

Common and modality-specific processes in the mental lexicon

KIM KIRSNER, D. MILECH, and P. STANDEN

University of Western Australia, Nedlands, Western Australia, Australia

Eight experiments were conducted to resolve: (1) empirical inconsistencies in repetition effects under intermodality conditions in word identification and lexical decision, and (2) an associated theoretical conflict concerning lexical organization. The results demonstrated that although more facilitation occurs under visual-prime/visual-test (VV) conditions than under auditory-prime/visual-test (AV) conditions, significant repetition facilitation also occurs under AV conditions. The results also indicated that: repetition effects observed for the VV and AV conditions apply to high- as well as to low-frequency words; they are insensitive to a variety of encoding tasks designed to emphasize different properties of words; and they are unaffected by differences in the ease of encoding of isolated auditory and visual words. The results are consistent with the existence of both modality-specific and common or modality-free processes in word recognition, in which word-frequency effects are restricted to the second and, by implication, lexical stage.

Research into modality differences has focused on a variety of empirical issues over the last 2 decades. For much of this time, interest has focused on episodic questions concerning, for example, the distinctions among sensory, short-term, and long-term memory systems. At the heart of this problem, however, is the modal specificity of the system responsible for the transition from sensory to categorical representation. If categorical information emanates from a single lexicon, albeit via distinct access routines, it follows that equivalent item information will be available following auditory and visual presentation. In this case, then, modality information must be stored dependently, as an associate of the item information. But if distinct lexical systems are involved in reading and in speech recognition, the item information available following auditory and visual presentation may not be equivalent. In this case, modality information may be provided automatically by the item information.

Research into the modal specificity of lexical processes has not yielded a definitive answer. One approach to the question involves observation of repetition effects in word-processing tasks. When equivalent repetition effects are observed under inter- and intramodality conditions, it is assumed that a common system is involved. But when the repetition effects are restricted to intramodality conditions, it is assumed that separate mechanisms are involved in text and speech recognition. Three positions, representing zero, partial, and complete transfer under intermodality conditions, may be identified in the literature. The zero-transfer position is repre-

sented in Morton's (1979) results and is reflected in his model. Using word recognition, with threshold-estimation procedures in vision and audition, he found no evidence of transfer under intermodality conditions. The model that he subsequently developed includes modality-specific word recognition units, or "logogens." According to Morton, these units are not connected except through a separate cognitive system, so intermodality transfer is absent in tasks that tap only word identification processes. Any task that taps semantic information, however, may yield intermodality transfer because the coordinating cognitive system is also involved. The lexical decision task is an obvious candidate for intermodality transfer because of its presumed dependence on semantic processes (James, 1975).

Kirsner and Smith (1974) reported an intermediate repetition effect under intermodality conditions. Using auditory and visual presentation of isolated words in a lexical decision task, they found reaction time (RT) facilitation relative to new words for visual-prime/visual-test (VV) and auditory-prime/visual-test (AV) conditions. This result can be accommodated by Morton's (1979) model if it is assumed that the lexical decision task reflects activation in the cognitive, as well as in the word recognition, system. The result is also consistent with Forster's (1976) account of lexical access. He assumes that there are both modality-specific and common components in word recognition. The modality-specific components in his model consist of peripheral access files that are coded in terms of orthography or phonology. The content of the entries in these files is assumed to be the address of semantic and other definitional information in a modality-free master file. Clearly, if it is assumed that entries in the peripheral and master files can be primed (regardless of the mech-

The authors' mailing address: Department of Psychology, University of Western Australia, Nedlands, Western Australia 6009, Australia.

anism or mechanisms concerned), it follows that, although there will be more facilitation under intra- than under intermodality conditions, some facilitation will also be observed under intermodality conditions.

The third position, involving complete transfer, does not enjoy empirical support, but has its adherents. Scarborough, Cortese, and Scarborough (1977) criticized Kirsner and Smith (1974) for using presentation conditions likely to enhance modality-specific encoding. The implication of their argument is that a common lexical system is invoked during identification of auditory and visual stimuli, and that intermodality transfer should be equal to intramodality transfer when less demanding viewing and listening conditions are used.

The first group of three experiments was designed to assess the hypothesis introduced above, namely, that relative to the control and intramodality conditions, the magnitude of intermodality transfer depends on depth of processing. The depth-of-processing argument has been invoked by Jacoby and Dallas (1981) and Morton (Note 1) in regard to retrieval rather than encoding, but both aspects are considered in the first group of experiments.

The retrieval argument is that intermodality transfer is observed in a lexical decision task because such task samples information stored in a modality-free semantic system, as well as in modality-specific systems. The sensitivity of this task to semantic variables has been documented (James, 1975). In contrast, it is claimed that the word identification task does not involve semantic information (Jacoby & Dallas, 1981), and that intermodality transfer should not, therefore, be observed when it is used.

The first two experiments examined the effect of depth of processing as manifested in lexical decision and word recognition. However, these tasks also involve a contrast in stimulus presentation conditions: In word recognition, the information available to the subject is limited by stimulus conditions such as duration or amount of masking, whereas in lexical decision information is available until a response is made. Here, these conditions are termed stimulus limited and response limited. Because this contrast represents a source of potential confounding in the interpretation of depth-of-processing effects, a third experiment involving a stimulus-limited lexical decision task was included. Thus, Experiments 1, 2, and 3 involved word identification, lexical decision^{SL} (SL = stimulus limited), and lexical decision^{RL} (RL = response limited), respectively. Given the relationship between task and representation level summarized above, it was anticipated that intermodality transfer would be observed in Experiment 3 and, depending on the effect of stimulus limitation on semantic access, in Experiment 2.

