DOCUMENT RESUME

ED 182 985 FL 010 837

AUTHOR Cowart, Wayne
TITLE Production, Comprehension, and Theories of the Mental Lexicon. CUNYForum, Numbers 5-6.
INSTITUTION City Univ. of New York, N.Y. Graduate School and Univ. Center. Program in Linguistics.
PUB DATE [79]
NOTE 22p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Auditory Discrimination; *Auditory Perception; *Cognitive Processes; Communicative Competence (Languages); Componential Analysis; Diachronic Linguistics; *Language Processing; *Linguistic Performance; *Listening Comprehension; *Memory; Semantics; Speech Communication; Syntax; Vocabulary

ABSTRACT

Problems related to the structure of the mental lexicon are considered. The single access assumption, the passive memory assumption, and the heterogeneous memory assumption are rejected in favor of the theory which assumes several active memories, each able to store expression based on only one homogenous set of abstract primitives. One lexicon provides a systematic phonemic description of each morpheme and another supplies a syntactic description of each item. The third contains the semantic description of every item in the phonemic lexicon and description of any syntactic elements that do not appear in the phonemic lexicon. Each lexicon stores only one kind and all of that kind of information. In comprehension, each lexicon compares the input to its contents. If it finds a mismatch, it withdraws as a possible analyzer of input and passes the item on to the next lexicon. This serial odering prevents errors in comprehension. Errors in production, such as malapropisms (which form the empirical motivation for the hypothesis) occur because output is planned simultaneously rather than serially. Through this theory of discrete components of the human language performance system, various disorders defined as the disruption of some subset of components may be predicted. (PMJ)

PRODUCTION, COMPREHENSION, AND THEORIES OF THE MENTAL LEXICON

Wayne Cowart
PhD Program in Linguistics
The Graduate Center, CUNY
Department of Linguistics
Queens College, CUNY

1. In this paper I will consider some problems related to the structure of the mental lexicon. Two recent theories will be examined and rejected. A new theory will be proposed. Three assumptions which are commonly made about the mental lexicon will figure importantly in the discussion.

According to the *single access assumption* there is only one point in the speech production process at which the speaker must make reference to the mental lexicon. Typically it is assumed that this one examination occurs at the point where a semantic and/or syntactic representation of the utterance is replaced by a phonemic representation. The speaker is said to look up the phonemic representation of each lexical item under the guidance of the semantic or syntactic "spelling" of the item.

According to the *passive memory assumption* the memory devices which store lexical information are unable to perform any sort of computational activity on their own. Rather, each lexical item is stored in a device with just two capacities: 1) to retain a permanent record of some lexical information and 2) to transmit that information to a controlling element when commanded to do so. On such a theory the memory elements all respond to the same kind of signal, and the controlling element cannot have any information about the contents of any particular location in the memory except by retrieving those contents and subjecting them to analytical or comparative processes. Therefore, the passive memory assumption implies that the controlling element has the capacity to systematically select a sequence of locations in the memory and to direct a signal along a unique path to each location in the memory. That is, the controller must be able to search the memory.

I will assume that human language processing devices employ at least three distinct sets of primitive terms. The first of these is a unique abstract alphabet in which the phonemic representation of an utterance must be expressed. Similarly, there are unique abstract alphabets for syntactic representations and semantic representations. The *heterogeneous memory assumption* embodies the claim that the memory which stores the lexicon is heterogeneous in the sense that it may store representations expressed in any of these three alphabets. This can be accomplished in either of two ways. We can find the total number

Cunyforum, Nos. 5-6.

of primitive terms in the three alphabets and assume that the memory can discriminate this number of physical primitive terms. We then assign each member of each of the three alphabets to a unique physical primitive in the memory. If the total of physical primitives is smaller than the total of abstract primitives in the three alphabets, we can resort to a code in which each abstract primitive is represented by a unique pattern of physical primitives. This is the approach used in computers; there sequences or patterns of the two primitives, "on" and "off", can be used to discriminate all the members of alphabets of theoretically unlimited size.

The single access and heterogeneous memory assumptions have been implicitly adopted in all discussions of the role of the mental lexicon in the production and comprehension of speech I have seen. The passive memory assumption, though frequently adopted without discussion, has been denied by some theorists. I will maintain that the best theory of the mental lexicon rejects all three. The theory I will propose assumes several active memories, each able to store expressions based on only one homogeneous set of abstract primitives. It also assumes that the lexicon must be accessed at three distinct stages of the production process. Much of the empirical motivation for this proposal emerges from an examination of two quite different recent proposals.

2. Fay and Cutler (1977) attempt to explain the facts about malapropisms. These they identify as a class of word substitution errors in which the error word (the word the speaker did not intend to use) is 1) a real word, 2) unrelated in meaning to the target (the word the speaker intended), and 3) closely related phonetically to the target word. For example, the utterance <u>John is looking for a uniform</u> includes a malapropism if the speaker intended <u>unicorn</u>.

Fay and Cutler make the following observations about a sample of these errors. First, the error word is almost always of the same grammatical category as the target. Second, the error word almost always has the same stress pattern as the target. Third, in all but a few cases, the target and error words have the same number of syllables. Finally, if we represent the target and error in terms of distinctive feature matricies and assess their degree of similarity working from left to right, we find that there is a marked tendency for the target and error to closely resemble each other at the segment at which they first diverge. Fay and Cutler show that another class of word substitution errors "semantic errors," does not have this property.

Fay and Cutler raise the question whether it is possible to account for these facts on a theory that posits just one mental lexicon. Implicitly adopting the assumptions that the lexicon is stored in a heterogeneous and passive memory they point out the problems that arise when we try to select some principle by which to order the items in the lexicon. If we order the items according to their phonetic properties, we facilitate the task of searching the lexicon during the lexical retrieval stage of comprehension. Simultaneously we exacerbate the problem of searching the lexicon during production when we must find phonetic representations that are identified in syntactic and/or semantic terms. We will likely be reduced to an exhaustive linear search in this aspect of processing. By contrast, if we order the lexicon according to the semantic properties of items, we facilitate searches during production but inhibit searches during comprehension. Fay and Cutler reject the possibility of an extravagantly redundant system in which there are two complete lexicons, one for production, ordered semantically, and another for comprehension, ordered phonetically.

Note that Fay and Cutler view the one-lexicon-or-two question strictly as a question of how many times each item is recorded in the brain. That is, they assume that any lexicon is heterogeneous and that the only question is whether there is more than one. They do not consider the possibility of splitting one lexicon into three (with different types of information in different lexicons) in such a way that each type of information could be ordered in the most appropriate way. Such a system is possible if we assume that each lexical entry includes a pointer directing the controller to entries in other lexicons containing other kinds of information about the same item. In such a system the phonemic entry for boy includes directions on where to find the syntactic and semantic facts about boy in other lexicons. We will come back to a related proposal after giving closer consideration to the passive memory assumption and the search procedures it entails.

