In. Perspectives on Mental Representation
eds. J. Mehler, E.C.T. Walker &
M. Garrett. Hillsdale, N.J.:
Lawrence Erlbaum.

# 5 Disintegrating the Lexicon: An Information Processing Approach

John Morton
*M.R.C. Applied Psychology Unit*
*Cambridge, England*

## INTRODUCTION

In this chapter, my intention is to give the reader a particular view of an information processing approach to cognitive psychology. First of all I will outline how I see the relation between psychology and other disciplines. This is just my own version of the anti-reductionist argument that many others have expressed.

Then I will trace the development of the model that I use to talk about a number of phenomena arising from the way we process words. I will show how the first model I proposed turned out to be incorrect and indicate how it was changed in response to a number of experimental findings. The resulting model is modular in form, with the primary assumption being that the individual processes that are indicated are substantially independent from one another. I will show that by experiment one can test assumptions about the detailed functioning of one of the processes, and will indicate how a model of this kind can be adapted to account for data from experimental paradigms other than those for which the model was constructed. In addition, I will show the relation between this model and other models.

In the final section I demonstrate how the model has been refined by use of data from brain damaged patients, and how it can be used in describing particular kinds of aphasic and dyslexic syndromes.

---

Paper delivered at CNRS Conference on Cognitive Psychology at l'Abbeye de Royaumont, 15 June 1980.

## GENERAL CONSIDERATIONS

Let me begin with a firm statement of belief. For some it will be so obviously true as to be not worth the saying. For others it will be so obviously wrong that it be not worth discussion by rational beings.

There are two (perhaps more), enterprises that psychologists engage in. One is that of finding cause; and the other has to do with representation; representation of the processes and structures that determine our behavior. It is with the latter that I occupy myself. Insofar as we are concerned with representation, then, I believe that psychology is independent. Its independence derives from the level of observation with which it is concerned. It is independent from anatomy and physiology on the one hand and from sociology on the other hand since these disciplines are concerned with observations at other levels.

Thus it seems to me that there is no physiological or anatomical fact that could in principle falsify or verify a purely psychological theory. Any possible counterexamples—and those of the Gestalt psychologists, Hull and Hebb come immediately to mind—will serve to illustrate what constitutes a purely psychological theory. The relevant error in those enterprises was to try to tie a psychological model to supposed physiological facts, perhaps in the belief that it would lend it support. The consequence was that the inadequacy of the physiology caused the demise of the psychology. However, investigation of brain processes and investigation of mental processes have only a slight relation. Once we establish the former and establish the latter we can go on to establish the mapping relationships between them. This is a perfectly respectable scientific enterprise, but must not be mistaken for the other two. Of course there is the temptation to try to make the link before it is known what is being linked, but, apart from the extreme of sensory and motor considerations, success is low. It is low because there appears to be no convenient mapping, at least at the cortical level, from structures and processes that interest cognitive psychologists, to other than the most gross localizations.

This position does not prevent me from using data from neuropsychology. Towards the end of the paper I will describe data from one or two brain damaged patients. However, I will only look at their behavior: the site and cause of their lesions is of no relevance to what I am doing though it is, of course, of great interest in its own right. In passing, we can observe that the terms *neuropsychology* and *neurolinguistics* are often used in two senses. In one sense they refer to the study of brain damaged patients; in the other sense they refer to the study of the mapping of mental functions onto brain functions. The two uses, I am claiming, are quite distinct although some authors, such as Arbib and Caplan (1979), switch from one use to the other as if they were indistinguishable.

Equally, I would not feel constrained to ignore results of experiments that involved "physiological" responses such as measures of electrical or other activity in the brain or galvanic skin responses. In as much as such experiments involved stimuli that were the concern of my model, the data would be usable. But note that I would be totally unconcerned as to the purely physiological mechanisms leading to the responses and that the nature of the model would be unchanged by the use of such data. Its form would be subject to change in exactly the same way as that forced upon it by experiments involving verbal responses.

My second preliminary concerns the nature of the representation that I use. A convenient label would be that of "information processing." The notation I employ is usually that of boxes joined together with directed lines. The boxes, in general, represent either a process that converts information from one code to another or a store of information of some kind. The model gives us a way of handling complexity by allowing us to partition the processes into functionally independent units of manageable proportion. It enables us to relate together a wide variety of experimental findings and everyday observations, and to make predictions about new findings. The method has flexibility and extensibility built into it. That is, the framework itself cannot be *wrong* though it is possible to be wrong *within* the framework, as we shall see.

To illustrate the non-triviality of the enterprise as I see it at the moment, we can see in Fig. 1 the extent of the systems that I am currently concerned with. The labels have been left off the individual boxes to avoid you trying to understand the model here and now. What I have represented here seems to be the minimum required to talk about word recognition in two modalities, object recognition, speaking and writing single words, together with a little concern for larger units of language and some concern for phenomena of memory for language materials over short time intervals. Virtually every one of these boxes and connections is supported by a number of observations. Lest you feel apprehensive, let me reassure you that I will be going to talk about only a subset of these boxes.

I would like to make a couple more comments about the nature of the model. Firstly, not all the lines mean the same thing. In some cases a process must finish completely before it can output anything. In other cases a process can pass on partial results. In further cases we want to indicate that one process just accesses another process or a store of information. Unfortunately we haven't yet worked out a notation or even the typology on which such a notation might be based. Secondly, do not become confused between form and content. There are other ways of expressing the same facts that would result in models isomorphic with this diagram. The form I use happens to help me to answer some of the questions I ask. The particular form predisposes one in particular ways. This form of model

