word, then in a second loop, all the words he has associated with those, and so on, recursively, until he has recalled every word he can reach by chaining through the associations he has developed during study. From this point of view it is not necessary that the *S* always aggregate the same words together during study, nor is it necessary that he trace out exactly the same search path through his associative network during recall. Indeed, anything that appears to be a unit is just a group of words which are highly interassociated and hence tend to occur together in the *S*'s introspections during study and together in his output during recall.

In this paper we will be interested in developing a model of the processes underlying standard free recall. By "standard free recall" we mean to reference those experiments where the list to be studied is composed of common nouns chosen in a fairly random fashion. The words are presented one at a time at a constant rate. If it is a multitrial experiment, study and recall phases are alternated. It is our contention that, under these circumstances, most *Ss* spontaneously adopt an associative method for free recall similar to that outlined above. FRAN (Free Recall in an Associative Net) is a computer program that simulates this associative strategy for free recall. In Section II, we describe the hypothetical mental mechanisms which FRAN was constructed to simulate. In Section III, we describe how these mechanisms are used to effect the associative strategy. Section IV of this paper compares FRAN's performance with that of humans in several free recall and recognition experiments. The final section evaluates our theoretical positions and discusses possible theoretical extensions that would create a more powerful theory.

II. The Mental Mechanisms

FRAN is a computer program implemented in LISP/360. We will not explain in detail the information processing occurring in the computer program itself. Rather we shall concentrate on describing the mental processes the program is simulating. In this section and

the next, we will indicate what we think are the important characteristics of the mental structures and processes that underlie free recall.

A. Structure of Long-Term Memory

FRAN's memory is structurally divided into a sensory register, a short-term store (STS), and a long-term store (LTS), somewhat along the lines suggested by Norman (1968). LTS is the memory system that contains all the individual's knowledge about the world. Therefore, in deciding what the structural characteristics of LTS should be, we wanted a structure that would be adequate not only for simulation of free recall, but also adequate to encode any linguistic fact. Beginning with such general goals offers the possibility that phenomena of free recall will be integrated in the future with other phenomena in human verbal memory.

1. Information, Associations, and Labeled Associations

If verbal memory simply consisted of a set of symbolic units (e.g., words), it would contain no information about the world. Information takes the form of connections among the symbolic units. A simple type of connection is the traditional association which is, formally, just an ordered pair of symbolic units (e.g., "dog–cat," "jewel–diamond"). Perhaps because of their simplicity, such associations have historically received the greatest attention in theoretical reconstructions of the mind. However, it has been frequently argued (e.g., Deese, 1968; Kintsch, 1970) that such simple associations are inadequate for representing certain elementary facts about behavior. For instance, this representation in terms of ordered pairs does not explain how it is that we know that the relationship between "dog" and "cat" is one of coordinates, while the relationship between "jewel" and "diamond" is one of superordinate–subordinate. That is, what is lacking in an elementary dictionary of associated pairs is knowledge of the relationship exemplified by the pairs; this lack is a serious deficiency since people obviously can search for a word, B, that bears a specified relation, R, to a stimulus word, A.

The inadequacy noted above can be circumvented by replacing the ordered pair representation by ordered triples or *labeled* associations in which the third member of the triple is a symbol identifying (labeling) the relationship between the first two members of the triple which are single elements. So far as we can determine, there are no simple demonstrations of the inadequacies of labeled associations as there are for simple associations. We should point out, however, that for every ordered triple there exists an equivalent ordered pair of a special kind. That is, the ordered triple < A B R > may be equivalently represented as the ordered pair <<A B> R>. Either notation is to be read as "A has relation R to B." Representation of such information in terms of ordered pairs is only possible, however, when we allow units in an association to include other associations (like the pair < A B > above). This move clearly contradicts conventional stimulus-response analyses because it permits, as units of description, elements which do not have a one-to-one correspondence to stimuli or responses [see Bever, Fodor, and Garrett (1968) for arguments regarding the inadequacy inherent in this restriction on stimulus-response (S-R) theorizing]. Whichever of these representations one adopts, restricted ordered triples or generalized ordered pairs, one goes beyond traditional associative analysis with its simple listing of associated pairs.

We do not want to leave the impression that associations have been abandoned in our model. Quite the contrary, we believe the "data base" of a person's memory is a complex associative network and our model for free recall consists of operations for marking subgraphs of that associative network. Where we differ from the traditional association doctrine is that we consider association to be a generic name for a whole class of different types of relations between concepts, and we believe that the name of the relation has to be stored with each associated pair. With this emendation, then, our model for free recall falls in the tradition of associationistic explanations of learning. Frankly, we have a difficult time imagining any radically different alternative type of theory of free recall.

2. *Semantic Markers?*

Following their critical analysis, Deese (1966, 1968) and Kintsch (1970) rejected traditional associationism, but turned, not to a more powerful associative representation as we have, but rather to a representation in terms of semantic markers. In this choice, they were undoubtedly motivated by the desire to establish common theoretical ground with recent linguistic analysis (e.g., Bierwisch, 1967; Katz & Fodor, 1963). However, upon close inspection, it is clear that there is nothing to motivate their choice over our own. The

claims made for such semantic marker theories are hardly noncontroversial in linguistics (e.g., Bolinger, 1965; Lyons, 1969; and Schank, 1970). Also is seems that many marker theories are structurally similar to associationist theories since, in the abstract, each concept is defined in terms of its pattern of "associations" to a base set of concepts, the alleged universal markers. Indeed, in Kintsch's analyses, it is not clear how he could encode his markers in a LISP memory structure efficiently and not adopt a representation essentially identical to that adopted for FRAN. Thus, it appears that the choice between the two conceptualizations may ultimately rest on heuristic considerations. We find the graph-theoretical representation, to which our associationist analysis leads us, much more congenial to our way of thinking (see also Quillian, 1969).

3. *Free Association Norms?*

The memory structure of FRAN, then, is a network of labeled associations between words. We think of the memory structure as representing what the *S* knows when he enters the laboratory and not something that is built up during the experiment. We have resisted the temptation to assign strengths to the various associations. In FRAN's memory, an association either does or does not exist between a pair of words; it is an all-or-none affair. The immediate question is how are we to account for the fact that some associations appear much stronger than others in free association norms.

There are at least three answers to this challenge. The first is to question the assumption that a person in generating an association selects only from the words which are directly associated to the stimulus. It is possible that the individual may chain through a number of associations in his network before generating the response. If so, the structural characteristics of an all-or-none network may "conspire" to favor some words as responses over others. In fact, Kiss (1967) assumes just this in fitting his associative net models to normative association data. The second tact is to argue that the graduated character of free associations is a consequence of pooling across *S*s in the construction of associational statistics. Each *S* has associations in an all-or-none manner, but only the strong associations are possessed by most *S*s. This analysis has the added utility of explaining in a natural way why weak associations are often bizarre – they represent the memory structure of only a few *S*s. The third response to the challenge is to question the relevance of free association norms to testing a theory concerning the structure of our knowledge of the world. Surely many responses appear in free association norms which have no semantic or conceptual relation to the stimulus word – included would be rhymes and all forms of "clang" associates. The point is that if we wished to find out whether people generally acknowledge a particular relation R between two words A and B, we would hardly set them the ill-defined and ambiguous task of freely associating to A in the hope that they might produce B. In any event, the force of these arguments is that the graduated character of free association norms need constitute no particular embarrassment to FRAN's network of all-or-none associations.

Another feature of FRAN's preexperimental associations is that they are symmetric; that is, if word B is an associate of word A, then A is an associate of B. The motivation for this decision is simply that if A has the relation R to B, then B has the inverse relation (R^{-1}) to A. For instance, if "cat" is a subordinate of "carnivore," then "carnivore" is a superordinate of "cat." In formal notation, if a person knows the relation <<A B>R>, he should also know the relation <<B A>R^{-1}>. To critics with a bent for parsimony, symmetric associations may appear as superfluous redundancies since one association could be inferred from the other. However, consider the problem of how a *S* could ever give the information contained in the association <<A B> R> if he were asked "what do you know about B?" Without the symmetric association <<B A>R^{-1}>, that information would not be accessible from B. For such reasons, it was deemed necessary that FRAN have this symmetric property in her associations.

4. Details of FRAN's Memory

Having discussed the general properties of FRAN's memory, we will now state the details of the actual memory structure that was used in the simulations to be reported. First, it should be confessed that, despite our arguments for labeled associations, FRAN's associations are, in the present computer program, unlabeled. We made provision in the memory structure for labeling, and this provision will be utilized to good advantage in the process of list learning in free recall. We did not prelabel the associations because the information that would be contained in the labeling established prior to the experiment is not now used by FRAN in our free recall experiments. She simply finds and marks any associations she can between list words, being indifferent to the type of relation she finds

bridging list words. We will argue in the next section that it is an empirical fact (although not *a priori* necessary) that the labelings are irrelevant to almost all predictions one would want to make in the standard free recall paradigm.

The memory structure chosen for FRAN in the simulation studies is less ambitious than our general theory in another way. Her vocabulary totals only 262 words, all nouns. The associations between words were chosen mainly by using dictionary definitions. Whenever a word was involved in the definition of another, two associations were created between the words, one in each direction. Additional associations were added whenever it was thought that a dictionary definition was deficient in some respect. We chose the words under the constraint that the associative network be connected; i.e., by chaining through associations it is possible to go from any word to any other word. Hobbes observed that "the mind can lead from anything to almost anything." In FRAN's connected memory, the qualification "almost" would not be necessary. An attempt was made to build in a number of superordinate concepts and their instances in order to facilitate simulation of free recall with categorized lists. In addition to the associations determined by these considerations, each word was given two random associations. This random component was an attempt to have FRAN's memory mimic some of the idiosyncratic information each *S* knows about the world.

Individual words are associated with as few as three or as many as 19 other words. The distribution of number of associations per word is given in Fig. 1. As can be seen, that distribution is quite negatively skewed; while most of the words have three, four, or five associations, the mean number of associations is 5.72. The words with a large number of associations are names of categories having many instances. This particular distribution is simply a consequence of the constraints under which the associations were chosen. It might seem that the particular associations chosen are irrelevant – that we might as well have used abstract symbols instead of words and imposed on these some artificial structure. However, we decided against such a move and we think for good reason. The functional consequences for free recall of various possible associative structures are not at all understood. In order to have a fair test of the model, it is necessary to begin with a graph structure that is as "humanlike" as possible. It is not unreasonable to suppose that the procedures just described did create a humanlike structure. It would be unreasonable to suppose this about an arbitrary associative network, since we do not understand the abstract structural characteristics of the

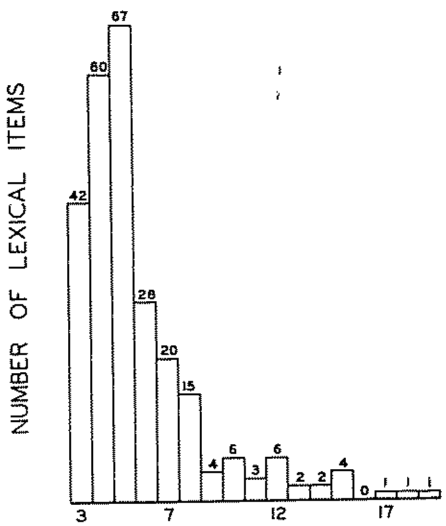

Fig. 1. Frequency of lexical items with each number of associations.

associative network that is human memory. Also use of actual words makes possible a comparison of the recall ordering FRAN gives to her words with the ordering human subjects give. Such comparisons would constitute an interesting test of our assumptions about the structure of human memory and the relation of that structure to free recall.

B. The Short-Term Store

The short-term store (STS) is an active subpart of the long-term store (LTS), the particular subpart varying from moment to moment. STS will denote that part of LTS which is currently in the focus of the subject's attention; i.e., it represents that part of his knowledge of which he is currently "conscious." As in Norman's model, the sensory features of the word presented activate and bring into STS the part of the associative structure corresponding to that word. It is assumed that STS has a small limited capacity of a few items. Therefore, once STS is filled, the act of entering a new item necessarily involves the loss of an old item from STS. The general rule for removing items from STS is that the item removed will be that one which has resided longest in the STS. Humans may be

capable of more flexible control of their STS, and that is a potential area for future development of the model. Items in STS have two characteristics distinguishing them from the remaining items in memory. First, the *S* has immediate access to the items in STS. Other items in memory can only be accessed by chaining associatively from those items to which the *S* has immediate access. Second, it is assumed that the structural elaborations that underlie "learning" are made to any part of the associative network only while it is part of the STS. The nature of those elaborations will be detailed shortly.

FRAN has been given a STS of five words. The decision is somewhat arbitrary, but not unreasonable in light of the literature on short-term memory. What is meant by a STS of five words in FRAN is deliberately left somewhat vague. We occasionally want more than just five words to be active since FRAN attempts to tag some of the associations surrounding the word currently under study. This operation is crucial because FRAN's associative strategy requires that she be able to tag associative paths between the words in STS as well as to tag the words themselves.

We will now describe how FRAN's memory structure is modified during free recall learning. Any structural modification will necessitate the development of new associations. Associations in LTS express facts the person has learned about the world. Therefore, we assume that in free recall learning the structural modifications must be encodings of (reflections of) experiences of the *S* during the experiment. In particular, we shall assume that most of the associations developed during free recall record that individual words or individual associations between words were encountered during the experiment. The recording of such facts involves the development of two different types of associations: one type of association goes from an individual word to a hypothetical entry called the "LIST-*N*" marker, where $N = 1, 2, 3, \ldots$ indexes the ordinal number of the list being learned in the experiment; the second type of association goes from an associative pathway (between two words A and B) to the LIST-*N* marker. This second type is an association that serves to label another association, and is the kind which traditionally S-R associationists have not used in their theorizing. But it is indispensible to FRAN's operation, since tagging associative pathways with LIST-*N* is the only way FRAN has of directing her search of memory at recall.

LIST-*N* is a hypothetical element introduced into LTS to index or stand for particular lists of words used in an experiment. The LIST-1, LIST-2, etc., entries may in fact be number words like "first," "second," or they may be different contextual cues that the person noticed or thought of during presentation of the separate lists. For brevity in the following, we will usually refer to the development of such indexical associations as "tagging" or "marking" the individual word or the associative pathway. It should be noted that tagging of a word establishes a one-way association *from* the word to the LIST-*N* marker and not vice versa. The reverse association, from LIST-*N* to a word, is rare, characterizing only those special words which serve as recall starters. These special words which are associated to the list marker are said to comprise the ENTRYSET, to be explained later.

Tagging with a list marker is assumed to be a stochastic process; specifically, the probability of successful tagging is exponentially related to the duration of the item in STS. The probability of success is described by the formula $1\text{-}a^T$, where T is the time in seconds that the item is in STS and a is a sort of "learning-rate" parameter that is allowed to vary from experiment to experiment. It may be that the processes underlying tagging are basically deterministic, but having no good ideas about them we will summarize the whole process with a probability.

The reader may proceed on the assumption that if a word is tagged, it very probably would be recognized as a list word from among a set of distractors on a later recognition test. Because of the necessity of handling false alarms or considering sophisticated guessing strategy, the actual model for recognition-test performance must be somewhat more complex than simply checking for presence or absence of an all-or-none list tag. This more sophisticated recognition model, developed along lines suggested by Bernbach (1967), will be discussed and used to interpret recognition results in Section IV, F of this paper.