Fay and Cutler resolve the problem of the ordering of the items in the lexicon by positing two independent search procedures, one for comprehension and another for production. The lexicon is assumed to be ordered phonetically so that it can be searched efficiently according to phonetic criteria during comprehension. The search procedure used in comprehension is, we gather, a more or less conventional iterative process which is able to take advantage of the phonetic ordering. In production, however, we assume that the lexicon is accessed via a tree-like network which defines a unique path into the lexicon for each semantic description of an item. Semantic descriptions, by hypothesis, consist of an ordered sequence of features, each of which can take two or more values. Each node in the network corresponds to some semantic feature. We can find any item, starting from the semantic description given by the production process, by traversing the network from its origin to its base, turning left or right at each juncture according to the semantic feature value for that node. At the end of each path we should find the item corresponding to the semantic description with which we started. If we also assume, as Fay and Cutler propose, that the items in the lexicon are grouped by grammatical category, number of syllables, stress pattern, and distinctive features, then we have a system which can account for the facts of malapropisms. We need only add that on some occasions the system will err in going from the lowest node of the semantic search tree into the lexicon proper. If this happens, and the production system retrieves not the target word but some near neighbor, we expect this neighbor to have just the properties already noted in malapropisms. The neighbor will be phonetically similar to the target but, in most cases semantically different.

Though this accounts for the facts of malapropisms, there are two important problems. First, there is a good deal left to the imagination with respect to the semantic feature system according to which the search tree is to be ordered. Many would doubt that it is possible to discriminate all of the senses represented in the lexicon of a natural language by way of some set of n-ary features. Certainly no one has yet come close to providing a detailed theory of such a system of features.[1]

More serious problems arise with respect to the single access assumption. Notice Fay and Cutler assume the production system constructs a semantic representation (presumably under the guidance of some higher-order conceptual representation) and parses this semantic representation into real lexical units without reference to the lexicon itself. That is, they assume the production system has information about what senses are available in the speaker's language, independently of the lexicon.[2] They are forced to this assumption because they have assigned the production system the task of sending lexical semantic representations into the top of their search net work. How might the production system know what concatenations of semantic primitives correspond to real words without using the lexicon? One possibility is to assume that each of the terms used in the conceptual system which drives the speech production system corresponds to a real word in the lexicon. In other words, we can commit ourselves to a strong form of the Whorfian hypothesis. The various arguments against the Whorfian hypothesis seem more than adequate to block this possibility. Another possible approach assumes that there is some recursive system which generates all and only the possible senses employed in the speaker's language. This suggestion ignores the highly irregular character of the set of senses. From another point of view, a generative system will almost surely have to incorporate a crypto-lexicon in order to constrain itself to the empirically prescribed domain of senses. In any case, the only motivation for such a proposal is to preserve a certain role for Fay and Cutler's network device.

The search network in Fay and Cutler's theory has in fact two functions. One is that for which it was proposed, the facilitation of the search for phonetic information in production. The other is the isolation of semantic information in the lexicon from the semantic information in the representation of the utterance which drives the speech production process. This second function is most apparent when we consider ways by which the system might find a lexical semantic interpretation of the conceptual representation of the utterance. We are here addressing the question: How does a speaker determine what words are most appropriate for the expression of some thought? One way is to examine the contents of the lexicon. The production system first performs a preliminary translation of the conceptual representation into a roughly sentential semantic representation. It then searches the lexicon for some combination of words which can appropriately express this initial structure. This will likely involve a good deal of trial and error. This approach, however, vitiates Fay and Cutler's account of malpropisms. If the search procedure selects a neighbor of the target at this stage, that error should be noticed when considering the usefulness of the item in the intended utterance. There are various ways we might overcome this problem. The system might retrieve only semantic information on the first pass, then execute another retrieval operation in which it extracts everything except semantic information. This is ad hoc and serves no purpose other than to preserve the present theory of malapropisms. We might have the network system run backwards after each successful contact with a real lexical item. The output at the origin of the path system will specify the lexical item contacted. This too will reveal the presence of malapropisms since the backward reading of the network will start from the error word in the lexicon, not the target. Thus Fay and Cutler's network system avoids direct contact between the semantic information in the initial representation of the utterance and the semantic information in the lexicon only at the cost of creating a problem with the lexicalization of the initial semantic representation. These difficulties are a direct result of the heterogeneous memory assumption.

The passive memory assumption also plays an important role in the Fay and Cutler theory. I will not, however, specifically consider their use of it other than to suggest that the arguments which are advanced against search theories in the next section also count against the Fay and Cutler proposal.

3. Marslen-Wilson and Welsh (1978) offer an interesting contribution to the theory of lexical access in language comprehension. The major questions in this area have to do with determining how the listener discovers which lexical meanings are intended by the speaker, how various kinds of contextual constraints become involved in this process, and at what point in the process particular kinds of context can be exploited. Marslen-Wilson and Welsh argue that the process of word recognition is the major locus of interaction between the sensory input and the constraints on the interpretation of that input provided by the listener's knowledge of his language. To put it in their terms, they regard word recognition as the point of convergence between bottom-up processing strategies based directly on the form of the acoustic input and top-down processing strategies based on the listener's prior knowledge of his language and the world.

One of Marslen-Wilson and Welsh's main concerns is to evaluate a class of theories of the mental lexicon which posit a passive memory which is to be searched by some serial procedure. These theories predict a strict temporal structuring of the kinds of information which can be employed at various points in the comprehension process. Marslen-Wilson and Welsh take a recent proposal of Forster's (1976) as representative of this class because his is possibly the most thoroughly and carefully worked out. Forster proposed a lexicon with one master file containing all of the phonological, syntactic, and semantic information about each lexical item. To overcome the ordering problems Fay and Cutler confronted, he proposes that several access files are associated with this master file. Each access file contains just one kind of information about each lexical item, together with a pointer indicating where in the master file the full entry for that item can be found. Thus to enter the lexicon during production we will search a semantic access file. There we find a semantic representation for every item in the speaker's vocabulary and a pointer directing us to a particular location in the master file. The advantage of this proposal is that we can now order each of the access files according to their own content, thus enabling efficient searches of the lexicon from any sort of initial description of the item wanted. In the comprehension process this theory predicts that no semantic or syntactic information can play a role in the initial phonologically guided search for a lexical item since none of this information is in the phonological access file.

