Cross modality facilitation in tachistoscopic word recognition

ROBERT CLARKE* AND JOHN MORTON††
*Birmingham Medical School, †M.R.C. Applied Psychology Unit, Cambridge

In its old form, the logogen model (Morton, 1969) would have predicted extensive cross-model facilitation of tachistoscopic word recognition. An experiment by Winnick and Daniel (1970) violated this prediction by showing no effects on subsequent visual word recognition of naming a picture or filling a definition. The latter result was replicated, and it was also found that pretraining with handwritten words transferred to printed words from 10-45 min later. A further experiment showed no significant transfer from auditory presentation to subsequent visual recognition. A third experiment, designed to minimize the role of conscious guessing strategies showed a significant cross-model transfer. The logogen model was modified substantially in the light of these data. There are now seen to be modality specific input categorization systems which are separate from the output lexicon and which are identified as the site of facilitation in the experiments.

Morton (1964a, 1969) has developed a functional model of word recognition which has been used to account for a large number of phenomena in word recognition. Various aspects of the model can be used to make both qualitative and quantitative predictions (Morton, 1969). Of particular interest here is the way the model treats the possibility of cross-modality facilitation: the study of this is important because the relations between types of input and output define the unit of recognition:

"Each unit is in effect defined by the information which it can accept and by the response it makes available" (Morton, 1969). The unit, called a logogen ("word producer"), is a hypothetical construct formulated as a working hypothesis to account for available evidence in the field of word recognition.

The logogen is the basic unit in the model. In the 1969 version of the model, shown in Figure 1, the logogen receives input from auditory (A) and visual (V) sensory analysers and from the Cognitive System (S). Each logogen functions as a signal detector which accumulates information from its inputs without regard to the source of the information. When the amount of information rises above a threshold value, the logogen responds, and the corresponding response is made available. In this way the logogen system effects a lexical categorization of the stimulus. One input from the logogen, supposed to be some phonological code, passes to the Response Buffer...
where it may elicit articulatory plans (for vocalization). In addition, the logogen sends some code to the Cognitive System where semantic analysis or other higher level processing can be effected.

![Diagram of the 1969 Logogen model](image)

**Figure 1.** A simplified version of the 1969 Logogen model.

The presence of a context produces a short-term facilitation in the identification of a degraded stimulus with tachistoscopic presentation (Tulving and Gold, 1963; Tulving, Mandler and Baumal, 1964; Morton, 1964b). This is accounted for in the model in terms of the summation of sensory and semantic information to the logogen system. That is, each logogen sums the evidence present which is relevant to itself without regard to its origin. This hypothesis received some support when Morton (1969) produced qualitative predictions concerning the benefits of context which agreed well with the data.

Production of the "available response" by a logogen is seen as causing a change in threshold of the unit. The result of this would be that for some period of time less evidence would be required to recognize the word. For this reason a word presented to a subject (for identification) is found to have a lowered tachistoscopic recognition threshold, if tested again some time after presentation.

This long-term facilitation (up to half an hour or more), which occurs when the identical word or at least the same root morpheme has been experienced, should not be confused with contextual effects. Inasmuch as context, sentential or other, feeds back from the Cognitive System to the Logogen System, its function would be to selectively affect those logogens corresponding to the set of words likely to occur. For this reason the effects of context would have to be of short duration, otherwise its discriminative function would be lost.

A crucial feature of the logogen, already noted, is its indifference to the source of stimulation. This was required for two reasons. At a phenomenal level, it seems that verbal stimuli are clearer in context than in isolation. The clearest demonstrations of this come in listening to speech where we are rarely aware of the relative contributions
of stimulus and context in our perceptual experience of individual words. More dramatic, perhaps, is the clarity with which we hear the words of a song, once they are familiar to us, words which had been indistinct on the first listening. Thus the model postulated that sensory and contextual information interacted directly. Secondly, for this interaction to be effective, sensory and contextual information relevant to a particular unit must be converted to a compatible form. Thus it was that logogens simply counted the number of their inputs which were currently active. All inputs to the system, then, simply reduce to numbers, and the origin of the numbers would be lost.

It was partly for these reasons that Morton postulated a single logogen system responsible for the categorization of both visual and auditory stimuli. An additional justification for the single logogen system was that, in this way, there would be a single source for a given response, regardless of the origin of the response. Thus, the same logogen would be involved following the visual presentation of CHAIR, the spoken "chair", naming a picture of a chair or producing the response "chair" to an incomplete sentence such as "He sat down on the ——" or as a free association to TABLE. In the last three cases the input to the logogen system would be from the Cognitive System. Since facilitation effects were seen as a consequence of a logogen response, and since logogens, by definition, were indifferent to the source of input, it followed that the model would predict equal facilitation following any logogen use. That is, if the response is produced, no matter how, then the effects on subsequent visual recognition for the appropriate word should be equivalent. The same would also be true for auditory recognition.

