COMPUTATIONAL DISTINCTIONS OF VOCABULARY TYPE

by

DIANNE CHRISTINE BRADLEY

B. Sc.(Hons.), Australian National University
(1973)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(July, 1978)

Signature of Author..
Department of Psychology, July, 1978

Certified by..
Thesis Supervisor

Accepted by..
Chairman, Departmental Committee
on Graduate Students

Hum.
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
FEB 9 1979
LIBRARIES

Computational Distinctions of Vocabulary Type
by
Dianne C. Bradley

Submitted to the Department of Psychology in July, 1978
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

ABSTRACT

This work pursues the hypothesis that psychologically plausible models of sentence processing must accommodate a gross cut in the vocabulary of a language. A contrast is drawn between the closed class vocabulary (consisting of elements of the minor lexical categories) and the open class vocabulary (made up of elements of the major lexical categories). The major computational consequence of that distinction, it is argued, is in terms of structural analysis, with the distribution and character of forms from the closed class providing the primary evidential base for syntactic inference.

We adopt the view that the lexicon (the mental inventory of words of a language) is best considered as an information-base in support of sentence processing, and explore the treatment given items from the two vocabularies at the level of word recognition. We suggest that when a distinction between the vocabulary types is manifested in recognition performance, it reflects mechanisms which lay the ground for the operation of higher level processes.

The first set of experiments reported examine two major properties of recognition performance with open class forms, and show that these properties do not hold in the recognition of closed class forms. Open class recognition is sensitive to the frequency of occurrence of forms in a reaction time task; closed class recognition is not (Experiments 1 and 2). An account of these results which turns on a special property of the closed class, its restricted size, is shown to be inadequate (Experiment 3). The left-to-right bias of open class recognition is reflected in the sensitivity of classification performance with nonword forms whose initial portions have word-status; Experiment 4 demonstrates that reaction time for the classification of nonwords is elevated by the presence of word initial portions only when those are from the open class, and not when those are from the closed. A countervailing account of these results, relying on the fact that closed class forms do not engage in word formation productively, is shown to be inadequate (Experment 5). We take the failure of these widely acknowledged trademarks of recognition performance as evidence that closed class forms are recognized by a mechanism which is independent of that serving the open class vocabulary.

Some support for the claim that a special treatment of the closed class vocabulary reflects its significance is provided by an examination of recognition performance in a population of aphasic speakers, whose impairment is apparently to be characterized as one of asyntactic sentence

processing. The performance of these patients contrasts with that of a control group of hospitalized normal speakers: the agrammatic aphasics show no distinction in recognition performance with items from the two vocabularies, so that all word forms induce frequency sensitivity in recognition (Experiment 6), and the vocabulary type of word initial portions of nonword forms no longer controls the elevation of reaction time in classification (Experiment 7). The contrast of "normal" and "agrammatic" recognition performance is mirrored in the relative success with which words can be identified by familially right-handed normal speakers, when presentation is lateralized to the visual hemi-fields (Experiment 8).

Thesis Supervisor: Merrill F. Garrett

Title: Associate Professor of Psychology

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I most gratefully acknowledge a debt to my advisor, Merrill Garrett, whose mind-boggling cheerfulness first prompted a suspicion that a thesis could be achieved, and whose support of all phases of the research allowed the circumstance where it would be.

To Mary-Louise Kean, Edward Walker and Molly Potter, I am also indebted, for their helpful discussions of the material.

The technical assistance of Dumont Billings-Dean, Martha Farah and Ninamarie Maragiolio has made the otherwise impossibly dull task of preparing films and running experiments almost bearable.

The studies with aphasic populations, laid out in Chapter 4, came out of a collaborative effort involving Merrill Garrett, Mary-Louise Kean, Herman Kolk and Edgar Zurif. We thank Michael Cicone and Suzanne Hanby, for their assistance in running subjects.

To the P.E.O. International Peace Scholarship Fund I express my sincere thanks for their support, financial and otherwise, during the past three years.

BIOGRAPHICAL NOTE

Dianne Christine Bradley was born on September 19, 1949 in Armidale, New South Wales. The transition from a rather small Australian country-town (in an area known as New England) to a rather large American metropolis (in an area known as New England) was most fortunately made gradually, with interim residences in Canberra (as an undergraduate at the Australian National University), in Sydney (as a mathematics teacher at the Kambala Church of England Foundation School for Girls), and in Melbourne (as a research assistant at Monash University).

CHAPTER 1

SENTENCE PROCESSING AND VOCABULARY TYPE

Perhaps the most striking aspect of normal language behavior is its extraordinary efficiency, given the demonstrable complexity of information processing which must underlie it: we speak, and understand when spoken to, rapidly, effortlessly, even under adverse circumstances. Faced with a normal speech rate of some 150 words per minute, both speaker and hearer have a problem: for the speaker, it is one of choosing and integrating a suitable linguistic vehicle for his intent, given a vast range of structural and stylistic options; while the hearer, coping with an evanescent stimulus, must uniquely recover an intended message. A research program (of some forty years standing) in the recognition of speech by machine testifies as much to the extent of the difficulties involved, as it does to progress in resolving them.

An encoding (producing) or decoding (understanding) of an utterance draws on information of many types: information about the grammatical and logical relations among sentence constituents (clauses, phrases, words), their morphological and phonological properties, as well as information about their reference or meaning. Much of this information needs be specified for individual words. That is, just as the vocabulary items of the language require a specification in terms of their sound and meaning, so must they be specified in terms of their control of structural properties of sentences. The sentence processing system (whether in comprehension or production) must note, for example, that the forms give and lose are, first of all, in the grammatical category VERB, and in more detail, that the first can take two objects, and the second only one; similarly, that the form fact is not only of the grammatical category NOUN, but that it can take complements. A lexicon, an inventory of the forms of a language, can be viewed as an information base in the service of sentence processing devices, arraying against each listing all relevant features of word form and distribution, as well as word meaning.

We might begin to consider the exploitation of lexically encoded information with a minimal characterization of sentence comprehension in terms of the following two processes: lexical access, whereby memory representations of forms are contacted[1]; and a subsequent parsing which yields a structural analysis upon which interpretation may be based, in whole or in part. The latter process will be a joint function of the structural information arrayed against lexical representations and information as to possible phrasal configurations of the language. Thus, in this rough plan, words are identified, and information is extracted from memory in virtue of which assignments of items to phrases, and phrases to larger constituents, can be made.

We have noted the several varieties of information listed against lexical entries. Precisely how those facts enter into the operation of parsing routines is an open question as yet. One thing is clear,

---

[1] Evidently there will be modality-specific problems of identification of lexical items; for example, the "segmentation" problem for acoustic representations of sentences differs from that of visual (orthographic) representations. However important such issues, we will not confront them here.

however, and that is that distinctions of grammatical category will play a fundamental role in the direction of initial parsing decisions. So, for example, in current formulations of ATN parsers (Kaplan, 1973; Wanner and Maratsos, 1978) grammatical category distinctions are the primary determinants of choice among the network paths which encapsulate the structural options of the language. It could hardly be otherwise: syntactic analyses are, in the end, defined in these terms.

This account of comprehension processes has a straightforward appeal. There are, however, two observations which must be borne in mind in our evaluation of its plausibility for models of human sentence processing.

First, most of the vocabulary (of English) is ambiguous as to grammatical category. This can be handled by an analysis schema of the sort we have sketched, but at cost. That is, the system can juggle competing analyses of an input string in accordance with lexically specified category assignments, filtering out inappropriate analyses in terms of encoded constraints on possible structural configurations. But to the extent that ambiguity of this kind is the usual case in the vocabulary, there is a combinatorial explosion of candidate analyses, and the efficiency with which inappropriate parsings are rejected becomes a matter of paramount concern. In the absence of a theory of the parser, of course, questions of efficiency are difficult to assess, but we can bring to bear a strong observation: there is evidently little pressure, in the evolution of a language like English, against the tendency for forms in one grammatical category to acquire usages in others. The very freedom with which there are zero morpheme category alternations (as in certain deverbal nominals, and denominal verbs) suggests either that the explosion of candidate structural analyses that we have noted is a problem readily solved, or, more radically, that it is not a problem at all, and our view of the comprehension system is mistaken.

The second observation concerns the "limiting case" of category ambiguity. It is clear that in situations where no grammatical category information is available for an item in a sentence, structural analyses can still be maintained. For example, in sentences like (1),

(1) We found an antique varbel abandoned in our garden

the fact that varbel is not a word of English does not block an analysis of the sentence.[2] Notice that this is not an unusual circumstance. People routinely converse with each other despite failures of correspondence in their working vocabularies, and distortions and failures of word identification are induced, occasionally, by noisy conversational settings. More striking, of course, is the success of "apprentice speakers" -- children and second-language learning adults must often analyze sentence form with local gaps of lexically specified information.

We take the lesson of (1) and related observations to mean that encoded constraints on structural configurations must go beyond a simple filtering among the recorded, lexically specified options. In this

---

[2] The more exotic constructed cases of nonsense sentences, dear to taxonomic grammarians, psychologists, and Lewis Carroll, are rather dramatic examples to the same point. Here, where all major category items are replaced by nonwords, an appreciation of sentence form is readily grasped by a native speaker.

limiting case, sequential constraints seem to specify assignment of grammatical category for unknown lexical items.

The crucial question is whether this performance is applicable in other than the limiting case. In particular, we might ask whether this points to the existence of an efficient routine which will avoid the costly proliferation of candidate phrasal analyses arising from the unchecked tendencies of the language to acquire categorial ambiguities. What form might such a process take? The obvious course is to search for "islands of certainty" in the string under analysis -- that is, for those particular cases for which categorial ambiguity is minimal, and which are locally determinate for specific phrasal configurations. Notice that one could attempt to isolate from the entire vocabulary just those items which idiosyncratically satisfy these specifications, for particular speakers on particular occasions in a particular language. If, however, our speculations about the generality of the problem and the significance of the role of categorial information in comprehension and acquisition are relevant, we would require a principled solution: namely, the isolation of a grammatically systematic subset of the vocabulary which sharply delimits incidence and type of categorial ambiguity.

There is a tradition in the psychological literature, undoubtedly rooted in the work of the taxonomic grammarians (for example, Fries, 1952, 1962), which proposes a gross cut in the vocabulary of just the kind that our considerations above require. That cut contrasts the "content" words (roughly, items of the lexical categories noun, verb and adjective) and the "function" words (minor category forms together with the affixal apparatus). Evidence is available from a variety of sources to suggest that the distinction of vocabulary type has marked consequences in several domains of language bevhavior: in comprehension, in production, in acquisition, and in the aphasias. We make a working assumption that the distinction is properly drawn in the same way in each of these cases and, in what follows, offer a formulation and briefly review the relevant evidence.

1.1 TWO VOCABULARIES

As a working hypothesis, we formulate a distinction of vocabulary type between elements of the major lexical categories (noun, verb and adjective) on the one hand, and elements of the minor lexical categories (determiner, e.g., _the_; auxiliary, e.g., _have_, _were_; preposition, e.g., _to_, _by_; quantifier, e.g., _all_, _some_; and so on) on the other. We shall refer to these vocabulary subsets as the _open_ and _closed classes_, respectively, consciously adopting a neutral terminology. The evidence at hand suggests that some such gross cut in the vocabulary will be part of psychologically plausible models of sentence processing, but the detail which might dictate exactly where the cut is to be placed is lacking, both in terms of our theoretical apparatus, and in terms of the evidential base. The more traditional distinction between "content" and "function" words, though having the flavor of the contrast we intend, suffers the problem that authors have differed in their account of the extent of the class of function words (see, for example, Fries (1952) and Miller, Newman and Friedman (1958)).

The open and closed classes differ in many respects, and we shall set out here some of those which are of the type which are often taken to be relevant in processing terms. Most obviously, the vocabulary classes differ in size. The open class has an enormous number of elements, so that Oldfield (1963) estimates that "the average, reasonably young, university-educated, kind of person" has control of some 75,000 such forms; and the class is 'open' in the sense that elements may be freely added as the uses of the language demand. In contrast, the closed class is a relatively small fixed set, with around 200 forms in English; and processes of change take place more slowly. Open class forms vary greatly in length, and frequency of usage, and ambiguity as to grammatical category is the norm rather than the exception; that is, forms like _fair_, _cast_, _store_ and _iron_, occurring variously as noun or verb or adjective, are in no sense unusual. Closed class items tend to be short, of very high frequency of occurrence; the relation between frequency of occurrence and length may be more marked for the closed class than the open (Miller, _et al_, 1958), but this seems unlikely, given the density of Latinate forms at the low end of the frequency range in the open class. Ambiguity as to grammatical category is much less frequent in the closed class than in the open, and certainly we do not see the systematic category alternations so common among nouns and verbs.

We shall be concerned, in our experimental evaluation, with contrasts among the free forms of the language. However, if within the class of bound morphemes it will be possible to draw a similar distinction, perhaps between the inflectional and derivational affixes, the appropriate systematic grammatical description might be better cast in phonological terms (after Kean, 1977) rather than in terms of lexical category.