The crucial evidence which Marslen-Wilson and Welsh bring to bear on Forster's theory is derived from the shadowing task and the listening for mispronunciation task. In shadowing an observer listens to a tape recording of spoken prose while simultaneously reproducing what he hears as quickly and accurately as possible. By making a single tape recording of the input which the observer hears and the output he produces, and by introducing occasional mispronounced words in the input, Marslen-Wilson and Welsh are able to extract a great deal of information about the time course of processing in comprehension. Their results strongly suggest that listeners can, at least some of the time, correctly identify a word in as little as 150 to 200 msec. That is, word selections are being made in about the time it takes for the observer merely to hear the first two or three segments in the word. The average duration of the words in one of these studies was about 370 msec. Nonetheless, even with these extraordinarily short identification times, there are significant effects of syntactic and semantic context. Scrambling the semantic context in a string seems to add about 60 msec to the identification process.[4]

As Marslen-Wilson and Welsh point out, these results suggest that syntactic and semantic context can play a role in the initial identification of a lexical item. Clearly there are many lexical items which cannot be uniquely picked out purely on the basis of phonological information about their first two or three segments. If correct discriminations are made before there is sufficient phonological information to justify them, it seems necessary to invoke other properties of the items. This Forster cannot do with an access file containing only phonological information. On the other hand, if Forster gives up the access files and attempts initial lexical identification in the master file, where there is a greater variety of information, all of the ordering problems discussed earlier re-emerge.

Though none of these considerations falsifies Forster's theory, they do seriously undermine its plausibility. Forster's only defense seems to be to claim that the initial search is done in a heterogeneous master file and that the search process is very fast; and even this does not answer every objection.

An alternative to search theories such as Forster's is available in the "logogen" model proposed by Morton and Broadbent (1967). The logogen model incorporates an active memory (in my sense of the term, but not in theirs) with a single memory element for each item in the lexicon.[5] This one element stores all of the phonological, syntactic, and semantic information about the lexical item it serves. During comprehension it rises to higher and higher levels of activation as it finds more and more similarities, of whatever kind, between the input and the lexical item the logogen stores. When the logogen reaches some threshold value it fires. Firing releases the descriptions of the lexical item the logogen contains to the rest of the system and is taken by Morton to constitute recognition of the item. There is a fairly wide range of evidence from comprehension studies which can be accounted for by one or another version of the logogen model. For example, the model readily accommodates the very early context effects which Marslen-Wilson and Welsh report. Nonetheless, in what follows we will be more concerned with the model's inadequacies than its successes.

Marslen-Wilson and Welsh discuss several problems in some detail. It should be sufficient to mention only the most severe. Since the logogen model provides no access to the incoming signal except through the logogens themselves, it seems unable to provide a convincing account of how speakers detect, store and reproduce deviations from the standard forms of words. If a word is mispronounced, the logogen model most naturally predicts there will often be a perceptual hole in the input stream, since there should be many such occasions when no logogen will fire. Similar problems arise when a person first hears a new word of his language. There seems to be no way to perceive such a thing, let alone determine whether or not it is a possible word or mere gibberish.

A further problem which Marslen-Wilson and Welsh do not discuss results from the serial order predictions which the logogen model seems to imply. Each logogen is independent of the rest and each has equal direct access to the incoming signal. Because context occurring at various distances forward of a lexical item in an incoming signal may contribute to the activation of the logogen, we may expect that the order in which the logogens fire will sometimes be different from that in which the lexical items are actually represented in the speech stream. Thus, the model predicts that it is possible in principle for a listener to perceive two different word orders on successive presentations of the same stretch of recorded speech. As far as I know, this never happens. Though errors of reordering do seem to occur with some reliability (though low frequency) in speech production, they do not seem to be characteristic of speech perception.[6]

Marslen-Wilson and Welsh propose a theory of word recognition which solves some problems of the search and logogen models. Their model is closely related to the logogen model in retaining the assumptions that there is a single memory element for each lexical item, that each element has complete information about all the properties of its lexical item, and that each element has independent direct access to the incoming signal. According to Marslen-Wilson and Welsh's theory, however, "...each memory element in the lexicon will be a computationally active processing entity," (p.56). Much recognition processing occurs in the lexicon itself. More specifically, each memory element can discover what the contextual requirements are for its own recognition. "...Given that a lexical memory element can be informed about the requirements of context, it can then determine whether or not the word it represents is either syntactically possible or semantically plausible at that point in the utterance" (p.58). The most marked difference between the logogen model and Marslen-Wilson and Welsh's model is in the means by which the individual elements respond to phonological properties of the incoming signal. Instead of having a few logogens (and ultimately just one) rising to higher levels of activation as they find themselves in closer accord with the properties of the input, Marslen-Wilson and Welsh propose just the reverse. That is, word recognition begins when a large group of elements are simultaneously raised to an active state on the basis of a loose match to the input. Then as the input continues, each of these active elements examines it. When an element finds a mismatch between the item it recognizes and the signal coming in, it shuts down and withdraws from the pool of active elements. As all of the active memory elements monitor the phonological, syntactic and semantic properties of the input we assume that all but one of them will very rapidly find some degree of divergence between itself and the item coming in. The last active element transmits its contents to higher order processors to effect recognition of the item and to conclude the recognition cycle.

Marslen-Wilson and Welsh's model provides a better account of many details of the comprehension data they review. In particular they provide a more elegant account of how words can be identified before they have been entirely heard. By having the input winnow the active elements to one, we avoid a variety of problems that arise when we assume rather trigger-happy logogens. Though the details needn't concern us here, the new model is also more successful at predicting the conditions of occurrence and the properties of instances where the listener detects or is able to reproduce a non-word or a mispronunciation. Nonetheless, the model does not incorporate any very clear account of how such cases can be perceived, because, like the logogen model, it confines the listener's knowledge about the phonological analysis of acoustic inputs to the memory elements. Again, because of the ablative character of the recognition procedure, the model also provides a somewhat better account of how the serial order of the lexical items is maintained. It successfully avoids the danger with the logogen model of having either too many words recognized or none at all; though the account in the latter case is still none too clear. Nonetheless, Marslen-Wilson and Welsh's model does not seem to provide an answer to the question why the one serial order assigned to the items recognized is so reliably the right one.

