

Surface form and the spacing effect

DENISE DELLAROSA and LYLE E. BOURNE, JR.
University of Colorado, Boulder, Colorado

Deficient processing theories of the spacing effect attribute poor recall of massed-repeated items to a failure to process one or both of the presentations fully. An implication of this approach is that anything that increases the probability that a repetition will receive full processing, or conversely, anything that decreases the probability that the item will be recognized as a repetition, should improve memorability of the item. The present set of experiments tested this prediction by manipulating the surface structure of repeated sentences. On the basis of previous research, it was assumed that memory for surface structure of sentences decays rapidly, and hence can contribute to initial identification of repetitions only at short spacings. Because this manipulation should hinder recognition of repetitions as repetitions, it was expected to induce full processing of massed repetitions, and thus facilitate recall of these items. This prediction was supported. When sentences were repeated verbatim (Experiment 1) or by the same speaker (Experiment 2), the typical spacing effect was obtained. However, when the surface structure or speaker changed at time of repetition, massed repetitions were recalled nearly as well (Experiment 1) or as well (Experiment 2) as their spaced counterparts.

Although the spacing effect is one of the best known and most researched memory phenomena, it has yet to be satisfactorily explained. Recent explanatory attempts have fallen into two classes. The first class includes those explanations that attribute the spacing effect to increasing independence of encoding events with increasing intervals between repetitions. The best known and most widely studied of these is the encoding-variability hypothesis, which attributes higher recall of spaced repetitions to a greater likelihood that a repetition will be encoded in a different subjective context at longer intervals than at shorter ones. The greater the number of retrieval routes to, or encoding contexts for, a given item, the greater the probability that the item will be retrieved. However plausible, this class of theories has not fared well empirically. Ross and Landauer (1978), for example, pointed out that theories in this class also predict that spacing should improve the probability of remembering at least one of two different items that were each studied once, a prediction not upheld by their data. Postman and Knecht (1983) tested the encoding-variability hypothesis by systematically increasing the number of explicit contexts in which a to-be-remembered item was embedded. They found recall levels to be lower after variable than after constant encoding, leading them to conclude that the availability of multiple retrieval routes was not a sufficient condition for improved recall.

The second class of spacing-effect theories includes those that appeal to deficiency or attenuation of processing as a consequence of massing items. These theories

postulate that massed repetitions receive less processing than their spaced counterparts, and that recall is a function of the amount of processing an item receives. Several mechanisms are purported to be responsible for variations in processing. The robustness of the phenomenon has prompted some theorists to implicate hard-wired, neurological mechanisms, such as consolidation and habituation, as its basis (Cornell, 1980; Hintzman, 1974). Consolidation theory holds that consolidation of the second massed presentation interferes with consolidation of the first presentation, whereas habituation theory posits that encoding of the first presentation temporarily habituates encoding processes until a sufficient recovery period transpires. Although attractive, these theories have received equivocal empirical support at best (see Hintzman, 1974, for a review). Moreover, recent evidence indicates that the spacing effect in recall memory is correlated with development, thereby introducing the awkward constraint that such neurological mechanisms are not present at birth, but rather emerge with maturation (Toppino & DiGeorge, 1984).

Other deficient-processing accounts hold that subjects adopt a voluntary strategy of not attending to, or processing fully, massed repetitions (Rundus, 1971; Shaughnessy, Zimmerman, & Underwood, 1972; Underwood, 1970; Waugh, 1970). However, manipulations that should induce subjects to attend to massed repetitions have failed to attenuate the spacing effect (e.g., Hintzman, Summers, Eki, & Moore, 1975).

Recently, researchers have suggested that accessibility of previous encodings may underlie the spacing effect (Cuddy & Jacoby, 1982; Jacoby, 1978; Rose, 1980; Rose & Rowe, 1976). The idea is that, when an item is repeated, an attempt is made to retrieve the previous encoding of that item. The accessibility of the item will vary directly with spacing between repetitions, and the less ac-

This article is Publication No. 140 from the Institute of Cognitive Science, University of Colorado. The research was supported by a grant from the Martin Marietta Corporation. Requests for reprints should be sent to D. Dellarosa, Department of Psychology, Campus Box 345, University of Colorado, Boulder, CO 80309.

cessible an item's encoding is, the greater the likelihood that it will enjoy reinstatement of full encoding processes.

Direct support for this account has been obtained from psychophysiological response data taken during processing of items repeated at varying intervals (Magliero, 1983; Silverstein, 1977). For example, Magliero found that processing of items repeated at long intervals produced greater pupil dilation than processing of items repeated at short intervals, pupil dilation being an indicator of increased processing effort or memory load (Kahneman & Beatty, 1966).

An implication of the accessibility hypothesis is that anything that increases the probability of full encoding processes when an item is repeated should also improve recall. Spacing of items is one way to increase this probability, since forgetting or fading of memory traces is greater over longer intervals. Jacoby (1978) and Cuddy and Jacoby (1982) manipulated other factors, such as ease of encoding (via a problem-solving procedure), similarity of intervening material, and cue effectiveness during learning, and found conditions less conducive to retrieving prior encodings of items to produce higher subsequent recall of those items. Moreover, the spacing effect was significantly attenuated under these conditions (Cuddy & Jacoby, 1982, Experiments 2 & 3). Thus, when full processing of repeated items was induced by rendering original encodings more difficult to retrieve, massed repetitions were recalled nearly as well as spaced repetitions.