1.2 EVIDENCE FROM COMPREHENSION

Sentence meaning is much more than a matter of concatenation of the meanings of individual words, and in this light it is significant that the vocabularies are thought to diverge in what might be termed "interpretative burden". Broadly, the closed class seems to primarily support structure, being "the heavily overworked glue that holds our sentences together" (Miller, et al., 1958); the open class bears reference. This is often illustrated in the case of nonsense sentences, where all open class forms are replaced by nonwords, as in Fries' (now classic) case: The vapy koobs dasaked the citar molently. It is clear that with no information as to grammatical category (and, for that matter, meaning) in the open class "slots", syntactic analyses of considerable detail can be sustained. At minimum, the nonsense forms can be assigned to grammatical categories. The conclusion is offered, that the sequence and character of closed class forms in an utterance provides a powerful base for syntactic inference, defining the structural options which in large part determine sentence interpretation. But we might exercise some caution. In concentrating on the role of the closed class in guiding structural analyses, we set aside the information of similar type which some open class forms carry: namely, the distributional facts, as, for example, the fact that give is a verb which can take two objects.[3]

A psychologically adequate model of a parsing device for English will specify the contributions of open and closed class sequences to the construction of syntactic analyses. With no such model at hand, we rely here on what seems to be a strong (and rather common) intuition that the small, over-used closed class vocabulary has a special role in sentence interpretation, trusting that the compelling effects in simple nonsense sentences carry over to more complex cases.

Experimental investigations of learnability support our observations with "syntactic" nonsense sentences. So, for example, Epstein (1961) was among the first to report striking differences in ease of learning (trials-to-criterion in immediate recall) for sequences of nonsense syllables with a minimal suggestion as to structure, and for scrambled versions of such sequences. That such effects depend crucially on the availability of a structure is strongly suggested by a later finding, that (with auditory presentation) recall is enhanced only if the sequence both has the form of a possible sentence of English and is spoken under normal sentence intonation (O'Connell, Turner and Onuska, 1968): it is exactly the possibility of syntactic

---

[3] It is interesting to note, in this regard, that Thorne, Bratley and Dewar (1968) in constructing a "special dictionary" for their parsing machine, included verbs taking other than a single object.

analysis which distinguishes sentences from lists, Similarly, closed class words and open class words interact very differently in rote or paired-associate learning of short phrases (Glanzer, 1962); for example nonsense-English-nonsense triplets are recalled better with closed class items (TAH of ZUM) than with open class items (YIG food SEB). It is interesting to note that (by inspection of Glanzer's figures) the advantage to closed class forms is restricted to cases where a phrasal structure could be supported--(YIG who SEB) has no advantage over (MEF think JAT). Closed class forms per se to not enhance recall, but closed class items, as they mediate structural analyses, do.

1.3 EVIDENCE FROM PRODUCTION

It has been argued (Garrett, 1975, 1976) that certain types of naturally occurring speech errors show very distinct patterns for the two vocabularies. For example, those errors involving the complete interchange of any two segments of the intended output (examples (2) and (3) below) contrast with those involving a shift in the location of a single bound or free morpheme (examples (4) and (5)):

(2) That's just a truck backing out → That's just a back trucking out

(3) Older men tend to choose younger wives → Older men choose to tend younger wives

(4) Who else did you say was coming? → Who did you say else was coming?

(5) .., and what I would want to check.. → .., and what would I want to check

Open class forms engage in exchanges, but rarely seem to shift their location; conversely, closed class items are rarely involved in errors of exchange, and by and large make up the corpus of shift errors. Under an assumption that elements which interact in speech errors are simultaneously represented in the mind of the speaker at some level of the planning of the an utterance (see Garrett (1978) for some discussion), the lack of interaction between open and closed class forms, and their separation in terms of speech error types, suggests that in the production of utterances, the vocabulary types are independently entered. Perhaps the most striking evidence of a separation of vocabulary types comes from sound exchange errors involving final portions of words; syntactically active bound morphemes, by hypothesis of the closed class, may simply be stranded:

(6) a monkey's uncle $\longrightarrow$ a monkle's unkey

These are a variety of related facts have been taken to suggest (Garrett, 1975, 1976, 1978) that the role of closed class forms in the processes underlying production is to be interpreted as reflecting a crucial involvement in the formulation of the structural frames of utterances.

1.4 EVIDENCE FROM ACQUISITION

Part of the problem for the child, as language-learner, is that of mastering an enormous number of words, very rapidly, hearing only their casual use in normal contexts (Carey, 1978). He must learn, not only acoustic shape and meaning, but also distributional properties. Two sources of evidence are available to the child: the linguistic context in which a new word occurs, and the situation in which it is used. The child's productions suggest that, for him, the linguistic context might be a poor source of evidence: certainly the closed class words generally appear later in development than the open class forms, as attested by the "telegraphic" speech which has been so widely discussed. But the child's utterances may not be an altogether reliable index of his tacit knowledge. A seventeen-month old girl is able to interpret the presence or absence of an article to determine the type (common, proper) of a given noun (Katz, Baker and McNamara, 1974), though the child's utterances at seventeen months clearly do not dictate that she has such distinctions under control. Older children (at four years of age) are able to employ the distinctions among a sib, some sib, and to sib to hypothesize that the referent is to be taken as count noun, mass noun, and verb, appropriately (Brown, 1957). In some ways the child, as word-learner, is in much the same position as a mature speaker in the hands of an experimenter bent on constructing syntactically constrained nonsense sentences; and like the mature speaker, he is apparently able to exploit quite minimal information to make what he can of the sentences he hears. In this, the closed class vocabulary plays some role.

1.5 LEXICAL ACCESS AND SYNTACTIC PROCESSES

We take the evidence at hand to suggest that among the vocabulary elements of a language, there is a subset, the closed class, which has a privileged status for syntactic inference. Our view of the lexical component as an information-base for sentence processing leads us to anticipate that the lexicon and its attendant access procedures will be highly structured so as to facilitate retrieval of the computationally relevant facts. Accordingly, we are prompted to formulate questions of a particular kind about the lexical component: namely, given a distinction among words which is exploited in the construction of syntactic descriptions of sentences -- the contrast of open and closed classes -- to what extent does it serve as a basis for lexical organization, and hence for retrieval processes for lexical forms? From this standpoint studies of word recognition need not simply examine an isolated memory capacity, but might rather offer a characterization of a complex device, geared to the support of encoding and decoding of sentences,

We might (quite neutrally) consider a word recognition device as one which, given a preliminary analysis of an input, engages mental representations from the lexical inventory which are congruent with that analysis. Such devices will, we assume, exhibit stable properties which reflect the kinds of initial analyses that are made, and the ways in which candidate lexical forms are selected for examination.

An enormous literature in word recognition is almost entirely concerned with characterizing recognition devices based on performance with open class forms. The marked asymmetry of that literature allows the use, here, of the following genereal strategy for research: first, isolate a feature of open class word recognition which (a) has some acknowledged experimental support and (b) for which there is some plausible basis for expecting a contrast with the closed class , given the general properties we have pointed out above; second, experimentally evaluate the contrast between performances over sets of open and closed class items in the circumstances where generally acknowledged properties of open class recognition are evident. The research reported here examines divergences in recognition performances, for three properties characteristic of open class recognition of visually presented words:

(a) frequency-sensitivity in reaction time to classify forms

(b) sensitivity to word-status of intial portions of nonword forms

(c) sensitivity to location within the visual form.

These properties are not idly chosen, but are, perhaps, the most widely acknowledged tradmarks of word recognition. We shall evaluate a claim that the open and closed class vocabularies are served by independent recognition systems in terms of these divergences. We shall take into account also a contrast between the various recognition performances of normal speakers, and those of a population of aphasic speakers whose language impairment apparently involves a differential control of the vocabularies, in reflection by hypothesis of their asyntactic treatment of sentences.

Some cautionary notes are necessary. First, not only is there no model at hand of a parsing device explicating a distinction between the two vocabularies, but it is unclear just what form such a model might take, so as to make a principled distinction between the closed class vocabulary, as we have formulated it here, and certain open class forms (say, those verbs controlling complex structures). Though a distinction has traditionally been drawn, most notably in a literature concerned with the statistical properties of sentences, its theoretical basis is not obvious; moreover, rather little attention has been paid to direct experimental evaluation of its consequences. Under the assumption that there is some subset of the vocabulary which is to be set apart because it provides the primary evidential base for syntactic inference, the problem cuts two ways, On the one hand, knowing what an adequate model of parsing is like, we can identify a special vocabulary; and on the other, having identified a vocabulary, we might constrain the formulation of parsing models. It is important to note here that an examination of performance in impaired speakers holds a special significance; the correlation of asyntactic sentence processing with appropriately interpretable patterns of word recognition lends weight to the most basic claim, that the notable exploitation of closed class forms is in terms of structural analysis.

Second, the assignment of major category items to the open class, and minor category items to the closed, reflects no more than a well-founded suspicion that this approximates the computationally relevant contrast. To the extent that forms have been misassigned, our experimental evaluations are inevitably confounded: but these are the hazards of the enterprise.

CHAPTER 2

FREQUENCY EFFECTS IN LEXICAL DECISION

It is well established (in a variety of paradigms) that familiar words are, other things being equal, more readily perceived than rarer forms. They are identified at shorter durations in tachistoscopic presentation (e.g., Howes and Solomon, 1951) and at lower signal-to-noise ratios in auditory presentation (e.g., Savin, 1963). That advantage is maintained even where the input is not impoverished, so that more frequent words are named (e.g., Forster and Chambers, 1973), classified (e.g., Frederiksen and Kroll, 1976) and matched (e.g., Chambers and Forster, 1975) with shorter latencies than less frequent words. The effect of frequency of occurrence in word recognition has been seen as a special case of the general influence of probability on perception: overall, a more probable event is more easily perceived. Indeed, much of the early work investigating frequency as a variable predicting efficiency in word recognition was carried out with an eye to the general question of probabilistic effects in perception (Broadbent, 1967).

Broadbent (1967) argues persuasively that frequency effects cannot be attributed to any so-called "response bias"--roughly, an artifactual influence of the subject's expectations in an experimental setting--except under a reading which makes that bias an inherent property of central perceptual mechanisms. Current models of the processes underlying word recognition give an account of the effects of frequency in just this spirit. On the one hand, "strength models" (e.g., Broadbent, 1967; Morton, 1969, 1970) have word frequency determine the threshold of logogens (evidence-collecting devices), whose activation to threshold is necessary for the recognition of a word. On the other hand, "search models" (e.g., Rubenstein, Lewis and Rubenstein, 1971; Forster, 1976) have word frequency determine the serial order of examination of a set of candidate representations which are selected as being in some degree of accord with a preliminary analysis of an input. We shall not be concerned here with debating the merits of these two approaches (though see Forster 1976 for a review of relevant arguments) and note that, in general, an account couched in terms of one model has its translation to the other.[3]

In summary, we can say that effects of the relative frequency of occurrence of forms on word recognition are robust and ubiquitous. Frequency-sensitivity is evidently a fundamental property of the word-recognition system; correspondingly, an account of the role of frequency in the processes underlying word recognition has assumed some importance in the construction of plausible psychological models. We note, though, that experimental evaluations of such effects have been made over item sets made up, by and large, of open class forms: closed class forms have, at best, been avoided in the setting up of experimental materials; occasionally, they are included in item sets, without any distinction being drawn between types.

If the closed class vocabulary has a particular significance in the

---

3 It could be said, in fact, that the battle-lines of the "search vs strength" debate are not even well-drawn. Strength models focus on the processes which allow lexical candidates to be activated, while search models are directed at an account of the processes underlying examination of sets of lexical candidates, for their congruence with inputs.

construction of structural analyses of sentences, it is possible (though not necessary) that they will be recognized by a retrieval device which is independent of that which operates over open class forms. And, given that the closed class vocabulary is a small, restricted set, such a device might <u>not</u> be frequency-sensitive, but will allow all forms in its domain an equal status.

In Experiments 1 and 2, we contrast effects on frequency of occurrence on the distributions of item reaction times in a lexical decision task, treating open and closed class items sets independently. We shall depart from the common current practice of constructing item sets which draw items in discrete groups, from the very high and the very low ends of the frequency range. Instead, items will be selected which span the relevant frequency ranges of the two vocabularies, following the practice of earlier work (e.g. Howes, 1954). The small number of closed class forms does not permit any other course.

