Lexical Processing during Sentence Comprehension:
Effects of Higher Order Constraints and Implications for Representation

DAVID A. SWINNEY

Tufts University

A fundamental problem in language comprehension is to obtain an accurate characterization of lexical processing and of the influence of contextual information upon that processing. This paper will examine evidence concerning the effects of so-called 'higher-order' constraints upon the various stages of lexical processing and will outline a model of the representation and processing of lexical material during sentence comprehension.

The conclusions that will result from this examination derive from a particular point of view about the general nature of cognitive processing and mental representation. Thus, some preliminary discussion of this view is important. In particular, I will take it as a basic tenet that questions of representation about speech and language can only be sensibly discussed in the context of a well defined performance model. In order to make meaningful (and testable) statements about the nature of representation in language, and to evaluate the relative role of the various representational 'facts' that have been garnered from theoretical linguistics and experimental psycholinguistics, the assumptions about the processes operating on those representations must be detailed. This position is certainly not novel to psycholinguistic inquiry. However, it appears to be overlooked far too often.

Thus, in order to begin at least a partial list of such assumptions, it can be noted that the domain of interest in this paper is that which I will call 'normal sentence comprehension'. This is intended to include those situations in which at least the following conditions prevail:

1. The materials of interest (words) do not occur in isolation, but are part of structured, fluent, utterances in which the goal of the listener is to understand the utterance.
2. The computation of a number of different types of linguistic information takes place (in either serial or parallel mode) during sentence comprehension. Among these informational types is (at least) acoustic, morpho-phonemic, lexical, syntactic, and semantic information. (No claims of independence or interactivity are being posited here.)
3. The speech signal is relatively intact and accessible despite a certain amount of noise in the signal. This is taken to be the most common situation in sentence comprehension even though it clearly does not always hold.

A brief discussion of the motivation behind these points may be helpful. It is reasonable to assume that the cognitive operations involved in different task situations will, themselves, differ. While some underlying operations are undoubtedly shared across different performance situations, it is not a trivial chore to determine which of the inferences that we make about cognitive processes are best attributable to 'underlying' representations and operations and which are
attributable to superficial or task-related operations. Results from certain task situations may reflect more of how subjects can perform given certain requirements, than of what subjects normally do during language comprehension. Obviously, when sufficient evidence from different 'task' sources coincide, we have the basis for arguments about underlying representations. However, that situation is rare. Thus, because the task a subject sees himself as having – the query he poses to his/her cognitive system - may determine much of the cognitive strategies employed in that situation, a conservative position will be adopted here: we will not examine data about 'isolated' lexical tasks (such as isolated lexical decisions, same-different judgements to word pairs, category inclusion/exclusion decisions on single words) except when no other data is available. In doing so, some relevant evidence will undoubtedly be overlooked, but the model we sketch will also not be forced to account for irrelevant or inappropriate data.

The assumption that different types of information are generated during sentence comprehension follows at least in part from the extensive documentation of the existence, at some level, of these informational types (see, e.g. Fodor, Bever, and Garrett, 1974). Because the purpose of this paper is to examine evidence about effects of the 'higher-level' information on lexical processing, no constraints on the availability of such information at any point in the processing scheme are being presupposed. (Note that this position is not always held in the literature; see Marslen-Wilson (1976) for a presentation of the arguments.)

Finally, the degree to which the signal is available for 'bottom-up' analysis deserves brief discussion. While much research exists on the conditions in which sufficient 'bottom-up' information is not available to provide unique identification for speech sounds or words, it is arguably the case that an adequate signal is available more often than not. Further, it is only in the presence of an adequate and available signal that the underlying roles of higher order information and bottom-up analysis can be examined definitively; if 'top down' information is only shown to operate following the failure of a 'bottom-up' analysis (due to insufficient low level information) then arguments for the predictivity of top down constraints are weak indeed. To follow this point a bit further, a number of the more important studies which have supported 'top down' constraints on processing (e.g., Marslen-Wilson and Welsh, 1978; Cole, 1973) have utilized tasks which encourage (and often require) these top-down processes to operate. While no one denies that such knowledge sources can have effects on interpretation, these tasks may not actually examine the normal comprehension process, and thus the value of these results to the issue at hand is uncertain.