Turning now to the more severe difficulties, we confront the question of how much processing activity is going on inside each element of the memory. If we take Marslen-Wilson and Welsh at their word, there seems to be quite a lot. They are not very explicit on this crucial point, but they forsee at least some syntactic and semantic analysis going on inside each memory element. Since doing syntactic and semantic analysis presumes reference to the properties of particular lexical items, we may have an infinite regress. Each memory element will have to have access to and make actual use of, information about lexical items other than the one it stores and thus each memory element will have to have its own internal lexicon. We can avoid this problem easily enough if we posit a single system, independent of the lexicon, which performs various higher order analyses on the incoming utterance as it appears. The results of these higher order analyses are then coded in some fashion and fed into the lexicon. The individual memory elements then base their context decisions on this externally provided information rather than on internal computations. Though some such system might be made to work, there is little evidence that this is the kind of situation Marslen-Wilson and Welsh have in mind. If they do mean to make the memory elements dependent on external computational processes for information about context, they then have no reason to ascribe computational powers of any interest to the memory elements. The elements need only be able to compare the inputs provided by various external sources with their own internal records of the lexical items they represent. This is not much in the way of "a computationally active processing entity."

Another class of problems arises when we consider what role the memory model Marslen-Wilson and Welsh propose might play in the production process. At the outset we should note that this is not quite fair to them since they make no mention of the production problem. Nonetheless, the present aim is to find some unified theory of the lexicon for production as well as comprehension. Thus it seems legitimate to evaluate the Marslen-Wilson and Welsh model in this context. In order to employ their lexicon in this way, we add the assumption that during the production process the system is able to retrieve phonological or syntactic information by broadcasting approximate semantic descriptions into the lexicon, just as the acoustic representation of the incoming signal is broadcast to all memory cells during comprehension. If this assumption can be effectively implemented, the initial conceptual representation of the utterance will be translated tentatively into a semantic representation. Some procedure selects portions of this representation as candidate lexical items. These partial representations are then injected into the lexicon. If one of the memory elements happens to correspond reasonably closely to the input, the element transmits a comprehensive description of its lexical item. From this procedure the controlling element develops a pool of candidate lexical items which it then submits to further semantic and syntactic processing. When this is complete, the phonological information for the items employed is used to develop an articulatory description of the utterance. Such a system can successfully produce utterances.

Nevertheless, what we need to explain production errors such as malapropisms is not only a system that works, but one that fails as well, and in quite particular ways. To be exact, we need some mechanism by which the model naturally predicts a class of error words which are related to their targets phonologically but not semantically. Because the system needs to go into the lexicon just once to retrieve all of the information about any one item, and because this retrieval episode is guided by semantic criteria, we have no basis for predicting the generation of malapropisms. We also cannot predict a class of semantic errors because the retrieval stage is followed presumably by a comparison of the semantic properties of the lexical items retrieved and the properties of the initial semantic representation of the utterance. Thus the Marslen-Wilson and Welsh model does not seem adequate as a model of the lexicon in production.

4. Before we turn to consideration of a new theory of the mental lexicon it will be helpful to carry our examination of the speech error data a little further.

What is it about malapropisms and semantic substitution errors that makes them difficult to accomodate in a theory of the lexicon? Apparently, when these distortions occur the production system is searching for a word under the guidance of phonological or syntactic criteria. But why, one might reasonably ask, should these kinds of criteria have any role at all in the selection of words in production? Aren't words chosen according to semantic criteria in production? Surely they are, at least initially. Nevertheless, malapropisms suggest that there is another stage during the production process when words in the developing utterance are selected by quite different criteria. The theories we've already examined cannot accommodate these errors because they have all incorporated both the assumptions that the lexicon is a single heterogeneous memory containing all types of lexical information, and that the lexicon need only be used once per lexical item in each episode of production or comprehension.

Notice that semantic errors also suggest more than one use of the lexicon in production. Semantic errors are interesting because the choice of the error word seems to be under the control of some sort of semantic criteria but not under the control of the overall semantic or conceptual plan of the utterance. They suggest an instance of lexical retrieval by semantic criteria that somehow follows the stage at which candidate words are considered for inclusion in the utterance.

The solution I will propose to these puzzles is based on the rejection of all three of the assumptions sketched at the outset. I will outline an integrated theory of the mental lexicon that assumes that there are several homogeneous lexicons, that there are several distinct points in the production and comprehension processes at which these play a role, and that all of the lexicons are active devices.

In this theory there are three separate active lexicons which are linked together in specific ways. Each lexicon stores only one kind of information about every item in the speaker's vocabulary, and each lexicon has all there is of that kind of information. Each lexical item is stored in a cell. Cells are like the memory elements of the Marslen-Wilson and Welsh theory except that their computational powers are strictly constrained. The only thing a cell can do is compare some input to its own contents. If it finds a mismatch, it withdraws itself as a possible analysis of the present input.

One lexicon contains something like a systematic phonemic description of each morpheme (the "phonemic lexicon"). Another (the "syntactic lexicon") contains a syntactic description of each item represented in the phonemic lexicon, plus representations of additional elements that can play a role in syntactic analysis, e.g., labeled brackets and the like. The next lexicon (the "semantic lexicon") contains a semantic description of every item in the phonemic lexicon as well as descriptions of any syntactic elements that do not appear in the phonemic lexicon.[8]

Associated with each lexicon is a processor which can only receive information from that one lexicon and transmit information into that lexicon. The processor can only deal in the terms used in the lexicon it serves. It cannot recognize or manipulate any other kind of symbol. The processors can communicate with each other only via their respective lexicons.

Somewhere between the ear and the phonemic lexicon there is a phonetic analyzer which converts the acoustic signal into a sequence of distinctive feature matrices. This device must also provide word boundaries.[9]

The language system makes contact with other congnitive cognitive processes through the semantic lexicon. Some other processes may be lexically oriented and play some role in comprehension through the semantic lexicon.

During comprehension this entire ensemble of lexicons and processors (see Figure 1) is initialized by the appearance of the first word boundary provided by the phonetic analyzer. At this point all of the cells are essentially announcing their availability as possible analyses of the incoming stimulus. As phonetic information comes in, each cell withdraws as soon as it detects any mismatch between itself and the input. The cells in the phonemic lexicon are linked by a logical OR function; when all but one have been suppressed, this one transmits its contents to the processor.[10] The three cells in the three lexicons that represent one item are directly linked. Thus whenever the cell for woman in the phonemic file transmits to its processor, the syntactic and semantic cells for this will also transmit to their respective processors. The next word boundary reinitializes the entire system and a new cycle begins.[11] This continues until the syntactic system declares closure on the clause. As each new item enters each processor, the processor applies all its resources to discover as much as possible about what sort of item will come next. The results of this analysis are communicated to the lexicon. One possible form for this contextual information is a specification of the types of items which cannot come next in the sequence of items. The effect of this information will be to withdraw all of the relevant items. That is, in processing an English sentence, if the syntactic processor receives a determiner, it may transmit a signal to the lexicon that the next item cannot be a verb. This will withdraw all of the verb cells within the syntactic system and through them all of their mates in the phonemic lexicon. In this fashion syntactic context can rapidly affect the processing of the incoming signal.[12]