Preliminary experiments (e.g. Bradley, 1977) indicate that the frequency effect is well described in the slope of the least-squares linear regression of reaction time on an estimate of relative frequency of occurrence which is expressed as a logarithm; or, equivalently, in the correlation of these variables. Roughly, reaction time (in lexical decision) decreases by some 40 milliseconds with every log (base 10) step in frequency per million, all other things being equal. Whaley (1978) notes that while frequency of occurrence is not the only variable affecting reaction time in lexical decision tasks, it clearly is the most powerful one (accounting for some 50% of variance). Note that an effective ceiling on the power of frequency estimates in predicting reaction time is imposed by the short-comings of the procedures by which relative frequency is estimated. A count of occurrences in text over approximately one million tokens (as, for example, that of Kucera and Francis (1967) which we make use of here) will itself suffer from sampling error; it is unclear, moreover, that counts in text are more appropriate than those in speech (or perhaps some combination of these); and, of course, estimates for the language at large only approximate the strictly relevant estimator, encounter rates by individual speakers. In this light, it is quite remarkable that estimated relative frequency of occurrence shows even the predictive power that experimental evaluations suggest.

EXPERIMENT 1

METHOD

SUBJECTS. Thirty adult native speakers of American English took part in the half-hour experiment. All were students at the Massachusetts Institute of Technology, and were paid for their participation.

MATERIALS. The item set was made up of three kinds of forms: closed class words (59), open class items (61), and legal nonwords (80). Frequency estimates for words are taken from the count of Kucera and Francis (1967) over approximately one million tokens in text, with occurrence rates being summed over syntactic inflection (tense and number); estimates are expressed here in units logarithm base 10. All item classes exhibit the same variation in length of forms (with mean number of letters being 4.7, 4.8 and 4.5 for closed class words, open class words, and nonwords, respectively).

The closed class items were selected to cover, as evenly as possible, the frequency range 1.5 - 4.8. The open class items fell in the frequency range 0.48 to 3.4.[4] In the frequency region in which the closed and open class sets overlapped (1.5 to 3.4) there were 46 closed class items and 43 open class items.

The items were arranged in two blocks of 100 items each in a pseudo-random order, under a constraint prohibiting runs of more than four items of the same response type (word/nonword). Additionally, sets of four "buffer" items headed each experimental block; these were intended to minimize warm-up effects on reaction times for the items of the experiment of interest. All subjects saw the entire item sequence, with order of presentation of the experimental blocks being reversed for half the subjects.

INSTRUCTIONS. Subjects were instructed to classify the individually presented items as words or nonwords, having been informed that all items would look as if they might be words of English. A button marked YES was pressed for word decisions ("yes, it's a word of English") and one marked NO for nonword decisions. The need for accuracy in classification

---

4 Words at the low end of the frequency range (for example, those estimated to occur once per million) are often referred to in the literature as "rare forms", suggesting that such words have only a marginal status in the working vocabulary of the speaker. But this is misleading, as the following (conservative) calculation shows: assuming (1) an average speech rate of 150 words/minute, and (2) that one speaks or is spoken to for some 8 hours/day, then one million words are encountered (one way or another) in every two weeks. It is unlikely that a speaker will not have control of forms which are encountered, on average, every two weeks.

It is more likely that "low frequency" closed class forms will present a problem. A count over text will over-estimate the relative frequencies of occurrence of forms which are restricted to a formal writing style; for example, thereby and despite are estimated in text to be of the same frequencies as healthy and monument; but do not occur in the spoken count of Howes (1966).

was emphasized, and subjects were instructed to permit themselves no more than 5% errors; assuming accuracy, they were asked to respond as rapidly as possible. A practice block of 30 items, representative of experimental items, was used to familiarize subjects with the task. Any subject whose average response time in the practice session was slow (roughly, above 800 msec) was encouraged, in that practice session, to respond faster.

PRESENTATION. The experiment was presented on film, using a variable speed projector running at two frames per second. The film was prepared using single frame shooting with a Bolex H16 reflex camera, the number of frames being determined to give the following projection times: fixation field (a row of asterisks in the position that an item would occupy), 0.5 seconds; blank field, 0.5 seconds; presentation field, 1 second, blank field, 2 seconds. In effect, the film simulates the cycling of a four-field tachistoscope.

Items were presented in lower case typeface, in negative (white characters on grey background). A light trigger in the first frame of the presentation field started a millisecond timer, which stopped with the subject's pressing of a response key. All reaction times, therefore, are from stimulus onset.

DATA TREATMENT. Only reaction times for correct "word" decisions were analyzed. Bad data were identified on a double criterion: a datum (reaction time for a particular subject for a particular item) is eliminated from the analysis if it exceeds <u>both</u> a cutoff for the subject (mean reaction time for item subset plus two standard deviations) <u>and</u> a cutoff for the item (mean reaction time plus two standard deviations).

A data comb of this kind identifies as aberrant some 2% of the data, and markedly reduces noise in the analysis (Harwood, 1978). Missing data (whether subject or experimenter errors, or bad data) are replaced with the subject's mean reaction time (for the relevant item subset) before item means are calculated.

An effect of frequency of occurrence is exhibited in the correlation of item reaction times with item frequencies, or, equivalently, in the slope of the (least squares) linear regression of reaction time on frequency. Items with error rates greater than 20% are omitted from such analyses.

RESULTS AND DISCUSSION

Table 1 summarizes the effects of frequency of occurrence (per million, in units logarithm base 10) on the distribution of item reaction times (in milliseconds) for open and closed class sets.

TABLE 1

Effects of relative frequency on item reaction time for open and closed class words

	CLOSED CLASS	OPEN CLASS
correlation coefficient	-0.22	-0.58
slope of linear regression	-7	-30
zero intercept	499	560
standard error of estimate	20	33
frequency range	1.5-4.8	0.6-3.4

The effect of frequency of occurrence on reaction time for open class items ($r = -0.58$, $t(57) = -5.36$, $p < .01$) is greater than that for closed class forms ($r = -0.22$, $t(55) = -1.71$, $p = 0.05$), $z = 2.27$. Both closed and open class item sets confound item length with item frequency, the problem being more acute in the closed class set ($r = -0.59$) than in the open class set ($r = -0.25$). This is, in fact, characteristic of the reaction time and frequency of occurrence (taking out variation due to length) suggests more clearly a contrast between vocabulary types. For the open class set, the effect of frequency on item reaction time is maintained, $r' = -0.54$, $t(57) = -4.86$, $p < .01$, while for the closed class set, it is not, $r' = -0.03$. These correlations differ, $z = 3.02$.

Thus, for open class items in the frequency range 0.6 to 3.4, we establish the frequency effect which is standardly taken as the foremost characteristic of normal word recognition performance (Whaley, 1978). And, in contrast, such effects are slighter for the closed class set, or perhaps absent altogether, in the frequency range 1.5 to 4.8. It should be stressed that this is an unusual outcome: frequency effects are robust and ubiquitous in the literature on word recognition.

We might note that although the ranges sampled in the open and closed class sets are equivalent, the end-points are not. This is unavoidable: by nature, closed class forms have high frequency of occurrence, relative to the open class. Restricting attention to that frequency range shared by open and closed class items,[5] we encounter two difficulties:

---

[5] This restriction is prompted by the suspicion that the linear regression of reaction time on log frequency is perhaps not the function of best fit. It has been suggested (A. Wingfield, personal communication) that the reaction time curve for open class forms "flattens out" at high frequencies. This observation would suggest that a lack of frequency effects for the closed class is attributable, simply, to its occupying that portion of the range for which there would never be

first, the number of relevant cases is reduced; and second, we are emphasizing the region in which the closed class give most uneven coverage of the frequency range.

In the frequency range 1.7 to 3.4 in which the item sets give relatively even coverage for both open and closed class sets, the contrast of frequency effects on reaction time fails: for open class items, the partial correlation of reaction time and frequency (taking out variation due to length) does not significantly exceed that for the closed class set, though the difference is in the appropriate direction: for the open class (with 38 cases) $r = -0.54$; for the closed class (with 44 cases), $r = -0.29$. These correlations do not differ significantly.

An absence of frequency effects in the distribution of reaction times for closed class items would stand at odds with a literature which reports frequency effects (for open class items) in almost every circumstance: in lexical decision, in naming tasks, in tachistoscopic recognition thresholds, for example. It is desirable, then, that such a claim about the closed class voacbulary be more firmly established than the somewhat mixed outcomes of Experiment 1 permit. In Experiment 2, item sets for the two vocabulary classes are more carefully selected with respect to length, and even coverage of their frequency ranges.

---

frequency effects. However, it does not seem to be the case that there is significant flattening of the curve for open class forms. A power curve fitted for the open class set along the entire range (0.60 to 3.44) has a maximum departure from the linear regression of 2 milliseconds.

EXPERIMENT 2

METHOD

SUBJECTS. Twenty adult native speakers of American English took part in the experiment, and were paid for their participation. All were students at the Massachusetts Institute of Technology, and were under 30 years of age.

MATERIALS. The item set was made up of 100 words and 80 legal non-words, with all forms being monosyllabic and ranging in length from 2 to 5 letters. There were 40 closed class forms and 60 open class forms. The three item classes exhibit the same variations in lengths (with mean number of letters being 3.7, 3.8 and 3.7, for closed class, open class, and nonword sets, respectively). Frequency estimates for words are taken from the count of Kucera and Francis (1967) over approximately one million tokens in text, with occurrence rates being summed over syntactic inflections (tense, number). Estimates of occurrence per million are expressed here in units, logarithm base 10. Closed class items were selected[6] to cover evenly the frequency range 2.2 to 4.5 (mean frequency = 3.11, s = 0.50); open class items cover the range 1.0 to 3.5 (mean frequency = 2.18, s = 0.70). These ranges give the maximal overlap of the two vocabulary classes, in which forms are monosyllabic and an even coverage of frequencies is possible: there were 32 items from each set in the frequency range 2.1 to 3.5. The item selection avoided the marked

6 Though item selection for Experiment 2 was independent of that for Experiment 1, the relative unavailability of closed class forms and high frequency open class forms forces overlap between the sets: 45% of the closed class forms and 15% of the open class forms in Experiment 2 were repeated from Experiment 1.

confounding of length and frequency which was a problem in Experiment 1; correlation of item frequency of occurrence and item length was, however, greater for the closed class set ($r = -0.25$) than for the open ($r = -0.06$).

The items were arranged in two blocks of 90 items each, in a pseudo-random order under a constraint prohibiting runs of more than three of the same response type (word/nonword). Additionally, sets of four "buffer" items headed each experimental block. Order of presentation of the two experimental blocks was reversed for half the subjects.

INSTRUCTIONS, PRESENTATION and DATA TREATMENT: As in Experiment 1.

RESULTS AND DISCUSSION

Table 2 summarizes the effects of frequency of occurrence (per million, in units logarithm base 10) on the distribution of item reaction times (in milliseconds) for open and closed class word sets.

TABLE 2

Effects of relative frequency on item reaction time for open and closed class words

	CLOSED CLASS	OPEN CLASS
correlation coefficient	-0.03	-0.75
slope of linear regression	-1	-35
zero intercept	484	566
standard error of estimate	22	22
frequency range	2.2-4.5	1.0-3.5

As in Experiment 1, there is an effect of frequency on item reaction time for open class words ($r = -0.75$, $t(58) = -8.60$, $p. < .01$) which exceeds that for closed class words ($r = -0.03$, $5(38) = -0.21$, $p < .05$), $z = 4.43$. The contrast in the effects of frequency over the ranges which are typical of the vocabulary classes sharply distinguishes recognition performance in the two cases. Figure 1 illustrates the different effects on frequency in recognition of closed and open class words, for both Experiment 1 and Experiment 2.

INSERT FIGURE 1

The different sensitivities of open and closed class sets to frequency of occurrence is, moreover, replicable over subject populations. Splitting the subjects who took part in Experiment 2 into two subgroups, and treating these as independent samples, we find that the effects are maintained: correlation between reaction time and frequency of occurrence for the open class are -0.58 and -0.77, and for the closed class, +0.01 and -0.07, for odd-numbered and even-numbered subjects, respectively.

Similarly, the contrast is robust when we restrict attention to those cases (32 in each set) which share a common frequency range, 2.1 to 3.5. Table 3 summarizes correlation coefficients, and the linear regression of reaction time on log frequency.