Before turning to the evidence about lexical processing and processing constraints, it will be worthwhile to briefly consider some points about the general nature of lexical representation. The mental lexicon has been the major focus for theory and research about context effects (the interaction of higher and lower informational types) precisely because the representation for a word must contain 'low level' form (acoustic, phonetic) information, as well as 'high level' syntactic (distributional) and semantic (meaning) information (see, e.g., Garrett, 1978; Foss, in press; Marslen-Wilson & Welsh, 1978). Because these informational types are represented in some (presumably common) code for each word, it has been natural to posit the lexicon as the major point of interaction for these various information types. Whatever the nature of this interaction, it is almost universally assumed that there are identifiable 'stages' (or, depending on your theorist, 'systems' or 'levels' or 'modules') in sentence processing. One of the goals of this paper is, then, to attempt to establish the relative independence or interactivity of each of these stages. Such stages need not only distinguish the major informational types, but may also constitute definable processing subsystems within each informational level. We will examine the manner in which three such subsystems of lexical
processing are constrained by higher order information: The process by which lexical candidates are segmented from (isolated in) the speech stream, the process by which lexical information is accessed, and the post-access processes which may be involved, for example, in the insertion of accessed lexical information into "higher" levels of sentence analysis. While different theorists have argued that these lexical processes operate on a number of different 'lower order' bases (such as the acoustic, phonetic, phonological, or syllabic form of the word; see, e.g., Klatt, this volume; Foss, in press; Liberman and Studdert-Kennedy, 1978; Pisoni, 1978), any and all of these characterizations are sufficient to support the representations and operations that will be discussed. Thus, as questions about the nature of this base are neutral with respect to what follows, they will be largely ignored.

DEVELOPING THE MODEL

Overall, it is undeniable that higher order information exerts some effect on lexical interpretation; evidence abounds that semantic and syntactic information interacts with lexical processing. The question is one of where and how such information exerts its influence on lexical interpretation. There are two general (and extreme) models of this process -each with a large number of variations in the Literature. One can be called the Contextually Predictive model and the other the Contextually Independent model. The Contextually Predictive model holds that higher order information acts to direct lexical processing. In the strong (and perhaps most common) version of this model, contextual information is used to predict subsequent information and, by doing so, to constrain and direct all stages of lexical processing. This is, then, a maximally interactive model in which semantic and syntactic information can each act to short-cut reliance on form alone in the retrieval and use of lexical information. The Contextually Independent model, on the other hand, holds that higher order constraints do not exist on all stages of lexical processing. Such constraints are seen as acting only on the output of the lexical retrieval process, following access. Thus, segmentation and access are viewed as independent and isolable subsystems that operate solely based on lower level information.

Examinations of the various forms of these models come from a number of sources. In the sentence comprehension domain, however, the overwhelming proportion of the evidence comes from studies which have examined the effects of contexts upon the interpretation of polysemous words (usually unsystematic lexical homographs/homophones). The reasons for this are straightforward. The meanings of lexical ambiguities can be differentiated rather easily, and thus the selective effects of higher order contextual constraints upon the functional activation of these meanings are relatively available to empirical examination. Further, not only are lexical ambiguities arguably as common as unambiguous words in language use, but nearly all words exhibit some type of indeterminancy in characterizations of their meanings anyway. Thus ambiguous words provide a useful and well founded vehicle for examining the effects of context upon lexical processing. Given a sufficiently sensitive task, one should be able to determine whether strongly predictive (higher order) contexts constrain the various stages of lexical processing by examining which meanings of the ambiguities are functional in the presence of these contexts.

LEXICAL ACCESS

The evidence from the large body of work in this area can be organized into a relatively coherent story. First, in the absence of any strongly biasing context it appears that all meanings of a lexical ambiguity are accessed. Such access is not available to conscious introspection and (typically) the listener eventually becomes aware of only one of the several meanings that was accessed for the ambiguity. More
importantly, even in the presence of prior constraining semantic and syntactic contexts (contexts which strongly bias interpretation of the word toward just one of its meanings) all interpretations for an ambiguous word are accessed during sentence comprehension.

