

Acoustic and semantic interference effects in words and pictures

MEENA DHAWAN and JAMES W. PELLEGRINO
University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Interference effects for pictures and words were investigated using a probe-recall task. Word stimuli showed acoustic interference effects for items at the end of the list and semantic interference effects for items at the beginning of the list, similar to results of Kintsch and Buschke (1969). Picture stimuli showed large semantic interference effects at all list positions with smaller acoustic interference effects. The results were related to latency data on picture-word processing and interpreted in terms of the differential order, probability, and/or speed of access to acoustic and semantic levels of processing. A levels of processing explanation of picture-word retention differences was related to dual coding theory. Both theoretical positions converge on an explanation of picture-word retention differences as a function of the relative capacity for semantic or associative processing.

One aspect of human memory performance which has received intensive study is the difference in retention between pictorial and verbal material. Paivio (1971) and others have shown a consistent superiority in retention of pictures relative to words in a variety of standard verbal learning paradigms. Such results have given rise to a dual coding theory of memory. Within this framework, differences in memory performance in a given task are explained by the number of probable memory codes existing for a given type of item. Multiple codes yield higher retention due to multiple sources of storage and retrieval. Both pictures and concrete words are assumed to have access to both a nonverbal and a verbal memory representation. However, pictures have a higher probability than concrete words of accessing both representations. This presumably results from an asymmetry in transfer between the two representations; that is, it is faster to access the verbal representation than the converse. The existence and high probability of accessing two memory codes produces superior retention of pictures relative to concrete words; concrete words, in turn, are retained better than abstract words which have only the single verbal representation (Paivio, 1971). In addition, Paivio (1975a) has proposed

that retention is a function of the relationship between the retention probabilities associated with given types of codes, where the relative value of each code can vary across tasks. As an illustration of this, Paivio (1975a) has presented data showing a functional independence of verbal and imaginal coding processes, with the additional result that the imaginal code is more beneficial for retention than is the verbal code. The assessment of any qualitative differences between codes depends upon the extent to which retention in multiple coding conditions matches the value predicted from an additive independence model. If two codes are functionally independent in memory, then retention of an item coded once verbally and coded imaginally on a second occurrence will equal the retention probability of a verbally coded item multiplied by the probability of an imaginally coded item minus the joint retention probability. Codes which are nonindependent, that is, not qualitatively different, will yield retention values smaller than that predicted by an additive model.

The question remains as to why the imaginal or nonverbal code is more beneficial for retention, and how to conceptualize the nature of such coding in relation to other coding processes for verbal stimuli. In particular, investigations manipulating the processing of verbal input (e.g., Craik & Tulving, 1975; Hyde & Jenkins, 1969) reveal that retention in both recall and recognition tasks is an increasing function of the depth or breadth of encoding across graphic, acoustic-phonemic, and semantic levels of analysis. Such investigations and resultant conclusions about the critical role of processing level have been limited to verbal stimuli, since coding of such stimuli is readily manipulable along these dimensions. In addition, the ordering of dimensions follows an assumed logical sequence for the serial or temporally overlapping processing stages of graphic verbal input. Research has yet to demonstrate whether the different coding operations show the same func-

This research was supported by funds provided by the Learning Research and Development Center, University of Pittsburgh, which is supported in part by funds from the National Institute of Education (NIE). The research reported is based upon a Masters thesis conducted by the first author under the supervision of the second author. The authors would like to express their appreciation to the other members of the thesis committee: Drs. James Voss, Alexander Siegel, and Judith Petrich for their helpful comments and criticisms. Thanks are also extended to Susan Goldman for her helpful comments on the present manuscript. Requests for reprints should be addressed to James W. Pellegrino, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania 15260.

tional independence that Paivio (1975) demonstrated for verbal (acoustic-phonemic) and nonverbal (imaginal) coding.