Syntactic properties of the clause which are implicit in the input will be made explicit by the syntactic processor. At the end of the clause, all of the terms in the elaborated string in the syntactic processor will be sent back into the lexicon in sequence. As each item is recognized by the syntactic lexicon, its mate in the semantic lexicon is transmited to the semantic processor. Thus some kinds of syntax-dependent information will never be available to the semantic processor before the clause boundary.[13]

The production process, while essentially the inverse of the foregoing, differs in certain respects (see Figure 2). Most importantly, production is taken to involve the simultaneous planning of outputs at different levels. Thus the semantic system may be preparing one clause while the syntactic system is working on the prior clause. In each case the processor communicates with the next stage of the procedure via its lexicon. When the semantic processor is finished with a clause, it communicates its results to the syntactic system by sending the elements in the clause into its own lexicon. As each element is recognized, this recognition is communicated to the syntactic lexicon where the corresponding element is transmitted to the processor there.

This model is empirically distinct from those discussed above. It predicts that reordering errors can occur in production but not in comprehension. This prediction arises from the assumption that the representation of the clause is retained in the phonemic system as syntactic and semantic processing on its proceeds. This then serves as a reliable guide to the serial order of the elements in the original input.[14] In production there is no similar guide to serial order. Each clause is processed at each level independently of all other levels. Thus there is no independent record of what went into the syntactic system by which to judge what comes out. Malapropisms and semantic errors in production are explained as occurring in the lexical access operations that follow semantic and syntactic processing.[15]

In comprehension, the model benefits from the same arguments that support the Marslen-Wilson and Welsh model, over the logogen and search models, though it remains empirically distinct from theirs. Unlike their model, the one just described predicts a definite time order effect for some kinds of semantic and other context effects. They argue that all varieties of context will be available at the same time. This issue should be experimentally decidable.

If the theory outlined above survives closer examination of the facts of comprehension and production, it may also provide useful predictions about possible language disorders. To get such predictions we adopt the view that the various processors and lexicons mentioned in the theory are discrete physical components of the human language performance system. On the further assumption that it is possible for disorders of the brain to destroy or degrade proper subsets of this set of components without affecting the remainder, the theory defines a class of possible disorders. Each possible disorder is defined as the disruption of some subset of components.[16] The difference between the present theory and others discussed earlier lies only in its provision of a more richly differentiated system of components. A theory which posits a heterogeneous lexicon and a correspondingly versatile processor will predict fewer possibilities for selective degradation of linguistic structure. To the extent that language disorders operate selectively on linguistic levels, the heterogeneous memory assumption makes them hard to explain.

5. A theory of the kind sketched above may be able to integrate fairly diverse facts about production, comprehension, and possibly language disorders. The model achieves this wide application as a result of three key assumptions about the mental lexicon. First, the model assumes that lexical information is stored in active memory devices which have a strictly limited capacity to compare their contents with a signal supplied from outside and to respond appropriately. Second, the model assumes that lexical memory and linguistic processing is subdivided into levels such that each level operates strictly in terms of a single homogeneous set of primitive terms. Finally, the model assumes that the various levels of the system communicate with each other only via their respective lexicons.

Notes

1. One might also wonder how the network proposal might fair if it were to turn out that in order to properly distinguish all of the senses in some natural language we were forced to resort to some system of semantic description that could only be defined by a Type 0 grammar, if the lexicon were infinite. In such a case the network proposal could remain viable only if a network could be constructed that would discriminate all of the lexical items. Since the lexicon is finite, there will be some Type 3 grammar that (weakly) generates all of the semantic descriptions needed. Further, for any Type 3 grammar there is a network of the kind that Fay and Cutler describe. Hence, for any semantic lexicon there is a network of the needed kind. The nodes in such a network, might, however have nothing whatever to do with the semantic properties represented in the lexicon. The existence of the Type 3 grammar guarantees only weak generation of the set of lexical items and thus does not guarantee that the paths through the network have anything to say about the lexical items at their end points.

2. I wish to thank Fred Katz for calling this important problem to my attention.

3. See, for example, the discussion and references in Fodor, Bever and Garrett [1974, 384-388] and Foss and Hakes [1978, Chapter 13].

4. Some of these claims are based primarily on earlier research. See especially Marslen-Wilson (1973, 1975) and Marslen-Wilson and Tyler (1975).

5. Morton and Broadbent (1967) describe their logogens as passive elements. They wish to emphasize that there is no search procedure by which some controlling element searches the lexicon. Rather they envision an access procedure in which there is no higher level activity, but only a passive acceptance of the information provided by the logogens. The logogens themselves are also passive in that they are not searching or examining the input, but merely responding to the presence of specific acoustic patterns.

Though Marslen-Wilson and Welsh (1978) attribute computational powers to their memory elements (corresponding to logogens) they still seem to regard their system as passive since the computational activity that corresponds to a search procedure is in the lexicon itself and not in some sort of controller.

Since I am assuming that any model of speech perception is going to have to posit some active processing system which uses the output of the lexicon, whether or not it actively searches the lexicon, I will use passive only with reference to properties of individual memory elements in the lexicon. If the element, by whatever means, makes the "decision" to recognize its word, then it is an active element. If all of the elements respond to the same signal and must be searched by a controller, they are passive (see §1).

6. The logogen model also seems to predict that there will be occasional instances of the simultaneous recognition of two or more items, another phenomenon which seems not to occur.

7. The empirical motivation for the homogeneous memory claim comes entirely from the speech error data. There are also several plausibility arguments that might be advanced on its behalf.

1) Suppose that the memories involved in speech production and comprehension are all heterogeneous and that each of the primitives of the different levels of analysis in the system is represented by a sequence of the primitives of the memories themselves (see §1). In such a system the critical information to which the processor must respond lies in the structure of some concatenation of primitives, as well as in the identities of the various primitive terms appearing in it. This entails that the processors will have to be able to deal in two levels of structure simultaneously; they will have to be able to maintain the sequence of their physical primitives in order to maintain the record of the higher order primitives used in some representation, and they will have to be able to maintain the sequence of the higher order primitives as well. We already know from Fromkin's (1971) error data that the processors involved are not error free in this respect. We do get rearrangements of both segments and words from time to time. Thus, imposing the requirement that the system keep track of two levels of structure simultaneously considerably increases its vulnerability to a type of error we know it makes.