Because the extent of intermodality transfer may also reflect the depth or type of processing invoked at encoding, an appropriate manipulation was included in the

design of each experiment in the first group. Given the notion of a supramodal semantic system, a plausible contingency is that intermodality facilitation depends on the conjunction of: (1) semantic processing at encoding, and (2) a semantically sensitive task, such as lexical decision, at retrieval. The selected encoding variable involved the contrast between labial and semantic judgments. In the labial conditions, the subjects were required to decide whether or not each word (seen or heard) normally contained sounds that required fully closing the lips. The subjects were not given direction to particular sounds, but, in practice, full closure of the lips is restricted to the letters p, b, and m. The labial task was chosen because, intuitively at least, it provided a "shallow" judgment that did not depend directly on input modality. In the semantic conditions, the subjects were required to decide whether or not each word involved an "animate" or an "inanimate" concept. Recognition memory judgments were collected in a separate block of trials at the end of the experiment in order to evaluate the success of the depth-of-processing treatment.

Experiments 1, 2, and 3 were also designed to meet Scarborough et al.'s (1977) objection to tasks that emphasize modality-specific encoding. Thus, visual stimuli were presented under conditions that could not be construed as degraded, as Scarborough et al. claimed of Kirsner and Smith's (1974) use of a television monitor. Scarborough et al. also criticized the use of orthographically and phonologically illegal nonwords, which may have enabled lexical decisions to be made without recourse to semantic information. In the following experiments, all nonwords were legal in this sense.

Finally, the following experiments were designed to examine another criticism of Kirsner and Smith's (1974) results, that recognition of words encoded auditorily is inferior because isolated words in that modality are more difficult to recognize. One way to remove the effect of this difference is to look at results conditionalized on correct recognition of the priming stimuli.

EXPERIMENTS 1, 2, AND 3

There were three blocks of trials for all subjects in the first three experiments. The first, or encoding, block involved either labial classification (lips-open vs. lips-closed words) or semantic classification (living vs. man-made words). The second block of trials defined the experiment and consisted of either a word identification^{SL} (Experiment 1), a lexical decision^{SL} (Experiment 2), or a lexical decision^{RL} (Experiment 3) task. The third block of trials involved a recognition memory judgment (old vs. new) for all subjects.

Independent groups of subjects were used in the six conditions resulting from the factorial combination of experiment (3) and encoding treatment (2). Repeated measures were used to evaluate three repetition con-

ditions: VV, AV, and CV (control condition involving new visual items) in the second and third trial blocks.

Method

Stimuli. Thirty-six words were chosen from the four categories resulting from the combination of the living/man-made and the labial/nonlabial classifications. The items were chosen with the following constraints: (1) Thorndike-Lorge (1944) frequency values were between 1 and 26 (inclusive); (2) the number of letters ranged from four to nine (inclusive); and (3) homophones and homographs were excluded. Sixty-four false words were constructed from a similar set of words by changing one or two letters while observing the phonological and orthographic constraints of English.

Ninety-six words (24 from each of the four subsets defined by the combination of the semantic and labial criteria) were presented in the first block of trials. Twelve words from each subset of 24 were presented in each modality, which provided the basis for the VV and AV comparisons that were made on the data from the second and third trial blocks. Words from the auditory and visual subsets were selected and presented randomly.

The same words were used for the labial and the semantic classification conditions: only the instructions determining which of these was diagnostic for each subject. Apart from the constraints imposed by the labial and semantic classifications, the words were allocated randomly to the VV, AV, and CV conditions and, within each of these, to repetition in the second or third block.

The second block consisted of 96 words: 8 from each combination of the semantic, labial, and repetition (VV, AV, and CV) variables, together with 48 false words in the case of the lexical decision experiments. These stimuli were all presented visually.

The third block of trials involved a recognition memory task with 32 old words and 16 new words. There were 8 old words, 4 per repetition condition (VV and AV), from each of four word subsets. These stimuli were also presented visually.

Procedure. In the visual condition, the words were presented on a BWD 594 cathode ray oscilloscope (P31 phosphor) in the center of the visual field. Each visually presented word was presented in lowercase letters, subtended 2-3 deg on the horizontal axis, and remained on display for 500 msec. In the auditory condition, the words were presented via an intercom from the experimenter, who read each word as it was presented on a video monitor in a separate room. The two levels of modality (auditory and visual) and orienting task (living and nonliving, or labial and nonlabial) were sampled in random order.

In the second block of trials, all of the stimuli were presented visually. In the word identification^{SL} and lexical decision^{SL} experiments, each word was presented for 40 msec and was followed by a pattern mask consisting of letter fragments for 100 msec. The words were presented in uppercase letters to preclude physical matches. The mask occupied a fixed region covering 0.6 and 3.0 deg on the vertical and horizontal axes, respectively, and completely "covered" nine-letter words. The mask for Experiments 1-3 consisted of 52 20-point letter fragments (one per lower- and uppercase letter) drawn from a Horatio Light character set in which each character occupied between 30 and 165 points.