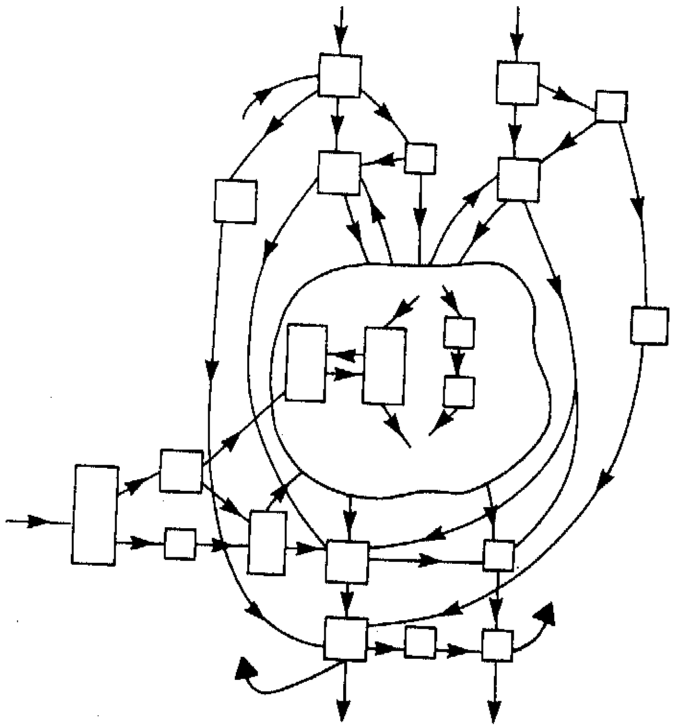


FIG. 1 An indication of the degree of elaboration necessary to express the relationships among the processes involved in word recognition, picture recognition, and naming and writing. All boxes and connections have empirical justification.

predisposes one to think in terms of the separability of processes rather than the way in which they interact. This might be a problem. However it is not a part of the theory.

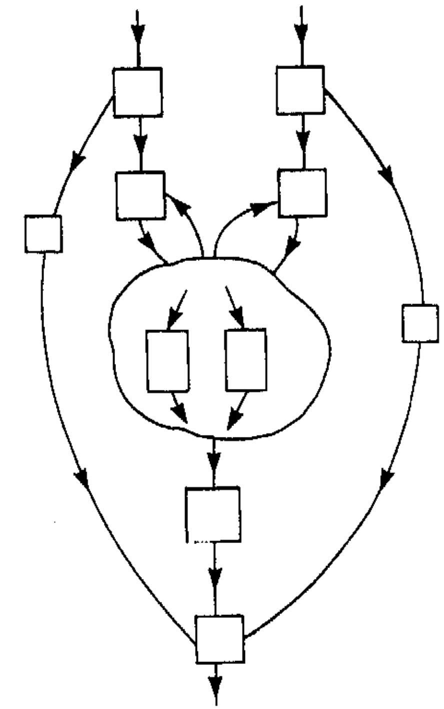
Let me say again that the form I use is not a part of the theory. The same relationships could be expressed in rewrite rules or in prose. They could be implemented in any number of ways—in augmented transition networks or in production systems, for example. I chose these two examples since there are those among the users of both these systems who seem to claim that the form has content (Wanner & Maratsos, 1978; Anderson, 1976; Newell, 1980). This seems to be in error. Any notation that is rich enough will be able to represent the things we are interested in. What is crucial in the choice of a particular form is whether the generalizations are interesting ones, whether the ideas are communicable; whether the model is extendable in novel ways so as to predict novel findings.

The full model, even if properly labeled, will only serve as a means of communication for a certain range of questions. And, in general, at any one



time we will only want to consider segments of the model. Thus, since we are not, at the moment, interested in object recognition, in writing, or in segments larger than the word, we will be left with the much more manageable structure shown in Fig. 2. This will enable us to look at questions of the organization of the internal lexicon. Note that even now we are faced with a problem of selection. What will be represented are what seem to be the most important processes and interconnections. Such an evaluation will, of course, be subject to modification.

FIG. 2 A subset of the model in Fig. 1. This is required to account for data on word recognition. The origins and nature of this part of the overall model will be described in the following text.



As a final preliminary, we will see that the boxes themselves may be more or less specified. The intention is that the way in which a particular process functions does not affect the operation of the rest of the system. Inasmuch as this modularity fails, then the model will be in need of modification and the form may be unsuitable.

## THE NATURE OF THE INTERNAL LEXICON

One extreme view of the internal lexicon is that all information about an individual word; its acoustic, visual, semantic, articulatory, graphic, and other descriptions, is to be found in the same "place"—where place is to be defined in terms of whatever model form is used. My own work has found itself concerned with the progressive separation of individual elements of what might otherwise be seen as a unified lexicon.

The first separation was that of the Logogen System from the Cognitive System as in Fig. 3. Each logogen was associated with a word or morpheme.

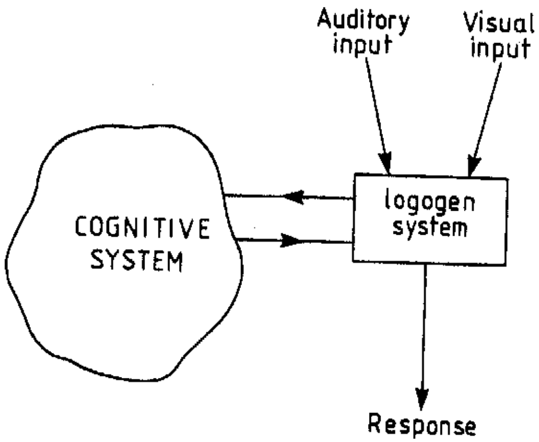



FIG. 3 The initial division of the lexicon. The logogen system is a device for categorizing verbal stimuli. It is the site of the facilitation effects described in the text. It gives outputs that can be interpreted either by the speech output processes or by the Cognitive System. The latter includes all syntactic and semantic processing.

Given an input from the visual or auditory analysers, the logogens collect evidence that a particular word has occurred. When the decision is made, then the phonological code is made available to the response system and some code is sent to the Cognitive System, where computations relevant to the meaning or reference of the word could be made. Most of the interesting work went on in the Cognitive System, which was, and still is, relatively unspecified (but see Morton, 1968; Morton & Patterson, 1980a). The logogen was also the means whereby words were produced in spontaneous speech and object naming, for example. The Cognitive System produces an appropriate code—let us loosely call it semantic—and the logogens collect evidence from it in just the same way. For the recognition of words in context, there would be both sensory information and information derived from the context. The logogens would regard all this information as equivalent and simply access the evidence.