At present, the depth or breadth of encoding analysis for verbal stimuli clearly indicates the utility of some form of associative or semantic processing. Given this framework, a post hoc consideration of picture-word retention differences would assume that pictures reach a greater depth or breadth of associative or semantic encoding, relative to their verbal counterparts, under typical constrained and equal rates of presentation. This is clearly a restatement of the results without a direct test, and no more elucidating than a multiple code explanation. However, there is evidence that pictures and words differ in the relative ease of semantic analysis. In both single and multiple item presentation conditions, pictures yield faster decision latencies than do words for semantic category judgments (Potter & Faulconer, 1975; Rosch, 1975; Pellegrino, Rosinski, & Siegel, Note 1; Rader, Note 2). The data clearly indicate that pictures have the potential for more rapid access to associative or semantic levels of analysis. The data further suggest that differences in the capability for rapid and extensive analysis at the level of meaning may underlie picture-word performance differences in typical memory tasks.

The previously mentioned decision-latency tasks require a subject to engage in a semantic level of analysis. However, differences in latency do not necessarily imply that the capability for faster semantic analysis is also a factor in the typical unconstrained memory situation. An explanation of picture-word retention differences, in terms of differences in processing level, requires that the difference in speed of semantic analysis be reflected in typical learning and memory situations. The present research attempted to specifically examine the assumed processing difference by using an interference methodology. In such a methodology, one infers the types of processing of a particular item from the nature of the stimuli which can interfere with its subsequent retention. In the case of verbal stimuli, it has been extensively demonstrated that there is a time course of item processing which controls the type of interference effects disrupting performance (e.g., Shulman, 1971). The nature of the interference is a function of both the material presented and the demands of the task (Shulman, 1971, 1972). Kintsch and Buschke (1969) employed a probe-recall procedure to examine interference effects and found that items from the initial portion of a 16-item word list were only affected by semantic interference, while words from the terminal portion of the list were only affected by acoustic interference. While the data have been interpreted primarily in terms of the separate coding processes in long-term and short-term memory, they are also compatible with a levels of processing position (Kintsch, 1974). The present experiment employed the probe-recall task to directly compare interference effects for pictorial

and verbal stimuli, with the expectation that verbal stimuli would show interference effects identical to those previously obtained by Kintsch and Buschke (1969). However, the interference effects obtained for pictorial stimuli should parallel those for the verbal stimuli only for the initial portion of the list. If pictures are more rapidly processed with respect to attributes of meaning, then items in the terminal portion of the list should also show semantic interference effects, in contrast to the general lack of such effects for verbal stimuli.

Such a result can also be predicted on the basis of free recall data investigating negative recency effects. The occurrence of negative recency for verbal stimuli has been related to a lower level of processing for items in the terminal portion of the list (e.g., Maskarinec & Brown, 1974; Watkins & Watkins, 1974). However, negative recency does not occur for pictures in the terminal list positions (Madigan, McCabe, & Itani, 1972), implying that the items have attained a processing level sufficient to sustain long-term retention. The latter is presumably a function of associative or semantic coding, and thus pictures in the terminal list positions should be primarily affected by semantic interference. Acoustic interference effects should be small relative to the level of semantic interference, since any processing of acoustic attributes should follow semantic processing. Thus, specific picture and word interference effects and differences in interference effects can be predicted from changes in coding processes over time. It must be noted, however, that the changes in coding processes do not necessarily conflict with assumptions of dual coding theory (Paivio, 1975a, b). A more complete discussion of this will follow presentation of the results.

METHOD

Design

This experiment was a 2 by 3 by 4 mixed factorial design, manipulating two between-subjects factors and one within-subjects factor. The between-subjects factors were the stimulus type (picture vs words) and type of distraction (acoustic, semantic, or control). Probe-target positions (3-4, 6-7, 10-11, or 13-14) were tested as a within-subjects factor.

Materials

A list consisted of 15 items, including a distractor item, which was always located in Position 8. The distractor was related either acoustically or semantically to the probed item, the position of which varied from trial to trial. Twenty lists of pictures or their corresponding verbal labels were prepared from a 75-item pool. Sixty of the items represented 20 different three-item sets. In all 20 sets, one item always served as the probe and the other two items were semantically or acoustically related to the probe. The acoustic distractors rhymed with the probe, while semantic distractors were chosen on the basis of membership in common conceptual categories and/or a high associative relationship. Examples of such three-item sets were: Star-Jar-Moon; Bed-Sled-Dresser; Lock-Clock-Key; Ring-Swing-Watch. No item appeared more than five times in the set of 20 lists. Also, no two items from the same conceptual category

appeared in the same list under acoustic or no-distractor conditions, and no two items from the same conceptual category appeared in the lists with semantic distractors, except for the probe and the distractor items. Thus, each list contained one probe-distractor pair, either acoustically or semantically related. The control group received identical lists, except that the related distractor was replaced by an unrelated item.