TABLE 3

Effects of relative frequency on item reaction time
for open and closed class words in a shared frequency range

	CLOSED CLASS	OPEN CLASS
correlation coefficient	+0.14	-0.55
slope of linear regression	+10	-34
zero intercept	452	562
standard error of estimate	23	20
frequency range	2.1 - 3.5	

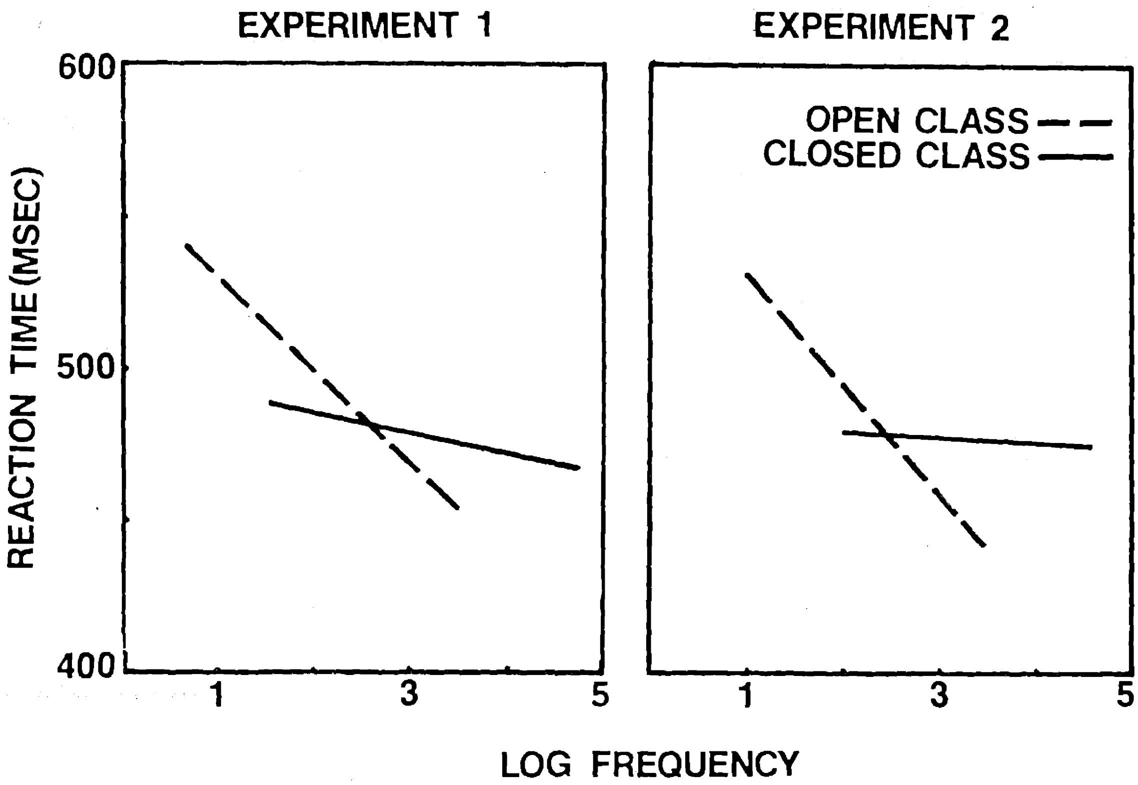


FIGURE 1: CONTRASTS IN FREQUENCY-SENSITIVITY OF OPEN AND CLOSED CLASS RECOGNITION

Linear regressions of item reaction time (in milliseconds) on item frequency of occurrence (per million, expressed in units logarithm base 10). Data are for populations of normal speakers.

The frequency effect for open class words ($r = -0.55$) significantly exceeds that for closed class words ($r = +0.14$), $z = 2.92$.

It appears, overall, that recognition of open and closed class vocabulary items is to be contrasted as follows: recognition performance over sets of open class items is frequency-sensitive, while over sets of closed class items it is not.

Given that frequency-sensitivity is a robust property of a word recognition device (and perhaps its most striking property (Whaley, 1978)), its inapplicability in a description of performance with closed class forms is not to be taken lightly. Unless we invoke some factor which will cleanly cancel effects of frequency, we are prompted to conclude that the retrieval mechanism which acts over the closed class vocabulary is <u>not</u> the same one which is employed in the recognition of open class forms. That is, we take it that the experimental effects reflect a privilege accorded the closed class vocabulary, in line with its hypothesized special contribution to sentence processing.

We might, however, appeal to an alternative account of the lack frequency effects, noting a feature of the closed class which is only indirectly related to its role in sentence interpretation; namely, the restricted size of the set. Even the loosest criteria for closed class membership puts the number of elements (for English) at about 400 (Miller, Newman, and Friedman, 1958); more plausibly, the set is restricted to some 200 forms. The relatively small number of closed class forms raises the possibility that some kind of task-specific inter-item facilitation distorts reaction times, by overriding (or obscuring) effects of frequency of occurrence.

An effect of possibly the right kind, a so-called "semantic facilitation" is well established. A response in lexical decision is made more quickly when an item (for example, nurse) is "primed" by a preceding associate (for example, doctor) (Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt and Ruddy, 1972). It is not necessary that the priming item immediately precede the test item (Schvaneveldt & Meyer, 1973); nor are facilitation effects restricted to cases where an associative connection holds between prime and test items (Bednall, 1976; Fischler, 1977). An item (for example, mirth) can be primed by a word chosen from the same semantic field (for example, fun) even though they have zero associative connectivity in terms of word association pairs. Further, facilitation effects may obtain when prime/test pairs are related, not because of meaning, but in some other way. Taft (1978) found an effect of priming using item pairs sharing a root morpheme, where pairs had been explicitly constructed to avoid any clear meaning relation (for example, corrupt/rupture, depend/pendulum). A rhyme relation between prime and test items may also decrease reaction time for the test item: Meyer, Schvaneveldt and Ruddy (1974) report a trend, at least, for pairs like bribe/tribe.

We should note that although there is no doubt about the existence (and robustness) of facilitation effects, the locus of effects, in terms of the processes underlying recognition, is unclear (Forster, 1978). On the one hand, the magnitude of semantic facilitation effects is dependent on the quality of the visual stimulus, being greater with poorly defined stimuli (Meyer, Schvaneveldt & Ruddy, 1975). This suggests the conclusion that priming and stimulus quality influence a common stage; possibly (though not necessarily) an encoding stage. On the other hand, the magnitude of facilitation effects depends also on the overall proportion of semantically related pairs in the stimulus set, being greater with more related pairs (Tweedy, Lapinski & Schvaneveldt, 1977). Tweedy et al. comment that "contextual effects are at least partially produced by the observer's increased ability to take advantage of the predictability of the stimulus sequence when it contains frequent instances of semantically related word pairs" (p. 80). That is, the facilitation effects are to be placed, at least in part, in an evaluation of candidate mental representations that may be engaged independently of the specific input offered by the experimenter, subject only to the observer's generating of predictions. Along the same line, we might note that the same trend to facilitation in words primed by rhymes is seen in nonwords primed by a nonword rhyme (for example, culse/gulse) (Meyer, Schvaneveldt & Ruddy, 1975). So these facilitation effects might be placed in post-access evaluations, where, by hypothesis, task-specific strategies have their effects.

But whatever the location of facilitation effects, it remains possible that these are contributing to the lack of frequency effects in the distribution of reaction times for closed class items. We should note, though, that the closed class is not a homogeneous set, *except* in terms of the syntactic functions which we have assumed. It seems unlikely, then, that closed class items (for example, some and when) are "semantically related" to each other in the ways that item pairs in facilitation experiments typically are.[7] Though subsets of the closed

---

[7] We note, though, that closed class forms are offered as responses to closed class prompts, with high probability, in tests of association; there is less agreement between subjects than for open class prompts, and responses are produced with greater latency (Glanzer, 1962).

class are semantic objects of roughly the same kind--for example, elements within a minor lexical category, say quantifier, bear a relation to each other--the set as a whole is semantically heterogeneous.

If inter-item priming among closed class items is solely responsible for the lack of frequency effects, then we should not observe frequency effects in the distribution of item reaction times for any well-defined item set when sufficient elements are presented in a single lexical decision experiment. The category "animals" is well-defined, available to all mature speakers, and a set of about the same size as the closed class vocabulary.[8] Certainly the homogeneity of semantic relations is at least as great as that among the closed class forms. One's suspicion that it is undoubtedly greater among animal names than among grammatical words will guarantee a strong test of the claim that inter-item priming is sufficient to abolish frequency effects.

EXPERIMENT 3

METHOD

SUBJECTS. Twenty adult native speakers volunteered for the experiment.

---

[8] We can estimate the size of the set of readily available animal names by counting the types given in the categories "four-footed animals", "birds", "insects", "fishes" and "snakes" in the normative (associative) data collected by Battig and Montague (1969). Restricting the count (a) to single-word names (to avoid essentially descriptive forms like diamond back rattlesnake, yellow-bellied sapsucker) and (b) to names given by at least 1% of the subject pool (to avoid, on the one hand, forms like Woody Woodpecker, man-on-all-fours, and on the other, regional variants such as wombat, goanna), we arrive at an estimate of about 260 as the number of elements in the set. The fact that there are, in fact, millions of animal species, and correspondingly millions of species names, need not exercise us here. We are concerned with the cases which are, in some interesting sense, part of the standard vocabulary of the native speaker.

All were employees or students of the Massachusetts Institute of Technology, and were under 30 years of age.

STIMULI. The item set was made up of 72 words and 72 nonwords. Within the word set, half were animal names, and half (non-animal) nouns. The word sets were matched, pairwise, for frequency of occurrence (mean frequency = 1.02, for both sets), for length (mean number of letters = 4.3, for both sets), and for syllabicity and orthographic structure (for example, otter/ulcer, shark/chalk). The words covered a frequency range 2.3 - 0.0 (log transforms of estimates of frequency per million, from the count of Kucera and Francis (1967), summing over frequencies for singular and plural forms). The nonword set matched the word sets in length (mean number of letters = 4.3) and in number of syllables, and were possible words of English.

PRESENTATION. Following a practice block of 32 items (in all respects representative of experimental items) subjects made WORD/NONWORD decisions in two experimental blocks of 72 items each. A psuedo-random order (allowing runs of no more than three of the same response type) was prepared for half the number of trials, and used for both trials. Elements of each of the matched animal name/noun pairs were randomly assigned one to each trial block, in identical positions of the within-block trial order. Order of presentation of the two experimental blocks was reversed for half the subjects.

PROCEDURE. As in Experiment 1.

RESULTS AND DISCUSSION

In an item analysis, mean reaction time to animal names was 518 msec, and to nouns, 534 msec. This 16 msec advantage to animal names was significant under analysis of covariance (with relative frequency of occurrence as covariate), $F(1,69) = 5.50$, $p < .05$. In a subject analysis, mean reaction time to animal names (518 msec) was less than that to nouns (538 msec), $F(1,19) = 19.6$, $p < .05$. Thus the presentation of a sufficient number of words from a relatively homogeneous set is sufficient to establish inter-item priming, $\underline{minF'}(1,88) = 4.29$, $p < .05$.

Does the presence of a priming effect abolish effects of frequency of occurrence on the distribution of item reaction times? It evidently does not: Table 4 summarizes effects, for animal names and nouns.

TABLE 4

Effects of relative frequency on item reaction time, for animal nouns and nouns

	ANIMAL NAMES	NOUNS
correlation coefficient	-0.77	-0.84
slope of linear regression	-61	-64
zero intercept	582	599
standard error of estimate	30	25
frequency range	0.0 - 2.3	

The frequency effect for animal names ($r = -0.77$, $t(34) = -7.04$, $p < .01$) does <u>not</u> differ from that for the list of mixed nouns ($r = -0.84$, $t(34) = -9.03$, $p < .01$), $z = 0.82$. Figure 2 illustrates the presence of an effect of frequency in both word sets, despite a significant priming effect for the set of animal names.

INSERT FIGURE 2

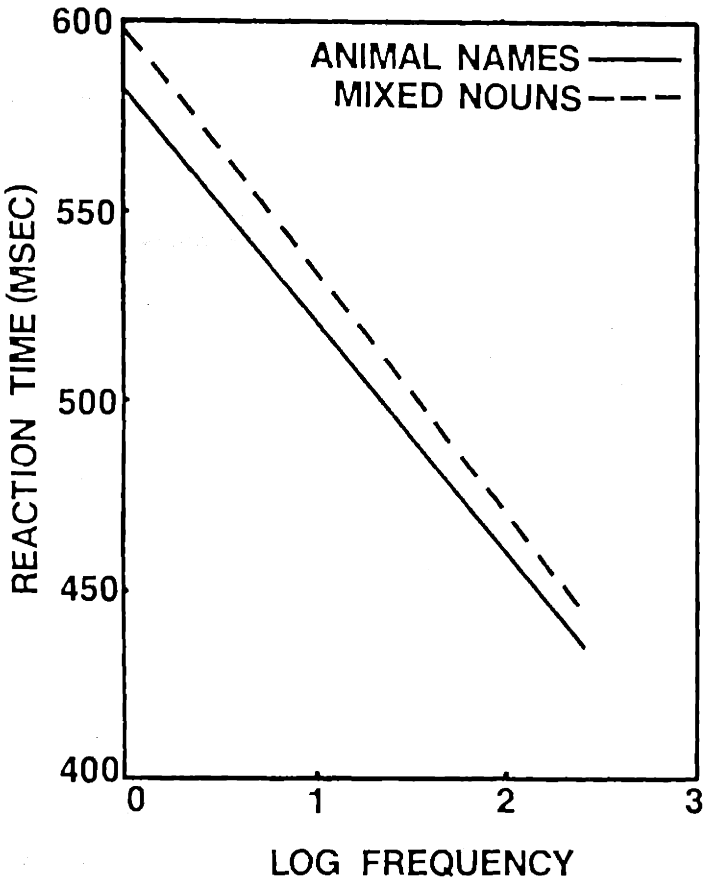


FIGURE 2: FREQUENCY-SENSITIVITY IN THE RECOGNITION OF ANIMAL NAMES AND NOUNS

Linear regression of item reaction time (in milliseconds) on item frequency of occurrence (per million, expressed in units logarithm base 10). Data are for a population of normal speakers.