2) We might avoid the two levels of structure by having a system which can map each of the linguistic primitives onto one physical primitive term. Thus the processor could keep track of the identity of the linguistic primitives by reference to the identity of the physical primitives. The structure of the sequence of physical primitives would be significant at only one level. Unfortunately, this implies a very large set of physical primitives. This too imposes a particular burden on the system. We may regard a memory as a communication channel that links a processor with itself. As with any other channel, memories will have a certain bandwidth. Roughly, the more primitives the system recognizes the more slowly it will operate (imagine a telegrapher who has to distinguish two varieties of "dot" and two varieties of "dash", as opposed to one who need only distinguish the usual one of each). What is optimal can only be determined by examination of the characteristics and task of the processor and memory in question. In general, however, we can say that systems employing very large sets of primitives are somewhat less likely than those employing relatively few.

3) Returning to the coding possibility, we should note that there is less inherent constraint on a system that concatenates physical primitives to represent linguistic primitives. In principle, such a system could represent n^m different primitives, where n=the longest sequence of physical primitives the system may regard as representing a single linguistic primitive (this is its machine level "word length") and m=the number of physical primitives the system discriminates. If n^m is greater than the total of linguistic primitives

required, we must explain why the system doesn't sometimes err and create too many linguistic primitives, or too few, and result in the speaker having a significantly different grammar than other members of his linguistic community. In other words, the coding possibility complicates the language learning problem somewhat.

4) Finally, since at least some language disorders have been shown to be associated with damage to specific parts of the brain, the assumption of a heterogeneous memory, especially one based on the coding notion, moves us somewhat further away from an account of such disorders. If linguistic primitives are represented by some code on physical primitives, we might reasonably expect that similar physical primitives would be available in other parts of the brain. Thus one might expect somewhat greater flexibility in the assignment of linguistic function to other parts of a damaged brain. Why couldn't linguistic processing, especially given its vital role in the life of the organism, be undertaken somewhere else?

8. I do not wish to claim that this list is exhaustive, nor that it is not exhaustive. At this point I have no idea whether or how a cut can be made between, for example, representations of stylistic facts, or representations of pragmatic facts, and the representations of semantic facts in the semantic lexicon. It may well be that we'll ultimately need a theory with one or more additional lexicons to cope with these dimensions of lexical use. Though I have no doubt that a cut can be made, and must be made, between representations of linguistic knowledge and representations of practical knowledge, some relation must be provided between these two. Unfortunately, I see virtually no way to constrain speculations on this relation at this time.

9. Some sentences, when produced normally, are virtually impossible to segment into words, e.g., "Mares eat oats, and does eat oats, and little lambs eat ivy." The example is typically heard as something like "Marsey doats, and doesy doats, and little lamsey divey." This shows, it seems to me, that the existence of an item in the lexicon is not sufficient to make that item recoverable from the speech stream. Thus we seem to have need of a procedure, external to the lexicon, which performs at least a preliminary segmentation into morphemes or words. Others have drawn quite different conclusions from the same data.

10. Obviously we cannot assume that all cells but one will have been suppressed by the end of every word. Any theory will ultimately have to provide some mechanism for resolving problems of this sort. For the moment I see no basis on which to select an appropriate mechanism within this theory. It is possible, however, to construct several plausible mechanisms.

11. The word boundaries carry a special burden for coordinating the work of the three processors. Since the three processors will communicate only by way of their respective lexicons (see below) we will need to assume that word boundaries are maintained by all three processors and that the boundaries themselves are indexed. Thus the material falling between boundaries _i_ and _i+1_ in each of the three processors will refer to corresponding portions of the input string.

12. A detailed proposal about the means by which syntactic context may have an effect has not yet been worked out. When the syntactic categories are directly named, the above proposal has the disadvantage that it requires a different pattern of response from the cells than that used in other phases of the operation of the syntactic lexicon. Where ordinarily a cell withdraws on discovery of a mismatch with input, here it would be required to withdraw when it matched. The inverse of this proposal is also unattractive. If we have the syntactic system transmit a list of the syntactic categories which can appear in the next position, we will have to have those cells that do not match this description withdraw. If more than one category is named in the context statement, the entire lexicon will withdraw. For example, the context signal following receipt of a determiner might specify that the next item could be a noun or an adjective. This would cause all non-nouns as well as all non-adjectives, i.e., everything, to withdraw.

A more attractive proposal seems feasible if we take account of Chomsky's suggestion (1972) that syntactic categories be regarded as feature bundles rather than as primitives of the syntactic system. Using Jackendoff's feature system (1977, 33) we can specify the properties of what may follow each of the elements of a simple sentence in something like the following manner. On recognition of an initial proper noun, the syntactic processor transmits into the lexicon the information that the next element must be marked [α Subject, α Object] assuming canonical order. This suppresses all items having a contrast in marking between these two features e.g., nouns, prepositions, etc. Similarly, recognition of a main verb induces the system to specify the next element as ([-Subject] [-Object]) thus suppressing all verbs and models. The appearance of an article leads to the prediction that the following element will be marked [- Object, - Determiner], thus suppressing everything except nouns, adjectives, adverbs, and certain quantifiers.

This proposal must be further refined to deal with the occasions where canonical form is violated. This can be done in the course of dealing with another problem which the syntactic processor must handle. The primary function of syntactic analysis in comprehension is the delineation of the phrasal constitutents of the incoming sentence and the hierarchical relations among these. One way to deal with this problem is to posit a set of elements which the syntactic processor may insert into a string during comprehension. These elements correspond to brackets and there must be several varieties of them. Furthermore, the system of brackets can be discriminated by an extension of the syntactic feature system.

Given assumptions such as these, we can allow for non-canonical sentences by giving priority to the assignment of brackets rather than the anticipation of the syntactic character of upcoming elements of the string. The syntactic system tentatively assigns a bracketing to the incoming constituent as soon as its syntactic properties are recovered from the lexicon. Violations of canonical order of the kind that occur with conjoined strings of nouns or verbs, or with various adverbial phrases, are typically phonologically marked in the final syllable before the divergence occurs. We posit a hierarchically ordered system of bracket types with the type implied by canonical order at the top of the hierarchy invoked for any particular syntactic environment. The phonological information associated with non-canonical boundaries can then be used to cue the syntactic system to replace the canonical boundary type with some type that occurs at a lower level of the hierarchy. The syntactic context signal transmitted into the syntactic lexicon during receipt of a new lexical item would then be determined by the character of the immediately preceding boundary marker, not by the immediately preceding lexical item.

13. Note that the empirical distinction between the present theory and that of Marslen-Wilson and Welsh remains even if the syntactic system communicates with the semantic system at the ends of units smaller than the clause. If information passes from one level to the next more frequently than suggested, the two levels of analysis can still never be fully simultaneous. Some aspects of semantic analysis will always have to await syntactic analysis.