Most of the studies in this area have examined the effects of semantic contexts, and the majority of these have utilized the so-called 'ambiguity effect' in their investigations. The principle behind this effect is that the presence of a lexically ambiguous word causes increased processing difficulty (when compared to an unambiguous word) in sentence comprehension. The presence or absence of this 'effect' is thus taken as an indication of whether context is simplifying the access process. Use of the ambiguity effect has occurred in a number of experimental paradigms. Holmes, Arwas and Garrett (1977), for example, demonstrated that time to classify a sentence as meaningful is increased by the presence of a lexical ambiguity and, also, that the number of words recalled in a rapid serial visual presentation task is fewer in the presence of an ambiguity. Chodorow (1979) has shown that recall of material under compressed speech presentation is poorer when lexical ambiguities are present. MacKay (1974) demonstrated that the presence of ambiguous words in sentence fragments causes sentence completion times to be slower. Foss and his colleagues (Foss, 1970; Foss and Jenkins, 1973; Cutler and Foss, Note 1) as well as Cairns and Kameron (1975) and Cairns and Hsu (1979) have demonstrated that monitoring for a phoneme takes longer following an ambiguous word, even in the presence of a prior, biasing context.1

In addition to this work (which, overall, has argued in favor of the Contextually Independent model largely by virtue of failing to discover reductions in the 'ambiguity effect' in the presence of biasing semantic contexts) there is work that has reflected the access and activation of the several meanings of an ambiguity somewhat more directly. Lackner and Garrett (1972) provided such evidence using a task which combined dichotic presentation of sentences and a shadowing/paraphrasing task. Conrad (1974) also provided such evidence using a modified stroop task. She found interference for words related to each meaning of an ambiguity that appeared at the end of a sentence containing a biased semantic context. Similarly, Swinney (in press) and Swinney, Onifer and Hirshkowitz (Note 2) found that, even in the presence of strongly biasing semantic contexts, lexical decisions to visually presented words which were related to each meaning of an ambiguity were facilitated when presented immediately after the ambiguity was heard. Thus, in summary, there is a fair amount of evidence which argues that even in the presence of prior, strongly biasing, semantic contexts all meanings of an ambiguity are accessed, at least momentarily, during sentence comprehension.

Before examining this evidence in more detail, it is important to reconsider a methodological concern that was raised earlier. While we might hope that the tasks mentioned above (phoneme monitoring, paraphrasing, semantic priming, stroop interference, lexical decision and sentence completion) all reflect the same underlying lexical (access) process, that assumption must be carefully examined because each task clearly adds its own set of required strategies to those needed for sentence comprehension. And, in addition, the tasks may be differentially sensitive to the lexical access process. While it is clear that we cannot have the ideal situation - one in which the task would add none of its characteristics to the basic sentence comprehension operation and would also be maximally sensitive to underlying processes - some of the tasks that have been used appear to approach that ideal more closely than others. One relevant factor appears to be the relative point of application of the experimental task to the comprehension process. Experimental techniques which examine lexical processing only following comprehension of the sentence containing the word of interest (such as occurs in some recall, paraphrase, and sentence classification tasks) appear to reflect different aspects of lexical
processing than do tasks which occur temporally close to the critical word during comprehension of the sentence (such as in phoneme monitoring, lexical priming, and certain click detection and error detection tasks). Garrett (1970) was the first to point out this fact, and many others have since demonstrated its essential truth (see, e.g., Foss, 1970; Forster, 1979; Cutler and Norris, 1979, for discussion.). In all, it appears that the on-line or ‘perceptual’ tasks more accurately reflect the nature of the processes used to gain access to the representation of lexical items, and that the other ‘post-perceptual’ tasks appear to reflect only the final outcome of such processes. Another methodological concern is the degree to which subjects are presented with predictable experimental situations which allow and encourage the development of specialized processing strategies that may affect (change) the processes under investigation. In general, it is obvious that the less any experimental task permits such special strategies to develop, the more certain one can be that it is a ‘normal’, basic, process reflected in the data obtained with that task.