The display conditions in the lexical decision^{RL} experiment and in the third, recognition memory block of trials followed those described above, except that each word was presented for 2 sec (3 sec in recognition memory) and no mask was presented at all. The subjects were instructed to press one of two designated buttons for the labial and nonlabial (or living and nonliving) conditions. Depending on the task, different response rules were used in the second block of trials. In the word recognition experiment (Experiment 1), the subjects were instructed to name the stimulus, if possible, or to say "don't know" if they could not recognize it. The experimenter, who

was located in a separate room, checked each response as correct or incorrect. In the lexical decision experiments (Experiments 2 and 3), the subjects pressed a designated button whenever a genuine English word was presented. They were instructed not to respond whenever a false word was presented. In the third block of trials, the subjects in all three experiments were asked to indicate whether or not each word was old (had been presented in Block 1) or new (had not been presented in Block 1) by pressing designated buttons. Independent subsets of the words used in Block 1 were repeated in the second and third blocks of trials.

Stimulus selection, stimulus presentation (apart from the auditory condition in Block 1), and data acquisition were controlled by a PDP-11/10 computer.

Subjects. Fourteen undergraduate students at the University of Western Australia volunteered to participate in each of the six conditions. Their ages ranged between 19 and 40 years. In the stimulus-limited experiments, any subject with an accuracy score outside the expected range (for $\alpha = .05$, $p = .5$, and 24 trials) in the control condition was replaced.

Results

Block 1: Priming. The results from the first block of trials were analyzed in terms of experimental group (a dummy variable in Block 1), task (labial or semantic classification), and modality (visual or auditory). Accuracy in the labial condition (84.5%) was significantly less than in the semantic condition (94.6%) [$F(1,78) = 25.9$, $MSe = 167.7$, $p < .01$], and the same disadvantage was evident in the RT data as well (2,293 msec vs. 1,583 msec). However, the main effect of modality ($F < 1$ in the accuracy data) and the interaction between modality and the between-subjects variables were not significant. The actual accuracy values for the auditory and visual conditions were 89.6% and 89.5% correct, respectively.

The auditory RT values included the experimenter's reading time (RT was calculated from visual onset), but this variable was orthogonal to task in Block 1 and was of no relevance to the rest of the study.

Block 2: Word identification and lexical decision. The results from the three experiments are summarized in Table 1. Because the three sets are qualitatively comparable, they are described in one section.

The results for the labial and semantic groups are identified separately in Table 1, but when consideration is restricted to the facilitation effects of the VV and AV conditions relative to the CV condition, it is apparent that the encoding manipulation had no consistent effects on performance. In the statistical analyses on the data from each experiment, the crucial term, the interaction between repetition status and encoding treatment, was not significant in any instance [$F(2,52) = 1.65$, $MSe = 69.3$, $p > .05$, in the word identification^{SL} experiment; $F(2,52) = 2.09$, $MSe = 24.9$, $p > .05$, in the lexical decision^{SL} experiment; and $p > .05$ in the lexical decision^{RL} experiment]. The assumption that the effects of repetition modality are independent of encoding treatment cannot be rejected in any one of the experiments. Encoding treatment was associated with a significant main effect in the word identification experiment [$F(1,26) = 4.18$, $MSe = 959$, $p < .05$], but be-

Table 1
Experiments 1, 2, and 3: Mean Accuracy (Percent Correct) and Reaction Time (in Milliseconds) Values
as a Function of Task, Type of Limitation, Priming Condition, and Repetition Condition

Task	Type of Limitation	Encoding Treatment	Repetition Condition		
			CV	AV	VV
Experiment 1					
Word Identification	Stimulus	Labial	49.0	61.2 (61.1)	67.9 (68.1)
		Semantic	67.4	73.2 (74.3)	78.8 (79.4)
		Mean	58.2	67.2 (67.7)	73.3 (73.7)
Experiment 2					
Lexical Decision	Stimulus	Labial	59.1	65.8 (66.6)	70.5 (70.8)
		Semantic	58.7	70.8 (71.3)	73.7 (73.4)
		Mean	58.9	68.3 (68.9)	72.1 (72.1)
Experiment 3					
Lexical Decision	Response	Labial	707	678 (677)	622 (624)
		Semantic	666	645 (637)	625 (612)
		Mean	687	662 (657)	623 (618)

Note—The accuracy and RT values for conditionalized data (correct in Block 1) are shown in parentheses. The repetition condition was visual in Blocks 1 and 2 (VV), auditory in Block 1 and visual in Block 2 (AV), or a control with only a visual presentation (CV) in Block 2.

cause the biggest advantage for the semantic priming condition occurred in the new-word condition, it must be attributed to subject rather than to priming differences.

The repetition effect was consistent: The VV, AV, and CV conditions ranged in that order from efficient (fast or accurate) to inefficient in all three experiments. Comparisons using the Newman-Keuls procedure indicated that the differences between the VV and the AV conditions, as well as those between the AV and CV conditions, were statistically reliable in each experiment.

The statistical analyses presented above are based on data that are not conditionalized on Block 1 accuracy. However, as shown in Table 1, the above summary does not depend on this. The results are quantitatively, as well as qualitatively, similar when the calculations for each subject are conditionalized on correct judgments in the first block of trials.

The RT analyses in the lexical decision^{RL} experiment were conducted on means calculated from correct observations for each condition. The accuracy values for the VV, AV, and CV conditions in this experiment were 97.3%, 94.2%, and 93.5%, respectively, and an analysis of variance showed that the main effect of repetition was significant [$F(2,52) = 8.23$, $MSe = 14.6$, $p < .05$].