14. There is no mechanism for direct coordination of the three processors. Rather coordination is achieved at the end of each clause when the contents of the three processors are read out into their respective lexicons. This is presumed to be coordinated by reference to the indexed morpheme boundaries provided (see Note 11). Roughly, the three processors must "select" coordinate forms of the same lexical item with the material each transmits between boundaries i and i + 1, or some sort of reprocessing routine is invoked.

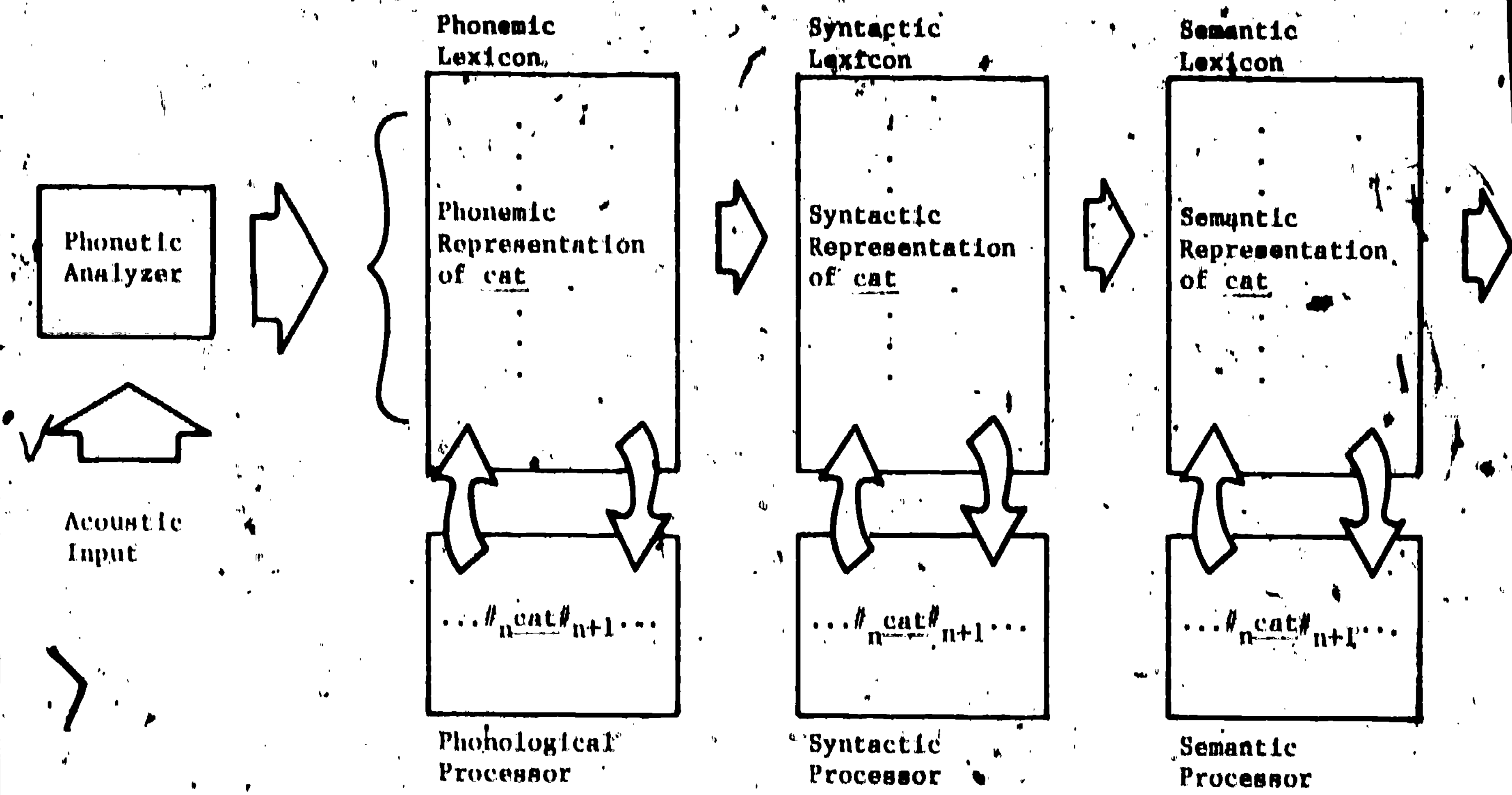
15. More specifically, the suggestion is that semantic errors occur when the semantic system is communicating with the syntactic system after the semantic integration of the utterance is complete. The original conceptual representation is no longer available. The semantic processor reads out its contents into the semantic lexicon. As each item is recognized, it causes its mate in the syntactic lexicon to be relayed to the syntactic processor. If an error were to occur at this stage it should reflect semantic constraints. The initial description of each item is in semantic terms and the cells in the semantic lexicon will, of course, assess inputs in semantic terms.

In the syntactic lexicon the situation is a little more complex. Many items in the syntactic lexicon will have the same syntactic properties as some one or more other items. In order to keep these items distinct it will be necessary to supplement each syntactic description with a quasi-phonemic description of the item. These quasi-phonemic representations must have two properties, a) they must represent at least part of the information about distinctive features provided by a systematic phonemic description of an item, and b) they must function as indivisible units at the syntactic level. They are best regarded as a kind of complex primitive term; they are complex because the system recognizes their internal structure, but they are primitives since their parts cannot be operated on in isolation. On these assumptions malapropisms are predicted as errors occuring in the lexical selection process occuring after the syntactic integration of the utterance. Syntactic descriptions of items are transmitted into the lexicon. Since both the properties of the incoming descriptions and the content of the lexicon are described in syntactic and quasi-phonemic terms, the errors that occur will reflect these kinds of constraint. That is, we expect error words selected at this level to resemble their targets in both syntactic and phonemic structure.

By treating phonemic segments as complex primitives rather than feature matrices, we get a further prediction that there will be a potential class of segment-sized reordering errors such as those reported by Fromkin (1971). Since the phonemic system clearly must deal in feature matrices, we cannot as readily predict this sort of error at that level since the unit on which that level operates, the feature, is not the unit of this error type.