Winnick and Daniel (1970, Experiment II) investigated response priming in tachistoscopic recognition of words. In this experiment, subjects were primed with typed words (W), pictures (P), or definitions (D) presented on cards. In each case, they had to speak the word appropriate for the stimulus card. This was followed by a free recall trial, in which items presented in condition W were poorly remembered by comparison with those from conditions P and D, presumably due to the relative automaticity of cognitive processing associated with W. The subsequent tachistoscopic recognition of typed words, however, yielded a very different result. Items presented initially as W were recognized with significantly lower exposure durations than words from conditions P and D. The P and D conditions were not different from the Control condition which consisted of words which had not occurred in the priming phase of the experiment.

The recall result makes it less plausible to attribute the differential W priming to a "sophisticated guessing" hypothesis. By this is meant the situation where the subjects see some stimulus information and use it to search their memory of the priming words in order to generate their response. For this explanation to work there would have to be either two memory sources or two kinds of search, one used in free recall, which favoured the P and D stimuli, and one used in the guessing procedure, which only benefited the W stimuli. One way for such a guessing hypothesis to work is to assume a modality-specific component to what is stored. In the recognition task then we would have to suppose that the visual information derived from the stimulus exposure was used directly as a means of accessing material from the store. The W group of words, being themselves in a visual code, either "sensory" or "abstract", could be accessed by
the stimulus-derived information. The P and D groups of priming items, however, would not be accessible by these means. This analysis makes some fairly strong assumptions about the content-addressable nature of the store required for a guessing hypothesis to be tenable. If the free-recall data is also going to be accounted for in terms of the same store, then some other means of access would also be necessary.

Winnick and Daniel restrict themselves to a discussion of the role of an (undifferentiated) "response availability". They used their recall data to reject the idea that their results could be accounted for by response factors and concluded that "perceptual sensitization" had occurred and the W priming "indicates an advantage where words to be recognized appear in the same form on both initial familiarization trials and subsequent recognition tests" (Winnick and Daniel, 1970, pp. 77–79). This is closer to a description of the results than to a theoretical interpretation: but in fact the perceptual sensitization cannot be due to visual correspondence alone, since Murrell and Morton (1974) found that facilitation between words requires a morpheme in common between priming and recognition stimuli: SEED did not facilitate SEES, although SEEN did facilitate SEES. In terms of physical similarity SEED and SEEN are equally close to SEES. This result also weakens the guessing hypothesis which was analysed above. However, Winnick and Daniel's result can be seen as posing a serious challenge to the logogen model since differential facilitation in favour of printed words has been shown. The logogen model would have to predict that W, P and D, each of which elicits a vocalized response, would be equivalent primers.

The first experiment aimed to replicate the essential features of Winnick and Daniel's study. In addition we aimed to test the notion of "perceptual sensitization" by seeing the extent to which the facilitation depended upon the precise visual form, or whether we should think in terms of some more abstract representation. Note that this experiment is not concerned with testing any kind of guessing hypothesis.

**Experiment I**

The purpose of this experiment was to test whether the W priming, demonstrated by Winnick and Daniel, was due to the exact correspondence of pre-training and recognition stimuli. A new pre-training condition was therefore introduced: H—untidy but legible handwriting. The other conditions used were: W—typed words; D—definitions. The picture condition was omitted since both Winnick and Daniel and a pilot study had indicated that it produced qualitatively similar results to condition D: no detectable facilitation. If H causes similar priming to W, the facilitation can be described as lexical, whereas if it causes significantly less, the explanation implied by Winnick and Daniel for W priming would be upheld; direct visual priming due to exact correspondence between pre-training and recognition stimuli. The option that the priming operates on a visual non-lexical representation not coded for format is thought to be ruled out by the Murrell and Morton experiment already referred to.

**Design**

In this and the following experiment, a list of words was compiled, all members having similar recognition thresholds, as tested upon five volunteers. The list was
subdivided into four groups of words approximately equivalent in terms of these thresholds.

The experimental design was a Latin Square with repeated measures. This allows statistical division of within-subject variance into that due to any word group effects remaining after list compilation, and that due to treatment condition. A fourth treatment condition was a Control (C) which consisted of words which were not presented in the priming phase.

Four subject groups were used, the differences in treatment between them reflecting the rotation of treatment conditions among the word groups. In this way, the four replications control for any remaining differences in recognition threshold due to "intrinsic" properties of the item groups.