We conclude that the presentation of many items from a restricted set, while sufficient to permit a priming effect, has no detectable impact on the frequency-sensitivity of recognition performance for open class items. It is clear, then, that the frequency-insensitivity of recognition performance over closed class words cannot be attributed to the fact that the closed class vocabulary is a rather limited set.

The situation might be summarized as follows: frequency-sensitivity, perhaps the most striking property of recognition performance over open class forms, is absent in recognition performance over closed class forms, suggesting that the recognition device which engages mental representations of open class forms is not the same device which engages representations of closed class forms. We note that there is some plausibility to the suggestion that a device, which is specialized for the retrieval of mental representations of closed class forms, might not be frequency-sensitive, given the small number of cases which fall in its domain. This account of frequency-insensitivity in closed class recognition, in terms of a restricted set, is to be distinguished from the one which was evaluated in Experiment 3: restricted set size *per se* will not induce frequency-insensitivity in the operation of a broadly based recognition system, but a restricted domain for a special purpose recognition system may have that effect, in virtue of the processes which permit mental representations to be engaged.

CHAPTER 3

NONWORD INTERFERENCE EFFECTS IN LEXICAL DECISION

Given some input (whether visual or acoustic) a word recognition system can address a mental inventory of forms, and engage mental representations, based on some (as yet unspecified) preliminary analysis of input features. There has been considerable interest, chiefly in studies of reading, in a so-called "code" for lexical access (for a recent example, see Coltheart, Davelaar, Jonasson and Besner, 1977)--that is, in the nature of the information which must be extracted from the input in order that access can take place. We note that, since the mental inventory is finite, it is not necessary that preliminary analyses take account of all the information which can be gleaned from its input, since an analysis which does not recover all the structural features might be sufficient to specify which (or which set of) mental representations could be in accord with the stimulus. That is, what the access routines might require is the minimal analysis which allows successful contact with lexical elements.

And indeed, there is evidence that some parameters of word form "count more" than others in lexical access: in particular, initial segments seem to permit candidate representations to be engaged more readily than do final or medial segments (Bruner and O'Dowd, 1958; Eriksen and Eriksen, 1974; Forster, 1974). If a mental representation can be contacted when only a partial analysis of input is at hand, it is clear that there must be some post-access procedure which can evaluate the accord between a candidate form from the inventory and a more fully specified analysis of input. A system which is inclined to be rash must be able to detect its errors, routinely (though it may not always be successful, Marslen-Wilson and Welsh, 1978). Candidate forms _need_ to be rejected.

A convincing case has been made (Taft and Forster, 1975, 1976; and in more detail, Taft, 1978) that the initial portion of an orthographic form, alone, may be sufficient to engage mental representations. Nonwords which are first syllables of word-forms (for example, _ath_ from _athlete_, _chim_ from _chimney_) take longer to classify in lexical decision than matched nonwords which are not first syllables; similarly, nonwords which are constructed to have initial portions which can be construed as occurring word-forms (for example, _strandlan_ with _strand_, or _scoundlan_, with _scound_ from _scoundrel_) take longer to classify than nonwords which have no such analyses (_spoardlan_). Nor are these effects restricted to nonword cases: words whose first syllable is an occurring form (for example, _bigot_ with an inappropriate sub-analysis _big_) of higher frequency of occurrence also take longer to recognize than matched forms without an initial portion which can lead to error (_pivot_). It cannot be the case that a nonword is more "wordlike" (in terms of digraph frequencies, for example) if its initial portion is an occurring form or its first syllable: initial portions of words which are not first syllables (for example, the _splot_ of _splotch_) do not show elevated reaction times in nonword classification, either standing alone or in construction (_splotpim_).

We may (after Taft and Forster) give an account of the way in which the recognition processes extract a description from the input (in terms of which candidates are selected from the mental inventory), as follows: the initial recognition process does not take account of the entire input form form, but rather only the initial segments are considered in achieving a description sufficient to engage lexical elements; those representations are then examined for their congruence with the entire input form. Consequently, the retrieval system can be led into temporary "error" by the presentation of nonword forms which have, as their initial portions, forms which are in fact words (or first syllables of words). For example, in the recognition of nonwords like padlato, where an analysis as pad+... can be made, the decision that the form is not a word is hampered when the retrieval system inappropriately engages a representation headed [PAD]. That representation must be rejected, of course, since the remainder ...lato is not a possible (that is, occurring) continuation of pad...; but the post-access evaluation procedure operates at time cost, and we observe an interference effect.

If, as we argued earlier, the closed class vocabulary is typically retrieved by a mechanism which is separate from that which operates over open class forms, and if the vocabulary, in consequence, is not available to that system which relies on a partial analysis of inputs in engaging mental representations (or is marked in some way as privileged), then open and closed class words might contrast in their inducing of interference effects in nonword constructions. Such a contrast would speak strongly to a distinction between the vocabularies: Taft (personal communication) reports that interference effects are to be seen in quite improbable nonword constructions (for example, penxptfl). The retrieval system is evidently rather insensitive to parameters of form in other than initial position.

Experiment 4 examines the pattern of interference effects in nonword classification, contrasting constructions with closed class forms (yetitude) with those with open class forms (setitude).

EXPERIMENT 4

METHOD

SUBJECTS. Twenty students at the Massachusetts Institute of Technology took part in the half-hour experiment. All were native speakers of American English, and were paid for their participation.

MATERIALS. The item set was made up of 64 nonwords and 80 words. Within the nonword set, items were constructed in quadruples, having one of each of the following nonword types:

Baseline Nonwords: nonwords whose initial strings cannot be construed as occurring word-forms (for example, nacherty)

Closed Class Nonwords: nonwords whose initial strings are closed class forms (for example sucherty, with such/erty)

Open Class (High) Nonwords: nonwords whose initial strings are open class forms of relatively high frequency of occurrence (for example, worderty, with word/erty)

Open Class (Low) Nonwords: nonwords whose initial strings are open class forms with relatively lower frequency of occurrence (for example, casterty, with cast/erty)

As the examples above suggest, forms used as initial strings were matched, within quadruples, in length and (as far as possible) orthographic structure. Words used as initial strings were selected so that, within quadruple, frequency of occurrence was greatest for closed class forms, and decreasing through open class (high) and open class (low) forms: mean frequency of occurrence for these initial strings was 3.24, 2.32 and 0.82. Two sets of open class nonwords, of different frequency, were constructed because it is not possible to find appropriately matched open and closed class pairs for which frequency is held constant. But if there is an effect on frequency on the magnetitude of interference effects, it is a positive one: interference effects tend (though not significantly, Taft, 1976) to increase with frequency. That is, closed class forms should induce greater interference effects than open class forms, if they are treated in the same way by the recognition system.

The 80 items in the word set were similarly selected in structurally matched quadruples, and exhibit the same variation in length as the nonword set to which they serve as foils in the experiment.

A pseudo-random order, containing no runs of more than three of the same response type (word/nonword) was prepared for 16 nonwords and 20 words, and repeated four times for a presentation sequence over all 144 experimental items. In assigning the experimental items to the prepared sequence, one of each of the 36 nonword and word quadruples was assigned to each of the quarters of the sequence in equivalent positions. Thus the elements of a given quadruple (which will be compared with each other) were presented in equivalent environments, in terms of the structure and type of preceding and succeeding items.

The sequence was seen by all subjects, in two blocks, with the block order being reversed for half the subjects. In addition to a practice block of 20 items (representative of experimental items) which was used to familiarize subjects with the task, each experimental block was headed by a small practice block of four items, to minimize warm-up effects on reaction times.

INSTRUCTIONS, PRESENTATION AND DATA TREATMENT: As in Experiment 1.

RESULTS AND DISCUSSION

Mean reaction times and error rates for the four nonword types are presented in Table 5. Figure 3 illustrates contrasts among word-headed nonword types with the baseline control.

TABLE 5

Mean reaction time (in milliseconds) and error rates as a function of nonword type

NONWORD TYPE	EXAMPLE	RT	%E
Baseline	vostner	588	1.3
Closed Class	mostner	595	3.1
Open Class (High)	lostner	629	4.4
Open Class (Low)	mistler	618	7.8

There is a significant effect on reaction time of nonword type, minF' $(3,62) = 2.76$, $p < .05$; a test of all types against the baseline nonword control (Dunnett, 1955) locates the effect in the elevated reaction times of the two open class nonword conditions. Closed class nonwords take no longer to classify than baseline nonwords. The failure of the interference effect when closed class forms head nonwords cannot be attributed to their higher frequency of occurrence;[9] although open

[9] Note also that the failure to find an interference effect in closed class nonwords undermines any general interpretation of effects of this sort in terms of the higher sequential letter probabilities of nonwords whose initial portions can be construed as words. Because closed class forms control the high end of the word frequency range, nonwords with closed class initial strings necessarily have high sequential letter probabilities--and yet show no interference.

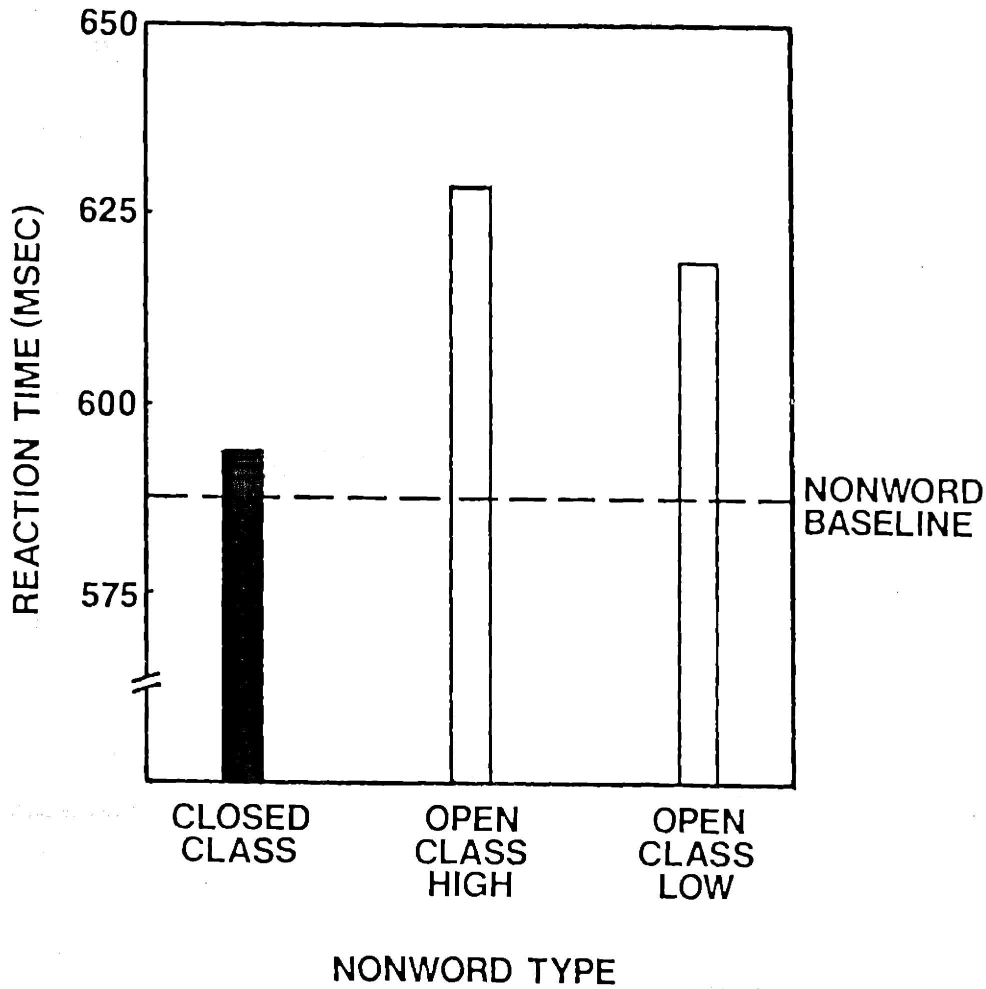


FIGURE 3: NONWORD INTERFERENCE EFFECTS IN LEXICAL DECISION

Mean reaction time (in milliseconds) for nonword decisions as a function of vocabulary class of initial portions. Data are for a population of normal speakers.

class (high) nonwords do not take more time to classify than open class (low) nonwords (minF' < 1), they surely do not take less time (629 vs 618 msec). The frequency trend runs against the differences observed between open and closed class nonwords.