Block 3: Recognition memory. The results from the third block of trials showed a consistent accuracy and speed advantage for the semantic condition. The accuracy (RT) values after labial processing were 60.9% (1,061 msec), 61.8% (1,069 msec), and 78.9% (1,195 msec) correct for the VV, AV, and CV conditions, respectively. The corresponding values after semantic processing were 78.4% (1,008 msec), 77.1% (999 msec), and 81.7% (1,162 msec) correct, respectively. The overall accuracy difference between the groups was significant [$F(1,78) = 34.5$, $MSe = 259.2$, $p < .01$], but the RT difference was not ($F < 1$).

Discussion

The results of Experiments 1-3 are readily summarized. The depth-of-processing treatment had a clear effect on the orienting tasks. The semantic classification task was completed more rapidly and more accurately than the labial classification task, and both speed and accuracy were superior in the recognition memory test following semantic classification. However, the depth-of-processing manipulation had no consistent effect on the relative positions of the VV, AV, and CV conditions in the word identification and lexical decision experiments. In all cases, the $VV > AV > CV$ order was maintained, and the differences were statistically significant.

The results are consistent with the claim that there is a visual word recognition system (Morton, 1979), the units of which are primed selectively by visual experience of the appropriate items. However, a stronger claim, that word presentation primes only modality-specific representation, was not supported. The presence of intermodality facilitation (cf. Kirsner & Smith, 1974) suggests either that some intersystem transfer occurs or that there is also a common, modality-free system (Forster, 1976).

The results do not support Scarborough et al.'s (1977) imputation that when appropriate caution is observed in regard to the equivalence of the initial perceptual conditions for isolated auditory and visual words, equivalent inter- and intramodality facilitation effects will be observed. Classification accuracy in the encoding block of trials was equal for auditory and visual stimuli, and the overall outcome was qualitatively unchanged by restricting consideration to items that were classified correctly in the first block of trials.

Finally, the results do not support the view that the word identification and lexical decision tasks sample qualitatively different levels of representation. Assuming that the lexical decision task sampled information in a

semantic system (James, 1975), whereas the word identification task depended on processes in modality-specific word recognition systems, it was expected that intermodality transfer would be greater, or complete, under lexical decision conditions. No such contingency between task and transfer was observed.

EXPERIMENTS 4, 5, AND 6

The experiments in the second group were designed to explore the generality of our conclusions about the pattern of facilitation effects in word identification in regard to three additional variables: word frequency, mask type, and the form of the subjects' responses in the priming phase.

Previous studies involving the repetition effect have shown that it is smaller for high-frequency words than for low-frequency words. But interpretation of this contingency is qualified by the different performance levels observed for the two word classes in both word identification (Jacoby & Dallas, 1981) and lexical decision^{RL} (Scarborough et al., 1977). That is, the larger facilitation effects observed for low-frequency words may simply reflect the fact that there is more room for improvement in speed and accuracy for this word class. This problem can be overcome for accuracy, at least if word presentation is controlled so that the level of performance is equated for unprimed low- and high-frequency words. Experiments 4, 5, and 6 included high- and low-frequency conditions and, by using an interactive psychophysical procedure, PEST (Pentland, 1980), to obtain separate thresholds for high- and low-frequency words for each subject, ensured comparable performance levels for each word type.

Marcel (1980) demonstrated that mask type can be an important determinant of priming effects when semantically related words are presented in the second temporal position in a three-item series. Specifically, he found semantic facilitation when a pattern mask was used to restrict processing in the second item, but not when an energy mask was used. Our experimental procedure does not provide for priming in this sense, but it could be claimed that the pattern mask does not completely terminate processing and that reference to a supramodal lexical system accounts for the transfer that was observed in the AV conditions in Experiments 1 and 2. To evaluate this account, Experiments 4 and 5 used energy and pattern masking, respectively.

Experiments 4, 5, and 6 involved a different approach to the priming phase of the experiment. The subjects in Experiments 4 and 5 were simply asked whether or not they were familiar with each auditory and visual word as it was presented. Their manual (familiar/unfamiliar) responses were recorded to provide a further check against the possibility that the lesser magnitude of facilitation in the AV condition reflects peripheral difficulties in identifying isolated spoken words. In Experiment 6, a further check on comprehension was

provided by asking subjects to repeat familiar words (whether heard or seen) and to respond "don't know" to unfamiliar words. The experimenter then recorded a familiar response if the subject's response matched the stimulus word. This naming procedure also permitted confirmation of Morton's (1979) claim that, even when subjects name the target stimuli in the AV condition, less transfer is observed than in the VV condition.

Method

Design. There were three blocks of trials for each experiment in this group. The first block involved an encoding task in which subjects indicated by a manual response (Experiments 4 and 5) or an oral response (Experiment 6) whether each word was familiar or not. One hundred and twenty words were presented in this block (30 in each condition resulting from the factorial combination of the two modalities and the two frequency levels). In the second and third blocks, only visual presentation was used, and the trials in the two blocks were contiguous. The second block, of between 20 and 55 trials per frequency condition, was used to establish threshold values for new high- and low-frequency words. The third block was designed to assess repetition effects. There were three repetition conditions: VV, AV, and CV, with 30 words per frequency level in each of these conditions.

Stimuli. One hundred and fifty-five words were chosen from each of two frequency bands in the list described by Gilhooly and Logie (1980). The frequency counts for the high- and low-frequency items were 50 or more per million and between 1 and 26 per million (inclusive), respectively. All of the items were chosen from moderate concreteness (C) levels ($2.91 > C > 4.31$), with four to nine letters (inclusive). Homophones and homographs were excluded.