One other feature of the logogens was that there were after-effects of their having been active. The crucial experimental paradigm is one where subjects have some pretraining task followed by the main task in which words are presented in a tachistoscope for identification. The duration of exposure required for a correct response is found to be influenced by the pretraining task. Perhaps the first experiment using this paradigm was Neisser (1954) who showed that if subjects had seen the word PHRASE in the pretraining, then their visual duration threshold for that word was lower than for subjects who had not seen it. Prior experience with the homophone FRAYS, however, had no effect on the subsequent recognition of PHRASE.

One variation on this paradigm served to better define the nature of the unit concerned. Murrell and Morton (1974) showed that so far as recognition of the word SEES was concerned, prior experience with SEEN was almost as effective as SEES itself. Having read SEED in the pretraining, however, had no effect on the subsequent recognition of SEES. Thus it seemed to be the morphology and not the physical resemblance that was crucial. Note that the time interval we are concerned with is anything from 15 to 45 minutes, and may, indeed, be orders of magnitude longer. In this respect, as we shall see later, these experiments are very different from those showing facilitation over time intervals of a few seconds.

By themselves such experimental findings would be subject to a number of possible interpretations. In the model, however, and thus in the context of all the experimental findings that had served to form the model, the only account was that the effect was to be located in the logogens, the effect of prior exposure serving to reduce the amount of sensory information required for subsequent identification.

This account, as found in Morton (1969, 1970), led to certain predictions. The only event that could have such consequences was the firing of a logogen. But it was a part of the model that the firing of a logogen would have the same consequences regardless of the cause of the firing. The model thus predicted that if one named a picture in the pretraining, then there would be consequences in the subsequent identification task. This turned out not to be the case. Winnick and Daniel (1970) showed that naming the picture of a butterfly or saying the word in answer to a definition had no detectable consequences on the subsequent visual duration threshold for the printed word BUTTERFLY. These data were replicated by Clark and Morton (unpublished) who also found that there was good transfer from reading a word written in a casual cursive script to the subsequent recognition of the printed word (see Morton, 1979).

Thus the model as it stood was incorrect. The site of the facilitation and the source of the phonological code could not be the same. Thus the original single Logogen System was split into two, one for input and one for output. The Input system maintained all the properties required for word recognition. It would remain unaffected when a picture was named or a word produced spontaneously. The new form of the model is shown in Fig. 4.

Note that there are alternative solutions such as supposing separate

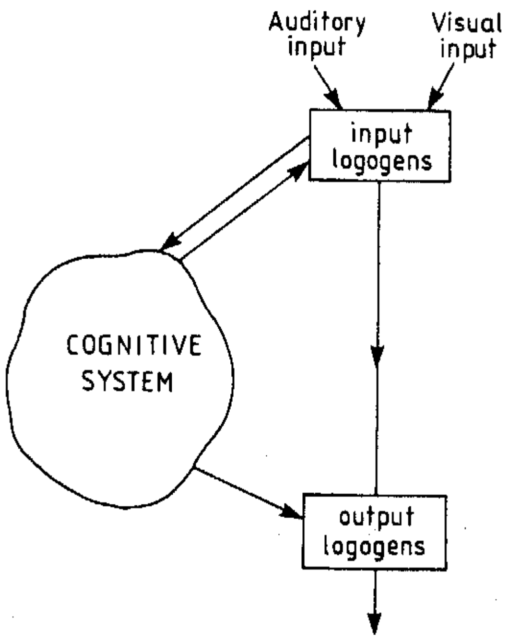



FIG. 4 The second division of the lexicon. This differs from Fig. 2 in that the logogen system has been divided into two parts. The input logogen system serves to categorize verbal stimuli and is the site of facilitation. The output logogen system is a store of phonological codes. This model permits an account of the lack of effect of picture naming on the subsequent visual recognition of the name.

sources for the phonological code following reading and naming. However, there appear to be a number of objections to such alternatives.

The next stage was that the input modalities had to be separated. Clark and Morton had also found very little transfer from hearing a word in the pretraining to the subsequent visual recognition of the word. Next, Jackson and Morton (unpublished) found that there was relatively little effect of having made a semantic judgment on the visually presented word on the subjects ability to recognize the word when it was spoken in a background of noise (see Morton, 1979). This compared with strong effects when the words were presented auditorily either in the same or in a different voice. The latter experiment has been repeated in unpublished experiments by Peter Gipson and Andrew Ellis, in different laboratories, and with tighter controls and under the right conditions, it turns out that there is no cross modal transfer whatsoever. The resulting model is shown in Fig. 5 with auditory and visual categorization systems (input logogens) kept separate. The facilitation effects referred to above take place in these systems.

Let us next consider briefly a couple of additions to the model that will come in useful later. In the first place nothing has been said about the



nature of the information fed into the input logogen systems. We can finesse this problem for the moment by adding a couple of processes to symbolize the transformation of sensory information into a suitable form. We can simply label these as analysis systems, for convenience of reference. The important point here is that so far as the functioning of the rest of the model is concerned it doesn't matter how these devices work nor what is the nature of the information they pass on to the categorization systems. Thus the information from the visual analysis could be in the form of letters. There are a number of models of word recognition that suppose this step as necessary (Massaro, 1975). I happen to think there is good evidence against this, not least of which is our ability to read carelessly written script. But the assembly of this evidence and the subsequent decision have no consequence as far as the rest of the model is concerned. Equally, the output from the auditory analysis system could be syllables, or phonemes, or simple acoustic features. This decision, too, is separate.