III. The Associative Strategy

A. A Variety of Possible Strategies

There are many conceivable strategies for doing free recall. While many *Ss* report using a strategy similar to that we have called associative, we have found a few *Ss* who reported very different strategies. Consider some of the spontaneous strategies that we have learned from postexperimental interviews with *Ss*:

(1) One *S* decided to study only every other word in an attempt to reduce memory load and focus attention on a subset of the list.

(2) Another *S* formed one image linking together the first three words he studied, another image linking the next three and so on. In recall he attempted to retrieve these images of triplets.

(3) A number of *Ss* have attempted to do free recall (usually just the first trial) as a serial learning experiment, rehearsing serial linkages in STS.

(4) One *S* tried to organize his list alphabetically, associating each word with its first letter, then cueing recall by reciting the alphabet to himself.

(5) One *S* learned the list by paired-associate means, associating each word with a person he knew at Stanford by means of an image. Then, in recall, he "mentally walked along" a predetermined path through the campus and recalled the words as he met the people he had used as stimuli for the paired-associate task.

The last *S* is myself. My method is only a variant of the "method of loci" or method of mnemonic pegwords (see Bower, 1970). It is the optimal free recall strategy, far superior to the others mentioned including the associative strategy adopted by FRAN and by most of our *Ss*. It would not be too difficult to reprogram FRAN so that she used the mnemonic pegword strategy. By so doing we would make her a more efficient free recall learner – roughly speaking, because the pegword strategy avoids the loss of access to associative clusters of words as does happen with the associative strategy. In any event, we have run FRAN on free recall with an associative strategy in the belief that this simulates how most of our *Ss* are learning free recall lists most of the time.

In free recall learning, when presented with a word to study, most of our *Ss* seem to just "let their mind wander" (their expression) or "free associate to the word." Inevitably, they associate a word they are studying to another word from the list. Our *Ss* soon recognize that associating together words from the list is a powerful strategy in that a few words recalled from the list can serve as retrieval cues for recall of all the words associated to the first words, then all the words associated to these associates, and so on recursively. Once the value of this associative strategy is understood *Ss* make more deliberate use of it. FRAN is a model that attempts to simulate the *S* when he has adopted the associative strategy unstintingly. Thus, comparison of her output to that of *Ss* is only fully valid when the *Ss* commit themselves to this free recall strategy from the outset of learning the list. One way to increase the chances that *Ss* will start with such an associative strategy is to inform them prior to learning of the utility of this strategy, as we sometimes do with our *Ss*. For instance, in the experiment reported in Section IV,B we instructed our *S*s to use the associative strategy.

Note that it is a characteristic of the associative strategy that the particular labels of the associations (i.e., the types of associative relations) are unimportant. Therefore, the associations have not been prelabeled in FRAN's current memory structure. Some plausible alternative strategies for free recall would require the information provided by labeled associations. For instance, if an *S* elected to only associate list words with their superordinates, then he clearly needs this relational information to guide his search. Such a strategy is probably quite efficient when the list has been explicitly organized into categories. Evidence presented in Section IV,E indicates that *S*s do deviate from the general associative strategy when the list has been organized into categories.

It is our position that no matter what free recall strategy an *S* adopts, (e.g., associative, pegword, superordinate category) it could be implemented by the mental mechanisms described in Section II. Different strategies correspond to different configurations of the mechanisms. In this section we shall describe the organization of the mental mechanisms underlying the strict associative strategy. At present writing, FRAN is unlike people insofar as she always uses this associative strategy for doing free recall. This is unobjectionable so long as we are comparing FRAN's behavior to that of college student *S*s doing "standard free recall" with unrelated word lists, since it is our conjecture that most of our *S*s eventually stumble upon the associative strategy for learning such lists. Where the strategy variation becomes problematical is when we try to fit FRAN to experiments deviating in certain ways from this standard testing ground (e.g., use of categorized lists or hierarchically organized lists). The unique manipulations of such experiments may lead *S*s to abandon the associative strategy in favor of others, like superordination, which the current FRAN does not pretend to simulate. A quaintly academic question is whether a model which simulates learning under strategy A can be faulted for misfitting results of experimental conditions that induce *S*s to use alternative strategies B, C, or D. For our own purposes, we have found FRAN's behavior in such "nonstandard" experiments to provide a firm anchor, standard, or reference point useful for giving contrastive descriptions of the behavior of humans. By assessing FRAN's deviations from the data, we may gain insights into limitations of the associative strategy as well as noting what strategies people do adopt in these situations.

B. The Study Processes

During study, FRAN is occupied with three separate jobs in pursuing her associative strategy: these are the tagging of words, searching out and tagging of associative pathways between list words, and updating of good recall "starters" in ENTRYSET. We will elaborate on each of these processes.

1. Word Tagging

Whenever FRAN enters a word into STS an attempt will be made to tag that word as a word to be later recalled. FRAN will enter a word into STS if it has been presented for study or if it is found to be an associate of a word that has been presented for study and it is recognized as a member of the list under study. The probability of tagging a word that resides in STS for T seconds is $1\text{-}a^T$. Since FRAN will attempt to tag any word entered into STS, she is restricted to only entering those words which are from the list. If she were to enter nonlist words into STS she might wrongly tag them as list members. Apparently human *S*s can enter nonlist words into their STS without tagging them, since this is what characterizes the superordination strategy which uses category labels as entry words for cueing recall even though these labels are not overtly recalled.

2. Association Tagging

The second process involves the examination of the associative surround of the word under study to see if there are any other list words to which it may be associated. It is difficult to determine how rapidly FRAN should examine associates of the word under study. One obvious assumption is that FRAN should study a constant n associates for each second of study time. However, arguments may be given for the alternative assumption that FRAN should examine associates at a decreasing rate over time. For one thing, as FRAN exhausts the first-order associates to the word being studied and resorts to second-order (i.e., associates of associates of the target word) and higher-order associates, it should take longer to find them. Also, there should ideally be some "processing capacity" trade-off between the number of associates FRAN finds and tries to tag *versus* the rate at which she continues her search for new associates. The formula that has proven useful in our simulations is $N = 5 + T$, where N is the number of associates examined in time T. Our interpretation of this formula is that within moments after exposure to the word, a rush of five associates comes to mind, but thereafter new associates are examined at the slower rate of one per second. This particular formula is highly dependent upon the fact that FRAN has only 262 words in her memory. We suspect that with a larger network, the search rate would have to be increased if FRAN is to maintain an accurate simulation of human recall results.

FRAN selects associates to examine in the following manner: she randomly selects one of the associates of the word she is currently studying. If she has already examined it (this is indicated by a temporary "check mark"), she randomly selects one of its associates. If that second-order associate has also been examined, she randomly selects one of its associates. This recursive process continues until a word is found which has not yet been examined. This random search may be viewed metaphorically as a blind excursion into the associative vicinity of the word being studied. That excursion is brought to a halt when something new (unchecked) is found. She will usually halt after one or two "steps" from the target word.

After finding a word that she has not previously examined during the current search, FRAN then checks to see if it is from the list she is studying. The word is identified as from the list if it satisfies at least one of the following three criteria: (a) it is currently in STS; or (b) it has been successfully tagged with an association to LIST-N; or (c) it is currently a member of "ENTRYSET." As will be detailed later, ENTRYSET is the subset of those list words which have become directly associated to the list marker.

If the word associate being examined fails all three of these criteria and therefore is not recognized, this examination ends with a failure, and FRAN returns to the target word being studied and commences another search out from this target. If the associate being examined is recognized, it is entered into STS. Consequently, another attempt is made to tag this associate. It may not have been previously tagged; FRAN could have recognized the word on the basis of its membership in ENTRYSET or STS. FRAN also attempts to tag the pathway she has found from the target word studied to the associate. Since FRAN's associative structure is symmetric, a path also exists from the associate to the studied word. An attempt is made to tag this backpath also. In the recall phase of the experiment, FRAN will use these marked associative paths to profitably guide her search for words in the list.

If the associative path FRAN has found between two list words involves several links, then she will try to tag each of these separate

links. The tagging of each association is a probabilistic process independent of the tagging of every other association in the path. In the formula for tagging probability, $1\text{-}a^{T}$, T will equal the time to study the presented word minus the time taken to find the associate. Even after finding and tagging a list associate, FRAN will continue looking for new associates for the word under study until she runs out of time or until the five slots in STS have been filled by the word under study and its associates.

We can make more understandable this process of searching the associative surround of a word for other list words by reference to Fig. 2. The word A in that figure is under study for one second. According to the formula relating number of paths examined to study time, FRAN will examine $5 + T$ or six associates. The six panels of Fig. 2 each describe one of the six paths examined by

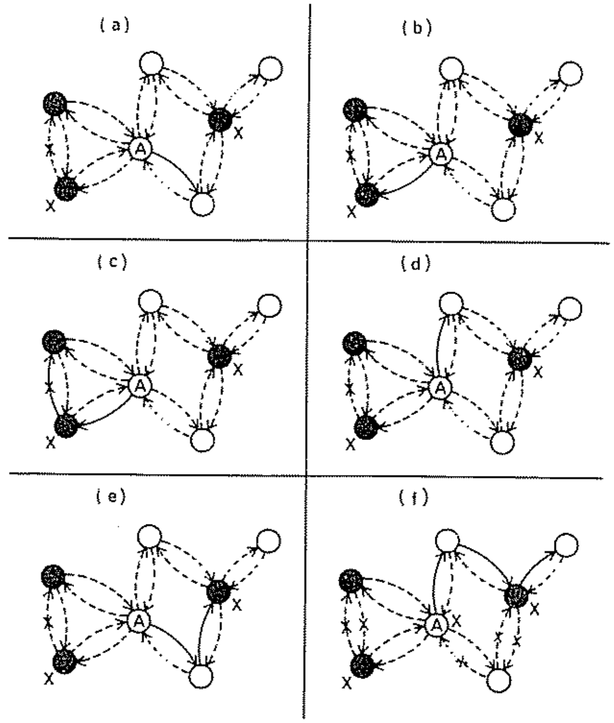

Fig. 2. A possible search of the memory structure surrounding Word A during study. See text for explanation.

FRAN, the path under examination being indicated by a solid line in the appropriate panel. FRAN randomly selects an associate of *A* and examines it. If it has been examined [as in the case of panels (c), (e), and (f)] she randomly selects a second-order associate. In panel (f), that second-order associate had been previously examined in panel (e). Therefore, in panel (f) FRAN looks at a third-order associate. The solid nodes in the figures indicate words FRAN can recognize by one of her three criteria for word recognition. When she encounters such a word during search [as happens in panels (b), (c), and (e)] she would enter that word into STS, attempt to tag the word and the pathways in both directions between it and the target word, *A*. Such tags are indicated by *X*'s in Fig. 2. The *X*'s in panels (a) through (e) indicate paths and words tagged prior to this study trial on word *A*. The new *X*'s added by panel (f) indicate tags that were developed during the study of *A*. Since tagging is a probabilistic event, FRAN does not succeed in tagging all the words and paths she tries to tag.

For simplicity, it is assumed that FRAN will be busy tagging a target word, its associates, and the paths between them only when that word is under active study. As soon as a new word is presented for study, all this activity surrounding the previous target word ceases. However, the word and its associates may reside in STS for some time after a new word comes under study. Two reasons motivated this decision to restrict the tagging period. First, in recorded introspections, our *S*s appeared to stop thinking about the former word and its associates when a new word was presented. Thus, it seemed natural to assume no further modifications were being made to this part of LTS although it was still part of STS. Secondly, the computational task of determining the probability of tagging a word or associate is much reduced if the duration for which the tagging will be attempted is known at the outset of the tagging process. This duration can be determined only if it is known in advance when the tagging process for a given target word will terminate.

3. *Selection of Entry Words*

The third and final process in which FRAN is engaged during study is the creation of ENTRYSET. This involves the selection of a limited number of "entry words" into the network which can serve as initial words from which FRAN may chain associatively during recall. In the program, words in ENTRYSET are those which have an association from the LIST-*N* marker. It is assumed that the number of words which can be associated in an experiment is very limited — specifically three. It is also assumed that the ENTRYSET is composed initially of the first words seen. Thereafter, changes in ENTRYSET involves replacing present members by new members. A member of the ENTRYSET will be replaced under two conditions:

1. If an item appears that is perceptually more distinctive than any current member of the ENTRYSET, the distinctive item may be put on the ENTRYSET. While we thus acknowledge in theory a Von Restorff effect and effects of instructional emphases, the LISP functions to simulate this feature are not implemented. FRAN is not a model of perceptual processing and has no way to determine if one word is perceptually more distinctive than another. This is another area for future development of the model.

2. FRAN has certain crude heuristics for determining during study which words are more central in the list; that is, for selecting those list words which lead associatively to the most other words from the list. Whenever a word is found which is more central than another in the ENTRYSET, the old one is deleted and the new one is added to ENTRYSET.

Thus, the three words in ENTRYSET constitute intuitively a very rough description of the entire list — its unique perceptual features and the major associative clusters in the list. Subjects commonly make remarks suggesting this ENTRYSET assumption. In rationalizing why they group words together, *S*s often say something on this order: "I can only remember a few things at once. So I tried to reduce the number of things I had to remember." The concept of an ENTRYSET is one explication of what *S*s mean by this remark. The principal motivation for introducing an ENTRYSET in the model was the belief that recall would be quite unstable and variable if the only way of accessing LTS was by random entry or with the words that happened to be in STS at the beginning of recall. It was felt that recall would proceed well only if FRAN had some way of directing her search to the relevant parts of LTS.

C. The Recall Processes

In recall FRAN is simultaneously engaged in two processes: recalling list words and studying the words she is recalling.

1. *The Recall Algorithm*

FRAN, is first of all, engaged in recalling the list words. FRAN always begins by recalling those words to which she has immediate access, viz., those in STS. Available evidence indicates that humans usually begin by recalling the current contents of STS. At the end of list input, the contents of STS are the last one to five items in the list and some of their associates (total sum is five). Postman and Keppel (1968) and Shuell and Keppel (1968) both report that *S*s tend to recall first the last items studied. This output strategy accounts for the well-known recency effect, viz., that items presented in the last few input positions are better recalled than earlier items. Interfering tasks interpolated between study and recall eliminate this recency effect (Glanzer & Cunitz, 1966; Postman & Phillips, 1965). This effect is predicted by any model which assumes that the effect of such interfering tasks is to remove from STS the words filling it immediately at the end of the study trial.