Although results like these are consistent with the accessibility account of spacing effects, they are not without criticism. Because subjects in these studies were required to perform extraordinary processing of repetitions (e.g., problem solving), the relevance of these results to spacing effects found in typical list-learning studies has been questioned (Glenberg & Smith, 1981). What is required is an attenuation of spacing effects under conditions that strongly resemble typical list-learning conditions, but that induce full processing of massed repetitions. This is no small task, as evidenced in the work of Shaughnessy et al. (1972), which suggests that manipulating subjects' intention to encode is not sufficient to ensure full processing of critical items.

The present study addresses this issue by capitalizing on the numerous representational levels attributed to complex stimuli, such as sentences. Like recall of other linguistic units, recall of sentences shows the typical spacing effect; that is, repeated sentences tend to be better recalled if the two occurrences are separated by other sentences than if they are not (Rothkopf & Coke, 1963; Underwood, 1970). Unlike other units, however, sentences contain information that is generally believed to be of at least two types, surface information (e.g., wording, modality, etc.) and semantic information (i.e., the sentence's meaning). Memory for one of these two types of information has been shown to be independent of memory for the other, each type of information requiring a separate memory representation (Begg, 1971; Kintsch, 1975) and suffering a different rate of decay (Anderson & Paulson,

1977; Garrod & Trabasso, 1973; Jarvella, 1971, 1973; Sachs, 1967; Wright, 1969). Memory for surface information generally tends to be more volatile than memory for meaning. For example, Jarvella (1971) required subjects to repeat clauses of sentences that were read to them, and found that the intervention of a single clause between presentation and recall was sufficient to significantly decrease memory for surface information. Similar results have been obtained in sentence-verification tasks, in which savings in response latencies due to verbatim repetition of sentences disappear with the intervention of a single sentence between repetitions (Anderson & Paulson, 1977; Garrod & Trabasso, 1973).

Because the evidence suggests that only an abstract gist representation is likely to be present after the intervention of a single sentence, one would expect little difference in the memorability of sentences repeated verbatim as opposed to those repeated with changes in surface structure after a lag of one intervening item. Moreover, these spaced items should exhibit high retention levels, as they are more likely to enjoy a reinstatement of full encoding processes, having suffered a delay in repetition. Massed repetitions, however, present a very different picture. Since an item's surface structure is highly likely to be present in memory when a massed repetition of that item occurs, it should contribute to the recognition of that item as a repetition, and hence facilitate retrieval of the item's prior encoding. Changing the item's surface structure, however, should hinder recognition of the item as a repetition. As a result, full encoding processes should be initiated on the item. The result of that processing, however, should yield the same gist representation. And because only this type of representation is likely to be present in memory at time of recall, the additional processing that such items receive should benefit recall.

The purpose of the following experiments was to test this prediction. Subjects were presented with lists of sentences in which certain items were repeated. Repetitions occurred at different lags, and the surface structures of the sentences were either repeated exactly or changed while leaving meanings intact. In Experiment 1, the wording of visually presented sentences was changed during some of the repetitions; in Experiment 2, the speaker of aurally presented sentences sometimes changed. Both of the manipulations were expected to induce full processing of the repeated items, even at massed presentations, because such changes were expected to hinder the recognition of repetitions as repetitions. However, the final product of comprehension processes was expected to yield equivalent representations of the repeated items, since the meanings of the items did not change. As a result, sentences with changes in surface structures at time of repetition were expected to be given full encoding processes regardless of the lag at which they were repeated; sentences whose surface structures remained unchanged at time of repetition, however, were expected to be given full encodings only when a lag had transpired that was sufficient to allow individuating sentence characteristics,

such as surface structure, to fade, and hence hinder repetition recognition. As past research suggested, a single intervening item was expected to constitute a sufficient lag.

EXPERIMENT 1

In Experiment 1, sentences were repeated verbatim or with meaning-preserving changes in their wording at each of four lags. We hypothesized that recall of massed repetitions would benefit from changes in surface form, since the items would not be recognized as repetitions until they had undergone substantial processing for meaning. The benefits of full processing were expected to appear in higher recall probability of these items when compared to recall of verbatim massed repetitions.

Method

Subjects. Sixty-four students enrolled in introductory psychology courses at the University of Colorado, Boulder, served as subjects in the experiment as a partial requirement for course credit.

Materials and design. Forty sentences, chosen from a variety of sources and representing a variety of semantic contents, were chosen as stimuli. They are presented in Appendix A. The sentences to be repeated ranged in length from 9 to 13 words, with an average length of 12.1 words. Gist versions of the sentences (also presented in Appendix A) were formed by rearranging clauses, rewording clauses, substituting synonyms for original words, and substituting propositions of equivalent meaning. The 40 sentences were divided into two lists (List A and List B). Each list contained 22 sentences, including 2 once-presented sentences, 8 twice-presented sentences, a primacy buffer (5 sentences), and a recency buffer (5 sentences). The middle part of the list was divided into two halves; within each half, one position was reserved for the second occurrence of a sentence at each of four lags (0, 1, 3, and 8 intervening sentences), and one position was reserved for a once-presented sentence. Two lags in each half were further reserved for a verbatim repetition, and two for a gist repetition, such that across the two halves all four lags received both a verbatim and gist repetition. A second list-presentation version also was constructed, such that all verbatim-repetition positions in the first version became gist-repetition positions, and vice versa. Thus, any differences in recall that may have been due to more favorable positions within the list were equally distributed across the verbatim- and gist-repetition conditions. The two list-presentation versions are presented in Appendix B. Finally, to control for differences in memorability among sentences, eight sentence-assignment orders were used, ensuring that each sentence was tested in each repetition-type condition at each lag. Each assignment order was presented to 2 subjects. All subjects saw both lists, list-presentation order (List A first or second) was counterbalanced across subjects, and list-presentation version (Version 1 or Version 2) served as a between-subjects variable.