Sixty-five words from each set were randomly chosen for use in the second block of trials (PEST). This set was used to establish high- and low-frequency threshold values for all subjects. These words were never used in the first and final blocks. The remaining 90 words in each frequency set were divided into three 30-item subsets. Each subset was subsequently used in the VV, AV, and CV conditions for four subjects in each experiment. Thus, over all subjects in each experiment, a common word set was used in each of the three repetition conditions.

Items were selected in random order from the repetition and frequency conditions in each block.

Procedure. The display procedures used for auditory and visual presentation in the first block followed the procedures used for the first block in Experiments 1-3. The subjects were instructed to classify, using designated buttons, each item as either familiar ("known") or unfamiliar ("unknown").

From the point of view of the subject, the second and third blocks of trials were contiguous, and consisted of between 248 and 328 word identification trials. Each trial involved a brief presentation of a word followed by either a pattern mask (Experiments 4 and 6) or an energy mask (Experiment 5). The subjects were instructed to name each word, if they could, or to respond "don't know," if they could not. The stimuli were presented in uppercase letters on a Burwood 594 cathode ray oscilloscope (P31 phosphor) under computer control, and the experimenter entered the responses via a terminal in a separate room.

From the point of view of the experimenter, the first 20 trials in the second block were conducted to provide the subjects with practice at comparatively long stimulus onset asynchronies (SOAs), and to establish upper SOA boundaries for high- and low-frequency words for the subsequent psychophysical analysis. The next 48-120 trials (24-60 at each frequency level) were conducted to estimate the 50% threshold SOAs for high- and low-frequency words for each subject. The threshold-estimation procedure was similar to that used by Pentland (1980), in which,

on every trial, a maximum-likelihood estimate of the threshold was calculated and then used as a test value for the next trial. In the high-frequency case, for example, the subjects were tested at each of 12 SOA values calculated by PEST to converge on the best solution. A lower boundary of 3 msec was used in all cases. The upper boundary for PEST was typically reduced from 100 to about 40 msec during the practice trials. In practice, between two and five trials were given at each SOA before a threshold estimate was made in order to minimize the impact of item and subject variability on threshold estimation. Following the practice and threshold-estimation trials, another 180 trials (2 frequency levels x 3 repetition conditions x 30 words per condition) were then presented at the SOAs established for low- and high-frequency words.

The pattern mask was constructed in the same way as that used in Experiments 2 and 3. The energy mask consisted of 1,040 points randomly distributed in the area (0.6 x 3.0 deg of visual angle at about 40 cm) used for the pattern.

Subjects. The subject selection procedure was identical to that used in Experiments 1, 2, and 3. Twelve undergraduates subjects were run in each experiment. The same criteria for acceptable performance in the control condition of the word recognition task were applied as in the previous experiments.

Results

Because the pattern of results was qualitatively similar for Experiments 4, 5, and 6, and because there were no significant main effects or interactions in an analysis of variance that included experiment as a factor, the results of the three studies are considered together.

Block 1: Encoding. The mean familiarity scores for the high-frequency/visual, high-frequency/auditory, low-frequency/visual, and low-frequency/auditory conditions were 98.3%, 97.1%, 83.2%, and 83.5% familiar, respectively. An analysis of variance indicated that the effect of frequency was significant [$F(1,33) = 105.1$, $MSe = 70.5$, $p < .01$], but that the effect of modality and the interaction between frequency and modality were not significant.

Block 2: Thresholds. The means of the PEST-derived threshold values were 11.3 and 20.3 msec for the high- and low-frequency word sets, respectively. As suggested

by the correct identification values for new (CV) words in the third block, this manipulation was effective (the means approached 50% correct in each condition). However, 15 subjects were rejected, and replaced, because one or both of their CV values were outside the range given by the binomial test for $p = .5$, $n = 30$, and $\alpha = .05$.

Block 3: Word identification. As shown in Table 2, the results of Experiments 4, 5, and 6 are qualitatively similar to those observed in the first three experiments. The percentage of words correctly identified followed the $VV > AV > CV$ order for both word frequency conditions in all three cases.

The statistical analysis was conducted in two stages. The first stage was designed to establish whether or not significant facilitation was occurring in each of the four repetition conditions. This comparison was conducted on unconditionalized data because it involved comparisons with the CV condition, for which conditionalization on Block 1 performance was not possible. A three-way analysis of variance (3 experiments x 3 repetition conditions x 2 word frequency levels) indicated that the main effects of repetition [$F(2,66) = 88.92$, $MSe = 86.1$, $p < .01$] and frequency [$F(1,33) = 3.46$, $MSe = 293.9$, $p < .05$], and the interaction between these variables [$F(2,66) = 5.77$, $MSe = 74.3$, $p < .05$], were significant. Newman-Keuls tests indicated that: (1) the differences between the VV and AV conditions were significant in the high and low word frequency conditions, (2) the differences between the AV and CV conditions were significant in the high and low frequency conditions, and (3) the differences between the high and low frequency conditions were significant in the VV and AV conditions, but not in the CV condition. This final outcome is, of course, expected under PEST.