Method

(a) Subjects
Thirty-six undergraduates served as subjects, nine in each replication. Participation was on a voluntary basis, psychology students were excluded, and all were inexperienced in tachistoscopic recognition. All subjects were normal sighted without glasses.

(b) Materials
The sixteen items used were: anchor, banana, bicycle, butterfly, feather, giraffe, microscope, mountain, necklace, parachute, propeller, submarine, television, thermometer, umbrella and window. Each was stencilled in upper case letters and handwritten in lower case script to give two sets of stimulus cards for presentation. Examples of the handwritten stimuli are shown in Figure 2. It can be seen that the cursive nature of this writing does not allow easy separation into the component letters. Definitions for Condition D were adapted from the Pocket Oxford Dictionary, e.g., window: "Opening in wall provided with glass for admission of light". These were typed onto cards. The definitions did not include any words used in the experiment. The items were grouped and counter-balanced across replications so that each served in each of the three presentations and as a control word as well.

(c) Procedure
Subjects came to the experimental sessions under the impression simply that their "reaction times" were to be tested. To reinforce this idea, and to prevent anticipation of the recall experiment, subjects were encouraged to respond quickly to the cards in presentation and to speak clearly so that a "dummy" tape-recorder could record the reaction times. The surprise registered by most subjects when recall was demanded indicates that any mnemonic strategies were effectively avoided.

The instructions for the original presentation required the subject to pronounce the appropriate word as each card was presented. The total time for presenting the 12 cards (4 each
from conditions W, H and D) was about 2 min. Four “dummy” cards at the beginning and four at the end of the fixed random presentation order were used to minimize possible primary and recency effects.

In the recall condition, the subject was asked to “remember as many of the words you have just spoken as possible, taking as long as you want”. This phase took up to 5 min.

In the final recognition stage, the subject was told that the experimenter wanted to determine how rapidly he could name words appearing in a tachistoscope. No indication was given that the same words would be used as in the pre-training task, and several (at least five) filler words were given first so that the subject could accommodate to the apparatus. Filters to cut down illumination were available for those who had low recognition thresholds for the practice words. This is justifiable since differentials within a subject’s performance are being sought, not absolute values. A fixed random sequence of recognition cards was used with items from the four conditions equally distributed through the list. The instructions included: “I will say ‘Now’ and you press the button, when you are ready, to flash up a word inside the machine. If you think you recognize the word, say immediately what you think the word said. If you have no idea what the word was, say ‘No’... I will continue increasing the length of the flash until you identify the word correctly.” The tachistoscope had a textured background for fixation before the subject pressed the exposure button.

Increments of 10 ms were used, starting with an exposure of 10 ms and increasing until the correct response was produced, at which point the exposure was recorded on the results sheet. All words were correctly recognized by 100-ms exposure. Incorrect identification of words was rare below 20 ms from correct recognition. Thus, although no specific mention of “guessing” was made in the instructions, subjects tended to wait until they were fairly confident in their identification.

Each experimental session lasted a little over 30 min.

Results

(a) Recall

The percentage of words recalled in each category was Typed—24.3%; Handwritten—31.3%; Definitions—67.4%. A Wilcoxon matched-pairs signed-ranks test showed that the differences between the D condition and the other two were highly significant (D vs. W, t = 7, P < 0.01; D vs. H, t = 12, P < 0.01). There was no difference between the W and H conditions.

(b) Recognition

Mean recognition times are given in Table I. These represent the exposure duration of the first correct response. An analysis of variance was carried out on the raw scores.

The between subjects variance can be analyzed into variance due to groups of subjects and variance due to subjects within groups (error). The F ratio is not significant, indicating that there were no significant differences between subject groups.

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean visual duration thresholds in Experiment I (in ms)</td>
</tr>
<tr>
<td>Words (typed)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Threshold</td>
</tr>
</tbody>
</table>
Within subjects variance can be divided into that due to word groups, treatments, word-treatment interaction and error. The only significant F ratio is that due to treatments: $F(3, 96) = 14.3, P < 0.001$.

The Newman-Keuls multiple comparison of means was therefore undertaken. No significant differences were found between W and H or between D and C. All other comparisons were significant: in the usual notation, $WH < DC \ (P = 0.01)$.

**Discussion**

This experiment replicates the Winnick and Daniel finding that words corresponding to previously seen definitions are much better recalled than those seen previously in printed form. This result was anticipated and is presumably due to the greater amount of “cognitive work”, degree of elaboration by the subject of “depth of processing” (Craik and Tulving, 1975) involved in the identification of an item from a definition. It would follow then, from the recall result for condition H, that no significant extra “depth of processing” is required to read handwriting, as compared with printed words! The apparent “reversal effect” whereby the worst recalled condition (W) is best recognized (lowest thresholds) has also been replicated.