Procedure. Presentation of stimuli was controlled by a PDP-11 mini-computer. Sentences were presented on a CRT screen at a 4-sec rate. Subjects were run in groups of 2 working at independent stations. They were instructed simply to read each sentence as it appeared on the screen. They were told that the lists would contain repetitions, and that the wording of some of the repetitions would change. They also were informed of the recall tests that followed the presentation of each list. Following presentation of each list, subjects were given sheets of paper and asked to write down as many of the sentences as they could remember in any order they liked. They also were told to write down sentence fragments if they could not recall an entire sentence. Subjects were given as much time as they needed to complete their recall.

Results and Discussion

A sentence was scored as correctly recalled if its gist was correctly reproduced. For example, if the actual sentence seen was, "It was an hour before breakfast and the house was silent," then the following variations would count as correct reproductions: "There was still an hour before breakfast, and the house was quiet"; "It was an hour before breakfast, and everything at home was quiet"; "The house was silent in the hour before breakfast." The following examples were not treated as correct reproductions: "Things were quiet while we ate breakfast"; "It was early in the morning, and everybody was still asleep"; "When I woke up before breakfast, everything was silent." The sense of the major part of each clause—"an hour before breakfast" and "the house was silent"—had to be present in the reproduction in order for it to be treated as a correct reproduction of the sentence. The mean proportion of sentences correctly recalled is presented in Table 1.

An analysis of variance performed on the number of repeated sentences correctly recalled included the following variables: repetition type (verbatim and gist), lag (0, 1, 3, and 8 intervening sentences), list-presentation order (List A presented first or second), presentation version (Version 1 or 2), and sentence-presentation order (Orders 1 through 8), with repeated measures on the first two variables. Significant interactions were further analyzed via simple-effects tests. Significant simple main effects involving more than one mean were further analyzed using Tukey's HSD test for comparisons among pairs of means.

The main effect of lag was significant [$F(3,96) = 21.13$, $MSe = .45$, $p < .001$], as was the main effect of repetition type [$F(1,32) = 15.36$, $MSe = .34$, $p < .001$]. More importantly, the significant results included the repetition type \times lag interaction [$F(3,96) = 3.73$, $MSe = .41$, $p < .025$]. Recall of verbatim repetitions showed a clear spacing effect [$F(3,192) = 19.79$, $MSe = .43$, $p < .01$]. Recall of massed repetitions was statistically inferior to recall of all spaced repetitions. (The critical range for Tukey's HSD test of pairwise comparisons was .301; comparison of recall of massed repetitions and each spaced repetition yielded differences of .641, .829, and .757 for Lags 1, 3, and 8, respectively.)

Recall of gist repetitions also exhibited an effect of spacing, albeit in an attenuated form [$F(3,192) = 5.84$, $MSe = .43$, $p < .01$]. Massed repetitions were found to be recalled less often than repetitions at Lags 1 and 8, but not at Lag 3. (The differences between mean recall of massed repetitions and each spaced repetition were .375,

Table 1
Proportion of Sentences Whose Gist Was Accurately Recalled, Experiment 1

Repetition Type	Number of Intervening Sentences				Once-Presented
	0	1	3	8	
Verbatim	.179	.500	.594	.558	.183
Gist	.406	.594	.554	.633	

.296, and .453 for Lags 1, 3, and 8, respectively.) However, recall of gist massed repetitions was clearly superior to that of verbatim massed repetitions [$F(1,128) = 16.71$, $MSe = .39$, $p < .001$], whereas recall of the two types of repetitions was equivalent at all other lags (all $F_s < 1$). These latter results indicate a clear attenuation of the spacing effect when sentences are repeated in gist-only form. Moreover, they are consistent with the hypothesis that changes in surface form trigger full processing activities of items so changed, thereby enhancing their memorability.

EXPERIMENT 2

The results of Experiment 1 were consistent with deficient-processing explanations of the spacing effect. However, one argument that could be leveled against this interpretation is that altering the wording of a sentence changes the item itself. A sentence comprises both wording and meaning; a sentence whose wording is changed is simply not the same sentence, and it is therefore not clear in what sense a gist-only repetition counts as a true repetition. However, there are other types of surface form that may be changed and allow wording and meaning to remain untouched. For example, Geiselman and Bellezza (1976, 1977) found that subjects could reliably identify which of two speakers had spoken a given sentence from a list under both intentional and incidental learning conditions. In contrast, subjects could remember the physical location of the source of a spoken sentence only if explicitly instructed to encode such information. The authors interpreted these results to mean that voice or speaker information is processed during sentence comprehension without requiring an allocation of processing resources beyond those used to encode linguistic aspects of the sentence. Since speaker information appears to be reliably encoded during comprehension, identification of an item as a repetition should be hampered if the speaker of the item is changed at time of repetition. Therefore, it is more likely that an item affected in this way will enjoy fuller processing than an item whose speaker is not changed at time of repetition, and reinstatement of comprehension processes should benefit recall of items of the former type.