FRAN uses these words from her STS and those from her ENTRYSET as retrieval cues or points at which to enter LTS. From these points she chains along the associative paths she has marked out during study. FRAN's memory may be represented by a symmetric connected graph like Fig. 3. As a consequence of study, certain words and associations have been tagged, and these are marked by X's in Fig. 3. In this way FRAN has marked out a subgraph as relevant for recall of this particular list. The subgraph

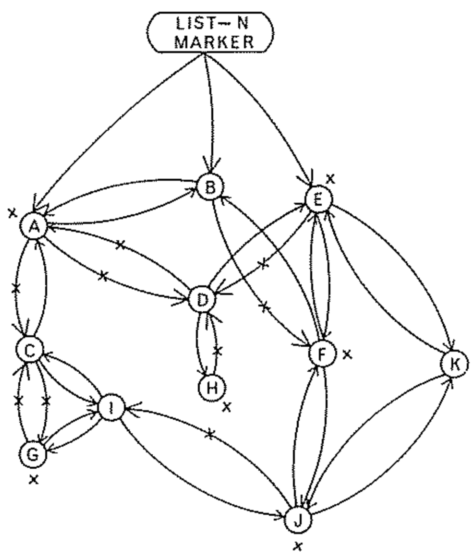

Fig. 3. A hypothetical memory structure that might exist after a study phase. See text for explanation.

marked out in this hypothetical case is given in Fig. 4. The subgraph need not be connected nor symmetric as was the original graph. Several entry points into the subgraph (i.e., words in ENTRYSET and the initial contents of STS) are available; in Figs. 3 and 4, arrows lead to these entry words. Any part of the subgraph that is not strongly connected (i.e., by arrows in the correct direction) to an entry point will not be recalled since FRAN will only examine that part of the subgraph which she can reach by following marked associative paths from the entry points. Therefore, although words H and J are marked in Fig. 4, they will not be recalled. FRAN also will not recall a word unless she can recognize it as from the list (she uses the same three criteria for word recognition as in study). Therefore, although the word C is part of the strongly connected subgraph, it will not be recalled. FRAN regards such words as mediators in an associative chain. In the case of C, it mediates recall of the marked word G.

We have defined which words FRAN will recall; we will now describe the order in which these words are recalled. FRAN searches the subgraph by randomly selecting one of her entry words and then searching that part of the subgraph that can be reached by following associations in a depth-first manner. Depth first means that FRAN will completely search a left-going branch of a node before searching any other branches that may lead from the node. It was trivial to

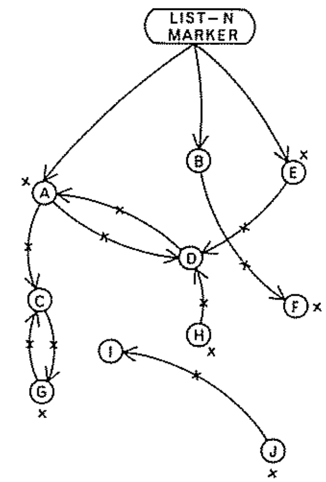

Fig. 4. The marked subgraph that is embedded in the memory structure of Fig. 3. See text for explanation.

define what was meant by "left-most" in FRAN's LISP memory structure. However, it is another matter to say what is meant by "left-most" in a human structure. Perhaps, it is the association that has the most distinctive tag.

FRAN leaves a temporary marking on every word encountered during her search. Occasional failure at temporary marking, something not possible in the current implementation of FRAN, would explain why some *S*s repeat items during recall. After FRAN recalls all words she can by following associative paths from one entry word, another entry word is randomly chosen. This process continues until FRAN has recalled all the words that can be reached from her entry points. Thus, the order of recall is partly determined by the random selection of entry words and partly by the associative chains cued off by these words.

It might be thought that, having searched the part of her memory accessible from her entry words and still having more to recall, FRAN might set out and start randomly searching her memory for tagged words like H and J. The problem is that if FRAN is to have any reasonable chance of finding such lost words, she must be prepared to search a large portion of her memory. With a human, although not with the current implementation of FRAN, that would involve search of thousands of words. Moreover, associative memory is badly organized for an exhaustive search, and it seems certain that there would be a tremendous amount of looping in which the organization of memory would "conspire" to bring one back again and again to the same place without breaking out into new regions of memory. This difficulty is just what FRAN's memory strategies are directed against, i.e., these strategies constitute an attempt to build up a structure that will lead to the relevant parts of memory and avoid time-consuming random searches. It may well be that the human does recall an occasional word by something analogous to a random search, but we hold that, overall, this possibility is not important.

From this description of the associative retrieval process, only a small theoretical step is required to handle latencies and interresponse times (IRTs) in free recall protocols. The general characterization of the sequences of IRTs (cf. Pollio, 1968) is that there are "bursts" of two to seven items recalled with short IRTs followed by a longer IRT pause, then another burst of responses, another pause, etc. The items within a burst tend to be those which the *S* has grouped together in a "subjective unit." The pauses between successive recall bursts grow progressively longer as though *S*s were exhausting the accessible units in memory.

To coordinate FRAN's retrieval concepts to this IRT description, a "burst of responses from a subjective group" corresponds to FRAN's recalling all items accessible from a given starter (taken either from the initial contents of STS or ENTRYSET). Assume that the retrieval-search process operates rapidly in this respect, following tagged associative pathways out from the starter word and from intermediate items recalled. By definition, the final word in such an associative recall chain is one from which *extensive* searching turns up no new tagged pathways or tagged items, so it is followed by a silent "deadtime." After such fruitless searching, FRAN then gives up on the final item of a given chain, and goes to fetch the next starter item from the initial STS or ENTRYSET, entering it into STS as the target for beginning another associative search. As a consequence of these several operations, FRAN would have relatively longer IRTs between "groups" that within groups of items. The fact that between-group IRTs grow progressively longer as the memory store is exhausted may be coordinated to the following facts about FRAN's operation: first, the items in STS may be recalled first with extreme rapidity, yielding short IRTs at the beginning of the protocol; second, the selection of a new starter either from ENTRYSET or earlier STS items may be governed by a "sampling-with-replacement" scheme (with editing and rejection of previously used starters). This means that the more entry words that have been used on a trial, the longer it will take the Executive to find another entry word it has not already used on this trial.

These several ideas about retrieval IRTs in free recall need further specification of details and implementation in the program. This is possibly a fruitful line for development of the model, since clearly the structural assumptions made above already suffice to explain the main features of free recall IRTs.

2. *Learning during Recall*

It is assumed that when FRAN recalls a word, she also studies it for a period of two seconds. Therefore, whenever FRAN recalls a word, at the same time she enters the word into STS and attempts to tag it, searches its associates for other list words, tries to mark the paths to and from these words, and tests to see whether she can place a more central item on ENTRYSET. Although a study time of two seconds was arbitrary, when a word is recalled the *S* clearly does have an opportunity to study it. The sole difference between seeing a word in study and seeing it in recall is the agent responsible for its presentation. There is no reason to assume that this difference should affect how the S reacts to the word. The fact that FRAN is simultaneously studying while she is recalling implies that the subgraph to which she has access from her entry points (e.g., that in Fig. 4) does not remain constant during recall.

This concludes the description of the processes which occupy FRAN during recall. The study and recall processes may be repeated any number of times and in any order. Improvement in recall with repeated study and recall trials is due to two factors. First, with later repetitions FRAN has further opportunities to successfully tag words and their interassociations. Second, the composition of ENTRYSET becomes more and more optimal.

The reader will note that several parameters control the model's processes. All but one parameter are held constant in the simulations to be reported in the next section. The constant parameters are: the size of STS, equal to five words; the size of ENTRYSET, equal to three words; the formula $5 + T$, relating the number of associates examined to the time T for study of a target word; and the study time for words recalled in the test phase, equal to two seconds. In addition, the associative structure of LTS constitutes a constant parameter of the simulations. The single variable parameter is a, the probability of not tagging a word in a second's worth of trying. Presumably, this parameter should also be constant from experiment to experiment for the same Ss. However, because FRAN's memory structure is very much smaller than that of a human, we have let a vary across experiments as a "slop factor" to take account of the possibility that the exact interaction between the experimental paradigm and memory structure may not be identical for FRAN and humans. By raising or lowering a for a simulation, we only change the overall level of performance and we do not alter any of the within-experiment comparisons that we wish to make. Consequently, the fact that FRAN performs at about the same overall level in any experiment as do humans should not be taken as a significant result. The more discriminating test of the model is whether the within-experiment comparisons of human performance are reproduced in the simulations.

IV. An Evaluation: FRAN *vs.* the Humans

A. Evaluation Strategy and Preliminary Details

The test of our model is whether it can simulate, at least qualitatively, the effects of various experimental manipulations in free recall. Several such simulations are reported in this section. An unfortunate practical limitation on this research is that FRAN has become a very expensive S from whom to gather data. For example, in compiled form, in a partition of the IBM System 360/67 that has 700,000 bytes of core, it requires about three seconds of CPU time to run one input-output cycle with one simulation S on a list of 30 words. This long time is due to the inherent slowness of LISP and our extensive use of randomization functions (e.g., for simulating tagging probabilities). Although three seconds may appear brief to some readers, it in fact represents a tremendous amount of computing for a 360 machine; and the three seconds per S per trial mounts rapidly as we simulate up to eight conditions in a given experiment, each condition with 10–20 Ss who are run perhaps five trials.

To make very accurate statements about what our free recall model predicts for a particular experiment would require prohibitively long and expensive computer runs. Therefore, in the following experiments only a small number of simulation Ss have been run for any condition. These artificial Ss produce a recall data differing quite a bit from one S to another; the variability is caused by list randomization, by random search activities, and by the stochastic tagging operations. Consequently, FRAN's data are subject to random fluctuation as are average data taken from a small number of Ss.

Furthermore, simulation data have been obtained only for a single choice of parameters. It would be exorbitantly expensive to map out all possible parameter combinations and then to report that one yielding the best-fitting Monte Carlo data. For this reason, the following simulations give us a good idea of the character of the model's predictions, but the predictions can in no sense be construed as "theoretical best fits."

A few remarks are appropriate regarding our theoretical strategy. We take it for granted that a model like FRAN in its initial formulations will be wrong in some details. The strategy for scientific progress, then, is to try to pinpoint those inaccuracies. Hence, we tend to focus attention and discussion on FRAN's failures rather than her successes. By minimizing and by playing down those details which FRAN fails to match in human behavior, the model could have been presented in a more favorable light. However, our concern is not with propaganda but with understanding human performance. Because we understand FRAN's behavior, those aspects of human behavior which she matches are not so interesting as those aspects which she cannot match.

B. STANDARD MULTITRIAL FREE RECALL

The following experiment was designed to obtain data to evaluate FRAN. Eighteen Stanford students participated for one hour in the experiment for a wage of $1.75. They were tested in groups of from one to seven. Three different lists of 32 words were used. They had been selected from FRAN's vocabulary randomly except that no two lists contained any words in common. The words were typed on slides and presented to the *Ss* by a slide projector at the rate of two seconds per word. Immediately after studying a list, the *Ss* had three minutes to recall the words they had just seen. Such study-recall cycles were repeated five times for a given list of words, with the word order being rerandomized on each study trial. Each *S* learned two of the three lists over the experimental session. With 18 *Ss* each studying two lists, there were 36 recall protocols in all.

The *Ss* had been informed ahead of time about the utility of an associative strategy like that adopted by FRAN and were instructed to restrict their recall efforts to such a strategy. That is, they were told that during study they should only look for conceptual relations between the words in the list and during recall they should use these relations to chain associatively through the list. As noted earlier, some *Ss* spontaneously adopt quite different strategies, and, while most *Ss* adopt strategies very similar to FRAN's, it is usually not until some point after the beginning of the experiment that they completely take up the associative strategy. We wanted data representing use of the associative strategy in a pure form over all trials of the experiment. In postexperimental interviews, our *Ss* unanimously reported that they had no difficulty using the associative strategy as instructed; they also thought it was very useful. While instructing *Ss* to use an associative strategy provides data less "contaminated" by random strategy variation, it complicates interpretation in a different way. To defend the assertion that FRAN is a plausible model of standard multitrial free recall, it is not enough to show that her behavior matches that of the humans in this experiment. It may be that, all our introspective reports to the contrary, the learning strategy spontaneously adopted by most *Ss* is not associative. Therefore, it should be shown that the data gathered from our *Ss* are similar to that gathered from *Ss* not instructed to use a particular learning strategy. We attempt to do this by accompanying our data with references to similar data in the literature.

From FRAN we also have 36 recall protocols. Each *S*'s exact sequence of presentation of words was simulated in one of the 36

computer runs. The value of the parameter *a* was set at .55 for these simulations.

1. The Mean Learning Curve

A comparison of FRAN's learning curve with that of the humans is shown in Fig. 5. FRAN's increase in recall across trials very closely parallels that of the human *Ss*. On each trial FRAN's recall is slightly

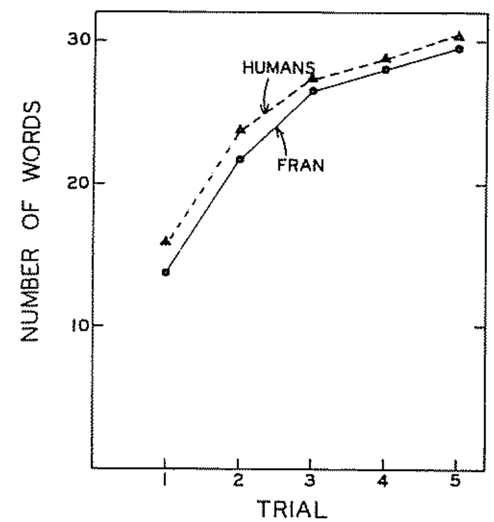

Fig. 5. FRAN and human *Ss* compared with respect to mean number of words recalled as a function of trial.

less, but this discrepancy could be rectified by a lower value of the parameter a. In summarizing a considerable quantity of data, Murdock (1960) concluded that the free recall learning curve was exponential with an asymptote equal to the number of words in the list. This description characterizes the learning curves of both FRAN and the humans quite well.

2. Recall Conditional on Prior Recall or Nonrecall

The success of FRAN in matching the overall recall performance of *Ss* hides one important difference that appears in Fig. 6. Figure 6 depicts, for FRAN and for our *Ss*, their success in recalling a particular word conditional upon whether they had recalled that word (R) or not (N) on the previous trial. The points plotted for Trial 1 are just the unconditional probabilities of recall for that trial. The humans and FRAN are quite close for the curves labeled $P(R_n|R_{n-1})$, the probability of recall conditional on recall on the prior trial. FRAN is somewhat better, probably reflecting her greater efficiency in searching her marked subgraph. Human *S*s may

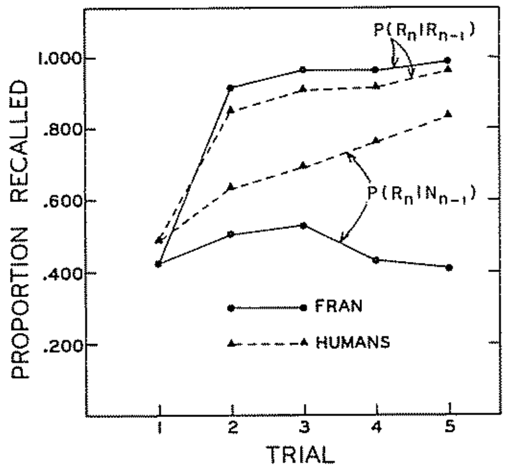

Fig. 6. FRAN and human *S*s compared with respect to mean proportion of words recalled (R_n) conditional on recall (R_{n-1}) or nonrecall (N_{n-1}) on the previous trial.

occasionally "forget" to check a word for associates leading from it, which is a failing that FRAN currently does not have. It would be trivial to introduce this into the program as random noise in retrieval. Therefore, the slight discrepancy in $P(R_n|R_{n-1})$ is not considered serious.