One addition that is necessary later is a way of reading and of repeating

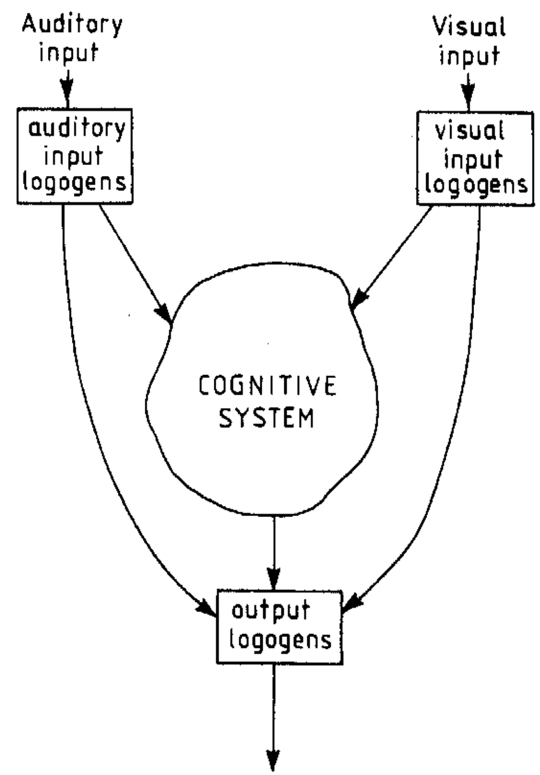



FIG. 5 The third division of the lexicon. The lack of cross-modal facilitation necessitates separate input logogen systems. The direct connection between input and output logogen systems is speculative

nonwords. This I indicate by the addition of grapheme-phoneme and auditory-phonemic routes. The boxes symbolize the processes, and while they are diagrammed as being separate from the lexical routes, the model is not discomforted by the possibility that we read nonwords by analogy with real words; it is merely complicated. These processes feed into the Response Buffer, which is the process responsible for converting all phonological code into speech output. The resulting model is shown in Fig. 6.

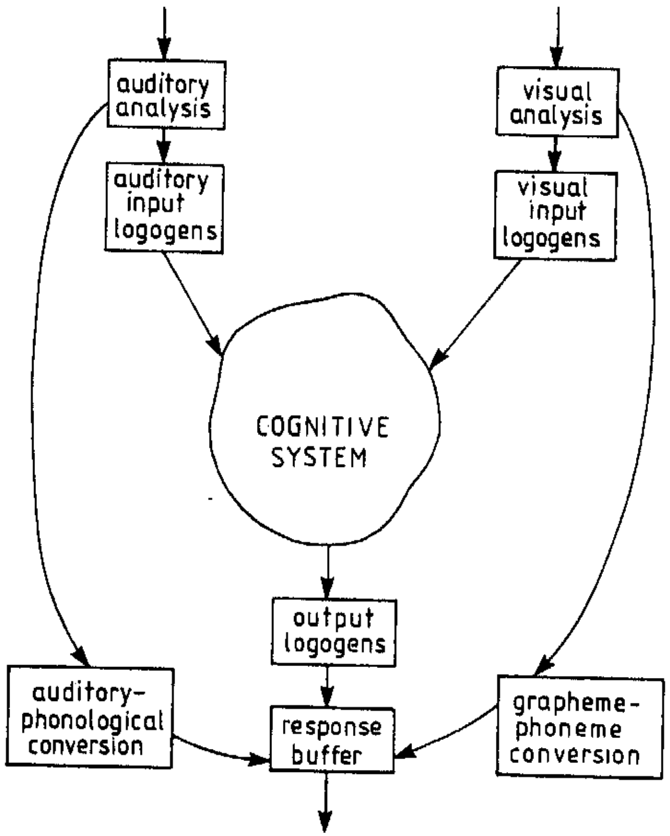



FIG. 6 An elaborated version of the model showing the sensory analysis systems, a response buffer, and routes connecting the two that bypass the lexical processes. These latter are necessitated by our ability to read or repeat nonwords.

Next we can consider the kind of decision we have to make about the detailed functioning of the processes. In the original formulation of the logogen model I supposed that each logogen collected evidence relevant to itself from the context and from the current stimulus. It was supposed that the first logogen to accumulate an amount of evidence greater than a particular threshold would then fire. The effect of this would be to stop any other logogen from firing. I took the simplest assumption for the interaction of stimulus and context information; that the two added together. It was then possible to test the assumptions by comparing three situations: recognizing a word presented in isolation; recognizing a word presented in context; and producing a response to the context in the absence of any stimulus information. It is clear that the underlying processes are not deterministic with respect to a particular task. Even the same subject does not always make the same response to complete an open-ended sentence. A usual assumption is that the system is noisy with respect to the current task—that is, that there is concurrent processing unconnected with what the experimenter is interested in. With this assumption I was able to use Luce's Choice Axiom in order to derive predictions about performance in the three situations. The resulting equation was:

$$\text{Logit } P_{SC} = \text{Logit } P_S + \text{Logit } P_C + K$$

where logit $P$ is given by $\log(p/(1 - p))$, and $P_C$, $P_S$ and $P_{SC}$ are the probabilities of being correct with context alone, with stimulus alone, and with both stimulus and context. I took data from Tulving, Mandler, and Baumal (1964) who had looked at the effect of context on the visual recognition of words. The equation, then, and the processing assumptions behind it, would apply to the visual input logogen system. I found the relationship as in Fig. 7. The straight lines have a slope of unity, predicted by the equation, and were fitted by eye (Morton, 1969). The fit seems to be good, and still strikes me that way.

The point about this demonstration is that it illustrates how one can make precise predictions about the operation of the model when one element is specified in detail. The experiments then allow us to accept or reject particular assumptions about the way the processes operate. Note, however, that if the predictions had failed, it would not affect the overall structure of the model, just the assumption about visual categorization. In addition, if it turns out that the data I used was peculiar in some way and that, after all, the equations do not predict what happens in a series of replications, all that will happen is that the specification of the way in which the visual input logogen system functions will have to be changed. Again, nothing else need be affected.