The recognition results indicate that a handwritten word seems to be as effective a prime as a typewritten word. The Winnick and Daniel explanation in terms of direct perceptual sensitization has been tested, and found unsatisfactory. Instead we are led to conclude that any visual presentation of a word leads to the same facilitation effect.

The original logogen model, however, would predict that definitions, which also elicit response production, should cause similar facilitation. The present results do not support this prediction: no significant differences between conditions D and C were detected, and the D condition was significantly different from W and H. Definitions can be represented as acting on the logogen system in Figure 1 by providing certain semantic information via the Cognitive System. The difference in facilitation between the word stimuli and condition D can only be explained in terms of an effect at the input to the logogen, since, in Figure 1, there is only one, modality non-specific output. Our data agree with those of Winnick and Daniel in showing that production of an available response does not per se lead to facilitation of word recognition and thus that the model in Figure 1 will not account for the data.

These data thus force a change in the logogen model. The minimum change possible is that shown in Figure 3. In this figure, which is in a simplified form, the original single logogen system has been split into two. The two logogen systems share the functions of the original system. The input logogen system is responsible for the classification of sensory inputs. This process would be facilitated in the presence of context as indicated by the feedback from the cognitive system to the input logogens. It would be the input logogens which would be responsible for the facilitation effects described in the Introduction and shown in our experiment and that of Winnick and Daniel.

The output logogen system would take over the function of providing the phonological code. It would receive an input from the cognitive system (marked “semantic code”), and on receipt of such an input would produce the appropriate phonological code, which would then be sent to the response buffer. There could in addition be direct connection from input logogen system to output logogen system.
FIGURE 3. The logogen model adapted to account for the results of Experiment I.

This has been marked with a dotted line to indicate that it is not actually required by the current data, since it is, in principle, possible for words to be read out loud or repeated by the route passing through the cognitive system. We suppose that there is no facilitation effect in the output logogen system, at least under the conditions of the experiments reported. It remains possible that there would be facilitation effects here over much shorter time intervals. These could be revealed in cross-modality naming experiments. However, Durso and Johnson (1979) could only find very small (8 ms) facilitation effects on word reading time after having produced that word in response to a picture which had been named 3–150 s earlier. In our experiment, however, the interval for any given word between the pre-training experience and the tachistoscopic recognition test was between 10 and 40 min.

We can now interpret the results of Winnick and Daniel and those of our first experiment in the following way. We suppose that the processes of picture recognition have consequences in the cognitive system without affecting the input logogens corresponding to the names of the pictures. We have diagrammed “definitions” in an analogous way since, although the recognition of the individual words in the printed definitions would have to pass through the input logogen system, they would not affect the logogens associated with the test words. The result of presenting a picture or a definition, then, would be that a semantic code would be sent to the output logogen system prior to the production of the appropriate response. Since there was no facilitation from picture naming to word recognition we require that the picture naming has no lasting effect on the input logogen system. Thus we conclude that either the “context” route does not operate under these circumstances or that it is insufficient to trigger the input logogens in the absence of stimulus information.

The lack of difference between the typewritten and handwritten primes indicates
that at the logogen level the two are treated as equivalent. For the present purposes we
simply take this result as one more piece of evidence that the facilitation effect we are
considering is not due to overlap of visual form but reflects a combined sensory and
lexical representation. A similar finding has been reported by Warren and Morton
(1982) who showed that picture identification was facilitated by the prior naming of a
visually different picture which had the same name (e.g. a circus clown and a
Harlequin clown). Since there was no word-to-picture facilitation under the same
conditions, Warren and Morton concluded that there existed an abstract object
representation which was equivalent to the logogen.

We failed to find facilitation of the visual recognition of a word after that word had
been produced in response to a picture or a definition. It now seems essential to
consider another assumption of the original model, namely, that visual and auditory
inputs operate on the same logogen. The experimental consequence of such a
hypothesis, as shown in Figure 3, would be that the prior auditory or visual
presentation of a word should give rise to equivalent priming in a subsequent visual
word recognition test. If it does, then the auditory system would be seen as sharing a
lexical categorization system with the visual system as in Figure 3. If not, then the
model will have to be further altered.

The following experiment will introduce a condition A, where subjects are presented
with items played from tape recorder, in order to test this hypothesis.

Experiment II

The purpose of this experiment was to investigate the effects of auditory priming on
subsequent visual word recognition. A new condition A was introduced in which a
subject had to repeat words presented to him from a tape recorder. The conditions W
and C were retained as fixed points for comparison.

Two further conditions were introduced, Wo and Ao, where subjects were required
to respond to printed and auditory stimuli by vocalizing their opposites. These
conditions were to provide confirmation that facilitation is dependent solely upon the
stimulus and not upon the response as well. Thus the five conditions used were W, Wo,
A, Ao and C.