Procedure

Twenty lists of one type (either with acoustic, semantic, or unrelated control items) were successively shown to a subject, with each item visually presented for 2 sec. After each list of 15 items was presented, the probe item was re-presented after a short blank on the screen. The subject's task was to recall the item which had followed the probe item earlier in the list, with 10 sec allowed for oral recall. If the subject failed to make any response within the 10-sec interval, an omission error was recorded. All subjects were tested at each pair of probe-target positions five times across the 20 lists, and probe-target positions were counterbalanced across subjects and lists.

Subjects

The subjects were 72 University of Pittsburgh undergraduates participating to fulfill a course requirement. Twelve subjects were randomly assigned to each of the six groups, with subjects tested in individual sessions.

RESULTS

Total correct recall was analyzed in terms of the factors of stimulus type (pictures, words), type of interference (acoustic, semantic, control), and the probe-target position (3-4, 6-7, 10-11, 13-14). The minimum significance level was 5%. The analysis showed that all the main effects, except stimulus modality, and all the interactions were significant. The results will be presented in terms of the higher-order interaction involving all three variables and the lower-order interactions and main effects. The interaction involving all three variables is shown in Figure 1 [$F(6,198) = 92.5$, $MSe = .022$]. To simplify this interaction and all others involving the variables of probe-target position, the latter was always broken down into three orthogonal components involving: (1) the first half (Positions 3-4 and 6-7) vs

the second half (Positions 10-11 and 13-14) of the list, (2) probe-target Positions 3-4 vs Positions 6-7, and (3) probe-target Positions 10-11 vs Positions 13-14.

As can be seen in Figure 1, the major component of the overall interaction involved the first half vs second half comparison [$F(2,198) = 42.54$, $MSe = .022$]. This shows that in the first half of the list, semantic interference produced the poorest performance for both picture and word stimuli and this also applied for picture stimuli in the second half of the list. However, for the word stimuli, semantic interference had no effect in the second half of the list, while acoustic interference produced a large decrement in performance. The differential pattern of interference effects in the second half of the list was large enough to produce a Picture-Word by Type of Interference interaction [$F(2,66) = 20.81$, $MSe = .048$], an interaction involving Type of Interference by First-Second Half of the list [$F(2,198) = 30.47$, $MSe = .022$], and it contributed to an overall effect of type of interference [$F(2,66) = 37.23$, $MSe = .048$].

The remaining two components of the three-factor interaction were also significant. The component involving the comparison of Positions 3-4 vs 6-7 [$F(2,196) = 5.98$, $MSe = .022$] was primarily due to the high performance in the picture-control condition at Position 3. This high recall contributed to a significant Picture-Word by First-Second Half interaction [$F(1,198) = 11.14$, $MSe = .022$] and a Picture-Word by Position 3-4 vs 6-7 interaction [$F(1,198) = 4.09$, $MSe = .022$]. Finally, the component of the three-factor interaction involving the comparison of Positions 10-11 vs 13-14 [$F(2,198) = 7.60$, $MSe = .022$] reflected the differential gain from Positions 10-11 to 13-14. For both pictures and words, the poorest condition at Position 10-11 showed the least gain at Positions 13-14. The picture and word stimuli clearly differed as to which type of interference produced the smallest gain across these positions.