The second stage of the statistical analysis was conducted to evaluate the repetition effects in the conditionalized data. For this purpose, an item was included

Table 2
Experiments 4, 5, and 6: Word Identification (Percent Correct) as a Function of Response Mode, Mask Type, Word Frequency, and Repetition Status

Response Mode	Mask Type	Word Frequency	Repetition Condition		
			CV	AV	VV
Experiment 4					
Manual	Pattern	High	51.1	60.6 (61.1)	70.6 (70.4)
		Low	49.7	61.9 (65.5)	73.9 (78.0)
Experiment 5					
Manual	Energy	High	50.9	55.3 (55.9)	72.2 (72.1)
		Low	47.6	67.1 (71.2)	74.6 (85.9)
Experiment 6					
Oral	Pattern	High	54.7	60.3 (61.2)	65.4 (66.0)
		Low	56.1	72.1 (77.1)	77.1 (82.5)
Experiments 4, 5, and 6					
		High	52.2	58.7 (59.4)	69.4 (69.5)
		Low	51.1	67.0 (71.3)	75.2 (82.1)

Note—Accuracy values for conditionalized data ("familiar" in Block 1) are shown in parentheses. The repetition conditions are described in Table 1.

in the calculation of individual accuracy scores in Block 3 repetition conditions only if the item had been classified as "familiar" in Block 1. A three-way analysis of variance (3 experiments x 2 repetition conditions x 2 word frequency levels) indicated that the main effects of both repetition [$F(1,33) = 33.2$, $MSe = 118.5$, $p < .01$] and frequency [$F(1,33) = 20.4$, $MSe = 163.6$, $p < .01$] were significant, but that the interaction between these variables was not ($F < 1.0$).

Before the implications of these data are discussed, two further experiments are reported. The first, involving AV priming only, was conducted to examine the outstanding factor separating our experimental procedures from those reported by Morton (1979). The second, involving only the AV and VV conditions, was conducted to ascertain whether or not low-frequency words receive any additional benefit in the intramodality condition that cannot be accounted for by reference to a common source of variance in the intermodal condition.

EXPERIMENT 7

In all of the experiments described above, the auditory and visual words were presented together in a mixed sequence during the encoding phase. This situation is very different from everyday conditions in which attention is not rapidly switched between modalities. It is possible that the intermodal priming observed there reflects the operation of attentional or other task-specific processes invoked in response to mixed-modality presentation. Experiment 7 overcame this limitation by restricting priming to auditory words.

Method

The design, stimulus, procedure, and subject selection conditions were identical to those used in Experiment 4, except that: (1) no visual words were presented in Block 1, and (2) only 120 words (2 repetition conditions x 2 word frequency conditions x 30 items per condition) were presented in Block 3.

Results

In Block 1, 97.8% and 87.8% of the words in the high- and low-frequency word conditions, respectively, were classified as "familiar." In Block 2, mean threshold values of 12.8 and 21.9 msec were obtained for high- and low-frequency words, respectively. The accuracy values for the critical conditions in the third block of trials are summarized in Table 3, in which it is apparent that facilitation is present in the AV conditions. An analysis of variance conducted on the unconditional data confirmed this, with a significant main effect of repetition [$F(1,11) = 46.78$, $MSe = 72.4$, $p < .01$] and a significant interaction between repetition and frequency [$F(1,11) = 7.59$, $MSe = 51.2$, $p < .05$]. Planned comparisons indicated that significant facilitation occurred in the AV conditions for high-frequency, as well as for low-frequency, words.

Table 3
Experiments 7 and 8: Word Identification (Percent Correct) as a Function of Word Frequency and Repetition

Word Frequency	Repetition Condition		
	CV	AV	VV
Experiment 7			
High	53.6	64.7 (65.2)	
Low	46.1	68.6 (71.8)	
Experiment 8			
High		(48.2)	(57.7)
Low		(50.1)	(60.4)

Note—Accuracy values for conditionalized data ("familiar" in Block 1) are shown in parentheses.

EXPERIMENT 8

The data summarized in Table 2 show that low-frequency words are more sensitive to priming than are high-frequency words (see Scarborough et al., 1977), and they suggest that most, if not all, of this benefit is achieved under intermodal repetition conditions. Intramodal repetition yields additional facilitation, but the magnitude of this increment is actually smaller for low-frequency words (8.2%) than for high-frequency words (10.7%). Because the proposition that frequency effects influence the intermodal, rather than the intramodal, component runs directly counter to Forster's (1976) two-stage characterization of word recognition processes, Experiment 8 was conducted to reevaluate this aspect of the relationship between modality and frequency.

If consideration is restricted to the AV and VV levels of the repetition factor, the claim that the frequency effect influences only the intermodal component yields a hypothesis that the relationship between modality and frequency is additive. If anything, the relationship observed in Experiments 4-6 was underadditive, but both this and an overadditive relationship observed under the same conditions are or would be suspect because of the nature of the percentage scale. Experiment 8 overcame this problem by applying PEST to the AV, rather than to the CV, condition. With recognition values equated at or near 50% correct under AV conditions, a significant interaction between modality and frequency would demonstrate that frequency is affecting the intramodality, as well as the intermodality, component.

Method

The design, stimulus, procedure, and subject selection conditions were identical to those used in Experiment 4, except that: (1) there were 80 visual words (40 per frequency condition) and 210 auditory words (105 per frequency condition) in Block 1, (2) the words used in PEST in Block 2 were all old and "familiar" auditory words from Block 1, and (3) only the AV and VV conditions were tested in Block 3. Two subsets of words were selected randomly; six subjects used one subset for VV and the other for AV, and the remaining subjects used the subsets

in the alternative way. The same subsets of high- and low-frequency words were allocated to PEST (auditory in Block 1, PEST in Block 2, not presented in Block 3) for all subjects.