Let me add to this that there are likely to be differences in the mode of operation when the stimulus is clear compared with when there is stimulus impoverishment, as when the words are presented in a tachistoscope. In particular, with clear stimuli, the influence of context on the categorization process is likely to be minimized. In addition, it is certain that inhibitory processes operate from the stimulus, but not from the context. Thus, seeing the sentence,

She was wearing a bright red xacket

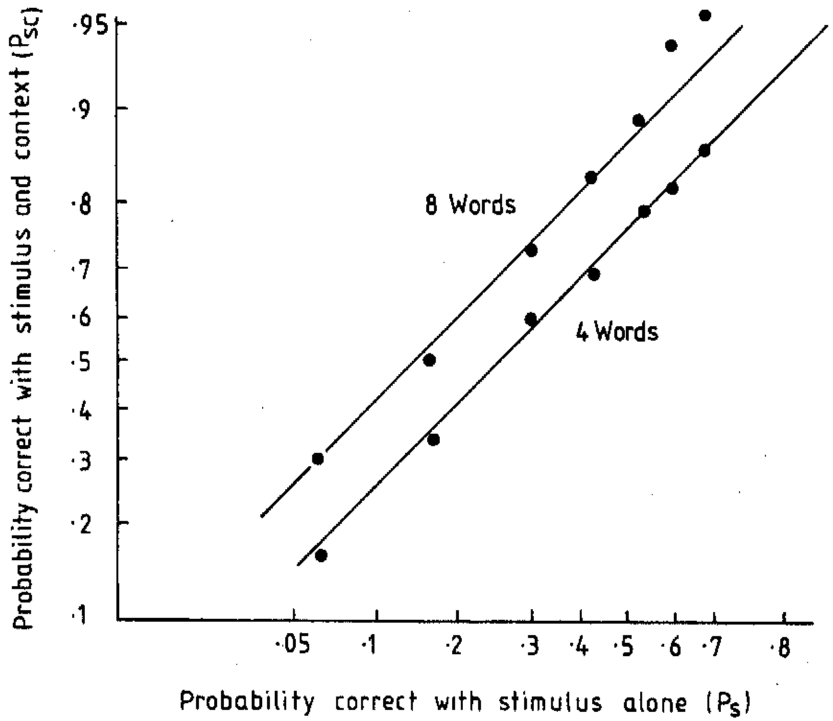



FIG. 7 A graph showing the relationship between the probability of recognizing a word with and without a context. The straight lines of slope = 1 were predicted by the model and were fitted by eye (from Morton, 1969).

the final word is not going to be read as "jacket" when it is presented clearly and read with caution. The reader may be able to work out that the word ought to have been *jacket* but would not think it *was* jacket.

The new form of the model has had an unexpected advantage. The separation of the auditory and the visual input logogen systems carries with it the consequence that the two may well operate in different ways. Thus, in the last few years, Marslen-Wilson and his collaborators have proposed a different model for auditory word recognition, which has been called the "cohort model" (Marslen-Wilson & Welch, 1978; Marslen-Wilson & Tyler, 1980). By this account the recognition process has two stages. After the first two or three phonemes have been analyzed, a set of possible candidates is isolated. Thereafter, subsequent inputs serve to eliminate members of this "cohort" set until only one candidate remains. He had produced some persuasive evidence in favor of such an idea, but it remains to be seen whether this model can also account for all the other data that has gone into the logogen model, as it remains to be seen whether an evidence-collecting



system, like the old logogen system, can be modified to account for his data. If it turns out that the cohort model wins, then the consequences are clear. The logogen system would have to be modified such that the auditory input categorization system operates on cohort principles and would thus be different from the visual system in this respect. Again, the modularity of the approach prevents the consequences from being more severe than this.

## LEXICAL DECISION

A number of other models of the lexicon have been proposed over the past few years. A number of these have started from a different set of data, the most usual being the lexical decision task. In this task the subject is presented with a string of letters and has to decide whether or not they constitute a word in the language. There are, perhaps, a hundred or two such studies published in the last ten years. In a number of them the logogen model has been rejected on the grounds that it made the wrong predictions. Unfortunately, in every case, this conclusion has been drawn in the belief that a lexical decision would have to be made on the basis of the operation of the logogen system, or, in the case of the revised model, as a result of the operation of the visual input logogen system. Logically speaking, this would be a possibility, since nonwords do not have a representation in the input system. Thus, if the letter string is successfully categorized it must, indeed, be a word. There are a number of problems with this account of what happens. In particular it would be impossible to account satisfactorily for responses to nonwords. The first difficult finding was that nonwords that were homophones of real words, such as BRANE, took longer to reject than nonwords having no real word equivalents, such as BLANE (Rubenstein, Lewis, & Rubenstein, 1971; Coltheart, Davelaar, Jonasson, & Besner, 1977). This effect could only arise if the letter strings are converted into the appropriate phonological codes and it would be difficult to involve the input systems in any subsequent decision.

A second problem relates to the strong word frequency effect in the lexical decision task. If the word frequency effect in this task is to be assigned to the logogen system then there should be an interaction of this effect with the effects of stimulus degradation. In fact, there are at least three experiments showing that word frequency and degradation are *additive* factors indicating that they act at different places (Stanners, Jastrzembski & Westbrook, 1975; Becker & Killian, 1977, expts. 3 and 4). Thirdly, Becker (1976) has shown that response times in the lexical decision task are subject to the attentional demands of a secondary task. This could not be the case if the logogen system was involved, since it is designed as a passive, non-attention demanding system.