The much greater and theoretically more significant discrepancy is between the curves labeled $P(R_n|N_{n-1})$, the probability of recall of a word conditional on its nonrecall on the previous trial. Human *S*s show a consistent increase across trials in their probability of recalling words missed on the previous trial; FRAN does not. The hardest words for FRAN to recall will be those that are "farthest away" from the rest of the words (where distance in the network is measured by the number of links intervening between two nodes). FRAN's study strategy, which mechanically allocates just two seconds study time to each word, will be inefficient in learning these "distant" words. Obviously these distant words constitute a larger proportion of nonrecalled words on later trials, and hence contribute

more on later trials to the curve labeled $P(R_n|N_{n-1})$. Therefore, it is not too surprising that FRAN's recall of these words does not increase across trials. The reason that it does not show a monotonic decrease is that there are compensating factors such as (*a*) the greater probability that on later trials the words will have been tagged and hence recognized, and (*b*) the more efficient organization of the marked subgraph which makes possible the freeing of room in ENTRYSET for adding a new word which accesses new parts of memory. Since FRAN's memory is connected, even "distant" words will be eventually linked into the accessible subgraph and hence will be recalled.

For humans also, the curve, $P(R_n|N_{n-1})$ should have a greater proportion of difficult words as the experiment progresses. Why, then, do *S*s show a monotonic increase in their ability to recall words previously not recalled? We think it is because human *S*s adopt the strategy of giving special attention or special processing priority to these difficult words. This supposes, as seems plausible, that *S*s can discriminate between easy (recalled) *versus* difficult (nonrecalled) items on a subsequent input trial. As one strategy, *S*s could give extra time to studying the difficult words at the expense of temporally adjacent items that are already well learned. Another strategy would be for the *S* to keep the difficult word in STS from the time it is studied until the time recall is initiated and, then to output that word immediately.

If *S*s were using this priority rating to replace items in STS, one would expect a greater than chance proportion of previously nonrecalled words to be recalled in the early part of the *S*s' output. This priority effect has been found in several recall experiments (Battig, Allen, & Jensen, 1965; Battig & Slaybaugh, 1969). There was a similar tendency in our experiment for those words omitted on Trial 4 then recalled on Trial 5. Of those items that were not recalled on Trial 4, were presented in the first half of the input list on Trial 5 and were recalled on Trial 5, 62% were recalled in the first half of the output protocol. Of previously omitted words presented in the last half of the Trial 5 input and then recalled, 67% were recalled in the first half of the output protocol. The chance expectation for these percentages, given a randomly selected output order, is 50%, so the data show a tendency toward early recall of previously unrecalled items. On the other hand, FRAN shows just the opposite result, with corresponding figures of 17% and 40% for proportion of words recalled in the first half of the Trial 5 protocol given that they were not recalled on Trial 4 and were presented in the first *versus* second half of the Trial 5 input list.

The reason that FRAN's percentages are below chance is that recall of a previously unrecalled word often depends on use of a single entry word that was added during the just-prior study trial. Late in learning the marked subgraph will have become so interconnected that most words in the subgraph will be connected by marked pathways to most other words. Therefore, if any other entry words are used before the one leading to the few nonrecalled words, those other entry words will probably lead associatively to the recall of most of the list before the new item is recalled. With eight entry words (five from STS and three from ENTRYSET), the probability is .88 that the first entry word chosen is not the one that leads to the recall of the previously omitted word. As a consequence, recall of these newly appearing words tends to be postponed until a late position in the recall sequence.

3. *Analysis of Complete Sequences*

A more complete analysis of the changes in recall across trials is given in Table I which presents the frequencies of the 32 possible

TABLE I

Frequencies of the Various Combinations of Recall and Nonrecall across Trials

Events	FRAN	Human	Events	FRAN	Human
NNNNN	59	4	RNNNN	8	1
NNNNR	29	14	RNNNR	1	1
NNNRN	4	9	RNNRN	0	0
NNNRR	66	39	RNNRR	4	24
NNRNN	9	1	RNRNN	2	2
NNRNR	10	23	RNRNR	0	5
NNRRN	0	6	RNRRN	0	2
NNRRR	148	115	RNRRR	22	46
NRNNN	5	3	RRNNN	1	2
NRNNR	9	6	RRNNR	0	3
NRNRN	2	6	RRNRN	0	3
NRNRR	6	37	RRNRR	3	21
NRRNN	1	3	RRRNN	0	4
NRRNR	6	24	RRRNR	4	17
NRRRN	0	11	RRRRN	0	5
NRRRR	305	285	RRRRR	448	430

recall sequences for a word over the five trials. In Table I, R stands for recall and N for nonrecall, so that RRRNR denotes recall on Trials 1, 2, 3, 5 and nonrecall on Trial 4. For both FRAN and human *S*s, there were 1152 observations (32 words times 36 recall protocols).

Careful inspection of Table I reveals no glaring discrepancies between the two sets of data beyond those noted in conjunction with the $P(R_n|N_{n-1})$ curves in Fig. 6. The correlation between the frequencies of the various sequences for FRAN and for human *S*s is .985. Therefore, if FRAN's frequencies had been used to predict the frequencies of the various events in the human data, FRAN would account for 97% of the variance. This "variance accounted for" is high despite the $P(R_n|N_{n-1})$ discrepancy because the latter statistic involves only a small percentage of the sequence data in Table I. Hence, one may justifiably point to the recall sequences in Table I as evidence that the major processes underlying free recall are being modeled, at least approximately, in FRAN. However, sufficient discrepancies exist to conclude that humans are capable of some complexities (e.g., giving priority to difficult items) which the current FRAN does not mimic.

4. *Serial Position Curve*

Another comparison of FRAN's performance with that of human *S*s comes from examination of serial position curves, relating recall probability to the ordinal position of an item in the input list. Figure 7 illustrates a comparison between the serial position curves of FRAN and our *S*s. The two sets of curves are quite similar to each other and to those reported in the literature (e.g:, Murdock, 1962; Shuell & Keppel, 1968). Both the humans and FRAN show a primacy effect for the initial four words but only on the first trial. The primacy effect in FRAN is caused by the fact that ENTRYSET is composed initially from the first words seen in the experiment. As a consequence, these first words have a good chance of being in ENTRYSET or retrievable from another word in ENTRYSET at the time of recall.

Both FRAN and the humans show a recency effect on all trials for the last four words. FRAN's recency effect is more pronounced than that observed in this experiment. The recency effect in FRAN is due to the fact that, at the beginning of recall, STS is composed of the last few words and their associates, and these items are sure to be recalled. The magnitude of the predicted recency effect could be easily reduced by decreasing the size of STS from its current

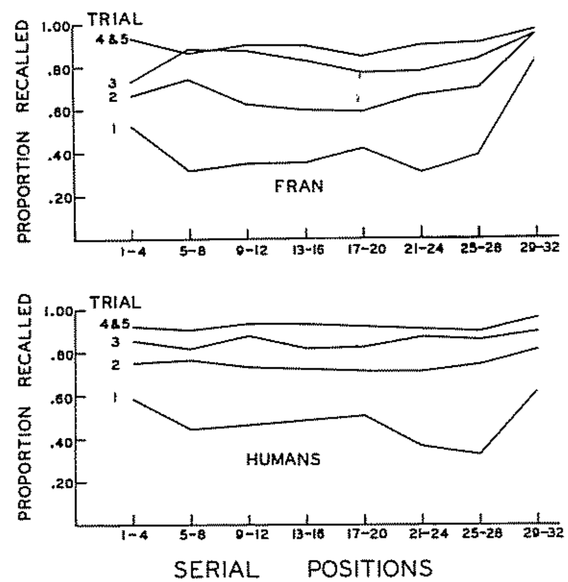

Fig. 7. FRAN and humans compared with respect to recall as a function of serial position of the word in the study sequence. Data for Trials 4 and 5 are pooled.

arbitrary size of five. Similarly, recency effects depend on a particular output strategy, of recalling items from STS before moving to those cued from ENTRYSET. Not all naive *S*s use this output strategy initially, though they tend to adopt it as they become more practiced on free recall tasks. If this were true, then the average curve for humans, pooling across different output strategies, will show less of a recency effect than does FRAN which always uses the "last in, first out" recall strategy. For such reasons as these, we do not consider the differing magnitudes of recency for FRAN *vs.* our *S*s to be a serious shortcoming of the model.

Except for the primacy effect on Trial 1, both the humans and FRAN show on all trials a stable level of recall for items preceding the last four recency items. While our assumptions regarding the use of STS and ENTRYSET can be seen intuitively to lead to a primacy effect on the first trial only and a recency effect on all trials, it might not seem obvious that FRAN would yield a constant level of recall for the remaining serial positions. All of these middle words have an equal probability of being tagged with a list marker; furthermore, it also appears that they would all have an equal probability of being

tied into the accessible part of the subgraph and hence retrieved during recall. On Trial 1, words studied early in the list will have the advantage of being searched for from more words, while later words will have the advantage of there being more tagged words to search for when they are studied. These two factors manage to nicely balance out; it can be proven that the expected number of associative paths that will be tagged to and from a given word is the same for all serial positions.

5. *Input-Output and Output-Output Correlations*

The free recall phenomenon which FRAN was primarily designed to explain was the fact that the order of recall is not random but rather is highly structured. We examined the degree to which the order of output on Trial n tended to maintain (*a*) the order in which the words were studied on Trial n, and (*b*) the order in which the words were recalled on Trial n -1. Figure 8 illustrates a comparison between FRAN and our *S*s with respect to these two types of recall stereotypy. For that figure the statistic, "proportion of repetitions," measures recall stereotypy and is defined as follows: for each *S*, let j be the number of pairs of words that occurred adjacently in the study list (or in the prior trail's recall output) and which were recalled not necessarily adjacently on Trial n. Let i be the number of these j pairs that are recalled adjacently in the same order as they occurred in the study sequence (or in the prior recall). The statistic, proportion of repetitions, is defined as $\Sigma i/\Sigma j$, where i and j were summed over the 36 recall protocols for that trial. This statistic estimates the probability of recalling two words in the same adjacent order as they were studied (or recalled on the prior trial). If the recall order were random, the expected value of this statistic would be $\Sigma j/\Sigma k(k-1)$, where k is the number of words recalled on Trial n. These expected values, shown as the lower lines in Fig. 8, are essentially identical for FRAN and our *S*s.

As Fig. 8 shows, on all trials the obtained values were greater than those expected for both measures of correspondence, between order of input and output, and between order of prior output and current output. With respect to the proportion of repetitions of prior output pairs, FRAN appears much like human *S*s. Both show an increase in this type of stereotypy across trials. Such an increase in output stereotypy has been reported by many investigators (e.g., Bousfield *et al.*, 1964; Rosner, 1970; Tulving, 1962). Turning to the other comparison in Fig. 8, both FRAN and our *S*s decrease across trials in

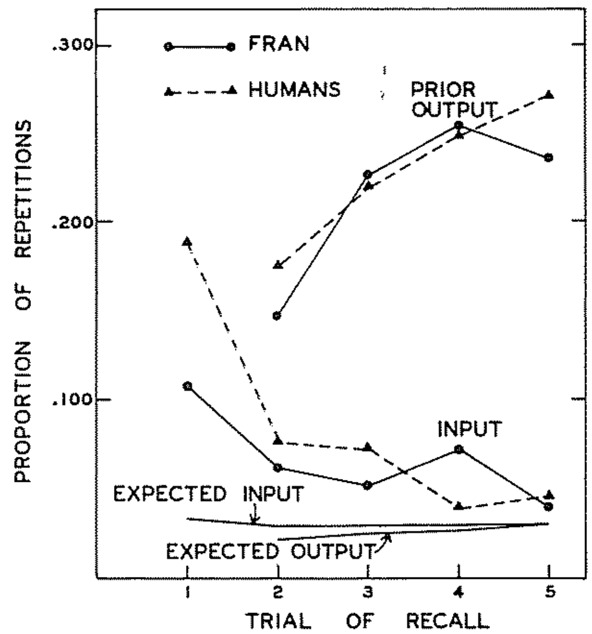

Fig. 8. FRAN and humans compared with respect to stereotypy of output order. Given in the figure are the probabilities with which adjacent pairs in the output of Trial n-1 are repeated in the output of Trial n and the probabilities with which adjacent pairs in the input of Trial n are repeated in the output of Trial n. See text for explanation of the statistic proportion of repetitions.

the frequency with which they recall word pairs in the same order as they were studied. For this measure FRAN fits the human data on Trials 2 through 5, but underpredicts considerably on Trial 1.

In trying to understand the discrepancy on Trial 1 it is well to understand why FRAN repeats input order with greater than chance frequency. Two factors operate to yield this latter result. First, the last few words in the input sequence are recalled immediately out of STS. Although they are recalled randomly from STS, they have a greater than chance probability of occurring adjacently, since the chance level is computed on the assumption that the words recalled are equally likely to come from any position in the list. Also the remaining words, because they cannot occur in the first five output positions, have as a consequence a slightly greater than chance probability of being recalled adjacently. The second factor yielding repetition of input adjacencies is particularly important on the first trial and it may be illustrated as follows: suppose FRAN, in searching out associates of a particular word A, comes upon the word B which had just preceded it in the study sequence. The word B is almost certain to still be in STS and therefore FRAN will recognize it as a list word; she will again attempt to tag word B (especially on the first trial, it may not yet be tagged), and she will attempt to tag the path leading from A to B, and the path from B to A. If she succeeds in tagging the path leading from B, she may use that path in recall to retrieve A from B. If she does, this will be one instance in which output order reflects input order. On the other hand, consider what would happen if FRAN, in searching out associates of A, came upon a list word C that had occurred much earlier in the study sequence. That word would probably not be in STS and, if it were not tagged, FRAN would not recognize it as a list word. Therefore, she would miss this opportunity to tag the associative path she had found from A to C. Comparing the A-B to the A-C case, it is seen that FRAN has a greater probability of marking associative paths between words which occur close together in the study sequence. Since this depends crucially on the fact that some words are not tagged, this tendency is especially strong on Trial 1 when many words are not yet tagged. In this way, FRAN tends to build a marked subgraph that reflects the input adjacencies. The results of the first trial shown in Fig. 8 indicate that these two factors are not enough because humans repeat input order more often than FRAN does.

There is another method for building a subgraph which would further emphasize input adjacencies and which a few *S*s have mentioned in intensive postexperimental interviews that we have conducted in other free recall experiments. There is a tendency to deliberately seek out associative links between the word under study and the previous one. This search strategy would, of course, particularly favor the reproduction of input order in recall, and it might be contributing to the discrepancy between FRAN and our *S*s on the first trial. FRAN currently only searches the associates of the word under study, but she could simulate this pair-wise searching method by doing a parallel search from both words, attempting to find an intersection. It seems likely that humans use some combination of these two methods.

It is interesting to consider whether FRAN would pass a Turing Test; i.e., whether we could successfully distinguish FRAN's recall protocols from those of humans. Ignoring irrelevant details like the format of output, it is clear that discrimination would not always be possible. Some protocols can easily be spotted as originating from a human. These are identifiable by such features as a tendency to give recall priority to previously omitted words and a tendency to preserve input adjacencies on the first recall trial. One could be fairly confident that such a protocol did not originate from FRAN. However, many of the human protocols are indistinguishable from FRAN's. The conclusion seems to be that, with respect to this standard multitrial free recall task, FRAN is like many humans, but that not all humans are identical. If this is so, it would be unrealistic to expect FRAN to give data identical to that averaged from 18 quite different individuals unless we want her to have a multiple personality.