The purpose of Experiment 2 was to test this prediction. Sentences were recorded on tape and presented to subjects auditorily. When a sentence was repeated, the speaker (female or male) of the sentence either was changed (female to male or vice versa) or remained the same. As in Experiment 1, we reasoned that if deficient processing of massed items does underlie the spacing effect, and if surface form participates in the determination of item identity, and hence, determines the degree of processing a given item should receive, then retention of massed items whose surface structures are changed when repeated should exceed that of massed items that are not changed. Moreover, the spacing effect itself should be attenuated by this manipulation. In effect, then, the

retention curves of Experiment 2 were expected to approximate those of Experiment 1.

Method

Subjects. Sixty-four subjects enrolled in introductory psychology courses at the University of Colorado, Boulder, served as subjects as a partial requirement for course credit. Because it was unknown whether subjects would differentially attend to voices of speakers whose gender matched their own, we decided to equate the number of male and female subjects in the experiment and test for interactions. Thus, 32 subjects were female and 32 were male.

Materials and Procedure. The original sentences used in Experiment 1 were used here; the gist versions were not. The two lists of sentences were combined into one, however, because pilot work indicated ceiling effects in recall of the individual lists when sentences were presented to subjects auditorily as opposed to visually. In combining the two lists, the recency buffer of List A and the primacy buffer of List B were removed so that only one primacy and one recency buffer remained for the now-combined list. Two list-presentation versions were constructed as follows: Version 1 consisted of the primacy buffer and repetition section from Version 1-Experiment 1 combined with the repetition section and recency buffer from Version 2-Experiment 1. Version 2 consisted of the primacy buffer and repetition section from Version 2-Experiment 1, combined with the repetition section and recency section from Version 1-Experiment 1. As in Experiment 1, the versions differed in that all verbatim-repetition positions in Version 1 were gist-repetition positions in Version 2 and vice versa, thereby ensuring that no repetition type enjoyed a more favorable presentation position than the other. Thirty-two tapes were constructed, two for each of the 16 sentence-assignment orders described in Experiment 1. For half of these tapes, the main speaker of the sentences was male, and for the remaining half, the main speaker was female. Sentences were repeated either by the male speaker or by the female speaker. The difference in speakers was clearly discernible on the tapes. Subjects therefore heard 32 sentences in all: 5 primacy-buffer sentences, 5 recency-buffer sentences, 4 once-presented sentences, 2 filler sentences, and 16 twice-presented sentences. Of the twice-presented sentences, 4 occurred at each of the four lags (0, 1, 3, and 8 intervening sentences). Of the 4 in each lag, 2 were repeated by the same speaker, and 2 were repeated by a different speaker. The various lag and repetition-type conditions were equally distributed within each quarter of the list across subjects.

Subjects were run in groups of 4, and were instructed as in Experiment 1. The sentences were presented on individual cassette tape recorders, and subjects listened to them over headphones. Subjects were instructed to remove their headphones when they no longer heard sentences being presented. At that time, they were given sheets of paper on which to write down the sentences they remembered, as in Experiment 1.

Results and Discussion

Recall was scored as in Experiment 1. An analysis of variance performed on the number of sentences correctly recalled included the following variables: sex of main speaker on the tape (male or female), sex of subject (male or female), list-presentation version (Version 1 or 2), repetition type (same and different speaker), and lag (0, 1, 3, and 8 intervening sentences), with repeated measures on the last two variables. As in Experiment 1, significant interactions were followed by simple-effects tests. Significant simple main effects involving more than one mean were compared, using Tukey's HSD test for comparisons among means. The mean numbers correct were

Table 2
Mean Proportion of Sentences Whose Gist Was
Accurately Recalled, Experiment 2

Repetition Type	Number of Intervening Sentences				Once-Presented
	0	1	3	8	
Same Speaker	.140	.328	.390	.375	.156
Different Speaker	.367	.351	.446	.398	

converted to proportions and are presented in Table 2 for ease of comparison.

The main effect of lag was significant [$F(3,168) = 6.90$, $MSe = .38$, $p < .001$], as was the main effect of repetition type [$F(1,56) = 10.15$, $MSe = .339$, $p < .001$]. More importantly, however, the interaction of lag and repetition type also was significant¹ [$F(3,168) = 2.69$, $MSe = .45$, $p < .05$]. Simple-effects tests were conducted on this interaction, and the results follow:

- (1) When sentences were repeated by the same speaker, the typical spacing effect was found [$F(3,336) = 8.1$, $MSe = .41$, $p < .01$]. More particularly, recall of massed repetitions was found to be significantly lower than recall of sentences at each of the other three lag conditions; recall levels among the three distributed conditions were equivalent. (The required difference among the means using Tukey's test for paired comparisons was .296. The obtained differences were as follows: 0 vs. 1 = .375; 0 vs. 3 = .500; 0 vs. 8 = .469; 1 vs. 3 = .125; 1 vs. 8 = .094; 3 vs. 8 = .031.) These results therefore mirrored those of verbatim repetitions in Experiment 1.
- (2) When sentences were repeated by a different speaker, however, the spacing effect was obliterated [$F(3,336) = 1.06$, $MSe = .42$, $p > .05$]. In fact, the levels of performance for all lags in this condition were nearly equivalent.
- (3) Although sentences repeated by a different speaker tended to be better recalled than those repeated by the same speaker at each lag, the difference was significant only for massed repetitions {same vs. different speaker at Lag 0 [$F(1,224) = 7.55$, $MSe = .87$, $p < .01$]; F_s at all other lags < 1 }. Thus, massed repetitions benefited most from a change in this type of surface structure.