Method

In order to accommodate the five conditions and to produce stimuli which could elicit
“opposites”, a new and longer list of twenty items was composed, and five approximately
equivalent sub-groups of items were formed as in the first experiment. The twenty items used
were: hard, under, rough, near, right, ugly, open, boy, wide, sharp, nice, easy, deep, light,
king, alive, winter, short, awake and start. The design of this five-condition, five-replication
experiment was identical to that used in the first experiment, except that the recall part was
omitted, main interest focussing on the recognition task. A further sample of forty
undergraduates was used.

Again, a fixed random presentation order was used, with four “dummy” items at either end of
the list. The two types of response (W, Wo and A, Ao) were distinguished in each case by the
orders either “repeat” or “opposite”. In the case of A and Ao, the orders were tape recorded just
before the stimulus, and with W and Wo, the orders were stencilled at the top of each stimulus
card. A remote control switch was hand-held by the experimenter and allowed the presentation
of a single item from the tape recorder, e.g. “Repeat . . . Winter”. After the correct response, the
subject was rapidly shown a visual stimulus, and then another auditory one, and so on until presentation was complete. The rate of presentation of these items varied a little from subject to subject but was about one item per second. The following instructions were given before a practice to demonstrate that the subject understood what was required:

"In the first part of this experiment, I am going to tape-record your responses to various stimuli, in order to measure reaction time. You will be presented alternatively with single spoken words from the tape recorder and single words typed on cards. With each word, you will be given the instruction, either "REPEAT", or "OPPOSITE". When a spoken or written word is preceded by the order "REPEAT", just say the word out loud as clearly as you can. With the order, "OPPOSITE", just say the opposite of the word, NOT the presented word itself. The instructions "repeat" and "opposite" will vary randomly from one word to the next, so it is always important to pay attention to the instructions before each word."

The recognition phase of the experiment was carried out using the same procedure as in Experiment I. This followed about two minutes after the training procedure.

Results

The same "by subjects" analysis of variance was used as for Experiment I. The mean recognition times are given in Table II. There are significant differences between the word groups \( F(4, 140) = 2.60, P < 0.001 \), but this does not complicate interpretation of the results since the word groups/treatment interaction is still non-significant: \( F(12, 140) < 1 \). The effect of treatments was highly significant \( F(4, 140) = 5.68, P < 0.001 \). The Newman-Keuls method was therefore used as a multiple comparison of treatment means and it was found that:

\[
\begin{align*}
W, W_0 &< A, A_0, C; P < 0.05 \\
W, W_0 &< C; P < 0.01
\end{align*}
\]

Thus although the auditory means were lower than controls, the difference could have occurred by chance. No significant differences were found between \( W \) and \( W_0 \) or between \( A \) and \( A_0 \).

Table II

<table>
<thead>
<tr>
<th>Pretraining condition</th>
<th>Mean visual duration thresholds in Experiment II (in ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W ) — read word aloud</td>
<td>39.3</td>
</tr>
<tr>
<td>( W_0 ) — read word; give opposite</td>
<td>40.2</td>
</tr>
<tr>
<td>( A ) — repeat spoken word</td>
<td>43.4</td>
</tr>
<tr>
<td>( A_0 ) — give opposite to spoken word</td>
<td>44.3</td>
</tr>
<tr>
<td>( C ) — control</td>
<td>46.2</td>
</tr>
</tbody>
</table>

Facilitation of visual word recognition following auditory pretraining has not been detected, although the means indicate a slight trend in that direction. This implies that the auditory and visual inputs cannot be regarded as equivalent. The implications of this will be considered at greater length in the Discussion. Finally, it seems that production of an opposite has no significant effect on the priming properties of either visual or auditory stimuli.
Discussion

These experiments on visual word recognition have shown consistently that facilitation is elicited by prior visual presentation of the words. In other conditions no significant facilitation was found in spite of the subjects having made a spoken response of the target word. This was true whether the subjects named a picture, filled a definition or repeated a spoken word.

The original concept of the logogen, as outlined in the introduction, was the unit for categorization and production of an item, with inputs from auditory and visual sensory analysers and "context" relevant to the item (see Figure 1). It was also the site of long-term (30 min or longer) facilitation. The model predicted complete cross-modality facilitation since each production of the response, however caused, should have led to an equivalent decrease in "evidence requirements" (or threshold) for the logogen unit. This system clearly requires modification to account for the differential facilitation in the present experiments.