C. Study Time and Recall

Waugh (1963, 1967) reported investigations relating study time to subsequent recall. A simple relationship between study time and recall was found, namely, that the number of words recalled from a list depended only on the total study time for the whole list. Variations in list length, presentation rate, and masses *vs.* distributed presentation times had no effect if study time was constant. Some of Waugh's conditions were simulated with FRAN to see to what extent FRAN could explain Waugh's results.

Waugh's two 1967 experiments will be of particular interest. In the first experiment, words were read to *Ss* at the rate of one per second. There were nine experimental conditions consisting of 120, 60, 40, 30, or 24 words each appearing once within a list; or a basic set of 60, 40, 30, or 24 different words permuted two, three, four, and five times respectively, to yield a total of 120 words in all. Figure 9 shows Waugh's results, with each point based on 72 observations of free recall.

Seven of the nine conditions were simulated, each with the parameter *a* equal to .65. To save money, the two lists with 40 words were omitted. Figures 10 and 11 compare the results of our simulations with those of Waugh's *Ss* for these seven conditions. Since only 10 to 12 observations (simulation runs) contribute to any of FRAN's points, the predictions are not as stable as are Waugh's data. In these figures the arrows point from the value obtained in the simulation to Waugh's observed value. An X indicates essentially identical points. The straight lines in Figs. 9, 10, and 11 indicate the linear relationships in Waugh's data between study time per word and probability of recall of any word in the list.

Clearly, FRAN has managed to simulate the general relations

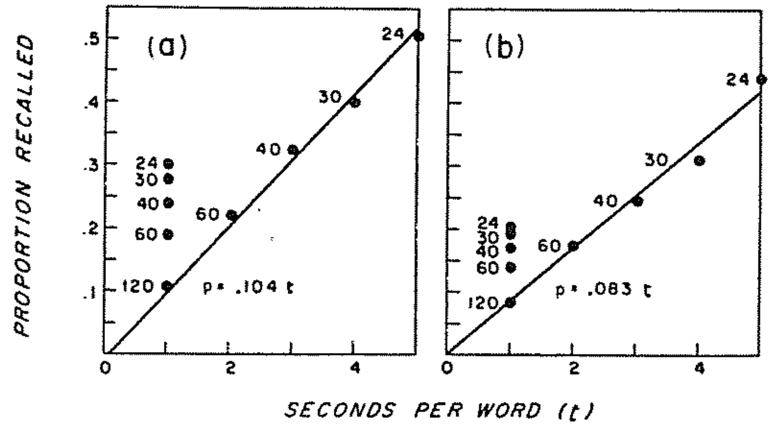

Fig. 9. Probability of recall for a given item as a function of its total presentation time. (The numbers denote the total numbers of items presented. The probabilities were estimated across all items in a list —Fig. 9a — or across items prior to the last seven — Fig. 9b.) The data are taken from Waugh (1967), by permission of the American Psychological Association, Inc.

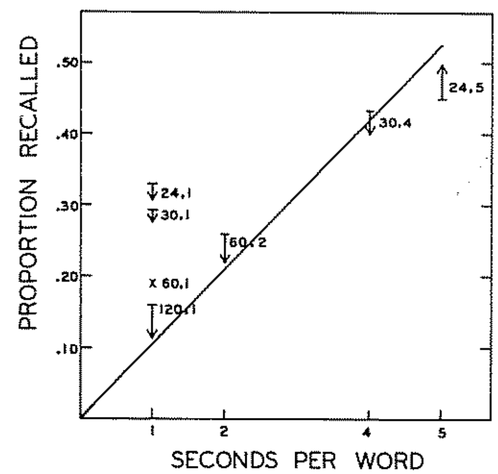

Fig. 10. FRAN and Waugh's *Ss* compared with respect to probability of recall of an item as a function of its total study time. Probabilities are pooled across all serial positions in the input list. The arrows point from the proportion recalled for FRAN to the proportion for the human *Ss*.

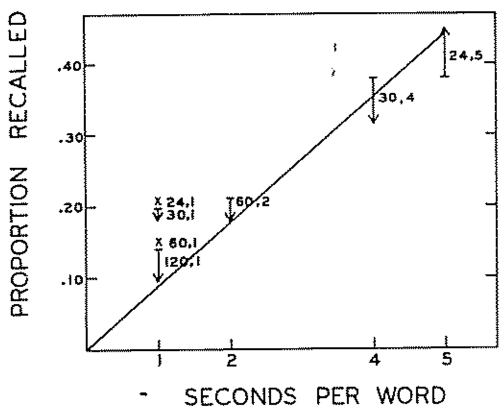

Fig. 11. FRAN and Waugh's *Ss* compared with respect to probability of recall of an item as a function of its total study time. Only serial positions prior to the last seven items contribute to the estimate of probabilities. The arrows point from the proportion recalled for FRAN to the proportion for human *Ss*.

among the observed points although her recall is somewhat higher overall. This could be rectified by a slightly lower value of the parameter *a*. The two most deviant predictions are for the list in which 120 words occurred once and for the list in which 24 words were presented five times. FRAN's excessive recall of the 120-item list may be related to the fact that 120 words is nearly half of FRAN's vocabulary, but is a trivial portion of an adult's vocabulary. Practically every word FRAN studied would have associates from the list. Her better-than-human performance may be due to the fact that retrieval becomes disproportionately easy in such circumstances.

In Fig. 12 we have replotted the data of Figs. 9 and 10 to show how total words recalled is related to the total study time. The point for 120 seconds is based on the average of the several points with this study time in Figs. 9 and 10. Murdock (1960) found a linear relationship between words recalled and study time, described by the equation $R_1 = 6.1 + .06t$, where t was the time in seconds. That predicted relation is shown in Fig. 12. As can be seen, the values estimated from Murdock's equation are fairly close to the values obtained from FRAN and Waugh. However, the relation between study time and words recalled, either for FRAN or for Waugh's *Ss*, is not linear although it is monotonically increasing. Deese (1960) also reports data in which this linear relationship was not upheld exactly.

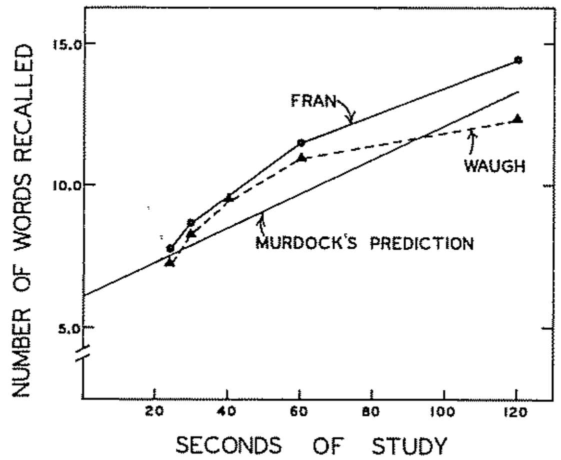

Fig. 12. FRAN and Waugh's *Ss* compared with respect to number of words recalled as a function of total study time. The predicted relation is taken from Murdock (1960).

In a second experiment, Waugh (1967) presented a list of 30 words for either one, two, three, four, or six seconds per word. In a massed condition, the word was presented for its total study time all at once. In a distributed condition, the words were presented only for a second at a time but appeared at several different positions in the input list. The results of her experiment are given in Fig. 13. The 100 observations contributing to each of her data points yield much more stable data than the 10 simulations (observations) that we have obtained from FRAN for each point. Only a subset of Waugh's conditions were simulated, those for one, two, four, and six seconds. FRAN and the humans are compared in Fig. 14 which is to be interpreted as were Figs. 10 and 11; i.e., the lines are taken from Waugh's data and the arrows point from our data to hers. Again, FRAN simulates the general relations in these data. Just how well FRAN does simulate the general relations can be seen in Fig. 15 which summarizes the data in Figs. 10, 11, and 14. It is a scatter plot in which each point corresponds to a single experimental condition. The value for a point on the ordinate is the proportion recalled by Waugh's *Ss* in that condition and the value on the abscissa is FRAN's proportion. If prediction were perfect all points would be on the diagonal line. Given that the proportions recalled are subject to random error, the result is very impressive.

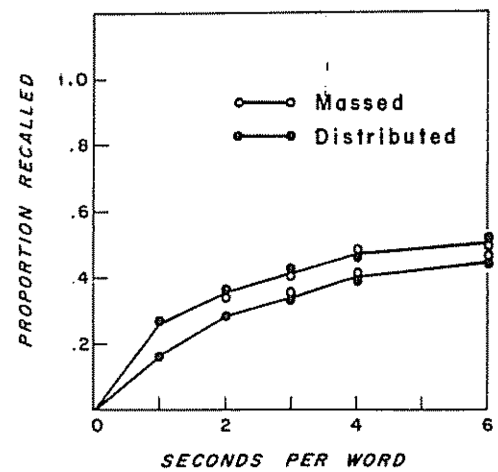

Fig. 13. Probability of recall for a given item as a function of its total presentation time, for all items in a list (upper function) or for items prior to the last seven (lower function). These data are taken from Waugh (1967), by permission of the American Psychological Association, Inc.

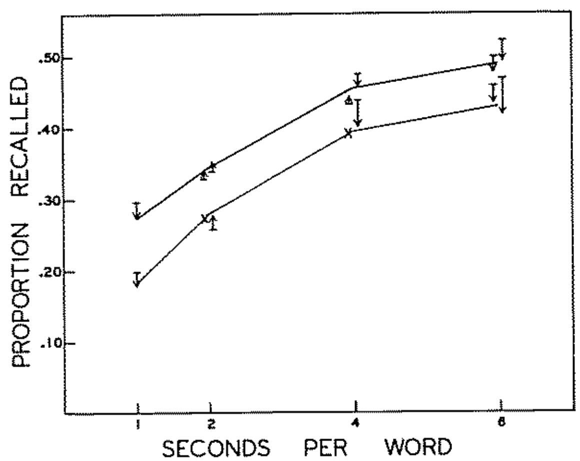

Fig. 14. FRAN and Waugh's *S*s compared with respect to probability of recall of a word as a function of its total presentation time, for all items in a list (upper function) or for items prior to the last seven (lower function). The arrows point from the proportion recalled for FRAN to the proportion for human *S*s. The data for the massed condition are given by the left arrow, for the distributed condition by the right arrow.

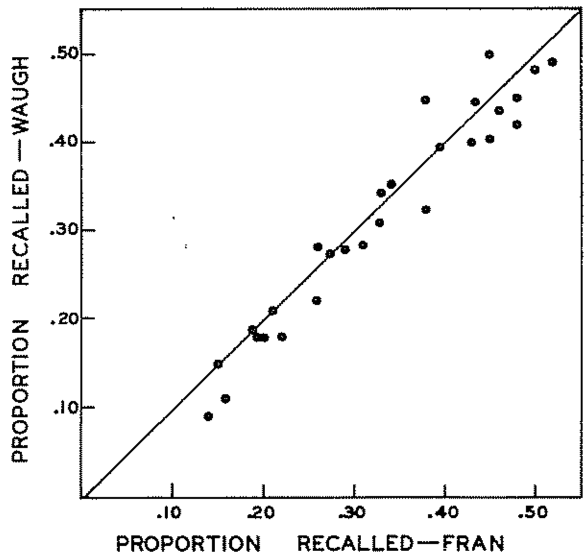

Fig. 15. Summary of data in Figs. 10, 11, and 14. Each point represents one list condition with the proportion either estimated from all items or just from the items prior to the last seven.

In an earlier study, Waugh (1963) examined free recall of 30 words presented at a one second rate, but with variation in the number of list words that were repeated. She described the procedure as follows:

> In order to control for serial position effects, we divided the 30-word lists into 14 early-late words, which occurred in positions 1-6 and 23-30 and 16 middle words, which occurred in positions 7-22. . . . We accordingly constructed five sorts of lists which differed in the number of repeated words that each contained. There were either 1, 2, 4, 6, or 8 words repeated in the middle group of 16. Each list represented an experimental condition, which I shall designate as Cond. 1, 2, 4, 6, or 8, according to the number of words that occurred twice in the middle segment of a list. In the early-late portion of the lists, the number of repeated words was always one less than the number repeated in the middle, so that the total number that occurred twice in a list was either 1, 3, 7, 11, or 15. In this last case, under Cond. 8, every word in the list occurred twice [Waugh, 1963, pp. 107–108, by permission of Academic Press, New York].

The results in this experiment are summarized in Fig. 16. Note that the total words recalled from the middle portion of the lists and from the early-late portion of the list are constant across conditions.

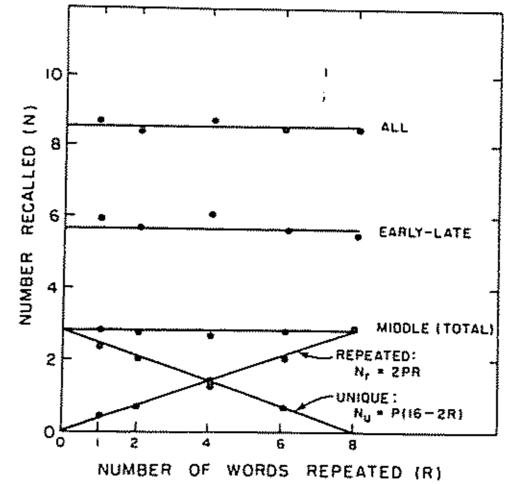

Fig. 16. Mean number of repeated and unique words recalled as a function of the number of words repeated within the middle 16 positions of a list. This figure is taken from Waugh (1963), by permission of Academic Press, New York.

Waugh's five experimental conditions were simulated with the value of $a = .65$. Figure 17 illustrates a comparison between the simulations and Waugh's results. As in previous figures, the arrows point from FRAN's prediction to Waugh's data. However, in contrast with previous figures, the straight lines are estimated from FRAN's data and not from Waugh's data. While each of her data points is based on over 300 observations, FRAN's rest on just 20 simulations, with correspondingly larger variances. While FRAN is recalling about the same number of early-late words as Waugh's *Ss*, FRAN's recall from the middle positions is about 25% lower. Except for this, the simulation data preserves the basic relations in Waugh's data. We probably could have matched recall both in the early-late portion and in the middle portion by simultaneously lowering the value of a and decreasing the size of STS. The number of words recalled from the middle portion would then increase because a (the probability of not tagging in a second) would be less. The increase in recall in the early-late portion due to the change in a would be cancelled out by the reduction in the size of STS. We have already noted in Section IV,B, in discussing the serial position curve, that STS should have been smaller in FRAN.