GENERAL DISCUSSION

Deficient-processing accounts of the spacing effect attribute poor recall of massed repetitions to attenuated processing of one of the presentations. A recent formulation (Cuddy & Jacoby, 1982) of the deficient-processing approach suggests that the likelihood of full processing for an item depends on the accessibility of previous encodings of the repeated item. If the previous encoding is readily accessible, as is the case in massed repetitions, full processing of the repetition is bypassed in favor of

simple retrieval of the previous encoding. If the previous encoding is not readily accessible, as is the case in spaced repetitions, the repetition is fully processed and hence is recalled more readily.

An implication of this approach is that anything that increases the probability of a repetition receiving full processing, or conversely, anything that decreases the probability of the item being recognized as a repetition, should improve memorability of the item. The present set of experiments tested this prediction by manipulating the surface structure of repeated sentences. On the basis of previous research, it was assumed that memory for surface structure of sentences decays rapidly, and hence can contribute to initial identification of repetitions only at short spacings. Because this manipulation should have hindered recognition of repetitions as repetitions, it was expected to induce full processing of massed repetitions, and thus facilitate recall of these items. This prediction was supported. When sentences were repeated verbatim (Experiment 1) or by the same speaker (Experiment 2), the typical spacing effect was obtained. However, when the surface structure or speaker was changed at time of repetition, massed repetitions were recalled nearly as well (Experiment 1) or as well (Experiment 2) as their spaced counterparts.

We claim that these results support a deficient-processing account of the spacing effect. Moreover, we argue that these results offer a clue to the mechanism responsible for the occurrence of deficient processing. Because the surface structure of our stimuli varied while the semantics remained constant, our results suggest that item identification in short-term memory (STM) is mediated primarily by surface-structure matches. When a surface-structure match is discovered in STM, processing of one of the items diminishes in favor of other items.

This interpretation is supported by the results of Glanzer and Duarte (1971), who found that repeating words in a different language (to bilinguals) attenuated the spacing effect relative to same-language repetitions. In their study, as in ours, surface structure was varied while semantics remained constant or nearly constant.

Similar results were reported by Glenberg and Smith (1981), who found that changing the modality of repetitions eliminated the spacing effect in word recognition. Their results are complicated, however, in that the same manipulation did not eliminate the spacing effect in word recall. The authors explain this contradiction by arguing that word recall, like frequency judgments (Hintzman, Block, & Summers, 1973), is mediated primarily by semantic factors, whereas word recognition may rely primarily on surface features. They further argued that such results are consistent with the components-level theory of memorability proposed by Glenberg (1979). This theory predicts the size of the spacing effect based on the number of study and test components that overlap. As the size of the space between repetitions increases, it is assumed that the number of item components that are encoded increases. The most important component is that

of context, which is assumed to be particularly useful on free-recall tasks, but also to drift over time. The theory accounts for Glenberg and Smith's (1981) results, but is at odds with ours and those of numerous other investigators who have reported spacing-effect attenuations in free recall using manipulations similar to those employed here (e.g., D'Agostino & DeRemer, 1973).

A modification of component-levels theory that would account for our results maintains the multicomponent representation of items in STM, but assigns differential weights to them in determining whether an item is to be fully processed or not. According to this theory, matches in surface structure components would lead to diminished processing of a newly received item.

There are three other interpretations of the spacing effect that might account for our results. The first is an encoding-context-variability hypothesis. On this view, one might argue that a change in surface structure represents a new context within which the gist of the sentence is embedded, thereby providing more retrieval routes/contexts for sentences repeated this way. The usefulness of this context at the time of retrieval is doubtful, however; long-term memory for surface structure, while often above chance, is nonetheless quite poor relative to memory for gist (Anderson & Paulson, 1977; Sachs, 1967). Indeed, memory for surface structure approximates memory for gist only when it is imbued with important pragmatic information (Keenan, MacWhinney, & Mahew, 1977; Kintsch & Bates, 1977). It is therefore unlikely that surface structure can serve as a useful retrieval cue or path when recalling an item. Moreover, as stated earlier, encoding context variability has been shown to be insufficient to enhance memorability of items (Postman & Knecht, 1983).

A closely related theoretical approach is that of semantic variability. According to this theory, the spacing effect is attributed to the greater likelihood of different semantic interpretations being assigned to an item at longer spacings due to "cognitive drift," much as in Glenberg's component-levels theory. This explanation has received most of its support from studies employing homographs as stimuli (Gartman & Johnson, 1972; Madigan, 1969). In these studies, two different senses of the homographs are biased by using two different presentation contexts (e.g., secluded-private, sergeant-private). This manipulation has been shown to produce an attenuation of the spacing effect. In the present study, a semantic-variability explanation would hold that changing the surface structure of the sentences produced a change in semantic interpretation. Although this explanation seems unlikely in our Experiment 1, where a wording change was employed, it is a plausible explanation in Experiment 2, where a change in speaker resulted in a much stronger attenuation of the spacing effect. Indeed, it has been suggested that speaker information may influence a sentence's interpretation (Geiselman & Bellezza, 1976, 1977; Geiselman & Crawley, 1983).

More to the point, however, is the contradiction that homograph results pose to our hypothesis concerning the importance of surface structure in item identification in STM. Homographs represent the mirror image of our manipulation. In the present study, surface structure varied, and semantics presumably remained constant. In homograph manipulations, surface structure is held constant, and semantics are varied. In the latter case, we would not expect a spacing-effect attenuation since we hypothesize processing-continuation decisions to be made primarily on the basis of surface-structure matches.