The model shown in Figure 3 allows us to account for the lack of effect of definition filling and picture naming on subsequent word recognition. However, it predicts complete cross-modal transfer and is thus incorrect. There seem to be three options open to us in modifying this model. In all cases the input logogen system would be split into two, one for visual inputs and one for auditory inputs. These modality-specific input logogen systems will serve the function of categorizing the sensory inputs. The categorization will be in terms of the morphological structure, as indicated by the experiment by Murrell and Morton (1974) referred to in the introduction. Kempley and Morton (1982) have reported similar findings in the auditory modality (see also Morton, 1981). In one, logically possible, model, the output from the visual input logogen system would map onto the auditory input logogen system. This would mean that prior visual stimuli would affect the subsequent recognition of auditorily presented words. This alternative was ruled out by Ellis (1982) and Gipson (unpublished manuscript 'some auditory priming experiments inspired by the logogen model') both of whom found no cross-modality facilitation at all in an auditory recognition task. In these experiments there was an interval of several minutes between prime and test, and both authors reported strong within-modality effects.

In the second alternative the visual input lexicon would be separated from an auditory-phonological lexicon. The latter serves both to categorize spoken input and to produce phonological code for output. The Auditory-phonological lexicon is a modality-specific version of the original logogen system. Models compatible with this version have been proposed by Seymour (1979). Seymour's model has the constraint that there is no lexical route from visual input to a phonological output without involving the Cognitive System. However, such a route appears to be required to account for patients such as that reported by Schwartz, Safran and Marin (1980). This patient could read irregular words without much problem but had virtually no intact semantics. A direct connection from the visual input logogen to the auditory-phonological lexicon is prohibited, because it would predict the cross-modal facilitation which is absent, as discussed in the preceding paragraph. The only options remaining seem to involve some duplication of lexical phonology. There is nothing prohibiting this.

The final alternative is shown in Figure 4. In this model the two sensory-specific
In all the models we have considered, perceptual priming is accounted for by facilitation at the input logogens. The operation of the logogen systems is seen as essentially “passive” (Morton and Broadbent, 1967). This is seen as the normal mode of recognition of words which are presented clearly to subjects as in normal reading or speech perception. However, we must not forget that in the conditions of our experiment there are alternative strategies open to the subjects. The stimulus is presented only briefly and on many occasions, under these conditions, subjects may recognize individual letters or get the shape of the word or its length correct. Under these conditions the subjects may be tempted to guess on the basis of such information as they have. Such accounts have been given by Soloman and Postman (1952), Newbigging (1961), Savin (1963), Neisser (1967), Catlin (1969, 1973) and Nakatani (1970) among others. These authors have usually supposed that this “sophisticated guessing” was the invariant method of word recognition. The mathematical implications of these models compared with a passive model have been tested by Broadbent (1967) and by Morton (1968). In both cases the data favoured the passive model over the guessing model. In addition, our discussion of identity priming in the introduction, concluded that this was an unlikely explanation of the main results.
However, it makes more sense to allow the possibility that under certain circumstances the subjects do guess.

Inasmuch as our subjects guessed, then, we would expect their guesses to reflect their memory of the words which were presented in the pre-training part of the experiment. It is in this way that we account for the trend, although non-significant, towards facilitation of words which have been presented auditorily in the pre-training session. There are a number of ways in which this guessing process could proceed. None of them is easy to represent in the model as illustrated in Figure 4 and all of them would implicate the cognitive system and, perhaps, the output logogen system. However, the existence of this possibility does not affect the main thrust of the current discussion. We are more concerned to demonstrate the primary cause of the facilitation, rather than give an exhaustive account of all possible forms of facilitation. The possibility of guessing, rather, leaves us with methodological problems concerning the conduct of recognition experiments, since the extent to which the subject consciously or unconsciously uses information in his memory will affect the outcome of the experiments.

Indeed, the problem of establishing exactly how much guessing is involved in the recognition task was only fully realized after the experiments were complete. No specific reference to this was made in the instructions to our subjects, but it would presumably be possible specifically to encourage or discourage guessing. For example, one could demand a response for every tachistoscopic exposure, or one could require the subject to be certain of the correctness of his response. One would then observe greater or lesser dominance of the facilitation which occurs in the input systems.

Experiment III

In an attempt to answer some of these questions, an experiment was performed to investigate auditory-to-visual priming in such a way as to minimise the effects of possible guessing strategies. This was done first of all by greatly increasing the number of words in the pre-training session, including 50 filler words. Secondly, in the test phase the subjects saw each word only once. In the first two experiments the repeated presentation of the test stimuli could, in principle, allow hypotheses to be formulated and, later, tested. The third precaution was by means of stronger instructions to the subjects not to guess. Finally we debriefed the subjects with the objective of excluding any who reported having used conscious guessing.

Method

(a) Subjects

Three groups of eight subjects were used. They were drawn from the APU Subject Panel, and were of both sexes and aged 21–60.