Waugh was interested in the relation between recall of middle

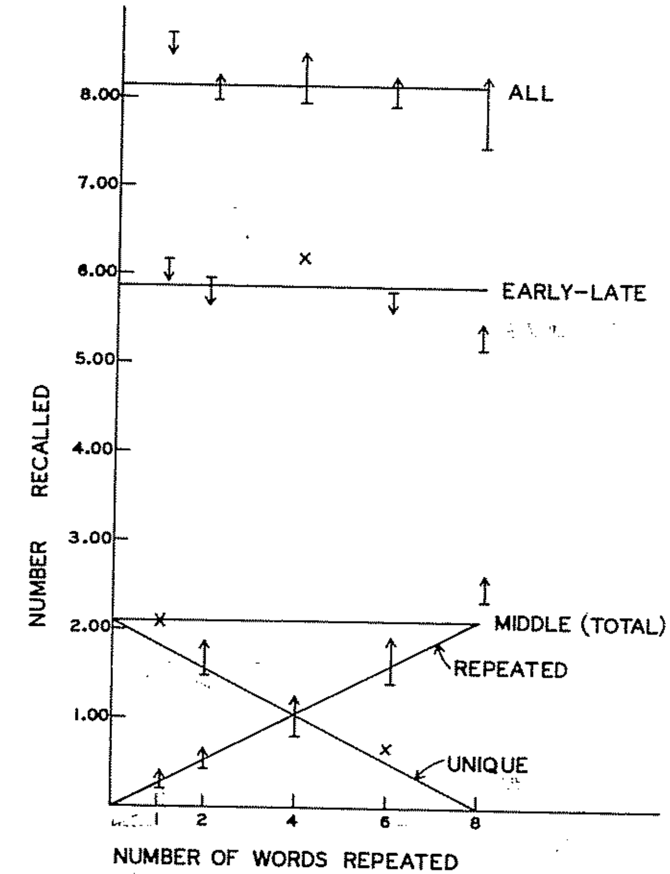

Fig. 17. FRAN and Waugh's *Ss* compared with respect to mean number of repeated and unique words recalled. The arrows point from the proportion recalled for FRAN to the proportion for the human *Ss*.

words that were repeated and those that were not. Letting P denote the probability of recalling a nonrepeated word, the issue was whether a repeated word would be recalled with probability P + (1-P)P or with probability 2P. She was able to discriminate statistically in favor of the latter hypothesis. We did not have enough simulated data to make such a slight distinction, but the probability of FRAN's recall for the nonrepeated words was .135 and for the repeated words, .257 (averaged over the five conditions).

Waugh was also interested in whether recall would be affected by the lag between repeated words. The lag refers to the number of items intervening between the first and second presentation of a given word. Waugh's lag data for the middle 16 words are presented in Fig. 18. There is no apparent effect of lag upon recall, a fact which Waugh found surprising. There appears no reason to expect an effect of lag for FRAN, and Fig. 19 confirms that there was no systematic

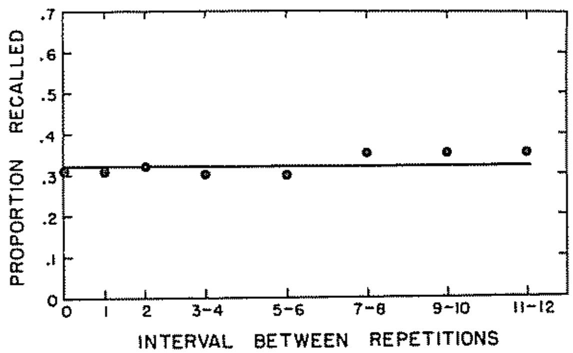

Fig. 18. Probability of recalling a repeated word as a function of the number of other words intervening between the first and second occurrences of a word. This figure is taken from Waugh (1963), by permission of Academic Press, New York.

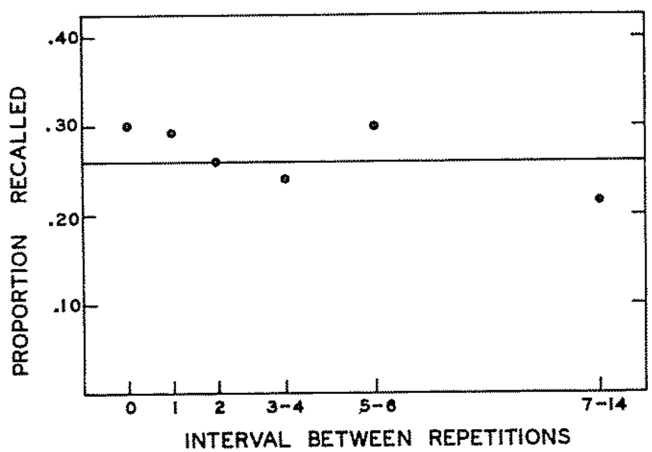

Fig. 19. FRAN's probability of recalling a repeated word as a function of the number of other words intervening between the first and second occurrences of the word.

effect for our simulations. In that figure, lags 7 through 14 were pooled because of small amounts of data for these positions.

It should be noted, in conclusion, that the experiments subsequent to Waugh's research (see Melton, 1970) have not always succeeded in replicating her findings that the effects of total study time are independent of variables such as massed *vs.* distributed presentation and the lag between repetition of items. These studies have involved various design differences such as rate of presentation. Whether FRAN will show similar departures from the simple relation between study time and recall when such experimental parameters are changed remains a question for further research.

D. Cued Recall

Slamecka (1968) reported six single-trial free recall experiments, each of which compared two groups of *S*s given different conditions of recall. During testing, one group was given part of the list and was required to recall the remainder. The other group recalled the whole list that they had studied without such cues. Slamecka was concerned with recall on that portion of the list which both groups were required to recall. The consistent finding was that the control group, which had to recall all the list, did as well as or better than the experimental group which had been cued with part of the list. Cueing FRAN with words from the list should give her additional entry points into her subgraph at retrieval, and this should increase recall. The following discussion is concerned, first, with procedural factors that might have affected the cued *vs.* noncued recall difference in Slamecka's experiment, and, second, with the magnitude of increased recall that an associative model like FRAN predicts for the cued group.

1. *Procedural Factors*

Slamecka noted one possible reason why his cued *S*s were often inferior to his noncued *S*s. The cued *S* tended to spend about the first 10 to 20 seconds reading over the list of cue words which they had been given and which they were instructed not to recall. He conjectured that, in the time spent initially scanning the cue words, these *S*s would lose access to the contents of STS from the prior study list. In an experiment in which the noncued group was given a comparably interfering task for the first 15 seconds, he found no difference in recall of the critical words by the cued and noncued groups.

This result suggested to us a reason why the cued group was not superior to the control group as FRAN would predict. The cued *Ss* have an additional task in recall which the control *Ss* do not have. Not only must they recall the words, but they must also check every word implicitly recalled against a list of words not to be recalled. Slamecka has shown that the second task may interfere with the first, but there is no reason to suppose that the interfering effects of the second task are limited to the first 15 seconds. What is clearly needed is an experimental design that equates the two groups with respect to this second task.

2. An Experiment

Thirty-two *Ss* had two study and recall tests on each of four lists of 40 common nouns. The words were presented on flashcards at a three second rate and *Ss* were given three minutes for recall. During the second recall test, all *Ss* were required to recall only a particular half of the list. That particular half was specified by the initial letters of the words to be recalled. For instance, a *S* might be asked to recall all the nouns that began with letters from A to M. The experimental manipulation was whether the *S* was given the list of words he did not have to recall on the second trial. By this procedure, all *Ss* would have to inhibit their recall of half of the list, but only one group would have the advantage of knowing what were the exact words on the inhibited half of the list. Comparison of the cued and noncued conditions was delayed until the second trial because it has been suggested that the list is not sufficiently organized on the first trial to take advantage of the cue words (although later data by Slamecka, 1969, dispute this point). Each of the 32 *Ss* recalled two lists under the cued condition and two lists under the noncued condition. They did not know in advance which lists would be tested with cueing. The variables of condition, list, and order of presentation were counter-balanced in a 4 x 4 greco-latin square. One greco-latin square was used with four blocks of four *Ss* and a different one with a second four blocks. In all there were 64 recall protocols for each condition.

On the first trial, when all *Ss* had to try to recall 40 words, 14.02 were recalled by those in the cued condition and 14.44 by those in the noncued condition. Under the cued condition, *Ss* recalled a mean of 10.33 words out of a possible 20 in the second recall; under the noncued condition, *Ss* recalled a mean of 9.63. Using as an error term the residual left when variance due to condition, subject, list, and greco-latin square is removed from the total variance, this difference is marginally significant at the .05 confidence level (*t* test, 1 tail). Even if this difference is of marginal significance, it certainly is microscopic; cued *Ss* recalled only 7% of what could not be recalled without cueing.

Allen (1969) reported a cued-recall experiment that was specifically designed to augment the advantage of the cueing group. Closely associated pairs of words were input contiguously. It was hoped that this would induce the *S* to form associations between the pairs. After the *Ss* had recalled all the words they could without benefit of cues, half the *Ss* were presented with a set of cues consisting of one member from each pair. It was hoped that, if neither word in a pair had been recalled, cueing with one member would facilitate the recall of the other. Allen found that 15% of the words that could not be recalled without the aid of cueing were recalled with its aid. This is a significant difference, but not particularly impressive considering Allen's manipulations to obtain an effect.

3. Cueing of FRAN's Recall

Intuitively, one would think that an associative model like FRAN would yield a much larger advantage for cueing than was obtained in our experiment or Allen's. These intuitions were checked by running 32 simulations of the cued and of the control conditions of our experiment. The parameter a was set at .65 for these simulations. In both the control and the cued conditions FRAN recalled only a specified half of the list as did our *Ss*. The cued condition was simulated by permitting FRAN to use as entry words the 20 cue words as well as the five from STS, and the three from ENTRYSET. This manipulation for the cued condition seemed reasonable given our intuitions that cueing a *S* with list words should give him extra points at which to access his memory.

In the control condition, after FRAN had recalled all the words she could, she studied those words she had recalled for an extra time. In studying these, she might succeed in finding paths to other list words and so recall them. The decision to let FRAN restudy the words she had recalled in the control condition was taken so that she would be engaging in some possibly beneficial activity to compensate for the time she was spending in the cued condition with the extra entry words. As it turned out, this was not a particularly rewarding additional activity; only three new words were recalled by means of this additional study in 32 simulations.

Turning to the results, 14.59 words were recalled on the first trial for simulations of the cued condition and 15.12 for the simulations of the noncued condition. This is slightly higher than the 14.02 and 14.44 reported above for the humans. On the second trial 14.59 of the 20 possible words were recalled in the cued condition and only 10.63 in the noncued condition, a difference of almost four words. Therefore, FRAN, with the aid of cueing, is able to recall 42% of the words she could not recall without the cues. This compares with the values of 7% in our experiment above and 15% in Allen's experiment. So, our intuitions were correct; an associative model predicts a much larger advantage of half-list cueing than is in fact found.

It would not have been difficult to have had FRAN behave differently in simulations of the half-list cueing task and thus reduce the difference predicted between the cued and the control conditions. The problem is to develop some independent motivation for such alternative assumptions. In the control condition, FRAN might be allowed to randomly search her memory looking for new words as another activity to compensate for the time spent with the extra cued words in the experimental condition. In a small memory like FRAN's, this would surely result in increased recall; but, as argued in Section III,C, it is doubtful whether this strategy is fruitful in a human-sized memory. Another assumption would be for FRAN to treat these half-list cue words with less careful consideration than is given to the other entry words. Perhaps, she would only look at a random subset of the associates of these words. It is not unreasonable to suppose that a human *S* when confronted with 20 cue words, most of which are of no help, may become somewhat negligent in his consideration of these words. While these two alternative assumptions have some plausibility, to our way of thinking, they are not as plausible as is FRAN's current behavior.

E. Recall of Categorized Lists

Much of the current interest in free recall surrounds research on the recall of categorized lists, i.e., lists composed of several instances from each of a set of categories. There has been considerable controversy as to the exact nature of the processes which underlie free recall of categorized lists. Shuell (1969) has outlined the three principal types of explanation that have been offered. First are

explanations relying on Miller's (1956) notion of recoding. The general idea is that the words from a category are recoded into a single chunk, and that the *S* recalls the chunk and not the individual words. This is a variant of the "unitization" hypothesis described in the Introduction. The basic difficulty with this explanation is that the details of how the instances are encoded into a chunk or how the chunk is decoded into its instances are left to the imagination. Because of this vagueness, it is not often contrasted with the other two theories about recall of categorized lists.

The second explanation of clustering is that offered by Bousfield (1953). His hypothesis was that the category name itself was implicitly recalled, and that recall of the category name facilitated recall of individual members. This is similar to the categorization strategy which we outlined in Section III,A. That strategy would consist of restricting search during study to those associative links of the subordinate-superordinate variety, trying to find superordinate concepts having several list items as instances. These superordinate concepts (whose names are usually not on the list) would probably form the members of ENTRYSET. Then at recall the *S* could search for marked paths from these concepts to subordinates that were on the list.

The third interpretation is that free recall of both categorized and noncategorized lists is determined on the basis of word associations like those uncovered in free association tests. Associations of a categorial nature are not viewed as having special status. This approach is quite similar to FRAN's present associative strategy. According to this view, phenomena like clustering are to be explained by the fact that words from the same category tend to be highly interassociated. Support for this explanation comes from findings that recall and clustering are higher in lists composed of high-frequency associates of the category name (e.g., Cofer *et al.*, 1966). An opposing result was reported by Marshall (1967) who found that both recall and clustering were higher for pairs of words that were from the same category than for pairs of words that had equal associative strength but were not coordinates.

Both Cofer (1965) and Tulving (1968) have criticized attempts to distinguish experimentally between the categorization and the associative hypotheses. The general point of their argument is that these are but two of several possible ways to organize word lists, and that one should not expect a person always to use just one and the same basis for organization. Although the claim that *S*s use many strategies in free recall is undoubtedly correct, it leaves unresolved the empirical question of how much humans tend to use a categorization strategy as opposed to an associative strategy. One way to decide the issue is to compare FRAN's performance on categorized lists to that of humans. FRAN uses a pure associative strategy. So far, we have presented evidence largely consistent with the hypothesis that FRAN correctly models the main aspects of the associative strategy which *S*s tend to adopt under standard free recall. If her performance on categorized lists deviates markedly from that of humans, we may have evidence that categorization strategies are quite prominent in such situations. We would also discover how characteristics of recall differ between the two strategies.

After searching the extensive literature on this topic, we decided to simulate an experiment by Dallett (1964) which manipulated two variables of interest in categorized recall, and which reported measures of overall recall as well as organization. Subjects in his Experiment IV studied lists of 24 words presented at a 1.4-second rate. The lists were composed either of two members from each of 12 categories, three members from eight, four members from six, or six members from four. Another variable was whether all the members of a category were presented contiguously (blocked presentation) or presented randomly scattered throughout the list (random presentation). There are eight combinations of the two methods of presentation and the four numbers of categories, and data were collected from 20 *S*s in each condition. It is not possible to have FRAN study at time intervals like Dallett's 1.4 seconds that are not in discrete seconds. As a compromise, data were obtained for six simulations at a one second rate and six simulations at a two second rate. The parameter *a* was set at .5 for these simulations.

Figure 20a shows that the recall of FRAN and of the humans closely matches when instances of categories were presented in random order. Both show an increase in words recalled as the number of categories decreases. FRAN improves with fewer categories because she has to develop access to fewer "regions" of memory. Unfortunately, this result of higher recall with fewer categories is by no means universal. Bousfield and Cohen (1956), Cohen and Bousfield (1956), Mathews (1954), Tulving and Pearlstone (1966), and Dallett (1964, Experiment I) report divergent outcomes of this comparison. The observed effect of number of categories appears to depend on list length and also on the *S*s sophistication with free recall.