Homograph manipulations, however, have been criticized for a variety of reasons (see MacFarland, Rhodes, & Frey, 1979, for a brief review). The most important criticisms for our purposes are the following. First, Madigan (1969) noted a spacing-effect attenuation only with cued recall; when retention was measured with free recall, the spacing effect remained intact. This result is consistent with the notion that processing was attenuated for one or more of the massed items, since items receiving deficient processing would require stronger cues for their retrieval than would their better-processed counterparts. Second, Gartman and Johnson (1972) noted an attenuation of the spacing effect in free recall of contextually biased homograph stimuli, but this attenuation did not depend on a change in semantic context at presentation time; biasing the same meaning twice was as effective as biasing two meanings on two different presentations. Thus, contextual biases seemed to facilitate processing of items overall. Finally, Slamecka and Barlow (1979) contrasted long-term memorability of homographs that were repeated in same semantic biases, different semantic biases, and literal repetitions. Homographs that received different semantic biases were recalled no better than those receiving same semantic biases, but literal repetitions were recalled better than both. On the basis of their results, the authors concluded that surface-feature matches in long-term memory contributed most to the successful retention of their items. These many and varied results of homograph studies probably best indicate that homographs constitute an awkward means of holding surface structure constant. From this standpoint, our manipulation provides a stronger and more straightforward means of comparing the roles played by semantics and surface structure in STM-processing decisions.

A third and final explanation of the spacing effect is based on semantic-feature variability. Here, words are conceived to be a nonexhaustive set of semantic features, or components, and recall is considered to be a function of the number of features activated during encoding. The semantic-feature-variability hypothesis, like all of the above, attributes the spacing effect to the greater likelihood that different features of a word representation will be activated after longer lags than after shorter lags. In support of this hypothesis, MacFarland et al. (1979) reported data in which the spacing effect was attenuated by repeating words in contexts that highlighted different

semantic features (e.g., knife: has a metal blade, is used to cut) relative to contexts that highlighted the same semantic features (MacFarland et al., 1979, Experiment 1). In contrast, repeating words in contexts that highlighted different nonsemantic features (e.g., rhyme, case), as opposed to the same nonsemantic features, did not attenuate the spacing effect (MacFarland, et al., Experiment 2).

The present results are consistent not only with a semantic-feature interpretation, but also with a deficient-processing explanation. From a deficient-processing point of view, MacFarland et al.'s (1979) manipulation ensured that both (Experiment 1) or neither (Experiment 2) word presentation(s) received full, elaborative processing. These results are similar to Elmes and Bjork's (1975) results, which contrasted the effects of maintenance and elaborative processing on retention of word pentads over varying lags. The spacing effect was observed only in the condition in which both word presentations received maintenance processing, corresponding to MacFarland et al.'s (1979) nonsemantic condition; if either or both word presentation(s) received elaborative processing, the spacing effect was attenuated, corresponding to MacFarland et al.'s semantic condition. Moreover, inducing elaborative processing on both items produced higher recall at all lags relative to a control condition, suggesting that, under normal conditions, maintenance processing is the strategy of choice by most subjects.

Elmes and Bjork's (1975) results suggest that inducing elaborative processing of at least one of the massed items enhances recallability of the item; this in turn implies that, in normal conditions, at least one massed item receives inadequate processing. Moreover, intrusion errors within the maintenance and control conditions overwhelmingly tended to be acoustic in nature, again suggesting the primary role played by surface structure in identifying items in STM.