The remaining significant effect was probe position [$F(3,198) = 95.46$, $MSe = .022$]. All three components of this main effect were significant. The first half was inferior

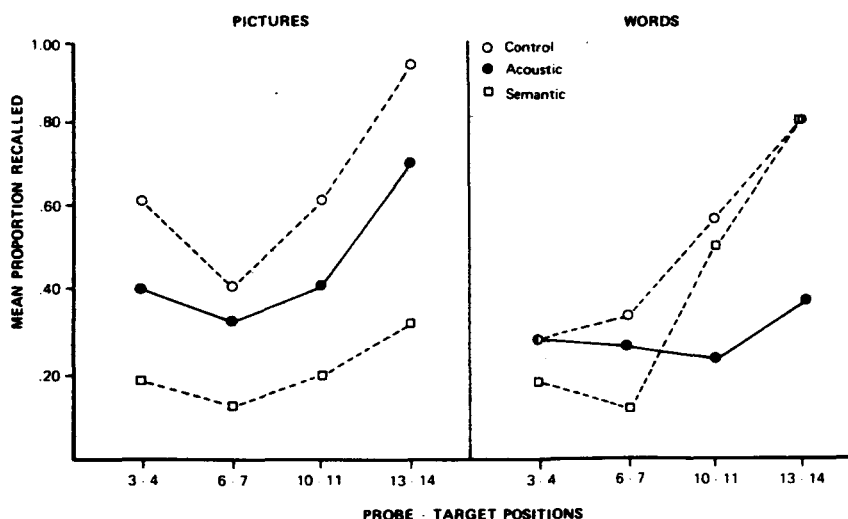
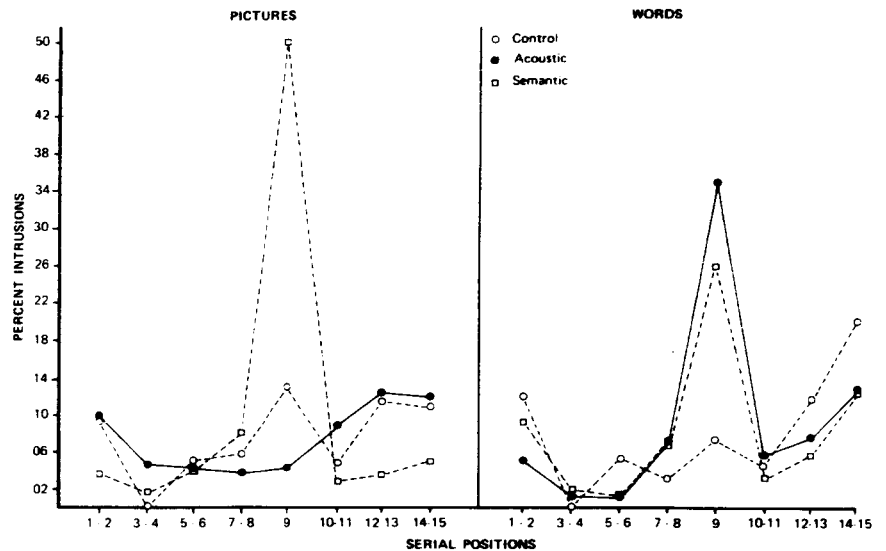


Figure 1. Mean proportion recalled as a function of stimulus type, interference condition, and probe-target position.

Figure 2. Intrusion proportions from serial list positions as a function of stimulus type and interference conditions.



to the second half of the list [$F(1,198) = 195.55$, $MSe = .022$]; recall at Positions 3-4 was superior to recall at Positions 6-7 [$F(1,198) = 5.05$, $MSe = .022$]; and, recall at Positions 13-14 was superior to recall at Positions 10-11 [$F(1,198) = 88.89$, $MSe = .022$].

Error Distribution

Total errors were divided into the errors due to omissions and the errors due to intrusions. A high proportion of intrusions was obtained in all conditions, with values ranging from .50 to .78. The intrusion errors were subsequently divided into intrusions from Positions 7-8-9 and intrusions from other serial positions. This was done because the distractor item always occupied Position 8, and, if the distractor was causing an error, the error should be localized in the area of the distractor. The proportion of intrusions at Positions 7-9 is shown in Table 1, together with values based upon chance. With the word stimuli, the proportion of intrusions from Positions 7-8-9 depended mainly upon the type of interference. When acoustic interference

was involved, more intrusions from Positions 7-8-9 were observed if the probe item was in the second half of the list. When the probe item was in the first half of the list, semantic interference produced more intrusions from these critical positions. With the picture stimuli, semantic interference produced a high proportion of intrusions from Positions 7-8-9, irrespective of probe position. Acoustic distraction had little effect on pictures and both control groups had a chance percentage of intrusions from Positions 7-8-9.