These, and some other objections, lose their force, however, if we assume that lexical decisions are based primarily on the operation of the processes in the cognitive system. The reason why the logogen system cannot be so used can be answered in a number of ways, the simplest being that we simply have no means of monitoring the system at the appropriate point. With the lexical decision based on the cognitive system we have a number of degrees of freedom that enable an account to be given of the problem results. These involve some elaboration of the available mechanisms, but the result is not noticeably more complicated than other accounts. Thus, the effect of the nonword homophones is accounted for by supposing that the phonological code is automatically produced and then used to access semantic information. This would succeed, but subsequent checks on the spelling would prevent an error response. While such checks are not a property of the logogen system, we have no reason to exclude them from the cognitive system. We would, however, hope to find evidence for them in other tasks. Secondly we can attribute the word frequency effect in the lexical decision task to the structure of the semantic and associative components of the cognitive system. In fact, I prefer to think of word frequency as being an index of some other property rather than having direct representation in the cognitive system. Thus, we could imagine that, in general, more frequent words have stronger or richer associations. The account one would give here would depend upon what processes were thought to be necessary for other reasons at this stage in the processing. In any case, given that the word frequency effect in this task is located in the cognitive system, we would then expect no interaction of frequency with stimulus degradation, since the latter variable must have its effect at the input logogen level or earlier. Finally, we now have no prior reason for supposing that the lexical decision should not be subject to attentional demands. Becker's result, referred to above, serves then to tell us something about the nature of the operations involved in finding information in the cognitive system. Again, we would demand that any other task whose explanation involved the relevant parts of the cognitive system would also show the same attentional effects. Such a task has not yet made itself known, but it surely will. There are degrees of freeedom remaining in our ability to account for data, but we have to expect that every new postulate is capable of a number of applications.

At this point I should acknowledge that Forster (1976, 1978) has put forward a model of lexical access that is superficially very different from the logogen model. The most striking difference is that the system uses *search* as the means of access, rather than automatic content addressing, which is the characteristic of the logogen model. A detailed comparison of the two approaches is in preparation. Two points might be made here. Firstly, the account given of some data is considerably more elegant in Forster's model than in mine. For some other data, however, he has to elaborate his system



well beyond the original intention, while the logogen model handles this data with ease. Such facts simply reflect the starting points of the two theories. The second point is that as the models become more elaborated, they begin to resemble each other. Thus, while I would reject search principles as a property of the input systems, there is no reason to exclude them from the cognitive system.

## READING IN JAPANESE

The question has been posed in discussion as to the predictions that would be made from the Logogen Model as to results in certain experiments when using Japanese script. In Japanese there are two writing systems, *Kana,* which is a syllabic script, and *Kanji,* which is an ideographic script. It is clear that the model has nothing to say about how these scripts are processed by native readers of the language. However, the methodology used in establishing the information processing account for alphabetic scripts should be usable for Japanese. How it will look is an empirical question.

## A FINAL CAUTION

As I have indicated, the model has been elaborated to account mainly for experiments involving single words. It remains open the extent to which the same processes will operate in the same way in continuous performance. Naturally, I would hope that only minimal changes would have to be made, but it remains possible that in continuous tasks some of the categorization decisions are made further into the system. One possibility is that in rapid speech, the segmentation problem cannot be satisfactorily solved without taking into account the kind of pragmatic constraint that it would be unreasonable to suppose could be fed back to the input systems. The operation of the overall system then might more resemble the HEARSAY program. It might be noted, however, that recent work by Marcus (in press) indicates that the segmentation problem may be soluble without the need of such added complexity.

In any case, a model that accurately represents the mental operations involved in the processing of individual words promises to have its uses. In particular the logogen model is currently being both tried and extended in neuropsychology. It turns out that a lot of the standard tests given to brain damaged patients involve the use of the single words. Patients are typically required to read them, categorize them, repeat them, write them to dictation, and use them to name objects that are seen, touched,or heard. In the final section of this paper I will outline some of the results of this endeavor.

## SOME NEUROLINGUISTIC CONNECTIONS

Recently, the logogen model or variants on the form have been used as a means of describing aphasic symptoms (Dérouesné & Beauvois, 1979; Goldblum, 1980; Morton & Patterson, 1980a, b; Patterson, 1980; Shallice, 1980; Schwartz, Saffran, & Marin, 1980). In addition, data from patients has been used to extend the model. One example is due to the patient P.W. studied by Morton and Patterson (1980a, b). This patient is known as a deep dyslexic (the relevant symptoms will be described later). In addition he is agrammatic. The agrammatism emerges in a number of ways. To start with, he finds function words much more difficult to read out loud than other words. Of a sample of 406 function words (taking the broadest definition of the class) he read correctly only 23%. This compares with 50% correct with a set of nouns. He also had a seeming inability to comprehend function words or use them. Since the patient is totally incapable of reading nonwords, and so, in terms of the model, must lack the grapheme-phoneme route, we originally entertained the hypothesis that function words were not represented in the logogen system but rather were read via the grapheme-phoneme route. This theory was abandoned when Patterson (1978) showed that P.W. was virtually perfect on a lexical decision task involving function words together with nonwords which closely resembled function words (e.g. THISE, THORE, WHAR, WETH). Next we discovered that he had access to the meaning of function words. We tested this by means of word triplets; the patient having to point to one word of a pair when "went with" a third word. Examples are

(over/under up) (every/few all) (except/with but);

these three representing space, frequency, and logical number. He performed well above chance on all these and others, but failed on judgments that were vaguely syntactic. Thus he could not judge the case of pronouns (he/him her), or relative pronouns (that/this these) though he could judge the number of these words (those/that many). We established his comprehension of prepositions in a number of ways. Then we tried him on sentences such as:

Put the cup on the saucer.

In one session he got 11 wrong out of twelve, making errors of the type of putting the saucer on the cup. We presumed that he was treating the preposition as if it was an adjective, and so attaching it to the following noun. Our conclusion was that while the semantic aspects of the function words were relatively unaffected, their structural aspects were missing (Morton & Patterson, 1980b). Overall, then, we came to the conclusion that



function words normally had two representations in the central part of the lexicon, one of which, the structural, was not operational for P.W. (Morton & Patterson, 1980a).