REFERENCES

- ANDERSON, J. R., & PAULSON, R. (1977). Representation and retention of verbatim information. *Journal of Verbal Learning & Verbal Behavior*, **16**, 439-451.
- BEGG, I. (1971). Recognition memory for sentence meaning and wording. *Journal of Verbal Learning & Verbal Behavior*, **10**, 176-181.
- CORNELL, E. H. (1980). Distributed study facilitates infants' delayed recognition memory. *Memory & Cognition*, **8**, 539-542.
- CUDDY, L. J., & JACOBY, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning & Verbal Behavior*, **21**, 451-467.
- D'AGOSTINO, P. R., & DEREMER, P. (1973). Repetition effects as a function of rehearsal and encoding variability. *Journal of Verbal Learning & Verbal Behavior*, **12**, 108-113.
- ELMES, D. G., & BJORK, R. A. (1975). The interaction of encoding and rehearsal processes in the recall of repeated and non-repeated words. *Journal of Verbal Learning & Verbal Behavior*, **14**, 30-42.
- GARROD, S., & TRABASSO, T. (1973). A dual-memory information processing interpretation of sentence comprehension. *Journal of Verbal Learning & Verbal Behavior*, **12**, 155-167.
- GARTMAN, L. M., & JOHNSON, N. F. (1972). Massed vs. distributed repetitions of homographs: A test of the differential encoding hypothesis. *Journal of Verbal Learning & Verbal Behavior*, **11**, 801-808.
- GEISELMAN, R. E., & BELLEZZA, F. S. (1976). Long-term memory for speaker's voice and source location. *Memory & Cognition*, **4**, 483-489.
- GEISELMAN, R. E., & BELLEZZA, F. S. (1977). Incidental retention of speaker's voice. *Memory & Cognition*, **5**, 658-665.
- GEISELMAN, R. E., & CRAWLEY, J. M. (1983). Incidental processing of speaker characteristics: Voice as connotative information. *Journal of Verbal Learning & Verbal Behavior*, **22**, 15-23.
- GLANZER, M., & DUARTE, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning & Verbal Behavior*, **10**, 625-630.
- GLENBERG, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, **7**, 95-112.
- GLENBERG, A. M., & SMITH, S. M. (1981). Spacing repetitions and solving problems are not the same. *Journal of Verbal Learning & Verbal Behavior*, **20**, 110-119.
- HINTZMAN, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, NJ: Erlbaum.
- HINTZMAN, D. L., BLOCK, R. A., & SUMMERS, J. J. (1973). Modality tags and memory for repetitions: Laws of spacing effect. *Journal of Verbal Learning & Verbal Behavior*, **12**, 229-238.
- HINTZMAN, D. L., SUMMERS, J. J., EKI, N. T., & MOORE, M. D. (1975). Voluntary attention and the spacing effect. *Memory & Cognition*, **3**, 576-580.
- JACOBY, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning & Verbal Behavior*, **17**, 649-667.
- JARVELLA, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning & Verbal Behavior*, **10**, 409-416.
- JARVELLA, R. J. (1973). Coreference and short-term memory for discourse. *Journal of Experimental Psychology*, **98**, 426-428.
- KAHNEMAN, D., & BEATTY, J. (1966). Pupil diameter and load on memory. *Science*, **154**, 1583-1585.
- KEENAN, J. M., MACWHINNEY, B., & MAHEW, D. (1977). Pragmatics in memory: A study of natural conversation. *Journal of Verbal Learning & Verbal Behavior*, **16**, 549-560.
- KINTSCH, W. (1975). Memory representation of text. In R. L. Solso, (Ed.), *Information processing and cognition*, Hillsdale, NJ: Erlbaum.
- KINTSCH, W., & BATES, E. (1977). Recognition memory for statements from a classroom lecture. *Journal of Verbal Learning & Verbal Behavior*, **3**, 150-159.
- MACFARLAND, C. E., RHODES, D. D., & FREY, T. J. (1979). Semantic-feature variability and spacing effect. *Journal of Verbal Learning & Verbal Behavior*, **18**, 163-172.
- MADIGAN, S. A. (1969). An intraserial repetition and coding processes in free recall. *Journal of Verbal Learning & Verbal Behavior*, **8**, 828-835.
- MAGLIERO, A. (1983). Pupil dilations following pairs of identical words and related to-be-remembered words. *Memory & Cognition*, **11**, 609-615.
- POSTMAN, L., & KNECHT, K. (1983). Encoding variability and retention. *Journal of Verbal Learning & Verbal Behavior*, **22**, 133-152.
- ROSE, R. J. (1980). Encoding variability, levels of processing, and the effects of spacing upon judgments of frequency. *Memory & Cognition*, **18**, 84-93.
- ROSE, R. J., & ROWE, E. J. (1976). Effects of orienting task, spacing of repetitions, and list context on judgments of frequency. *Journal of Experimental Psychology: Human Learning & Memory*, **2**, 142-152.
- ROSS, B. H., & LANDAUER, T. K. (1978). Memory for at least one of two items: Test and failure of retrieval theories of spacing effects. *Journal of Verbal Learning & Verbal Behavior*, **17**, 669-680.
- ROTHKOPF, E. Z., & COKE, E. U. (1963). Repetition internal and rehearsal method in learning equivalences from written sentences. *Journal of Verbal Learning & Verbal Behavior*, **2**, 406-416.
- RUNDUS, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Verbal Learning & Verbal Behavior*, **89**, 63-77.
- SACHS, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, **2**, 437-442.

- SHAUGHNESSY, J. J., ZIMMERMAN, J., & UNDERWOOD, B. J. (1972). Further evidence on the MP-DP effect in free recall learning. *Journal of Verbal Learning & Verbal Behavior*, **11**, 1-12.
- SILVERSTEIN, L. D. (1977). *Repetition and distribution effects on memory: A psychophysiological analysis*. Unpublished doctoral dissertation, University of Florida, Gainesville.
- SLAMECKA, N. J., & BARLOW, W. (1979). The role of semantic and surface features in word repetition effects. *Journal of Verbal Learning & Verbal Behavior*, **18**, 617-627.
- TOPPINO, T. C., & DIGEORGE, W. (1984). The spacing effect in free recall emerges with development. *Memory & Cognition*, **12**, 118-122.
- UNDERWOOD, B. J. (1970). A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning & Verbal Behavior*, **9**, 573-580.
- WAUGH, N. C. (1970). On the effective duration of a repeated word. *Journal of Verbal Learning & Verbal Behavior*, **19**, 587-595.
- WRIGHT, P. (1969). Transformations and the understanding of sentences. *Language & Speech*, 156-166.

NOTE

1. This interaction itself interacted with presentation order and sex of subject. Post hoc tests indicated that this modification resulted from higher recall levels by females on Presentation Order 1 at Lag 3. Thus, although the shape and directions of the lag curves were similar at each level of these variables, Lag 3 recall was elevated for females on Presentation Order 1. While no explanation of this effect is readily apparent other than random fluctuation, its presence does not affect the basic interpretation of the results.

APPENDIX A

Original and Gist(*) Sentences (Experiments 1 and 2)

LIST A

Repeated sentences:

- They turned up the street, where they had a clear view of the lake.
- * The lake became clearly visible to them as they turned up the street.
- Due to his inadequacies, his appointment as director was terminated.
- * He was fired from his job as director due to his inadequacies.
- A man tried to remove a fallen tree which was obstructing the road.
- * A tree fell across the road, and a man tried to remove it.
- The courses are designed as much for professionals as for amateurs.
- * Both professionals and amateurs can benefit from the courses.
- On the desk where she wrote letters were many objects coated with dust.
- * She wrote her letters at her desk, which was cluttered with dusty objects.
- He listened carefully because he had the impression that he knew the voices.
- * Hearing the voices, he listened carefully because they sounded familiar.
- There was still an hour before breakfast, and the house was silent.
- * The house was quiet and there was still an hour before breakfast.
- At the end of the musician's performance, the enthusiastic crowd applauded.
- * The crowd broke into loud applause when the musician's performance ended.