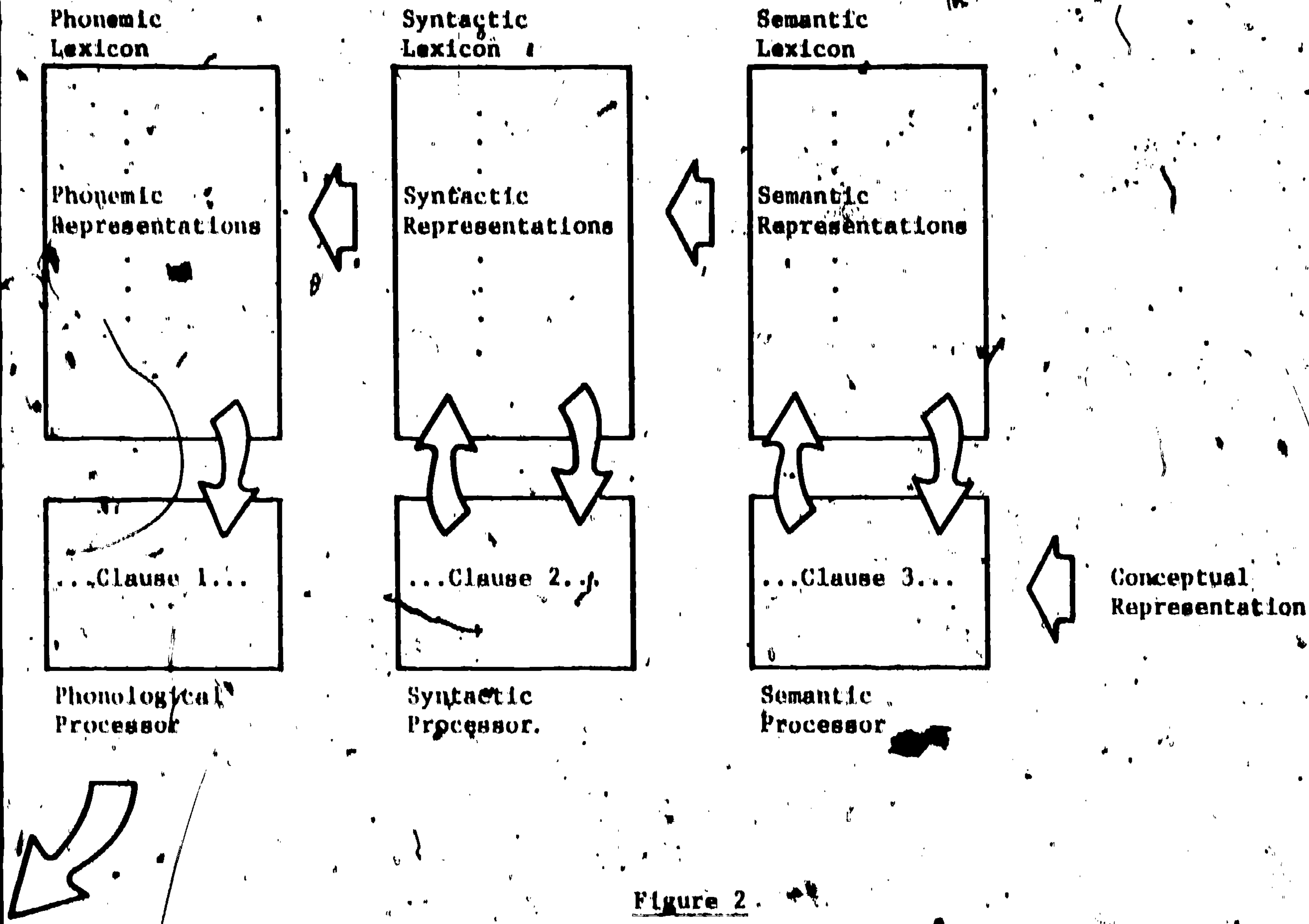
16. Zuriff (1978) has discussed some current work on Broca's aphasia that is particularly interesting in this connection. Though most investigators have claimed that agrammatic patients (those who produce "telegraphic" speech lacking most grammatical morphemes, inflections, etc.) have essentially unimpaired comprehension ability, Zuriff argues that these patients do in fact have a comprehension deficit, and furthermore, one that closely parallels their production deficit. Even where problems of effort and memory are controlled for, agrammatic patients were unable to correctly use auxiliary verbs, prepositions, etc., in several comprehension studies Zuriff discusses.

Something like this pattern of results is predicted by the present theory if the syntactic processor is removed from the system.



Phonetic Analyzer
Acoustic Input
Phonemic Lexicon
Phonemic Representation of cat
...#ncat#n+1...
Phonological Processor
Syntactic Lexicon
Syntactic Representation of cat
...#ncat#n+1...
Syntactic Processor
Semantic Lexicon
Semantic Representation of cat
...#ncat#n+1...
Semantic Processor

Figure 1



Phonemic Lexicon
Syntactic Lexicon
Semantic Lexicon
Phonemic Representations
Syntactic Representations
Semantic Representations
...Clause 1...
...Clause 2...
...Clause 3...
Conceptual Representation
Phonological Processor
Syntactic Processor.
Semantic Processor

Figure 2

References

Chomsky, Noam (1972). "Remarks on nominalization," in Studies on Semantics in Generative Grammar. The Hague: Mouton.

Fay, David and A. Cutler (1977). "Malapropisms and the structure of the mental lexicon," Linguistic Inquiry, 8, 505-520.

Fodor, Jerry A., T.G. Bever and M.F. Garrett (1974). The Psychology of Language. New York: McGraw-Hill.

Forster, Kenneth I. (1976). "Accessing the mental lexicon," in R.J. Wales and E. Walker, eds., New Approaches to Language Mechanisms. Amsterdam: North Holland.

Foss, Donald J. and D.T. Hakes (1978). Psycholinguistics. Englewood Cliffs, N.J.: Prentice-Hall.

Fromkin, Victoria A. (1971). "The non-anomalous nature of anomalous utterances," Language, 47, 27-52.

Jackendoff, Ray S. (1977). X Syntax: A Study of Phrase Structure. Cambridge, Mass.: M.I.T. Press.

Marslen-Wilson, William D. (1973). "Linguistic structure and speech shadowing at very short latencies," Nature, 244, 522-523.

Marslen-Wilson, William D. (1975). "Sentence preception as an interactive parallel process," Science, 189, 226-228.

Marslen-Wilson, William D. and L.K. Tyler (1975). "Processing structure of sentence perception," Nature, 257, 284-286.

Marslen-Wilson, William D. and A. Welsh (1978). "Processing interactions and lexical access during word recognition in continuous speech," Cognitive Psychology, 10, 29-63.

Morton, John and D. E. Broadbent (1967). "Passive versus active recognition models, or, Is your homonculus really necessary?," in W. Wathen-Dunn, ed., Models for the Perception of Speech and Visual Form. Cambridge, Mass.: M.I.T. Press.

Zuriff, Edgar (1978). Colloquium presented at the Graduate Center, CUNY, New York, N.Y., December.