We can now think about the lexical decision task again. For normal subjects we know that there is a strong word frequence effect, which is almost always tested with content words. In the one exception, Bradley (1978) found that there was no word frequency effect for lexical decision times using function words. The simplest conclusion, in terms of the model we have elaborated, is that function word lexical decisions can be based on the structural component of the lexicon. The rationale of this is that the number of structural possibilities (usually only one) of function words will not be affected by how common it is in the language. We would also deduce from the data that access to the structural output was in general faster than access to the semantic component, though this would not be a necessary conclusion. If the structural component was lost, however, lexical decisions would have to be based on the semantic component in which, as we have seen, there is a word frequency effect. Indeed, Bradley showed that this was the case. Needless to say, we didn't actually predict these data before we saw Bradley's thesis, but we might have done so if we had been really smart.

Let me now say a little more about P.W. who is a typical deep dyslexic. His problem with reading may be summarized as follows:

1. He is completely incapable of reading any nonword, and, indeed, appears to be unable even to get single phonemes except via a real word, which may be a part. At the same time he can repeat nonwords without much trouble.

2. In reading words he makes about 50% errors. All his responses are real words. The errors may be classified into three main types:

semantic: e.g. paddock is read as 'kennel'
derivational: e.g. edition is read as 'editor'
visual: e.g. narrow is read as 'marrow.'

In general he is aware of the semantic errors but not aware of errors of the other two classes. This is a simplification of the data, and a full account is given in Patterson (1978, 1979).

3. He makes more errors on function words than on other types of words. In this respect he resembles a typical agrammatic. Thus he can read *inn* but not *in*, bee but *be* and he can only read the word *over* when he is reminded that the word is also a term in the game of cricket (see also Gardner & Zurif, 1975).

Karalyn Patterson and I have given a detailed analysis of this and other data (Morton & Patterson, 1980a). What we have done is locate the 'lesions'

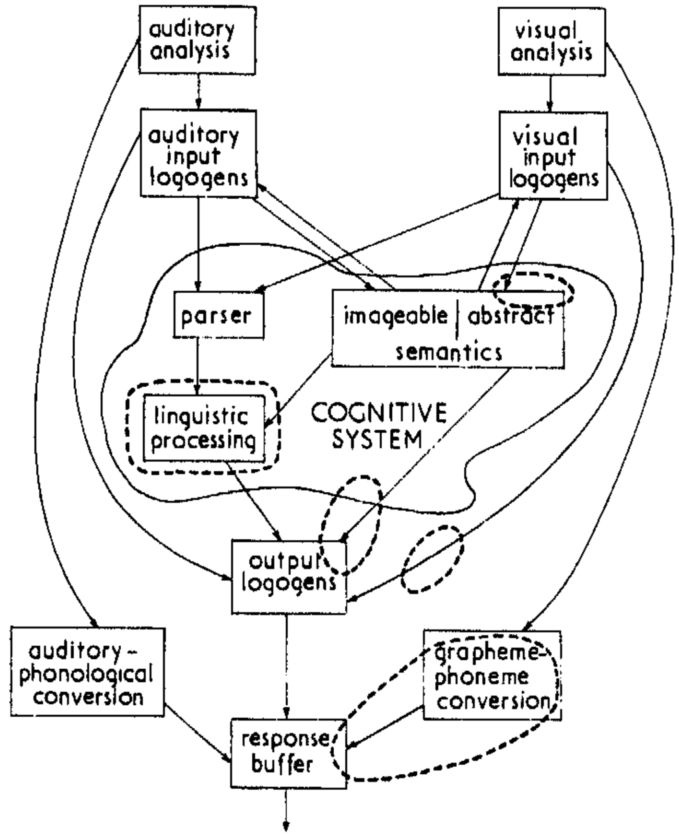



FIG. 8 An expansion of the logogen model indicating the location of the functional deficits in a deep dyslexic patient (from Morton & Patterson, 1980a).

in the model. This is illustrated in Fig. 8. Basically we attribute the problem with nonwords to a breakdown of that system which effects conversion from graphic to phonological forms. From the modeling principles outlined above, it should be clear that such an attribution is independent of the nature of this system. The semantic errors are due to a problem with the connection between the cognitive system and the output logogen system together with an absence of the direct input-output route. The visual errors are due to a problem with the connection between the input system and the cognitive system. Note that the visual errors cannot be attributed to a problem with visual analysis since the patient does not produce word responses to nonwords that very closely resemble words. Thus he would not read *narket* as "market" but, rather, would say that it was not a word. Awareness of the semantic errors arises from the fact that he has a representation of the target in the cognitive system that can be compared with his response. The same would not be true for the visual errors. The derivational errors we would associate



with his problems with function words and with syntactic function in general. This we represent as a problem with the module labeled "linguistic processing."

This patient contrasts readily with a number of others. Thus, the *phonological alexic* patients studied by Beauvois and Dérouesné (1979) had reading problems only with nonwords. The problem here would be solely with the grapheme-phoneme route. Then the class of patients termed *surface dyslexics* have equivalent problems with words and nonwords (Marshall & Newcombe, 1973). In such cases, we would say that the visual input logogen system was not functioning correctly. There is then a class of patient who has problems with repetition of spoken words, but who has unimpaired reading (Morton, 1980), or reading problems of an unrelated kind (Goldblum, 1980). Such patients reinforce the belief in the fractionation of the lexicon.

More detailed comparisons of these and other patients are made by Patterson (1981). From all these patients, we get a feel for the suitability of a model that stresses the modularity of the compenent processes. Within the model we can readily form contrasts between the patients.

It should be noted that the model does not permit us to state why there is a tendency for loss of the grapheme-phoneme route to be accompanied by agrammatism, or a number of other correlations. Such facts I take as being the province of activity I separated out at the beginning of this chapter—that of the mapping of psychological function onto brain function. This task becomes much simpler when we begin to get an adequate description of the component processes and the way in which they interrelate. The question might have occurred to you that the use of aphasic patients in forming the model presupposes localization of function. In fact this is not the case. A number of the critical patients have massive lesions, and it would not be safe to say any more than that most of the functions in question use a particular region of the left hemisphere. Beyond that we have the problem of not knowing whether the loss of function has to do with the destruction of the process or its disconnection. Such distinctions may well become important in the near future, but at the moment we get the feeling of progress in our work without worrying too much about this. Finally we have no guarantee that the identical lesions in the brains of identical twins will give the same result.