Once-presented sentences:

- Taxicabs used to be the main form of transportation to and from airports.
- * In the past, people usually took taxicabs to get to and from airports.

- When the teacher spoke to the boy, he became embarrassed.
- * The boy was embarrassed that the teacher had spoken to him.

Primacy buffer:

- Smokers were told to refrain from smoking until the end of the production.
- * Smokers were not allowed to smoke until the production ended.
- The rain and wind kept beating against the rattling window panes.
- * The window panes shook from the windy, rainy weather.
- It was shortly after this that I was called to town on business.
- * I was called to town on business after this occurred.
- He stuffed his denim jacket into his pack and fastened the straps.
- * He shoved his denim jacket into his backpack and fastened the straps.
- Digital electronics will play an important role in our future.
- * Digital electronics will strongly influence our future.

Recency buffer:

- To observe the effects of the drugs the doctor hospitalized the patient.
- * To monitor the drugs' effects the doctor hospitalized the patient.
- For one entire summer the child learned to swim.
- * The child learned to swim all summer long.
- The company won an award for its new invention.
- * As a result of its new invention, the company was given an award.
- The old fashioned room contained benches made of logs.
- * There were long benches made of logs in the old fashioned room.
- He had thought things over while the secretary was there.
- * He had been thinking things over while the secretary was with us.

Filler item:

- On the kitchen counter was a big jar that she kept filled with cookies.
- * She always kept cookies on top of the kitchen counter in a cookie jar.

LIST B

Repeated sentences:

- The roadbed over which the railroad tracks were built was dry and cracked.
- * The railroad tracks stretched across an arid and rugged terrain.
- Heat from the sun caused the hot air balloon to rise quickly.
- * The large balloon rapidly rose as the sun's heat warmed it.
- The police arrived too late; the burglars had left the room in a shambles.
- * The burglars ransacked the room and escaped before the police arrived.
- He stood at the edge of the crowd while they were singing.
- * He stood on the periphery of the crowd and listened to the singing.
- Looking at the building, I saw that it had not suffered from the fire.
- * The building appeared to have escaped damage from the fire.
- The entire town arrived to see the political candidate.
- * The political candidate was welcomed by all of the town's inhabitants.
- The old lady talked to her neighbor on her weekly walks from church.
- * Walking home from services, the old lady usually visited her neighbor.
- Without any hesitation, he plunged into the mathematics assignment.
- * He tackled the mathematics assignment with little hesitation.

Once-presented sentences:

The old dog reached his master's house in a matter of a few hours.

* It took the old dog only a few hours to find his way home.

The captain sat beside the man and stared at the muddy ground.

* Sitting beside the man, the captain never lifted his eyes from the ground.

Primacy buffer:

It took us over a month to find the kind of house we wanted to buy.

* After a month of looking, we finally found the house we wanted.

The favorite tree of many landscapers is the birch.

* Birch trees are popular among landscapers.

At the sound of the crash, some passers-by came to see what was happening.

* A small crowd gathered when the crash was heard.

When the first snow fell, all the stores closed and the workers went home.

* On the first snowy day, the store workers took a snow holiday.

The book had several interesting stories and colorful pictures.

* The books' stories were interesting and well illustrated.

Recency buffer:

Directions for using the equipment were in three languages.

* The equipment's directions were written in three languages.

In the early days, glass was usually opaque and colored.

* Years ago, glass used to be opaque and colored.

The guard looked for articles that might have been tossed from the car.

* The guard searched around the car for articles that might have been tossed.

Well-trained horses are hard to find and expensive to buy.

* Well-trained horses are both expensive and difficult to locate.

Camera equipment has become very sophisticated in recent years.

* Camera equipment increased in sophistication over the years.

Filler item:

As a young child, I loved grapes and I loved raisins best.

* Grapes and raisins were my favorite fruits when I was a child.

APPENDIX B

List-Presentation Versions--Experiment 1

Condition		Sentence
Version 1 (Version 2)		Original or Gist
	Primacy	O (G)
	Primacy	G (O)
	Primacy	G (O)
	Primacy	O (G)
	Primacy	O (G)
	V8 ₁ (G8 ₁)	O (O)
	V1 ₁ (G1 ₁)	O (O)
	G8 ₁ (V8 ₁)	O (O)
	V1 ₂ (G1 ₂)	O (G)
	G3 ₁ (V3 ₁)	O (O)
	Once	G (O)
	V0 ₁ (G0 ₁)	O (O)
	V0 ₂ (G0 ₂)	O (G)
	G3 ₂ (V3 ₂)	G (O)
	V8 ₂ (G8 ₃)	O (G)
	Filler	O (G)
	G8 ₂ (V8 ₂)	G (O)
	G0 ₁ (V0 ₁)	O (O)
	G0 ₂ (V0 ₂)	G (O)
	V3 ₁ (G3 ₁)	O (O)
	G1 ₁ (V1 ₁)	O (O)
	Once	O (G)
	G1 ₂ (V1 ₂)	G (O)
	V3 ₂ (G3 ₂)	O (G)
	Recency	G (O)
	Recency	O (G)
	Recency	G (O)
	Recency	G (O)
	Recency	O (G)

(Manuscript received December 4, 1984; revision accepted for publication August 15, 1985.)