The control data for FRAN represent recall of 24 randomly selected words. Ten simulations were run at both the one- and the

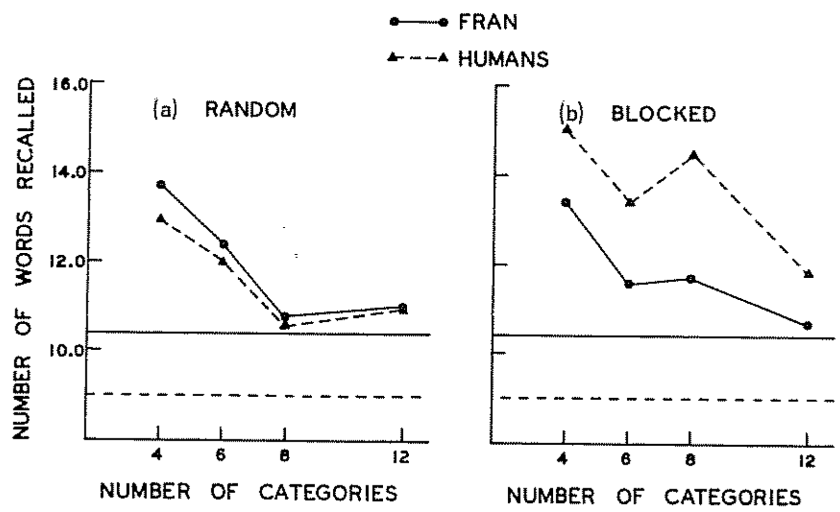

Fig. 20. FRAN and Dallett's *S*s compared with respect to number of words recalled out of 24 as a function of number of categories and mode of presentation (blocked *vs.* random).

two-second input rates. Categorized recall was higher than the control recall for FRAN at all numbers of categories, a consistent result in the literature. The control recall reported by Dallett is not comparable because it was based on recall of lists composed of nouns, verbs, adjectives, and adverbs. With such control lists, it is none too surprising that his control recall was lower than FRAN's. Verbs, adjectives, and adverbs are generally more abstract than nouns, and abstract words are recalled less well.

Figure 20b illustrates a comparison between the data and the predictions when presentation was blocked. Although humans and FRAN both improve in recall with fewer categories, the humans are considerably superior to FRAN. The recall of the humans in the blocked condition was one to four words more than in the random condition, a result that has been replicated by Cofer, Bruce, and Reicher (1966). In contrast, FRAN recalled a mean of 11.94 words in the random condition, averaged across all numbers of categories, and a mean of 11.86 in the blocked condition.

In the blocked condition FRAN has a better chance than in the random condition of marking associations between words within a category. If FRAN finds an associative path between a word under study and a previous list word from the same category, in the blocked condition she will probably recognize the discovered word because it is likely to still be in STS. The word is still likely to be in STS in the blocked condition because it would have appeared only a few items earlier. If FRAN does recognize it, it will be re-entered into STS, and she will attempt to tag the paths to and from the word and attempt to tag the word itself. On the other hand, in the random condition, there is considerable chance that the discovered coordinate word will not be recognized. Failure of recognition is possible because the word is not likely to be in STS at the time and it may not have been tagged when originally studied. This increased probability of tagging paths between contiguous coordinate pairs should somewhat facilitate recall in the blocked condition.

This fact, that associative paths are more likely to be tagged between contiguous words, was also used in Section IV,B to explain why the input order was maintained to some degree in the output order. Glanzer (1969) has shown that related word pairs are more likely to be recalled if they are presented close together – a result which is to be predicted by this same factor within FRAN. In the simulations for the random condition, we examined recall of pairs of words from the same category which had occurred in positions 4 through 19 in the input sequence. The examination was restricted to these middle 16 positions to avoid any complications due to primacy and recency effects. The relevant data concern recall of the second member of a pair conditional upon the recall of the first member of the pair. When the lag between the two words was between zero and four words, the conditional probability of recall of the second member was 54%; for lags between 5 and 14, it was 49%. This result, like Glanzer's, is to be explained in terms of the fact that useable associative paths between two list words are more likely to be tagged the nearer the items appear to one another. However, the difference in our data is not very large, suggesting that extrinsic contiguity conferred relatively little advantage to the formation of associations.

A second, opposing factor apparently cancelled out this small contiguity advantage of the blocked condition. The disadvantage in the blocked condition is that the members of STS at recall are likely to come from fewer categories than in the random condition because the last words FRAN studies in the blocked condition are all from the same category. Therefore, FRAN will have access to fewer categorical clusters at recall through the contents of STS.

Why, then, do human *Ss* recall more with blocked than with random input? We suspect it is because under the blocked condition humans are likely to adopt the categorization strategy outlined

previously. This strategy is effective with categorized lists, particularly where they are blocked. That humans were using this strategy can be inferred from Fig. 21 which reports clustering measures for FRAN and the human *Ss*. The measure used in this figure is the same as that Dallett reported, viz., deviation from expected clustering. This measure is calculated by counting the number of times one member follows another member of the same category in the output and substracting from that sum the expected number of such repetitions. The expected number is given by the formula: $(\Sigma m_i^2/\Sigma m_i)-1$, where m_i is the number of items recalled from the ith category. For both the humans and FRAN, clustering

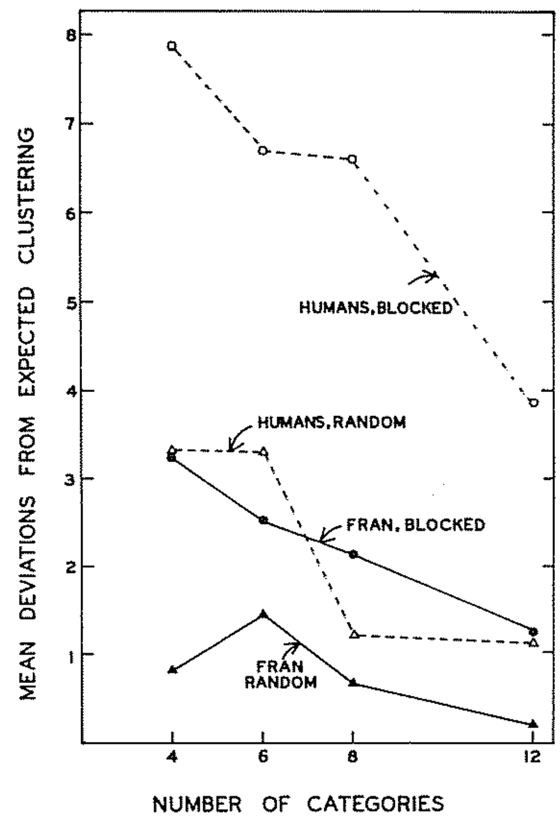

Fig. 21. FRAN and Dallett's *Ss* compared with respect to amount of category clustering. See text for explanation of clustering metric, mean deviations from expected clustering.

exceeds chance and the deviations from expectation increase as the number of categories decrease. Also, both humans and FRAN show much greater clustering with blocked presentation of the list. However, under all conditions, the humans show much more clustering than does FRAN, particularly in the blocked condition. Referring to the level of recall in Fig. 20b, it may be confirmed that the level of clustering shown by the humans under the blocked conditions was practically perfect. Such perfect clustering scores are to be expected, of course, if the human *S*s were using the categorization strategy as outlined. From their less than optimal clustering scores, *S*s in the random condition apparently were using some combination of the associative and the categorization strategy; or alternatively, some *S*s were using an associative strategy like FRAN and others, the categorization strategy.

F. Word Recognition

Embedded within FRAN is a simple model of word recognition. In a word recognition test, there are three possible bases for FRAN to identify an item as a list member – membership in STS, membership in ENTRYSET, or an association between the item and the list marker. FRAN would be able to recognize any word meeting any of these three criteria. For any other word from the list and for all the distractors, FRAN would have to resort to some pure guessing strategy to determine which items (if any) to call list members. Essentially, then, FRAN can classify words into only two categories, those she remembers as coming from the list and those she cannot remember being in the list. However, humans can discriminate words into more than two categories as is shown by their ability to give ratings of their confidence that a test word was in the list. These confidence ratings bear a monotonic relationship to the probability that the word came from the studied list. As currently programmed, FRAN does not yield such multiconfidence ratings.

However, a simple and natural elaboration of the model will handle such confidence judgments in recognition. The requisite assumption presupposes an imperfect decision process that interprets associations to the list markers. In a forthcoming paper this recognition model will be described in detail. Essentially, it is a two-state recognition model like that of Bernbach (1967). In the terminology of signal-detectability theory, it is assumed that there is one normal distribution of values on the decision axis for tagged words and a different normal distribution for untagged words.

According to this model, a *S* will recognize a word whenever the word exceeds an adjustable criterion value on the decision axis. Therefore, it will occasionally happen that tagged words will not be recognized and untagged words will be recognized. With this brief elaboration of the recognition model in mind, we can now discuss some of the data from word recognition experiments.

The primacy effect in FRAN's free recall has been attributed to the fact that ENTRYSET in initially composed of the first words in the list, and the recency effect to the fact that, at the initiation of recall, STS is composed of the last words in the list and their associates. If recognition judgments take account of the contents of STS and ENTRYSET as we postulate, one would expect to find primacy and recency effects in recognition judgments. The data of Waugh and Norman (1968) confirm that this is the case. They also found that the recency effect is much more pronounced for those items that are presented for recognition judgment early in the test phase, a result consistent with FRAN's recognition component. If an item late in the study sequence is tested early, there is a high probability that it will still be in STS when tested and therefore recognized. As more items are tested before the late-studied word, the probability increases that the word will no longer be in STS.

Another interesting phenomena is the occurrence of intrusions in the free-recall output. In searching her memory, FRAN must decide whether each word she encounters is a list word or not. If the decision model described above were governing these judgments, FRAN would occasionally incorrectly identify as a list member one of the nonlist words she encountered in her search. FRAN, in her search at the time of recall, only examines words that are on marked associative pathways. Therefore, the only nonlist words considered are on associative pathways that have been marked out between list words. As a consequence, one would expect most intrusions to be strongly related to some of the words in the list. Deese (1959b) demonstrated that this is the case; he found a strong positive correlation between the frequency with which the words in the list evoked a particular word in free association and the frequency with which that word occurred as an intrusion in free recall.

While many variables such as presentation rate and serial position appear to have the same effect on recognition as on recall, a few variables affect the two testing methods differently. These differences are easily interpreted in terms of the processes underlying FRAN's performance. For instance, it is to be predicted that any manipulation that increases the organization of the list should facilitate the retrieval processes but not item recognition. Therefore, such manipulations should improve recall but leave word recognition unaffected. Cofer (1967) and Kintsch (1968) report data confirming this prediction.

Also, *S*s should give better recall under intentional than incidental instructions because intentional *S*s are likely to actively search out relationships and retrieval pathways amongst the list words. On the other hand, recognition performance depends only upon tagging of the word which, in turn, depends only upon the entry of the word into STS. Presumably, any incidental task that directs the *S*'s attention to the word will insure that the word enters STS. Therefore, intentional learning instructions should not have any superiority over incidental if the method of test is recognition but will be superior on recall tests. Eagle and Leiter (1964) found that recall is better under intentional instructions, but recognition is actually better under incidental instructions. Both findings have been confirmed in other experiments (Dornbush & Winnich, 1967; Postman, Adams, & Phillips, 1955). Consistent with our interpretation of the facilitating effect of intentional instructions, Eagle and Leiter found that only those intentional *S*s who reported using some recall strategy were superior to the incidental *S*s. They also offered a plausible explanation of the unexpected superiority of the incidental group in recognition. Eagle and Leiter suggest that, under intentional instructions, *S*s focus on particular words at the expense of others; in contrast, under incidental learning instructions, *S*s distribute their time equally among the words. Focusing of attention may be useful in developing a successful system for retrieval, but the optimal strategy for recognition is to distribute an equal amount of time to each word. That recognition is optimal when *S* distributes his study time equally follows from FRAN's assumption that probability of tagging is exponentially related to the duration of time for which the word is studied.

V. Concluding Remarks

Throughout Section IV, we confined ourselves to the tasks of comparing FRAN's performance with that of humans, of explaining FRAN's performance, and of conjecturing why humans deviated from FRAN's behavior when they did. The time is now come to evaluate the theoretical positions developed in Sections II and III in light of the results of Section IV. To summarize, three basic claims

were made about free recall: (*a*) Most humans in conditions of standard free recall adopt the associative strategy of FRAN as described in Section III; (*b*) procedural deviations from the standard free recall paradigm may induce *S*s to adopt strategies quite different from the associative; and (*c*) these different strategies could be implemented with the same mental mechanisms as outlined in Section II. The question now is, to what extent have these three claims been substantiated?

As for the claim that FRAN's associative strategy is the rule for standard free recall, the data presented in Sections IV,B,C, and F, as well as the introspective evidence referenced throughout this paper, provide positive support. The data in Section IV,D, the failure of humans to show substantial improvement when cued with list words, provide the only major source of embarrassment for this hypothesis. As noted, it would be easy to introduce minor alterations to FRAN's behavior that would reduce the difference she now shows between the cued and noncued conditions. However, these changes would be posthoc, and would only mitigate the embarrassment of the theory. Nonetheless, it would be mistaken to overemphasize the importance of this negative result. Both we and Allen have shown that cueing with list words has beneficial effects. Beyond this, the data of Tulving and Pearlstone (1966) and of Tulving and Osler (1968) showed that cueing with extralist words could have very beneficial effects. Therefore, Slamecka's original thesis about the complete independence of memory traces is quite probably wrong. The difficulty only concerns why the memory traces are not as dependent as our associative model would seem to predict.

If it is accepted that FRAN approximately models the associative strategy for standard free recall, it is fairly clear from Section IV,E that the associative strategy is not always the major strategy in all nonstandard free recall experiments. Dallett's data indicated that adoption of the alternative, categorization strategy increased as the categorized nature of the list was made explicit. Therefore, this is a stronger result than just that a different strategy was adopted when the paradigm was changed from standard free recall; rather, the nature of the strategy change was predictably related to the nature of the change in experimental parameters.

We have outlined how several different strategies could be implemented in terms of FRAN's mental mechanisms. The fact that these various strategies can be formulated in terms of FRAN's machinery supports the assertion that FRAN models the structures and processes underlying human memory in general. However, this fact alone is hardly sufficient. To mention a counter example, all these strategies can be implemented on any standard computer. This would hardly lead one to conclude that the structures and processes which underlie the computer have much relation to those in the human head. The reason why we can reject the raw computer as a model of human memory, independently of physiological considerations, is that it is possible to have the computer perform feats of memory of which the human is just not capable. The point is that anything which FRAN can do, humans must be able to do also. To substantiate our third theoretical contention will require stronger results than given in this paper. We must present data that would adequately substantiate that FRAN can simulate practically all memory strategies that people use; we must also show that we can instruct humans to adopt any strategy of which FRAN is capable.

We must address ourselves to one remaining question in this evaluation; viz., to what extent do the simulation results depend upon the particular parameters? Might we have produced simulations showing little benefit of cueing? Could we have obtained a difference between blocked and random presentation of lists? It is, of course, completely unfeasible to do the sort of parameter search that would determine an answer to these questions with certainty. Given that this is impractical, one must rest content with the type of informal answers provided in Section IV. Sometimes, as with the too-large recency effect, we pointed out it seemed clear that a change of parameters would eliminate this discrepancy in the simulation. In other circumstances, such as with the too-large cueing effect, the result seemed inevitable given the associative strategy, and there was no apparent way to avoid it by juggling the parameters.