A more detailed analysis was also done across all serial positions. Because of the differential frequency with which intrusions could occur from any particular serial position, upper limits of error frequencies were computed. The latter were based upon the total opportunities for an intrusion to occur from a given position, including probe-target positions. The limits were then used to calculate the exact proportion of intrusion errors from a given serial position, and this is shown in Figure 2. It can be seen that, with pictures, semantic distraction produced about 50% of the maximum intrusions that could have occurred from Position 9, while control and acoustic distraction conditions did not show such an effect. Word stimuli also showed a large proportion of intrusions from Position 9 for both acoustic and semantic interference conditions, with no similar effect in the control condition. However, the proportions of intrusions from Position 9 were not as large as in the picture-semantic condition. This is due to the fact that with word stimuli, semantic and acoustic interference affected only half of the list. Therefore, the intrusions should be somewhat more evenly distributed across all positions.

DISCUSSION

The present results will be considered first in terms of the original predictions concerning the order in which

Table 1
Proportions of Intrusions Localized from Positions 7-8-9 as a Function of Stimulus Type, Interference Condition, and Probe-Target Positions

	Probe-Target Positions			
	3-4	6-7	10-11	13-14
Chance Baseline	.23	.15	.23	.23
Words				
Control	.15	.09	.17	.17
Acoustic	.07	.29	.64	.65
Semantic	.41	.44	.39	.00*
Pictures				
Control	.29	.15	.25	.00*
Acoustic	.15	.12	.09	.00*
Semantic	.53	.56	.68	.67

*Based on nine or less total intrusions

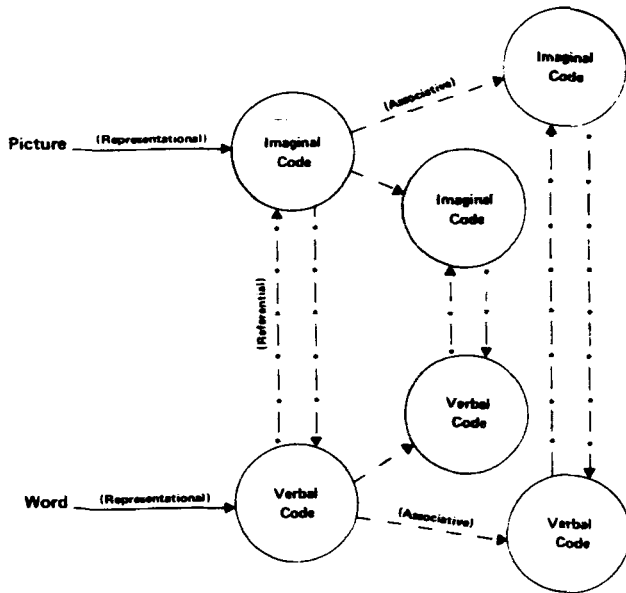


Figure 3. Representation of processing levels in a dual coding theory of memory.

various levels of processing may occur for pictures and words. On the basis of previous research by Kintsch and Buschke (1969) and others (e.g., Shulman, 1971), it was hypothesized that words require more time to reach a semantic as opposed to an acoustic-phonemic level of processing. Thus, words which occupy initial list positions would show semantic interference effects, while words occupying terminal list positions would show acoustic interference effects. The present experiment clearly showed this pattern and replicated the results obtained by Kintsch and Buschke (1969) in a similar probe-recall task.

A different pattern of interference effects was predicted for pictures from a hypothesis that pictures are rapidly coded at an associative or a semantic level, followed by possible coding at an acoustic level. Again,

the data clearly supported these predictions with substantial reductions in recall in the semantic interference condition at all probe-target positions. There was also evidence of acoustic interference at all probe-target positions, but the level of interference was decidedly less than that observed in the semantic interference condition.