The point of this discussion is that a number of reviews exist which indict many of the experimental techniques discussed above for violations of these methodological considerations. At least partially for the reason that it appears to be one of the studies least susceptible to these methodological problems, the work reported by Swinney (in press) will thus be examined here in some detail. In this work, a cross-modal lexical priming task was used. In this, subjects had two tasks. First, they were required to comprehend auditorily presented sentences which contained lexical ambiguities. In some conditions, a strongly biasing semantic context occurred prior to the ambiguity in these sentences. Subjects also performed a second task. They made lexical (word/nonword) decisions about visually presented letter strings. (In all cases, the relevant letter strings (words) appeared after the point at which the ambiguity was heard in the sentence.) This task employs use of the fact that automatic priming (facilitation) occurs for lexical decisions made to words that are related to material that has just been previously processed (Neeley, 1978; Swinney, Onifer, Prather, and Hirshkowitz, 1978). The rationale here is that such facilitation will only occur for lexical decisions made to words that are related to meanings of the ambiguity that were actually accessed and activated. That is, if strong semantic contexts constrain lexical access, one should only find priming for lexical decisions to words related to the contextually relevant meaning of the ambiguity; the contextually inappropriate meaning should never be accessed and thus will not provide a base for priming. It was found, however, that lexical decisions for visual words which were related to both the contextually relevant and the contextually inappropriate meanings of the ambiguity were facilitated (in comparison to unrelated control words) when they were presented immediately after the offset of the lexical ambiguity. Sentence 1 provides a sample of one condition of these materials (that containing biasing semantic contexts). Lexical decisions were required for the words ANT, SPY, and SEW (which constitute the contextually relevant, contextually inappropriate, and unrelated control conditions, respectively). These words appeared (in different conditions, to different subjects) at the point indicated by ‘Δ’.

1) The man was not surprised when he found several spiders, roaches, and other bugs in the corner of his room.

The data from this study clearly support the Contextually Independent model of lexical access; all meanings are accessed for a word in spite of the presence of strong ‘constraining’ semantic contexts. (It is worth noting that there were sufficient ‘foil’ trials so that subjects could rarely report any relationship between material in the sentence and the visually presented words. In addition the task did not draw attention to the presence of ambiguities in the materials.)

Using this same task, Swinney, Onifer and Hirshkowitz (Note 2) have shown that these same effects hold even for sentences containing polarized ambiguous words
(those ambiguities with one very likely meaning and one unlikely meaning). thus, even in the most constraining situation where both frequency-of-use and strong semantic context predict that only a single, highly frequent meaning for a word will be appropriate, all meanings appear to be accessed. Lexical access is apparently an exhaustive process, unconstrained by higher order semantic information.

A related question - that of whether access proceeds by means of a direct form-addressable system or by a search procedure - can be examined with these data. The most popular of these models is the Terminating Ordered Search hypothesis, in which it is argued that candidate forms are compared against internal representations in order of frequency in a search which terminates once some entry is encountered (see, e.g., Forster, 1976, for a well developed presentation of this argument). However, much of the evidence for this model comes from isolated word processing paradigms and, while the conclusions are undoubtedly valid for that processing conditions, they are not supported in the sentence processing experiments of Swinney, Onifer and Hirshkowitz (Note 2). In fact, the Swinney et. al. results argue against any type of terminating search. This, in effect, leaves us with two models which will be very difficult to distinguish functionally: Access might occur via an exhaustive, form addressable system (an hypothesis compatible with, for example, the Morton (1969) logogen model) or it might be accomplished by means of an exhaustive search procedure (which could be an ordered serial, a random serial, or even a parallel process). However it is accomplished, the access routine that occurs during sentence comprehension appears to result in the exhaustive retrieval of information stored for a word.

While semantic information does not appear to constrain lexical access, other candidates for this role certainly exist. For example, it is fairly obvious that information about the grammatical class of a word can be used to access the lexicon. Given that this is so, it may well be that any higher order information which is able to predict the grammatical class of subsequent words can be used to constrain the access process for these words. Just such a role has been assigned to the closed class vocabulary (the words in the minor grammatical categories which are often the so-called 'function words'). It has been argued that these words are actively employed in the development of the structural description of a sentence, largely by virtue of facilitating the assignment of appropriate grammatical classifications to open class vocabulary (See, e.g., Bradley, 1979; Garrett, 1974.) For example, a context of 'the' can clearly aid the disambiguation of words with multiple grammatical class entries, such as 'watch'. Prather and Swinney (Note 3) utilized these facts in an experiment designed to determine whether the grammatical information provided by closed class materials could constrain lexical access. They presented sentences such as (2) and (3) which contained unsystematic categorial lexical ambiguities (e.g., watch, cross) which were themselves preceded in the sentence either by 'the' (or 'the + adj') or by 'to' (or 'to + adverb').