Results

In the first block of trials, 99.6% and 99.2% of the visual and auditory high-frequency words were classified as familiar. The equivalent values for the low-frequency words were 95.4% and 92.6%. In the second block of trials, threshold values of 12.0 and 15.6 msec were obtained for the high- and low-frequency words, respectively.

The accuracy values for the critical conditions in the third block of trials are summarized in Table 3, in which it is apparent that the magnitude of facilitation under VV is virtually the same for the high- and low-frequency word conditions. Because the CV condition was omitted, the analysis of variance (2 modalities \times 2 frequencies) was conducted on conditionalized data. In this analysis, the main effect of repetition was significant [$F(1,11) = 10.6$, $MSe = 110.7$, $p < .01$], but the main effect of word frequency and the interaction between repetition and word frequency were not significant ($F < 1.0$ in each case).

DISCUSSION

The results of the eight experiments demonstrated that: (1) significant facilitation occurs under AV as well as under VV conditions, (2) more facilitation occurs under VV than under AV conditions, and (3) the additional facilitation effect generally observed for low-frequency words exerts its influence equally under inter- and intramodal conditions. The results also indicate that repetition effects observed for the VV and AV conditions are insensitive to: (1) a variety of encoding tasks designed to emphasize different properties of words and (2) differences in the ease of encoding of isolated auditory and visual words.

The results are inconsistent with Morton's (1979) claim that repetition effects are restricted to intramodality conditions. However, they do not disconfirm his theory that modality-specific recognition units are involved in word identification, and in the repetition effect in particular. Perhaps the most obvious reconciliation of the AV data and Morton's theory is to assume that repetition effects are determined at two levels. The first of these is modality-specific, and accounts for the difference between the VV and AV conditions. The second is modality-free, and accounts for the facilitation effects observed in the AV condition. The results are also consistent with the further hypothesis that frequency effects are restricted to the second, modality-free system. The results also suggest that the facilitation difference between the high-frequency and low-frequency word conditions is entirely revealed in the AV modality combination, indicating that the word frequency effect is determined in a common process.

The two-stage account summarized above is broadly consistent with Forster's (1976) claim that word recognition involves: (1) provision of an address in a modality-free master file directory by peripheral orthographic and phonological access files and (2) access to the appropriate entry in the master file directory. In one critical respect, however, our account must differ from Forster's. Whereas he proposes that the peripheral access files involve bins of frequency-organized items with a common stem or feature, our data suggest that word frequency has its effect in the subsequent, modality-free stage.

Two problems are raised in attempting to modify Forster's (1976) and Morton's (1979) models to account for the present findings. One is that identifying frequency effects with a modality-independent system is contrary to the spirit of both models, in which these effects are central to the nature of the logogen or peripheral file system and thus to the model itself. Related to this is the second problem of identifying the nature of the modality-free system presently proposed. Hypotheses concerning this can be grouped into two classes involving one- and two-level structures, respectively.

A single-level account can be invoked to explain intermodal transfer in terms of either auditory-to-visual translation during encoding or the use of a common phonological code during the encoding and test phases. There is little evidence of auditory stimuli's being translated into a visually based format, although Jacoby and Witherspoon (1982) showed that requiring subjects to spell a visually presented word results in substantial transfer to a visual test condition. There is, however, no evidence that this occurs in the absence of appropriate task demands. The alternative account, that intermodal transfer effects reflect access to a phonologically based code during the priming and test phases, is difficult to evaluate, given the discrepant findings in the literature on phonological recoding (see McCusker, Hillinger, & Bias, 1981, for a recent review). In neither translation explanation, however, is there any mechanism to account for the evident contingency between frequency and modality. On the contrary, assuming that auditory and visual frequency accounts are highly correlated, it follows that translation and frequency effects will be independent. The translation accounts may therefore be rejected.

The second explanatory class involves two-level models in which the frequency-sensitive mechanism is located in the second, modality-independent stage. This general characterization raises several issues about the type of information represented at each level.

With respect to access, information-gathering devices of the type proposed by Morton (1969) appear to offer a simpler, and therefore preferable, solution than the solutions offered by search models. The general problem for serial-search models of the lexicon is that if their structurally determined organization (e.g., into bins in

Forster's, 1976, model) and serial-search principles (based on frequency) cannot be confined to the modality-specific level, the economies provided by these features are lost. The modality-specific peripheral files cannot simply provide the address of lexical entry, but, rather, a serial search would be required through the master file as well, and unique organizational principles would need to be specified for each level. Furthermore, a preliminary review of the possible organizing principles suggests that there may be too many of them, since it is not only necessary to account for frequency effects but also for a variety of semantic, morphological, grammatical, and concreteness effects. Different organizations would be appropriate for different tasks, and a set of presumably parallel or substitutable search options would have to be specified.

Assuming, then, that information-gathering devices or logogens are located at two levels, reflecting modality-specific and modality-independent effects, respectively, what information is represented at each level? For convenience, we will draw on the characterization provided by Rumelhart and McClelland (1981). Ignoring preliminary stages of analysis, it is proposed that the modality-specific level consists of lists of perceptual features that are: (1) sufficiently general to handle variations in position, orientation, color, and case (see Scarborough et al., 1977) and (2) insufficiently general to handle variations across modality (speech and text, or text and objects). These units do not reflect frequency effects as specified by Morton (1979), although frequency effects might still be obtained at this level when letter cluster or phoneme frequency are not controlled.