As noted in the introduction, the present pattern of interference effects and related assumptions about processing differences between words and pictures are not necessarily in conflict with dual coding theory. Differences in speed of processing can be accommodated by such a theory, given the basic processing assumptions already inherent in the theory (Paivio, 1975a, b, c). In dual coding theory, the initial processing of pictures and words is at the representational level, followed by referential and/or associative processing, as illustrated in Figure 3. Paivio (1975c) has argued that associative processing in the imaginal or nonverbal system is more rapid than such processing in the verbal system. The argument was based upon the assumption of different organizational characteristics of the two systems. The difference in the speed of associative processing is consistent with the present difference between pictures and words in semantic interference effects for the terminal list positions, as well as the superiority of pictures over words in the speed of semantic category judgments (Potter & Faulconer, 1975; Rosch, 1975; Pellegrino et al., Note 1). If it is further assumed that associative processing in the nonverbal system is more probable and/or more rapid than referential processing between the nonverbal and verbal systems, then it is also possible to account for the present levels of semantic and acoustic interference with pictures.

The foregoing assumptions about the order and speed of activation of processing dimensions can be represented as in the right panel of Figure 4. The figure illustrates that the processing of pictures or words can involve several dimensions which operate in parallel and which may yield independent additive effects for retention.

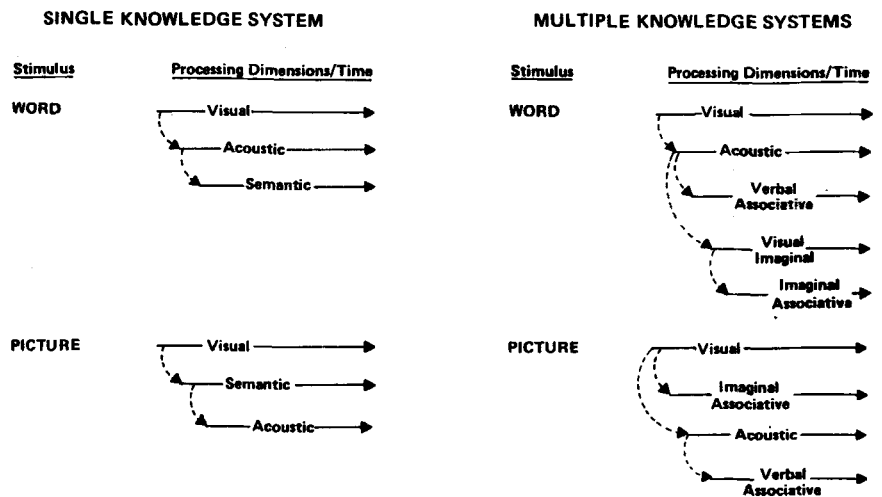


Figure 4. Relationships among processing dimensions for picture and word stimuli in a theory postulating a single meaning system vs a theory postulating multiple meaning systems.

The activation of a dimension is contingent upon activation of other dimensions of processing. When dual coding theory is viewed in this way, there exist two potential explanations of picture-word retention differences. The first is the typical explanation based upon different dimensions of processing (i.e., number of codes). The second explanation relies upon the speed of activating the associative dimensions of processing. The second explanation is also completely consistent with a levels or depth of processing interpretation of picture-word retention differences. The depth of processing position is clear about the relative retention benefits of semantic or associative processing, as well as the benefits of increased elaboration or spread of encoding along the semantic or associative dimension (Craik & Tulving, 1975). If, as argued above, pictures and words differ in the speed with which this processing dimension is activated, then they may also differ in the relative amount of processing or activation on this dimension, given constrained and equal rates of presentation. The difference in associative or semantic processing would be expected to produce superior picture retention relative to word retention. Thus, a depth of processing interpretation is not inconsistent with dual coding theory, nor is one a restatement of the other (see also Nelson & Reed, 1976). Either explanation may be applied to picture-word retention differences, or they may be combined to yield a composite explanation.

As represented in Figures 3 and 4, dual coding theory assumes two or more separate and independent but interconnected knowledge systems, in other words, the theory assumes multiple codes = multiple knowledge bases. While models of the human information processing system assume multiple sensory input channels, they need not also assume multiple knowledge systems corresponding to each form of sensory input (e.g., Anderson & Bower, 1973; Chase & Clark, 1972; Pylyshyn, 1973). Thus, it is also necessary to consider a processing model consistent with the multiple code, single knowledge position. This model can be offered in place of the processing model developed earlier based upon the multiple code/multiple knowledge position. The processing assumptions of the model are represented in the left panel of Figure 4. The first processing assumption is that a visual, nonlinguistic input rapidly activates analysis at the semantic or meaning level. The second assumption is that there may be processing of the acoustic-phonemic features associated with a pictured concept, but such processing must follow activation of the concept on the semantic dimension. The final assumption is that the opposite sequence of semantic and acoustic-phonemic activation occurs in the processing of words. The difference in the sequential activation of the semantic and acoustic processing dimensions is consistent with the interference effects obtained in the present study, as well as with all of the