(2) The seasoned old woodsman told the young boy the (battered) cross the indian gave him was made of silver.
(3) The seasoned old woodsman told the young boy to (quickly) cross the stream above the beaver pond.

Again, employing the cross modal lexical priming technique, access and activation of the noun and the verb meanings for the ambiguity were examined in each of these biasing context conditions. Under all context conditions the evidence indicated that all meanings of an ambiguity are accessed, even if the grammatical class of some of the accessed meanings are inappropriate.

Finally, in somewhat related work both Swinney (unpublished) and Holmes (1979) have evidence indicating that selectional restrictions provide no apparent constraining effect on lexical access for either unambiguous or ambiguous words (respectively). While the research available does not exhaust the inventory of
possible higher order constraints on the access process, that which is available
strongly suggests that the access of lexical information is unconstrained by higher
order information during sentence comprehension. Lexical access appears to be an
independent subprocess in the sentence comprehension routine.

POST-ACCESS PROCESSING

It is clear, however, that higher order information does affect lexical inter-
pretation. Thus, it seems reasonable to posit that the locus of higher order
constraints on lexical processing will be in a post-access decision process, a process
which will eventuate in the choice of a 'single' meaning of a word for insertion into
the ongoing sentential analysis. A basic question about this process concerns the
time course for the posited interaction of contextual information with the various
accessed word meanings. Does such interaction immediately follow access of each
word in a sentence, or must it wait until the end of a clause or sentoid (see, e.g.,
Foss, Bever, Silver, 1968; Fodor et. al., 1974, Bever, Garrett, and Hurtig, 1973)?

The study by Swinney (in press) also utilized the cross modal lexical priming
technique to examine this question. In particular, the degree to which the several
meanings for an ambiguity were (still) active at a point 3 syllables (about 900 msecs)
following occurrence of the ambiguity was examined. It was found that only the
contextually appropriate meaning of an ambiguity was active at this point.
Apparently the time course of access, activation and post access processing is
completed by the time that 3 additional syllables of information are processed. (Of
course, this result most likely underestimates the speed of the post-access decision
process; it may well have been completed before the point at which it was tested.)
While these data do not allow us to determine the manner in which the contextually
irrelevant meanings for a word are suppressed or discarded, they do support the
conclusion that a post-access decision process occurs immediately following lexical
access for each word, and does not wait for the conclusion of clausal processing.
Further, it seems reasonable to suggest that this post-access decision mechanism
may be a general process; only a subset of all the information stored and accessed
for any word (whether it comprises a single 'sense' of an unambiguous word or a
single 'meaning' of an ambiguity) will be selected by this process for integration into
ongoing sentential analysis.

As a final comment about constraints on access and post-access lexical processing,
the situation as noted that there is recent work by Blank and Foss (1978) that has been
interpreted as a demonstration that lexical access during sentence comprehension
is facilitated by the occurrence of prior semantically related words. In this it was
found that time to monitor for a phoneme target was faster when the word preceding
the target was itself preceded by a semantically related word. However, as the
phoneme monitoring task has been shown by Swinney (1979) to be relatively
insensitive to access processes (due largely to the time-window involved in its
application) it appears that the results could just as easily be attributed to the
workings of the post-access decision process posited here, a process which is aided
by semantic context. (Such an argument also applies to similar work by Morton and
Long, 1976). Considering all the data available, this interpretation appears to be the
more parsimonious one.