Closed class forms, used as initial portions of nonwords, fail to induce an otherwise robust interference effect in reaction time for nonword classification. What can that tell us about the processes underlying recognition of open and closed class forms? In the framework set out earlier, two interpretations suggest themselves:

(1) the closed class vocabulary is simply unavailable to a retrieval system which makes use of initial portions of input in engaging mental representations; and

(2) the closed class vocabulary, though available to that retrieval mechanism, is marked such that when mental representations of closed class forms are falsely engaged (as when [THAN] is encountered as the system deals with the input thanage), they are rejected at zero time cost (or at least a cost which is sufficiently small as to be undetectable with our experimental technique).

The first interpretation is attractive, given the claims made earlier about the existence fo a retrieval mechanism specific to the closed class: for, if closed class forms already fall in the domain of a special-purpose recognition device,[10] there is no necessity for them

---

10 The open and closed class forms which made up the initial portions of nonwords in Experiment 4 were presented (as words) in Experiment 4'. It is quite clear that they take part in the contrast we noted in Chapter 2. That is, recognition performance is frequency-sensitive for open class forms (r = -0.72), but not for closed class forms (r = +0.01).

to be dealt with by a second, more broadly based recognition system. But (to anticipate our discussion of word recognition performance in a population of aphasic speakers) it is probably not possible to make such a clean division of responsibilities, between two retrieval devices. We shall postpone a discussion of the relative merits of the two interpretations, and concentrate here on specifying, in more detail, the second.

Though proposals of zero-cost operations are generally unpalatable, there are considerations which we can bring to bear which make the suggestion more acceptable. Taft and Forster (1976), in their discussion of nonword interference effects of just the type we are concerned with here, place the interference (the elevation of reaction time) in the time taken to search through the contents of a lexical entry, to evaluate the "goodness of fit" of the remainder of an input to its initial portion. In these terms, the zero-cost proposal is to be cashed as follows: there is "something about" the mental representations of closed class items, such that this (costly) evaluation procedure is not carried out.

There _is_ a property of the closed class vocabulary, which is a ready candidate for being cast as that "something". While open class items enter quite freely into compounds (e.g., _birdcage_, _rollercoaster_, _pigheaded_) and undergo derivation (e.g. _fickleness_, _treatment_, _migration_), closed class items generally do not. There is some compounding (e.g., _nevertheless_, _whomsoever_) of closed class items, but these are fixed forms, and productive compounding is quite rare (cf., _somewhere_: *_somewhy_); and derivation is absent. If the retrieval system's use of initial portions of inputs to engage mental representations of words is closely tied to a treatment of the facts of word-formation (as Taft and

Forster seem to suggest it might be), then word-formation properties may indeed control the application of an evaluation procedure. A lexical entry might be marked as "non-engaging-in-word-formation", for example, blocking any further search through the entry contents.

We can consider patterns of word-formation in two ways. On the one hand, we can talk about words which may, and words which may not, engage in word-formation, abstracting away from irregularities in attested cases and stating a general principle; in these terms, open class items may engage in word-formation while closed class items may not. On the other hand, referring to attested cases of compounding and derivation, we can talk about forms which do, and forms which do not; this corresponds only approximately with a distinction between the two vocabularies, for, in addition to the closed class compounds cited above, there are accidental gaps in patterns of word-formation in the open class. That is, though an open class item may potentially enter compounds or undergo derivation, not all do so; for example, there are no attested compounds for the word dawn (though there could be, e.g. dawnmist), in contrast with the form drum, which is a rich source for compounds (drumbeat, drumhead, drumroll, drumstick, and so on).

If patterns of interference in nonword classification turn on whether a word used as initial portion of a nonword may engage in word-formation, then, in effect, it is vocabulary class which controls whether an evaluation of the contents of lexical entries is undertaken. But if interference patterns appear only when words do engage in word-formation, the effects established in Experiment 4 cannot be truly cast in terms of a distinction between open and closed class vocabularies, since the distribution of attested compounds and derived forms is closely related to but not identical with that distinction. A systematic comparison of reaction times to nonwords headed by forms which are inactive for word-formation (in terms of attested cases) with those to nonwords headed by forms which are active for word-formation will distinguish these alternatives. If it is vocabulary class which counts, then open class forms like dawnicle (with inactive dawn) as well as forms like drumacle (with active drum) should take longer to classify than a nonword like drimacle, lacking an initial portion which is a word. But if only the attested facts of word-formation apply, then nonwords like dawnicle will not show the elevated reaction times which identify an interference effect.

In drawing a contrast between open class forms which are active or inactive for word-formation, we are setting aside the fact that some derivation is available to all open class items (just as we set aside the similar facts of compounding). These are the forms which are productively derived under word-boundary affixation in English. Thus, for example, the suffix #like can be applied to most nouns, certainly concrete ones, to forms adjectives (for example, cowlike, pulpitlike, dawnlike); noun-forming #er takes verbs (for example, singer, caterwauler); noun-forming #ness takes adjectives (for example, blueness, squelchiness). This does not present a problem for the analysis being pursued here. The availability of such productive forms is by and large an automatic consequence of membership in some major lexical category (for example, adjective, in the case of #ness). To claim that the simple availability of productively derived forms is sufficient to indice interference effects in nonword classification is to claim, in effect, that being in a major lexical category (being in the open class) is sufficient. Again, this returns to the claim that vocabulary class controls patterns of interference.

In Experiment 5, we contrast performances in nonword classification for forms having two kinds of initial portions--open class forms which are active for word-formation, and open class forms which are inactive for word-formation--with performance for baseline nonwords.[11]

EXPERIMENT 5

METHOD

SUBJECTS. Twenty adult native speakers of American English volunteered for the experiment. All were employees or students, under 30 years of age, at the Massachusetts Institute of Technology.

MATERIALS. The item set was made up of 48 nonwords and 48 words. Within the critical nonword set, items were constructed in triples, elements of triples being one of each of the following types:

Baseline Nonwords: nonwords whose initial string cannot be construed as a word of English (for example, drimacle)

Open Class Active Nonwords: nonwords whose initial string are open class forms which enter attested compounds or derivations (for example, drumacle)

---

[11] We could equally well contrast nonword cases constructed with closed class words which do or do not enter into attested compounds. There are, however, difficulties, practical rather than theoretical: (a) we run the risk of an experimental outcome (no differences among nonword types) which is ambiguous between interpretation as an interesting result, and as a failed experiment; and (2) there are simply not enough cases among closed class forms which satisfy the requirements of such a contrast.

Open Class Inactive Nonwords: nonwords whose initial strings are open class forms which do not enter attested compounds or derivations, and which do not form first syllables of any other attested forms (for example, dawnicle).

As the examples above suggest, forms used as initial portions to nonwords were matched (with triples) in length and orthographic structure, and embedded in equivalent environments. In the case of open class initial strings, elements within triples where matched for frequency of occurrence (mean frequencies of 1.77 and 1.81 for open class active and inactive respectively).

The word set was made up of 16 triples of non-derived words, structurally matched within triple; the set exhibited the same variation in length as the critical nonword set.

A pseudo-random order, containing no runs of more than three of the same response type (word/nonword) was prepared for 16 nonwords and 16 words, and repeated three times for a presentation sequence over all 96 experimental items. Experimental items were assigned to the prepared sequence in the way set out for Experiment 4.

The sequence was seen by all subjects, in a single block. A practice block of 20 items was used to familiarize subjects with the task, and a small practice block of four items headed the experimental block.

INSTRUCTIONS, PRESENTATION AND DATA TREATMENT: As in Experiment 1.

RESULTS AND DISCUSSION

Table 6 gives mean reaction times and error rates for the three nonword types. Contrasts among nonword types are illustrated in Figure 4.

INSERT FIGURE 4 HERE

TABLE 6

Mean reaction time (in milliseconds) and error rate, as a function of nonword type

NONWORD TYPE	EXAMPLE	RT	%E
Baseline	tostilude	561	1.6
Open Class Active	postilude	594	1.6
Open Class Inactive	lostitude	589	1.6

There is a significant effect on reaction time for nonword classification of nonword type minF' = 7.9, p. < .01; both categories of open class nonwords have significantly elevated reaction times in comparison with the control nonword baseline. There is, moreover, no difference between open class types, minF' < 1.

The presence of an open class item as initial portion of a nonword induces interference in nonword classification, regardless of whether that form actively engages in word-formation.[12] That is (summing over Experiments 4 and 5), interference effects seem to pattern with the

[12] For the argument at hand, it is not necessary that interference effects be of the same magnitude for open class active and open class inactive nonword constructions. All that is necessary is that there be some interference in both cases. In Experiment 5, an open class active nonword header elevated reaction times by 33 msec, as compared with 28 msec for open class inactive forms; Taft (1976) found a larger (and more reliable) difference in an experiment using (roughly) equivalent cases. His experiment, however, failed to include the baseline condition which is crucial for the analysis we pursue here.

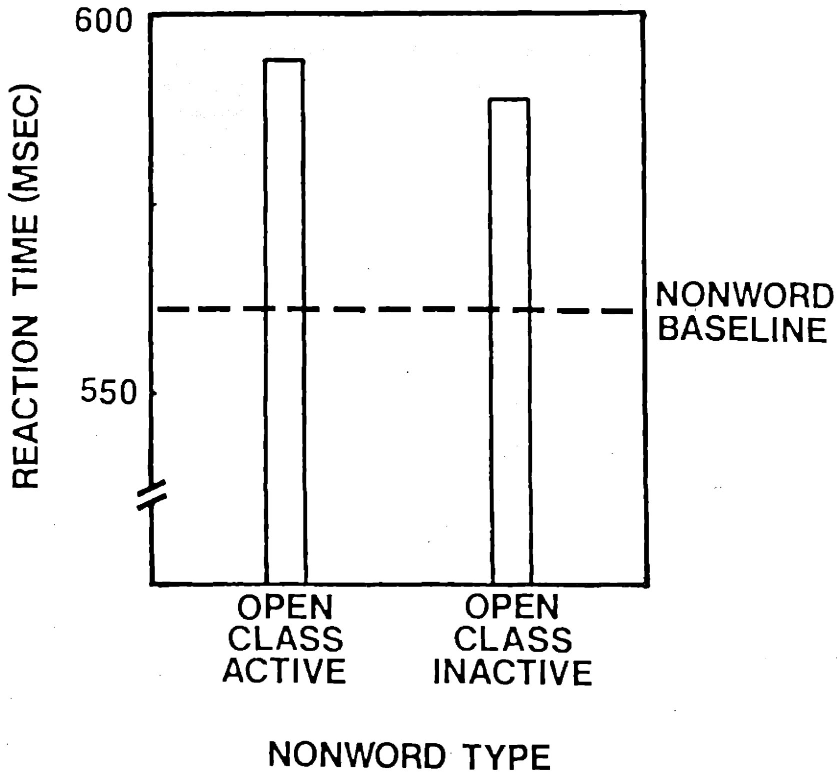


FIGURE 4: NONWORD INTERFERENCE EFFECTS IN LEXICAL DECISION

Mean reaction times (in milliseconds) for nonword decisions as a function of nonword type. Data are for a population of normal speakers of American English possibility of word-formation rather than with the distribution of attested cases of derivation and compounding. And since possibility of word-formation[13] rests with the major lexical classes (noun, verb, adjective) --that is, what we are calling here the open class vocabulary--it is vocabulary class which is picked out as the feature determining the patterns of interference effects.

We laid out earlier two interpretations of the lack of interference effects for nonwords headed by closed class forms: the first, that the closed class forms are not available to an open class recognition device which makes use of initial portions in identifying lexical candidates; the second, that the closed class forms, while in the domain of an open class recognition device, are marked such that post-access evaluation of the information arrayed against lexical representations is not carried out. Now these two interpretations converge in an important respect, namely, in pointing to the contrast between vocabularies as a determiner of lexical retrieval processes. But they diverge in a significant respect: in the parcelling out of responsibility to subcomponents (roughly, look-up and post-access evaluation) of those processes. This contrast is one which is difficult, at best, to weigh experimentally, and the litrature in semantic priming and/or context effects surely bears witness

---

[13] We ignore here a kind of "word-formation" whose distribution in English also distinguishes the closed and open class vocabularies: this is the variation in surface form created by inflection. It is an accidental property of English (related to the rather weak inflectional devices of the language) that there are no inflectional paradigms in the closed class. Experiments with English materials cannot determine the contribution of the distribution of inflection to the nonword interference effect; rather, we need to experiment in a language which extends inflection into the closed class.

of the problem. We shall return to this point in Chapter 4.

In summary, it appears that the trademarks we have associated with recognition performance over open class forms--frequency sensitivity and sensitivity to initial portions of input--are absent in recognition performance over closed class forms. This is evidence in support of a view that access to mental representations of a restricted vocabulary of closed class forms is accomplished by use of a retrieval system which differs from that which allows recognition of forms of the open class vocabulary.

CHAPTER 4

RETRIEVAL FAILURES IN BROCA'S APHASIA

If, as we speculated above, a close relation holds between the use of a specialized retrieval system for the closed class vocabulary and the manipulation of structural facts in sentence processing, then some ground is to be gained by an investigation of retrieval mechanisms in populations of speakers who are known to lack control of such facts: Broca's aphasics appear to be just such a population.