previously cited latency data on semantic and acoustic judgments with words and pictures. The differences in processing are consistent with the depth of processing explanation offered earlier for picture-word retention differences. This framework also allows one to view the beneficial effects of imagery processing instructions as the activation of meaning features on the semantic dimension, thereby producing a rich or elaborated memorial representation (cf. Anderson & Bower, 1973). Furthermore, none of the present assumptions are at variance with Paivio's (1975a) results showing the independence and additivity of imaginal and verbal (acoustic-phonemic) coding processes in retention, and the lack of additivity of imaginal and pleasantness coding processes.

It must be stressed that the preceding discussion is not intended as an argument for either a single or dual knowledge system. Rather, the purpose has been to present processing models derived from these positions. The specific processing assumptions of each model are necessary to account for the present data as well as latency data on picture-word processing differences. Regardless of whether one embraces a theory with single or multiple knowledge systems, adequate explanations for the retention of verbal and nonverbal input must deal more explicitly with the dimensions of processing, their interrelations in terms of the speed, order, and probability of activating particular dimensions, and the level of activation within each dimension.

REFERENCE NOTES

1. Pellegrino, J. W., Rosinski, R. R., & Siegel, A. W. *Semantic processing of pictures and words*. Paper presented at Psychonomic Society Meetings, Denver, Colorado, November 1975.
2. Rader, N. L. *Developmental changes in getting meaning from written words*. Paper presented at Society for Research in Child Development Meetings, Denver, Colorado, April 1975.

REFERENCES

- ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington, D.C.: V. H. Winston, 1973.
- CHASE, W. G., & CLARK, H. H. Mental operations in the comparison of sentences and pictures. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.
- CRAIK, F. I. M., & TULVING, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, 1, 268-294.
- HYDE, T. S., & JENKINS, J. J. Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 1969, 82, 472-481.
- KINTSCH, W. *The representation of meaning in memory*. New York: Lawrence Erlbaum, 1974.
- KINTSCH, W., & BUSCHKE, H. Homophones and synonyms in short-term memory. *Journal of Experimental Psychology*, 1969, 80, 403-407.
- MADIGAN, S., MCCABE, L., & ITATANI, E. Immediate and delayed recall of words and pictures. *Canadian Journal of Psychology*, 1972, 26, 407-414.

- MASKARINEC, A. S., & BROWN, S. C. Positive and negative recency effects in free recall learning. *Journal of Verbal Learning and Verbal Behavior*, 1974, **16**, 328-334.
- NELSON, D. L., & REED, V. On the nature of pictorial encoding: A levels-of-processing analysis. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 49-57.
- PAIVIO, A. *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston, 1971.
- PAIVIO, A. Coding distinctions and repetition effects in memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975. (a)
- PAIVIO, A. Perceptual comparisons through the mind's eye. *Memory & Cognition*, 1975, **3**, 635-647. (b)
- PAIVIO, A. Imagery and synchronic thinking. *Canadian Psychological Review*, 1975, **16**, 147-163. (c)
- POTTER, M. C., & FAULCONER, B. A. Time to understand pictures and words. *Nature*, 1975, **253**, 437-438.
- PYLYSHYN, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 1973, **80**, 1-24.
- ROSCH, E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975, **1**, 192-233.
- SHULMAN, H. G. Similarity effects in short-term memory. *Psychological Bulletin*, 1971, **75**, 399-415.
- SHULMAN, H. G. Semantic confusion errors in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1972, **11**, 221-227.
- WATKINS, M. J., & WATKINS, O. C. Processing of recency items for free recall. *Journal of Experimental Psychology*, 1974, **102**, 488-493.

(Received for publication April 19, 1976;
revision received July 13, 1976.)