Although the nature of the information represented at the modality-independent level is less clear, three points can be made. First, any representation of semantic information must be restricted to word-specific features. There is no evidence of transfer across synonyms or translations (Kirsner, Brown, Abrol, Chaddha, & Sharma, 1980; Kirsner, Smith, Lockhart, King, & Jain, in press), and it appears that the facilitation of semantically related words such as DOCTOR and NURSE (Meyer & Schvaneveldt, 1971) does not persist over the duration of the effects described here (Dannenbring & Briand, 1982). Second, to account for frequency effects related to the summed frequencies of all members of a family of morphophonemically related words (Bradley, 1979), and for priming effects between members of such families (Murrell & Morton, 1974; Stanners, Neiser, Herson, & Hall, 1979; Downie, Milech, & Kirsner, Note 2), it is proposed that units at this level represent word stems rather than whole words. Finally, the present data indicate that units at this level must be frequency sensitive.

When compared with similar experiments involving languages (Harvey, 1981; Kirsner et al., 1980; Kirsner et al., in press), the present results suggest that insofar as lexical organization is concerned, modality is a weaker

variable, providing only a secondary basis for organization. In all of the above studies involving bilinguals (English-Hindi, English-French, and English-Italian, respectively), no repetition effects were observed in lexical decision tasks and other response-limited tasks unless subjects were specifically instructed to consider translations during encoding. In terms of the two-stage description, then, these data presumably reflect access to language-specific units, in the general lexicon, or to a master file directory, in Forster's (1976) terminology. They do not, of course, demand the stronger conclusion that language-specific lexical systems are involved.

The implications of these data for models of visual word recognition may be briefly summarized. First, the proposition that there is a visual, modality-specific mechanism receives general support, confirming previous statements (Kirsner & Smith, 1974; Morton, 1979). Second, contrary to what Morton (1979) found, the results indicate that visual word recognition efficiency is sensitive to prior auditory exposure, indicating that recognition includes reference to a common or modality-free system. Third, the frequency results suggest that "lexical" status should be assigned to the modality-free, rather than to the modality-specific, visual component, an inference in conflict with both Forster (1976) and Morton (1979).

REFERENCE NOTES

1. Morton, J. Personal communication, November 18, 1980.
2. Downie, R., Milech, D., & Kirsner, K. *Unit definition in the mental lexicon*. Manuscript in preparation, 1983.

REFERENCES

- BRADLEY, D. Lexical representation of derivational relations. In M. Aranoff & M.-L. Kean (Eds.), *Juncture*. Cambridge, Mass: MIT Press, 1979.
- DANNENBRING, G. L., & BRIAND, K. Semantic priming and the word repetition effect in a lexical decision task. *Canadian Journal of Psychology*, 1982, **36**, 435-444.
- FORSTER, K. I. Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland, 1976.
- GILHOOLY, K. J., & LOGIE, R. H. Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, 1980, **12**, 395-427.
- HARVEY, R. *The structure of semantic representation in bilinguals*. Unpublished honors thesis, University of Western Australia, 1981.
- JACOBY, L. L., & DALLAS, M. On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology*, 1981, **110**, 306-340.
- JACOBY, L. L., & WITHERSPOON, D. Remembering without awareness. *Canadian Journal of Psychology*, 1982, **36**, 300-324.
- JAMES, C. T. The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, **104**, 130-136.
- KIRSNER, K., BROWN, H. L., ABROL, S., CHADDHA, N. N., & SHARMA, N. K. Bilingualism and lexical representation. *Quarterly Journal of Experimental Psychology*, 1980, **4**, 585-594.

- KIRSNER, K., & SMITH, M. C. Modality effects in word identification. *Memory & Cognition*, 1974, 2, 637-640.
- KIRSNER, K., SMITH, M. C., LOCKHART, R. S., KING, M.-L., & JAIN, M. The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior*, in press.
- MARCEL, T. Conscious and preconscious recognition of polysemous words locating the selective effects of prior verbal context. In R. S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, N.J: Erlbaum, 1980.
- MCCUSKER, L. X., HILLINGER, M. L., & BIAS, R. G. Phonological recoding and reading. *Psychological Bulletin*, 1981, 89, 217-245.
- MEYER, D. E., & SCHVANEVELDT, R. W. Facilitation in recognizing pairs of words: Evidence of dependence between retrieval operations. *Journal of Experimental Psychology*, 1971, 90, 227-234.
- MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, 76, 165-178.
- MORTON, J. Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolars, M. E. Wrolstad, & M. Bouma (Eds.), *Processing of visible language*. New York: Plenum, 1979.
- MURRELL, G. A., & MORTON, J. Word recognition and morphemic structure. *Journal of Experimental Psychology*, 1974, 102, 963-968.
- PENTLAND, A. Maximum likelihood estimation: The best PEST. *Perception & Psychophysics*, 1980, 28, 377-379.
- RUMELHART, D. E., & MCCLELLAND, J. L. Interactive processing through spreading activation. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading*. Hillsdale, N.J: Erlbaum, 1981.
- SCARBOROUGH, D. L., CORTESE, C., & SCARBOROUGH, H. S. Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 1-17.
- STANNERS, R. F., NEISER, J. J., HERNON, W. P., & HALL, R. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 399-412.
- THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

(Manuscript received September 24, 1982;
revision accepted for publication June 24, 1983.)