4.0 DEFICITS IN BROCA'S APHASIA

Broca's aphasia is a clinical syndrome resulting from relatively extensive anterior damage to the dominant (typically, left) hemisphere. Mohr (1976) has pointed out that damage restricted to what has traditionally been called "Broca's area" (the pars opercularis of the third frontal convolution) frequently produces no lasting disturgance in language function. Rather, the syndrome is associated with a major infarction involving most of the territory of supply of the upper division of the middle cerebral artery.

Clinically, the syndrome is distinguished by an initially total aphasia (roughly, mutism) which rapidly evolves to stabilize in the severe language impairment which is characterized as Broca's aphasia (or expressive, or motor, or non-fluent, or agrammatic aphasia, as it has been variously termed). The most widely disseminated description of Broca's aphasics is that they seem to know what they want to say, yet speak effortfully and in a "telegraphic" manner (Goodglass and Kaplan, 1972). The telegrammatic or agrammatic output is characterized by the use of simple syntactic forms (e.g. Goodglass, 1976), a relative omission of grammatical morphemes--what we have called here the closed class forms--whether bound or free (e.g. Howes and Geschwind, 1964) and a concomitant dependence upon nouns, and to a lesser extent, verbs in their uninflected form. This pattern persists even when the patient commands a considerable speaking vocabulary (Zurif and Caramazza, 1976).

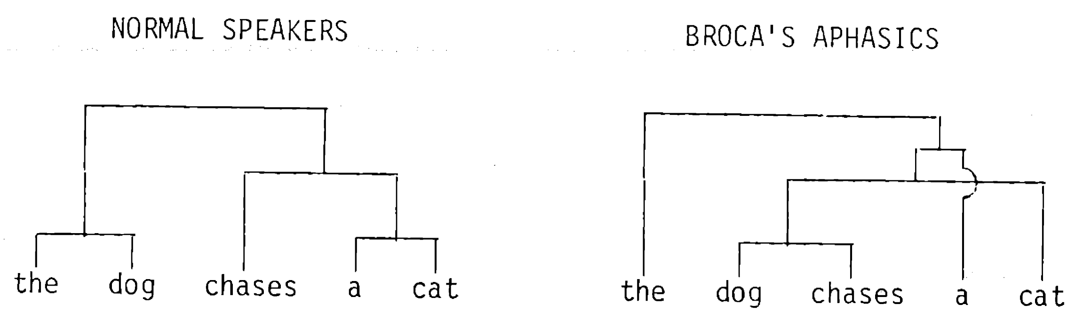
It has been suggested (for example, by Isserlin, 1922) that telegrammatic output represents a move on the part of the patient to "save effort" in the face of a problem in articulation which is part of the symptom complex. This cannot be an adequate account of the agrammatism. In a fascinating study with a single patient, Andreewsky and Seron (1975) report a reliable omission of closed class forms in reading from a <u>list</u> of words, except for cases of forms homographic with an open class word (for example, <u>will</u>, which in isolation can be taken as a noun rather than as auxiliary). Neglect of the closed class by agrammatic speakers in their spontaneous productions is paralleled by a neglect of that vocabulary when patients read or repeat sentences, though in automatic speech (for example, phrases such as "I don't know" or "I guess so" which seem to be available pre-packaged) well-intoned production is accompanied by reliable inclusion of the closed class forms. The problem seems to be one of the integration of information types in the planning and execution of utterances, rather than a simple motor difficulty. To the extent that the degree of agrammatism is not under the control of the speaker, it cannot usefully by characterized as an effort-saving strategy. And certainly to the extent that agrammatism is carried over into the domain of sentence comprehension, it is not well-described in these terms.

The current (common) clinical impression is that comprehension in Broca's aphasia is, relatively speaking, spared. The dissociation of a severely impoverished production in the context of an apparently adequate comprehension is traditional, stemming perhaps from the clinician's quite practical concern over a gross and very apparent "expressive loss". But there has not always been universal acceptance. Wernicke, for example, writing in 1908, states some doubts:

> "I no longer am of the opinion that in pure motor aphasia the ability to understand speech always remains unimpaired....There is almost always a certain inability to understand complicated constructions and the finer differentiations of speech..." (cited in Mohr, 1976).

Indeed, careful probing has begun to uncover deficits in comprehension which parallel those so evident in production. For example, when Broca's aphasics sort phrases ([den Jäger (the hunter)], [der Hase (the hare)], [schiesst (shoots)]) to form sentences, they do so in a way that allows a plausible reading ("the hunter shoots the hare"), *despite* the fact that this is ungrammatical: *Jäger* is marked as grammatical object with *den*, and *Hase* as grammatical subject (von Stockert and Bader, 1976). Similarly, Caramazza and Zurif (1976) report that comprehension fails, when such patients deal with sentences whose meaning violates "real world knowledge"; for example, *the boy that the dog is patting is fat*. Summarizing, we can say that a sentence seems to be understood only on the basis of a "likely meaning" or "what makes factual sense", inferred from a sampling of the lexical content, independent of structure. Asyntactic comprehension, involving (essentially) concatenation of references under pragmatic constraint, is undoubtedly a strategy available to normal speakers. What is striking in the case of Broca's aphasia is an apparently total reliance on the strategy.

Perhaps the most striking evidence of an asyntactic comprehension in Broca's aphasia comes from performance in a metalinguistic task, which makes use of a paradigm developed by Levelt (1970). A hierarchical cluster analysis (Johnson, 1967) over a matrix of relatedness measures for all possible word pairs from a presented sentence, recovers a hierarchical structure, which (for normals) is remarkably similar to a syntactic bracketing. Broca's aphasics, in contrast, simply concatenate open class items, treating closed class items inappropriately, even to the extent (on occasion) of grouping these together (Zurif, Caramazza & Myerson, 1972). A case is illustrated below:



The problems of Broca's aphasia, in comprehension, appear to be in many ways equivalent to those which are so evident in production, centering in a lack of control of syntactic structures and a presumably related neglect of the grammatical vocabulary. (This is not to claim that there are no deficits which are specific to either comprehension or to production--we are setting aside a consideration of, for example, the purely motor aspects of production performance in Broca's aphasia.) But to say that agrammatic speakers have a "problem" in their handling of the closed class vocabulary does not begin to specify the nature of the deficit in terms of the mechanisms which underlie sentence processing. In the following sections we explore the performance of Broca's aphasics, in word recognition, in terms of the computational distinctions between open and closed class vocabularies which we have established in populations of normal speakers.

4.1 FREQUENCY EFFECTS IN LEXICAL DECISION

EXPERIMENT 6

METHOD

SUBJECTS. Five patients presenting with Broca's aphasia were selected from the records of the Aphasia Unit of the Boston V. A. Hospital. These patients had been diagnosed, on the basis of clinical examinations and the results of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972) as classic cases of Broca's aphasia. Locus of lesion had been established from clinical indices (e.g. hemiplegia) and from brain scans. All but one patient had sustained damage in cerebro-vascular accidents; the fifth, from a tumor.

Five patients selected from non-neurological wards of the Boston V. A. Hospital matched with the aphasic patients in terms of age and educational background, served as control subjects.

All subjects were male, and native speakers of American English.

MATERIALS. As in Experiment 1.

INSTRUCTIONS. As in Experiment 1. Additionally, to make sure that aphasic patients understood the word/nonword classification, they were pre-tested informally with items (not used in the experiment) on flash cards, and coached where necessary.

PROCEDURE. As in Experiment 1, except for such modifications as are dictated by the use of Broca's aphasics as a subject population; namely, having WORD/NONWORD made with one hand (the non-preferred) on a telegraph key, and extending the intertrial interval for those patients with relatively long reaction times. Though there are no grounds for expecting these modifications to effect distributions of reaction times, control subjects made their responses in the same fashion.

DATA TREATMENT. Only reaction times for correct "word" decisions were analysed. Bad data were identified on a subject-based criterion only (the small number of subjects, together with striking differences among subjects in base-line reaction time, precludes the use of an item-based criterion): a datum is eliminated from the analysis if it exceeds a cutoff value for the subject (mean reaction time for all word decisions plus two standard deviations). Missing data (whether due to subject or experimenter error, or bad data) are replaced with the mean of the subject's remaining reaction times.

We should note that a data comb of this kind is quite conservative, and will (on average) reduce experimental effects.

As in Experiment 1, effects of relative frequency on item reaction time are evaluated in the correlation of item reaction time with item frequency of occurrence, and, equivalently, in the slope of the linear regression of reaction time on frequency. No item for which any data was available was eliminated from the analysis.

RESULTS AND DISCUSSION

We note, first, that Broca's aphasics readily identified closed class words (which they rarely utter and perhaps neglect in comprehension) as words of their language; classification errors for closed class forms are only 8%, and for open class forms, 5%. Table 7 summarizes effects of relative frequency of occurrence on item reaction time.

TABLE 7

Effects of relative frequency on item reaction time for open and closed class words for two speaker populations

	NORMAL SPEAKERS		AGRAMMATIC SPEAKERS	
	OPEN CLASS	CLOSED CLASS	OPEN CLASS	CLOSED CLASS
correlation coefficient	-0.37	+0.05	-0.37	-0.38
slope of linear regression	-26	+5	-40	-45
zero intercept	849	773	1032	1076
standard error of estimate	58	65	82	76
frequency range	0.6-3.4	1.5-4.8	0.6-3.4	1.5-4.8

Though the performance of hospitalized controls parallels that of our previous populations, that of the aphasics does not, and there is no reflection of the distinction between vocabulary classes that is evident with normal speakers. Frequency of occurrence influences reaction time for <u>both</u> open and closed class sets, with $r = -0.37$ for open class items ($t(59) = -3.01$, $p < .01$) and $r = -0.38$ for closed class items, ($t(56) = -3.07$, $p < .01$). Figure 5 illustrates the contrasting performances.

INSERT FIGURE 5

We must note that for reaction time experiments, five subjects is a tiny sample, and these findings are to be considered preliminary. The remarkable consistency of the obtained correlations allows some comfort.

In the light of the patterns established for normal subjects the

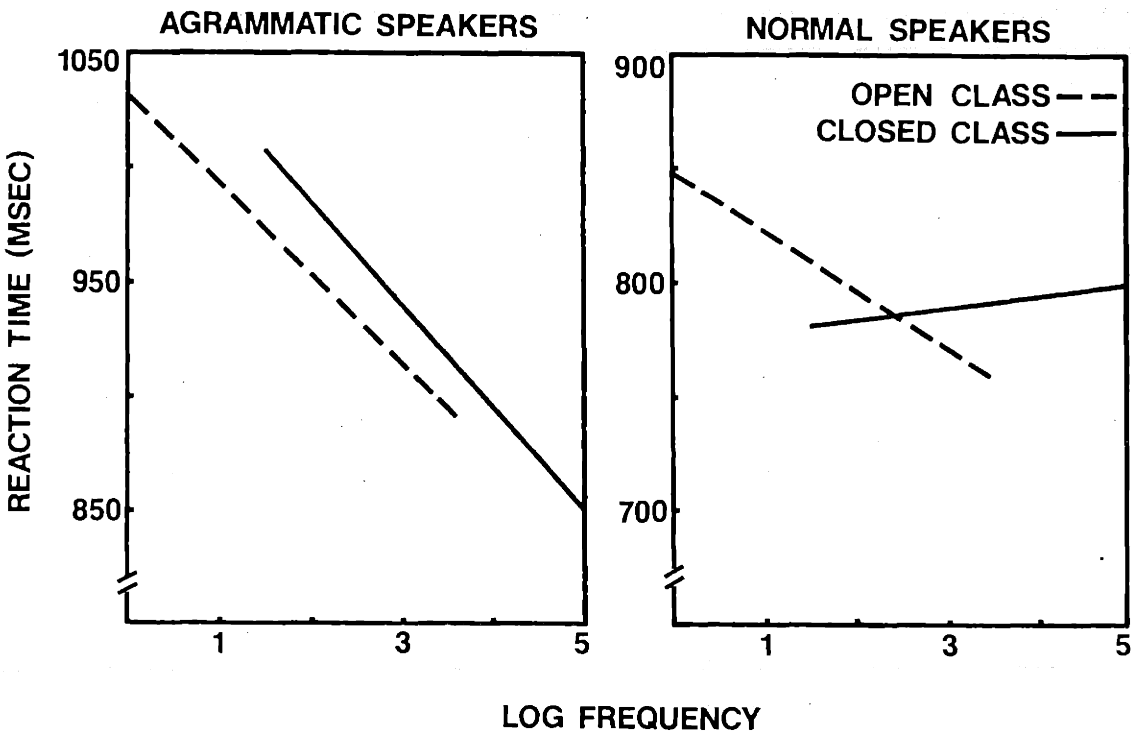


FIGURE 5: PATTERNS OF FREQUENCY-SENSITIVITY IN OPEN AND CLOSED CLASS RECOGNITION, FOR TWO POPULATIONS OF SPEAKERS

Linear regressions of item reaction time (in milliseconds) on item frequency of occurrence (per million, expressed in units logarithm base 10).

performance of Broca's aphasics in recognition of closed class words is to be taken as a retrieval failure. It is not, however, a failure to contact mental representations of closed class items at all (for the error rate is rather low); rather it is a failure to retrieve mental representations in the way that normal speakers do; that is, via a frequency-insensitive device.