(b) Materials

A list of 150 words was selected, ranging from 247 to 42 frequency of occurrence in the Kucera and Francis norms. Words were from 4 to 6 letters long. The words divided into 6 groups, balanced for frequency and word length. Three of these groups were designated as experimental words and three as filler items. Fifty further words were selected, matched to the others for frequency and word length, to act as practice and threshold items.
(c) Procedure

In the priming phase of the experiment the subjects were presented with 110 words and were asked to rate them for their imageability on a five point scale. Half of these words were presented auditorily and half visually, the two modalities alternating. Of these words, 25 of the auditory and 25 of the visual sets were experimental and the rest were fillers. Fillers occurred as the first 16 and the last 6 of the items. In the test phase, subjects were presented with words in a tachistoscope. There were 75 experimental words. Of these, 25 had been presented visually in the priming phase, and 25 had been presented auditorily. A further 25 had been neither seen nor heard, and constituted the Control stimuli. There were also 25 other words which were used as fillers, giving a total of 100 test items. These were ordered so that the four classes of words were equally represented in each set of 4 words.

When the subjects sat at the tachistoscope, they had 40 trials with further filler words, during which time the exposure duration was adjusted to give a performance level of about 40% correct. Then the 100 test stimuli were presented at this exposure duration.

The instructions were worded so as to encourage the subjects not to guess their responses. The wording was: "If a word comes immediately to mind, tell me that. You may sometimes think of a word but not be sure that you definitely saw all the letters. Just tell me the word. Sometimes you will see the word clearly, in which case tell me that. However, if you only see a few letters, just report those; don't try and make a word out of them, just report what you see. Sometimes you might have absolutely no idea what was presented; just say so. Basically, I want you to report exactly what you see, but if you do feel you know what the word was, but aren't sure you actually 'saw' it, then report what the word was. To reflect this difficulty, I'd like you to rate how sure you are of your response on a three point scale."

After the experiment the subjects were asked about their strategies. We wanted to check whether or not any subject had deliberately used his memory of the pretraining in an attempt to recognize the experimental words.

Results

The data from six subjects were not used. These subjects reported deliberately using constructive processes to help them produce word responses when the percept consisted only of letters. As we wanted to separate out such a possibility and leave as pure a facilitation effect as possible, these subjects were replaced.

The data were scored according to the confidence level of the responses. The cumulative scores are given in Table III. It can be seen that confidence levels 1 and 2 were well used. Confidence level 3, "guess", was not much used, but, then, the subjects had been instructed not to guess. At all three levels of confidence there was a significant main effect of conditions (level 1: $F(2, 42) = 10.52, P < 0.001$; level 2: $F(2, 42) = 10.41$, $P < 0.001$; level 3: $F(2, 42) = 5.41$, $P < 0.01$).

<table>
<thead>
<tr>
<th>Pre-training condition</th>
<th>Confidence of response (cumulative)</th>
<th>$2 + 1$</th>
<th>$3 + 2 + 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>0.396</td>
<td>0.641</td>
<td>0.651</td>
</tr>
<tr>
<td>Auditory</td>
<td>0.344</td>
<td>0.549</td>
<td>0.566</td>
</tr>
<tr>
<td>Control</td>
<td>0.276</td>
<td>0.500</td>
<td>0.510</td>
</tr>
</tbody>
</table>

A > C, $P < 0.01$  V > A, $P < 0.01$  V > A, $P < 0.05$
V > C, $P < 0.01$  V > C, $P < 0.01$  V > C, $P < 0.01$
TACHISTOSCOPIC WORD RECOGNITION

$P < 0.001$; level 3: $F(2, 42) = 9.268, P < 0.001$. There were no significant interaction terms. Newman-Keuls analyses were performed and the results can be seen in Table III. The effects of Visual pre-training were significantly different from the Control at all levels. The Visual condition differed from the Auditory at the lower levels of response confidence but the difference failed to reach significance at the most confident level. At this level the Auditory pre-training condition was different from the Control.

An attempt was made to perform a signal detection analysis of the data but the false alarm rates were too low to be reliable. Overall there were only 5 false alarms from the Visually pre-trained words, 6 from the Auditory pre-trained words and 2 from the Control words. These were out of a total of 28,000 opportunities.