I hope that I have shown how an information processing approach to cognition and language has the properties one demands of a scientific pursuit.

We can relate together a wide variety of facts.
We can predict new facts.
We can ask new questions and apply new descriptions in an applied area.
We can falsify individual elements of the structure of the model (though not the form)

I think that all this can only be done in an *interesting* way, if we tackle the evident complexity of the mental processes. I hope that the model I have described is sufficiently complex to be interesting and that the form is congenial enough to be intelligible.

## REFERENCES


Anderson, J. R. *Language, memory and thought.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1976.

Arbib, M. A., & Caplan, D. Neurolinguistics must be computational. *The Behavioural and Brain Sciences,* 1979, *2,* 449–483.

Becker, C. A. Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 1976, *2,* 556–566.

Becker, C. A., & Killian, T. H. Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 1977, *3,* 389–401.

Beauvois, M-F, & Dérouesné, J. Phonological alexia; three dissociations. *Journal of Neurology, Neurosurgery and Psychiatry,* 1979, *42,* 1115–1124.

Bradley, D. C. *Computational distinctions of vocabulary type.* Ph.D. Thesis, MIT, 1978.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI.* Potomac, Md.: Lawrence Erlbaum Associates, 1977.

Dérouesné, J., & Beauvois, M-F. Phonlogical processing in reading: Data from alexia. *Journal of Neurology, Neurosurgery and Psychiatry,* 1979, *42,* 1125–1132.

Forster, K. I. Accessing the mental lexicon. In E. C. T. Walker & R. J. Wales (Eds.), *New approaches to language mechanisms.* Amsterdam: North Holland, 1976.

Forster, K. I. Accessing the mental lexicon. In E. Walker (Ed.), *Explorations in the biology of language.* Hassocks: Harvester Press. (The M.I.T. work group on the biology of language), 1978.

Gardner, H., & Zurif, E. BEE but not BE: Oral reading of single words in aphasia and alexia. *Neuropsychologia,* 1975, *13,* 181–190.

Goldblum, M. C. *Auditory equivalents to deep dyslexia.* Paper delivered to the meeting of the European Psycholinguistics Association on Information processing models in aphasia, Paris, 1980.

Marcus, S. M. ERIS—context sensitive coding in speech perception. *Journal of Phonetics,* in press.

Marshall, J. C., & Newcombe, F. Patterns of paralexia: A psycholinguistic approach. *Journal of Psycholinguistic Research,* 1973, *2,* 175–199.

Marslen-Wilson, W. D., & Tyler, L. K. The temporal structure of spoken language understanding. *Cognition,* 1980, *8,* 1, 1–71.

Marslen-Wilson, W. D., & Welch, A. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology,* 1978, *10,* 29–63.

Massaro, D. W. Primary and secondary recognition in reading. In D. W. Massaro (Ed.), *Understanding language: An information processing analysis of speech perception, reading and psycholinguistics.* New York: Academic Press, 1975.

Morton, J. Considerations of grammar and computation in language behavior. In J. C. Catford (Ed.), *Studies in language and language behaviour.* C.R.L.L.B. Progress Report No. VI, University of Michigan, 1968.





Morton, J. The interaction of information in word recognition. *Psychological Review,* 1969, *76,* 165–178.

Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory.* New York: Academic Press, 1970.

Morton, J. Some experiments on facilitation in word and picture recognition and their relevance for the evolution of a theoretical position. In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *The processing of visual language I.* New York: Plenum Press, 1979.

Morton, J. Two auditory parallels to deep dyslexia. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep Dyslexia.* London: Routledge & Kegan Paul, 1980.

Morton, J., & Patterson, K. E. A new attempt at an interpretation, or an attempt at a new interpretation. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul, 1980. (a)

Morton, J., & Patterson, K. E. "Little words—No." In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul, 1980. (b)

Murrell, G. A., & Morton, J. Word recognition and morphemic structure. *Journal of Experimental Psychology,* 1974, *102,* 963–968.

Neisser, U. An experimental distinction between perceptual process and verbal response. *Journal of Experimental Psychology,* 1954, *47,* 399–402.

Newell, A. HARPY: Production systems and human cognition. In R. Cole (Ed.), *Perception and production of fluent speech.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1980.

Patterson, K. E. Phonemic dyslexia: errors of meaning and the meaning of errors. *Quarterly Journal of Experimental Psychology,* 1978, *30,* 587–601.

Patterson, K. E. What is right with 'deep' dyslexic patients? *Brain and Language,* 1979, *8,* 111–129.

Patterson, K. E. Derivational errors. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul, 1980.

Patterson, K. E. Neuropsychological approaches to the study of reading. *British Journal of Psychology,* 1981, *72,* 151–174.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior,* 1971, *10,* 645–657.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. Fractionating the reading process in dementia: Evidence for word-specific print-to-sound associations. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia.* London: Routledge & Kegan Paul, 1980.

Shallice, T. Case study approach in neuropsychological research. *Journal of Clinical Neuropsychology,* 1980, *1,* 183–211.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. Frequency and visual quality classification. *Journal of Verbal Learning and Verbal Behavior,* 1975, *14,* 259–264.

Tulving, E., Mandler, G., & Baumal, R. Interaction of two sources of information in tachistoscopic recognition. *Canadian Journal of Psychology,* 1964, *18,* 62–71.

Wanner, E., & Maratsos, M. An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality.* Cambridge, Mass.: MIT Press, 1978.

Winnick, W. A., & Daniel, S. A. Two kinds of response priming in tachistoscopic recognition. *Journal of Experimental Psychology,* 1970, *84,* 74–81.