SEGMENTATION

While the access of information associated with any particular word may not be
constrained once a 'sufficient' candidate form has initiated the process, it does not
follow that the manner by which such lexical candidates are selected from the
speech stream is also unconstrained. There are several types of constraining
influences that might obtain. One of these has to do with the intrinsic ordering of
the segmentation process. That is, there may be constraints on segmentation provided by the order in which potential lexical candidates are examined in the speech stream. For example, one possible model of the process could involve a minimal segmentation procedure in which the first encountered string of phonemes that constituted a word is accessed, followed by the access of the next string of phonemes that forms a word, etc. (Thus, all else being equal, the word "new" and then the word "display" would be accessed from the input / nudisple/. Clearly, other possible order conditions on the segmentation process can be envisioned. There is evidence which examines possible order constraints on the segmentation of lexical candidates using the cross modal lexical priming technique. This work examines the processing of two types of complex lexical items - idioms and nominal compounds. Swinney and Cutler (in press) have demonstrated that idioms (such as 'kick the bucket') are represented in the lexicon in the same manner as is any other word. The cross modal priming task was used to examine the relative time course of segmentation (and access) of the idiomatic and literal meanings comprising such idiomatic word strings. Lexical decisions were obtained for visually presented words which were related to the literal words comprising the idiom string ('kick' and 'bucket') as well as for words related to the entire idiom (e.g., DIE). The related words were presented in two places. One test point was immediately following the first literal word in the idiom (kick) and the other followed the last literal word in the idiom (bucket). All idioms were heard in sentences. The results indicated that at the first test point the meaning for the literal word (kick), but not that for the entire idiom, had been accessed and activated. However, at the second test point, only the idiomatic meaning gave evidence of being activated; the literal word 'bucket' was apparently not accessed. A similar effect was found by Penny Prather and myself for certain nominal compounds (e.g. 'boycott') when examined in isolation. In all, it appears that the first potential word encountered in the speech stream is momentarily accessed, and that immediately subsequent potential words are not accessed if this subsequent information can combine with the originally accessed material to form a (large) word. This is the case for both nominal compounds and idioms.

To summarize the observations and conclusions that have been made about lexical processing:

1. The segmentation of lexical candidates appears to be a modifiable left-to-right process that operates on what can be called a Minimal Accretion principle. By this, the first viable candidate (based on analysis of the underlying form) that is encountered is accessed. In addition, the form information for this accessed word is also conjoined with immediately subsequent information to determine whether the initial segment may be part of a larger word. The decision process which is posited to occur following access of any word seems likely to be involved in an evaluation of the viability of various accessed candidates in light of the contextual constraints which are incorporated in that decision process.

2. Once a candidate form is determined, all information stored for that form is accessed and made available for evaluation and further processing; no higher order constraints appear to operate at this level.

3. A rapid post-access decision process operates to evaluate the accessed information with respect to higher order contextual information. This process eventuates in selection of a subset of the available information (for further sentential processing) and in the suppression of inappropriate information.

4. These observations hold for the domain of sentence comprehension in which an experienced (adult) listener comprehends a reasonably intact sentential signal.
higher order information is not used by the lexical access process. However, on closer examination there is an externally motivated consideration that makes this claim more parsimonious in the larger scheme of things. This consideration relates to summary point # 4, above. The sentence comprehension routine is a highly practiced ('over-learned') process for the adult listener. Much of what has been discovered about such processes indicates that they develop a degree of automaticity in which 'set' routines operate rapidly and without interruption (see, e.g., Neeley, 1977; Yates, 1978; Posner & Snyder, 1975; Shiffrin & Schneider, 1978). It may be that in the context of such routines, extremely rapid and automatic' lexical recognition is best accomplished by an exhaustive retrieval of information which is then evaluated against all context. Evidence for such exhaustive routines in other practiced cognitive processes certainly exists in the literature' (see, e.g., Sternberg, 1966). Further, it seems likely that when the lexical processing routine is removed from relatively 'normal' (automatic) processing conditions different operations may be brought to bear. The more that a task situation differs from the normal comprehension situation, the more one may find that (because they are neither as automatic nor as rapid) the access routines are constrained by contextual information.

In all, it may be that the variable of 'processing condition' represents the largest constraint on lexical processing. Because the perceived situational demands form the basis for what can be thought of as higher order pragmatic information, specification of the effects of such pragmatic constraints appear likely to form a necessary condition for the goal of specifying a complete model of lexical representation and processing.

Notes

1. Note that while recent papers by Newman and Dell (1979), Mehler, Segui & Carey (1978), and Cutler and Norris (1979) cast some doubt on the phoneme monitoring work due to the presence of confounding variables, it is not certain that the confounds are critical to the interpretation; see Swinney, 1979 and Cairns and Hsu, 1979, for related discussion.

2. Either: "new display" or "nudist play".

Reference Notes