Discussion

In spite of all our precautions to eliminate guessing we have found an effect of Auditory priming on the visual recognition of words. We are satisfied that we have eliminated the possibility of deliberate guessing strategies. There seem to be two theoretical options available in the current framework. The first is that during the priming phase the visual input system is affected following the auditory presentation of a word. We still have cause to separate the visual and auditory input systems, however, since the difference between Auditory and Visual priming remains. The second option is that a part, at least, of the priming effect is to be assigned to the post-categorical phase. This could work either by means of "hypotheses" occasionally being produced on the basis of partial information as in the theories of Becker (1976), Newbigging (1961) or Savin (1963) or by the categorization system producing more than one output. In both these cases, final decisions would be made on the basis of some memory representation of the priming phase. These options had been rejected as possibilities by Morton (1968, 1969, 1979) on the grounds that they would predict a $d'$ account of the word frequency effect rather than a criterion account. However, the calculations which led to the acceptance of a criterion account were based on data from the auditory recognition of words. The different positions can be sustained now that the visual and auditory input systems have been separated.¹

It is not possible at this time to decide among the alternative positions. It should be noted, however, that if a post-categorical stage is to be implicated in these facilitation effects it cannot be equated with the stage responsible for results in recognition memory. This follows from an experiment by Jacoby and Dallas (1981) who compared the effects of three different priming tasks on subsequent performance in recognition memory and tachistoscopic word recognition. Recognition memory performance was greatly affected by the nature of the priming task, following a "depth of processing" prediction. Pure recognition, however, was primed to the same extent by all the tasks.

Also, as was pointed out in the Introduction, any post-categorical stage which was effective in producing facilitation would have to be different in kind or in access from

¹ Allport and Funnell (1981) have suggested an alternative solution which effectively maintains a single logogen system and locates the facilitation effects on the inputs to this system, or the processes feeding this system, thus avoiding cross-modal priming. However, the data by Murrell and Morton (1974) with visual stimuli and by Kempsey and Morton (1982) with auditory stimuli indicates that the facilitation is morphological in character. In this case the inputs would have at least two of the main properties currently ascribed to the input logogen systems and the distinction between the two accounts is no longer clear.
that used in recall. Winnick and Daniel (1970) found better recall of the items presented as pictures and definitions but no facilitation effects for these items.

The post-categorical stages which are currently subsumed under the Cognitive System will, in the end, have to be differentiated by means of converging operations. Among its other functions it is the Cognitive System which we suppose to be responsible for most of the effects found in lexical decision and in tasks involving the classification of words. According to the current hypothesis the only manipulation in these tasks which does operate at the logogen level is that of stimulus degradation, together with its interaction with semantic or contextual variables. The arguments concerning all this are complex and use the results of more than fifty studies (Morton, manuscript in preparation 'Task analysis in word recognition'). Among other arguments is the following. When the amount of stimulus information available is limited it will be the initial categorization system which is most affected. Word recognition in these circumstances will reveal the importance of particular variables to this system. The effects of word frequency on visual word recognition appear to be weak, particularly when subjects are instructed not to guess (Dogget and Richards, 1975; Erdelyi, 1974; Eriksen, 1963). On the other hand, the word frequency effect is large and easily obtained with the lexical decision task (Becker and Killion, 1977; Scarborough et al., 1977) and with picture recognition (Oldfield and Wingfield, 1963). These results are consistent with the idea that word frequency has major effects in the Cognitive System (by virtue, for example, of being correlated with the richness and diversity of the representations which are accessed by a word), but only minor effects in the input system. From this position it would be predicted that the effects of stimulus degradation and word frequency would be additive in lexical decision task. Such a result has been reported by Becker and Killion (1977, Experiments 3 and 4) and Stanners, Jastrzembski and Westbrook (1975). It is thought worth noting these arguments here, however, since a number of authors have argued against the current framework on the basis of data from studies involving lexical decision and classification (e.g. Forster, 1976; Sanford, Garrod and Boyle, 1977). In terms of the current hypothesis, these studies do not actually address the issues discussed in this paper.

Confused as the picture might be, the data reported in this paper guarantee the demise of the single logogen system. Neither the naming of pictures (Winnick and Daniel, 1970) nor the filling of definitions have any detectable effect upon subsequent visual recognition. In addition there is significantly less cross-modal facilitation than within-modality facilitation. These data demand a separation of the visual input system from the phonological lexicon and from the auditory input system. The modal which is currently preferred is that shown in Figure 4. This has been further expanded in Morton and Patterson (1980) in the context of deep dyslexia. It is clear, however, that it is in need of much more detailed specification before it can be counted as satisfactory. In particular the fact that Gipson (unpublished manuscript) and Ellis (1982) have succeeded in eliminating visual-to-auditory facilitation completely whilst we have failed to eliminate auditory-to-visual facilitation indicates that the categorization systems for the two modalities may be dramatically different from each other.

Experiments I and II formed part of a B.A. dissertation in the medical Sciences Tripos in the University of Cambridge, 1977. We are grateful to Lindsay Evett for running Experiment III and discussing the paper and Karalyn Patterson for much help and encouragement.
References


*Revised version received 4 March 1982*