Clearly, while some of the results are encouraging, FRAN is not a complete model of free recall. The model is not adequate even for all the details of standard free recall. Before FRAN would constitute an adequate "explanation" of free recall, we would have to augment the program so that she could adequately replicate all relevant relations found in all free recall experiments. This is clearly a tall order. Rather than attempting to achieve such a complete model, it will be more judicious to set our sights lower and seek a closer approximation of FRAN to the data.

There are a few theoretical changes that would transform FRAN into a more adequate model. One obvious improvement would be to give FRAN a larger memory. There is no technical reason why it could not be about 2000 concepts. Such a memory would at least be approaching the order of magnitude of a human's, and FRAN would

be faced with some of the problems a human has when he must search for lost needles in a mnemonic haystack. Since FRAN's current memory is at least a hundred times less than that of a human, it is impossible to assign much psychological reality to some of her processes such as those selected to govern her search for and marking of associative paths. With a memory that approached human size, it would be possible to explore interesting questions about the exact manner in which the memory is searched. For instance, we might try to determine the rate at which associations are searched from a particular word. It will be remembered we used the crude formula $N = 5 + T$ to relate number of associations (N) to total study time (T). With a larger memory it would also be possible to have FRAN occasionally resort to a random search of her memory, as humans probably do, without being in danger of quickly exhausting the memory.

The other important direction in which FRAN should be improved is to permit her to adopt a variety of strategies and to give her some heuristic principles by which to select a particular strategy for a particular free recall task. This is much more easily said than done. Essentially, a meta-program is needed that is capable of writing a set of different programs, each program reflecting a different strategy. The program that we have developed for execution of the associative strategy would be just one of many strategy-implementing programs. Although there is some research in computer science on programs to write programs, nothing relevant to our complexities is at hand. Although the task of programming such a meta-program is beyond our current abilities and ideas, any fully adequate simulation program for human memory will have to take the form of such a meta-program.

Appendix: An Introspective Report

The following is a transcription of the verbal introspective remarks offered by one student as she studied and recalled a set of 40 words presented at an eight-second rate. She had twice studied and recalled a previous list of 40 words. She had also studied and recalled this current list a previous time. In the first trial she had recalled in the following order the words tattoo, lieutenant, mercenary, destroyer, sideburns, skeleton, pumpkin, dignitary, city, pond, chestnut, mountain, kitchen, widow, wrist, student, and present.

Study. The word following each number indicates the word being studied.

Following it is a transcription of the *S*'s remarks. Each italicized word in the transcription is a word that was in the set of 40 to-be-recalled words:

1. garrison – *garrison, lieutenant, dignitary*
2. dignitary – *crown* queen, oh . . . *dignitary*
3. vulture – *vulture* . . . bird, there was a bird *present* . . . *vulture*, bird . . . *garrison*
4. disk – *disk* . . . *disk*, record, *disk*, can't remember statue
5. crown – *crown* queen, the *dignitary* visits the *crown* queen, the *lieutenant* is in the *garrison*
6. bowl – *bowl* . . . *bowl* of flowers . . . the *dignitary* visits the *crown* queen and gives her a *bowl* of flowers
7. present – I am *present*, I also am a *student*, I think . . . *student*
8. student – *student*, I am *present*, I also am a *student* . . . the *dignitary* is also a *student* philosopher
9. dragon – oh, I forgot all the fairy tales . . . *goose, dragon*, mother *goose* fairy tales
10. kitchen – *kitchen*, still the mother, the *widowed* mother
11. clergyman – *clergyman* . . . the *clergyman* visits the *widowed* mother who is in the *kitchen*
12. airport – *airport* . . . the *dignitary* arrives in a airplane in the *airport* to visit the *crown* queen with a *bowl* of flowers
13. lieutenant – *lieutenant* is in the *garrison* . . . and he is being attacked by a *vulture* who came through the window
14. pepper– *pepper* . . . the mother put the *pepper* on the food
15. chestnut – still the open fire, *chestnuts, pond*, next word is *pond*
16 pond – right . . . remember *garrison* and remember *dignitary*
17. wrist – *wrist*, still watch, *wrist* watch, *wrist* action, *wrist* watch
18. mountain – oh, instead of valley remember the *mountain* that created the valley
19. flashlight – you remembered *city* but you didn't remember *flashlight* and the reason you should remember *city* is because you have to use a *flashlight*
20. city – *city, flashlight* . . . *kitchen, dignitary, city, flashlight*
21. scorpion – *scorpion*, remember *vulture* with *scorpion*, the *garrison* is loaded with kooky animals
22. congregation – *congregation*, I am trying to remember a *congregation* of things, *congregation*
23. goose – *goose*, remember *dragon*, remember mother *goose* . . . a fairy tale . . . or you could remember happy times, *goose, dragon, pond, chestnuts*
24. cone – and an ice-cream *cone, goose, dragon, pond, chestnuts*, and an ice-cream *cone*
25. tub – tub . . . I still only remember bath and I'll never remember that and when you see *tub*, I remember rub-a-dub-*tub*, tune in *tub*, ah yes
26. pliers – *pliers*, they never will be remembered by me . . . oh, bicycle, to fix a bicycle you have to use some *pliers*
27. beard – *beard*, the *dignitary* had a *beard* and the *lieutenant* had the *sideburns* . . . Jesus
28. mercenary – the *lieutenant* was the *mercenary*, right

29. frontier – he is fighting on the *frontier*, let me see now . . . somewhere in Arabia . . . the *lieutenant* is in the *garrison* because he didn't fight
30. castle – *castle*, fairy tale . . . don't remember the sentry, remember the *castle* he guards
31. officer – the *lieutenant* is a *officer* in the . . . oh . . . he didn't obey the duties
32. destroyer – the *lieutenant* is an *officer, destroyer, mercenary* . . . the *lieutenant* is too much . . . he's a *destroyer*
33. jackknife – *jackknife*, in the summer you can use the *pond* to swim in and one of the things you can do is *jackknife* into it
34. widow – the mother is a *widow* in the *kitchen*, the *widowed* mother is in the *kitchen*
35. gambler – oh, another *city* you could think of is Las Vegas and Monte Carlo as opposed to New York
36. pumpkin – *pumpkin*, remember with *skeleton*, Hallowe'en . . . fun, *chestnuts*
37. sideburns – *sideburns*, the *lieutenant* has *sideburns*, the *dignitary* has a *beard*
38. skeleton – oops, oh, the *pumpkin* and the *skeleton*, right, *skeleton*
39. film – what I am looking at could be said to be a *film* or a *congregation* of objects
40. tattoo – *tattoo* . . . oh, 2 t's, 2 o's, spelled it wrong last time.

Recall. This is a complete transcription of the *S*'s remarks during the three minutes she had for recall. Italicized are those words that were written on the recall sheet.

Seeing as how I spelled it wrong last time – *t, a, t, t, o, o.* What I am looking at could be said to be a *film* of a *congregation* of objects. The *lieutenant* . . . lieu-ten-ant . . . is a *mercenary* with *sideburns* . . . *destroyer* . . . *officer* . . . who's in the *garrison* . . . and is being attacked by *vultures* and *scorpions* . . . and a *gambler's city* . . . called Las Vegas, Monte Carlo as opposed to a city where you have to use a *flashlight* like New York when it had the blackout. Let me see now . . . the mother, oops not mother but *kitchen* . . . kitchen, mother's in the kitchen using *pepper* . . . on the phone and she was visited by the *clergyman* . . . seeing as how she was recently *widowed* . . . she has many children who like to roast *chestnuts* . . . on the fire and tell fairy tales about *dragons* . . . mother *goose* fairy tales, that is, and ladies in *castles* being rescued . . . let me see now, afterward they go skating on a *pond*, that's during the winter, however they can use the *jackknife* during the summer . . . and they live in a valley between two *mountains* . . . and one time they were making the jackknife and they broke their *wrist*, one kid broke his wrist, right . . . and . . . let me see now, the lieutenant was a mercenary who had sideburns, the *dignitary* . . . was the guy who had the *beard* and he carried a *bowl* of flowers to the queen who was already in the castle . . . and . . . well, the lieutenant could be said to have a *skeleton* in his closet that he is reserving for Hallowe'en when he can use it with his *pumpkin* . . . oh, yes . . . let me see, what else can I remember about those nice people . . . the lieutenant could be said to have a skeleton in his closet . . . which he could use for Hallowe'en . . . let me see now, scorpion, gambler, city, flashlight . . . clergyman, widow, chestnut . . . goose, castle, pond, jackknife, wrist, dignitary, I am *present* and I am a *student* . . . oh, there is a word that I am not going to remember . . . because I remember it but I can't remember it, let me think . . . I am present and I am a student . . . *tub*, I remembered it . . . no, that's not the right one . . . tub, bath, pond . . . chestnuts roasting in an open fire . . . I'm not too sure I am not remembering the other 40 words.

References

Allen, M. Cueing and retrieval in free recall. *Journal of Experimental Psychology*, 1969, 81, 29–35.

Battig, W. F., Allen, M., & Jensen, A. R. Priority of the free recall of newly learned items. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 175–179.

Battig, W. F., & Slaybaugh, G. D. Evidence that priority of free recall of newly learned items is not a recency artifact. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 536–538.

Bernbach, H. A. Decision processes in memory. *Psychological Review*, 1967, 74, 462–480.

Bever, T. G., Fodor, J. A., & Garrett, M. A formal limitation of associationism. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Bierwisch, M. Some semantic universals of German adjectivals. *Foundations of Language*, 1967, 3, 1–36.

Bolinger, D. The atomization of meaning. *Language*, 1965, 41, 555–573.

Bousfield, W. A. The occurrence of clustering in recall of randomly arranged associates. *Journal of General Psychology*, 1953, 49, 229–240.

Bousfield, W. A., & Cohen, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *Journal of Psychology*, 1953, 36, 67–81.

Bousfield, W. A., & Cohen, B. H. Clustering in recall as a function of the number of word-categories in stimulus-word lists. *Journal of General Psychology*, 1956, 54, 95–106.

Bousfield, W. A., Puff, C. R., & Cohen, B. H. The development of constancies in sequential organization during repeated free recall. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 489–495.

Bower, G. H. Analysis of a mnemonic device. *American Scientist*, 1970, 58, 496–510.

Cofer, C. N. On some factors in the organizational characteristics of free recall. *American Psychologist*, 1965, 20, 261–272.

Cofer, C. N. Does conceptual organization influence the amount retained in immediate free recall. In B. Kleinmutz (Ed.), *Concepts and the structure of memory*. New York: Wiley, 1967.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology*, 1966, 71, 858–866.

Cohen, B. H., & Bousfield, W. A. The effects of a dual-level stimulus-word list on the occurrence of clustering in recall. *Journal of General Psychology*, 1956, 55, 51–58.

Dallett, K. M. Number of categories and category information in free recall. *Journal of Experimental Psychology*, 1964, 68, 1–12.

Deese, J. Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, 1959, 5, 305–312. (a)

Deese, J. On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 1959, 58, 17–22. (b)

Deese, J. Frequency of usage and number of words in free recall: The role of association. *Psychological Reports*, 1960, 7, 337–344.

Deese, J. *The structure of associations in language and thought.* Baltimore: Johns Hopkins Press, 1966.

Deese, J. Association and memory. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Dornbush, R. L., & Winnich, W. A. Short-term intentional and incidental learning. *Journal of Experimental Psychology*, 1967, 73, 608–611.

Eagle, M., & Leiter, E. Recall and recognition in intentional and incidental learning. *Journal of Experimental Psychology*, 1964, 68, 58–63.

Glanzer, M. Distance between related words in free recall: Trace of the STS. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 105–111.

Glanzer, M., & Cunitz, A. R. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 351–360.

Jenkins, J. J., and Russell, W. A. Associative clustering during recall. *Journal of Abnormal and Social Psychology*, 1952, 47, 818–821.

Katz, J. J., & Fodor, J. A. The structure of a semantic theory. *Language*, 1963, 39, 170–210.

Kintsch, W. Recognition and recall of organized lists. *Journal of Experimental Psychology*, 1968, 78, 481–487.

Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*, New York: Academic Press, 1970.

Kiss, G. R. A test of the word selection model using multiple stimuli in word association. Paper presented at the conference of the British Psychological Society, London, 1967.

Lyons, J. *Introduction to theoretical linguistics.* London and New York: Cambridge University Press, 1969.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* Vol. 1. New York: Academic Press, 1967. Pp. 328–372.

Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 126–131.

Marshall, G. R. Effect of total association and conceptual cohesiveness among words on recall, clustering, and recognition association. *Psychological Reports*, 1967, 20, 29–44.

Mathews, R. Recall as a function of number of classificatory categories. *Journal of Experimental Psychology*, 1954, 47, 241–247.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 596–606.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review*, 1956, 63, 81–97.

Murdock, B. B., Jr. The immediate retention of unrelated words. *Journal of Experimental Psychology*, 1960, 60, 222–234.

Murdock, B. B., Jr. The serial position effect in free recall. *Journal of Experimental Psychology*, 1962, 64, 482-488.
Norman, D. A. Toward a theory of memory and attention. *Psychological Review*, 1968, 75, 522-536.
Pollio, H. R. Associative structure and verbal behavior. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
Postman, L., Adams, P. A., & Phillips, L. W. Studies in incidental learning: II. The effects of association value and method of testing. *Journal of Experimental Psychology*, 1955, 49, 1-10.
Postman, L., & Keppel, G. Conditions determining the priority of new items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 260-263.
Postman, L., & Phillips, L. W. Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, 1965, 17, 132-138.
Puff, C. R. Role of clustering in free recall. *Journal of Experimental Psychology*, 1970, 86, 384-386.
Quillian, M. R. *The teachable language comprehender: A simulation program and theory of language*. Project No. 8688. Cambridge, Mass.: Bolt, Beranek & Newman, 1969.
Rosner, S. R. The effects of presentation and recall trials on organization in multitrial free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 69-74.
Schank, R. "Semantics" in conceptual analysis. Stanford A. I. Memo No. 122, May 1970, Computer Science Dept., Stanford University.
Shuell, T. J. Clustering and organization in free recall. *Psychological Bulletin*, 1969, 72, 353-374.
Shuell, T. J., & Keppel, G. Item priority in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 969-971.
Slamecka, N. J. An examination of trace storage in free recall. *Journal of Experimental Psychology*, 1968, 76, 504-513.
Slamecka, N. J. Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology*, 1969, 81, 557-560.
Tulving, E. Subjective organization in free recall of "unrelated" words. *Psychological Review*, 1962, 69, 344-354.
Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
Tulving, E., & Osler, S. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 1968, 77, 593-601.
Tulving, E., & Pearlstone, Z. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 381-391.
Underwood, B. J. The representativeness of rote learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 47-78.
Waugh, N. C. Immediate memory as a function of repetition. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 107-112.
Waugh, N. C. Presentation time and free recall. *Journal of Experimental Psychology*, 1967, 73, 39-44.
Waugh, N. C., & Norman, D. A. Stimulus and response interference in recognition-memory experiments. *Journal of Experimental Psychology*, 1968, 78, 551-559.

Author Index

Numbers in italics refer to the pages on which the complete references are listed.

A

B

C