We might make the (not unreasonable) assumption that damage incurred by the brain will not create new mechanisms for processing inputs. Our findings with agrammatic speakers suggest that closed class forms are recognized by these patients via a system indistinguishable from that which allows serves in the recognition of open class forms. Then, a version of our assumption (setting aside questions of mutual inhibition and the like) dictates that, at least in some circumstances, a broadly based frequency-sensitive retrieval mechanism must also be able to engage mental represntations of closed class forms in normal speakers. The clean dissociation of one recognition system, whose domain is the open class vocabulary, and another, whose domain is the closed class vocabulary is lost (with the caveat noted). Rather it would seem that a broadly based system operates over the entire vocabulary (open and closed), while a special purpose device operates over just the closed class vocabulary.

If, in Broca's aphasia, the closed class vocabulary is readily available to a broadly based recognition system, we might ask whether the same appreciation of the character (and significance) of membership in a restricted vocabulary is gained when access is not via a specialized recognition device. We argued (Chapter 4) that the pattern of interference in nonword classification, as it relates to the presence of an open or closed class word as initial portion is a nonword, is to be taken as evidence that the distinction between the vocabularies is marked (at least) at the level of lexical entries. The failure of a closed class retrieval device, in Broca's aphasia, may extend to a failure to mark lexical entries as to vocabulary type (or, perhaps in a way that is indistinguishable from this, a failure to be able to utilize information about vocabulary class even though it is arrayed against entries).

4.2 NONWORD INTERFERENCE EFFECTS IN LEXICAL DECISION

EXPERIMENT 7

METHOD

SUBJECTS. As in Experiment 6.

MATERIALS. As in Experiment 4. Briefly, the crucial nonword set is made up of four kinds of nonwords: baseline nonwords (e.g. *nacherty*); closed class nonwords (e.g. *sucherty*); and two classes of open class nonwords (e.g. *worderty*, *casterty*), differing in the relative frequencies of occurrence of the words used as initial portions in the construction of nonwords.

INSTRUCTIONS, PROCEDURE AND DATA TREATMENT. As in Experiment 6.

RESULTS AND DISCUSSION

The performance of hospitalized controls exactly parallels that of the normal speakers (M.I.T. undergraduates) with whom we had established a pattern of interference effects in Experiment 4. Mean reaction times and error rates for the four nonword types are presented in Table 8.

TABLE 8

Hospitalized controls: mean reaction time (RT, in milliseconds) and error rate (%E), as a function of nonword type

NONWORD TYPE	EXAMPLE	RT	%E
Baseline	dititude	953	0.0
Closed Class	yetitude	953	0.8
Open Class (High)	setitude	1024	1.7
Open Class (Low)	petitude	1024	1.7

The performance of hospitalized control subjects shows a significant effect of nonword type, minF'(3,41) = 3.12, $p < .05$, located in the elevated reaction times for open class nonwords: Dunnett t(4,45) = 2.98, $p < .05$, and Dunnett t(4,45) = 2.95, $p < .05$, in an item analysis, for open class (high) and open class (low) nonwords respectively. It is clear that closed class nonwords do not differ from the baseline.

The performance of Broca's aphasics is strikingly different: Table 9 summarizes effects.

TABLE 9

Broca's aphasics: mean reaction time (RT, in milliseconds) and error rates (%E), as a function of nonword type.

NONWORD TYPE	EXAMPLE	RT	%E
Baseline	thiddace	1112	12.3
Closed Class	thattice	1203	16.3
Open Class (High)	shuttice	1206	16.3
Open Class (Low)	chattice	1193	23.8

There is a realitively high error rate (that is, false acceptances as "word"), which is unevenly distributed over nonword types, though the effect fails to reach significance: by subject analysis, $F(3,12) = 2.25$, $p = 0.14$. Why this should be so is unclear: we note, though, that false acceptances can be seen as interference effects of the grossest kind. For a reaction time analysis, it is certainly the case that a high error rate (in combination with a small number of subjects) creates problems. Missing data are replaced, by subject, with the mean of correct nonword responses; this will tend to minimize differences among nonword types.

Nevertheless, it is clear that there is an interference effect in the distribution of reaction times: all nonword types which are headed by words (whether closed or open class) are classified at longer latencies than baseline nonwords; in an item-based analysis, $F(3,45) = 2.54$, $p < .07$, and, in a subject-based analysis, $F(3,12) = 2.88$, $p < .08$.[14]

In Figure 6, the contrasting patterns of interference in the nonword classifications of agrammatic and normal speakers are illustrated.

INSERT FIGURE 6

Thus, in addition to the finding that agrammatic speakers do not recognize closed class words in the way that normal speakers do, it appears that they do not have control of the consequences of membership

[14] As with Experiment 6, these results are best taken as preliminary, in view of the small number of subjects. It is to be expected that the addition of more subjects will, in overcoming difficulties of missing data, establish the reliability of the contrasts noted.

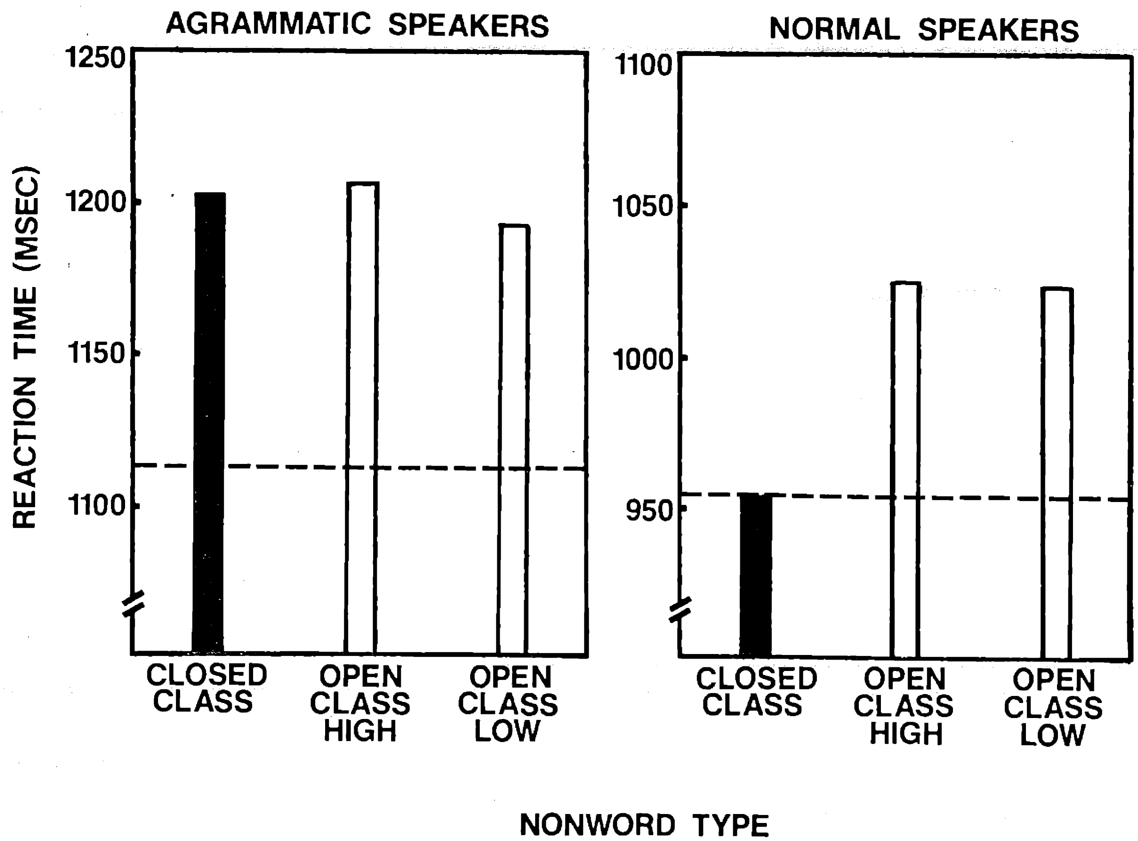


FIGURE 6: PATTERNS OF NONWORD INTERFERENCE IN LEXICAL DECISION, FOR TWO POPULATIONS OF SPEAKERS

Mean reaction time (in milliseconds) for nonword decision as a function of vocabulary class of initial portions, compared with baseline nonword decision time (dashed line).

of a vocabulary class in the way that normal speakers do. If we suppose (with Taft and Forster, 1975, 1976) that interference effects are to be located in the post-access evaluation procedures which match candidate mental representations of words with inputs, then, unlike normal speakers, agrammatic speakers cannot employ the fact of membership in the closed class vocabulary to block evaluation procedures when, for example, the representation [THAN] is engaged as the recognition systems deals with input thanage. Two accounts of the lack of control of class information by agrammatic speakers suggest themselves. On the one hand, we might suppose that for Broca's aphasics there are no grounds for distinguishing the vocabularies; that is, class information is simply unavailable to the sentence processing system, in every sense, and closed class words are treated in all respects as if they were open class forms. On the other hand, it might be that use of the fact of closed class membership at the level of word recognition rests entirely with the availability of a special-purpose retrieval device which picks out closed class forms. There is some evidence at hand which goes against the former account; namely, those instances where agrammatic speakers behave as if they were crucially aware of vocabulary type (e.g. the single patient of Andreewsky and Seron, 1976). Certainly the noun-ish (or nominalized) utterances of these speakers suggests that some grammatical category information (though here, within the open class) is under control (see Kean, 1977, for a discussion).

How closely the failure of Broca's aphasics to distinguish open and closed class vocabularies can be tied to their agrammatism (and apparently asyntactic comprehension) remains a fascinating problem. We can establish that it is at least correlated with their agrammatism, in some sense, by dismissing the possibility such failures are associated with aphasic syndromes in general, rather than agrammatism in particular. Patients with anomic aphasia share with Broca's aphasics a word-finding difficulty, in the context of fluent, grammatically well-formed speech (Goodglass and Kaplan, 1972). If our presumption is correct, such patients show not exhibit the word recognition performances which we have associated with agrammatism. And indeed they appear not to: tests (with a single patient, to date) indicate performances in recognition of the open and closed class vocabulary which exactly parallel those of normal speakers.

CHAPTER 5

LATERAL ASYMMETRIES IN WORD RECOGNITION

The left and right hemispheres in man differ in processing capability. Striking evidence of this arises in the course of clinical studies, where it appears that it is the left hemisphere which mediates complex language behavior, in the overwhelming majority of cases. Zangwill (1962) goes so far as to claim that "for practical purposes, the probability of right cerebral dominance [that is, right hemisphere mediation of language behavior--DCB] in a fully righthanded individual is so low that it may be disregarded" (p. 106). Zaidel (1978), summarizing studies of the residual behaviors of an "isolated" hemisphere (that is, cases resulting from commissurotomy or hemispherectomy), provides evidence that performance mediated by an isolated left hemisphere is by and large normal, while that controlled by an isolated right hemisphere shows impairment: speech is minimal, if present at all, and comprehension apparently asyntactic. A marked difference in the incidence of aphasias, following lesions restricted to right or left hemispheres, bears out these observations rather clearly: cases of aphasia with well-lateralized lesions of the right hemisphere are exceptional.

Studies of hemisphere differences in normal subjects exploit the restricted patterns of centripetal projection of peripheral sensory systems (both visual and auditory). There is an impressively large literature (reviewed in White, 1969) reporting superior performance in the recognition of words when presentation is to the right visual hemifield rather than to the left: the identification of words is faster and more accurate, and can be accomplished at lower exposure durations, when presentation is well-lateralized in that hemiretina whose direct projection is to the contralateral left hemisphere. Similarly, when competing speech (or speech-like) messages are presented simultaneously to right and left ears, performance is better with material to the right ear, whose strongest projection is to the contralateral hemisphere (Kimura, 1961). This pattern of so-called "lateral asymmetries" in the performance of normal subjects shows remarkable convergence with clinical observations.

A superior performance with verbal material directed to the right visual hemifield (or right ear) does not demand an interpretation in terms of a left hemisphere which is solely responsible for language processing (though it is one which is sometimes made, e.g., Ellis and Shepard, 1974). Indeed, such an interpretation is inconsistent with clinical findings: the isolated right hemisphere shows some capacity for handling language (Zaidel, 1978). In general, it seems appropriate to look to qualitative differences between the left and right hemisphere processors for an account of observed patterns of asymmetry in performance (e.g., Patterson and Bradshaw, 1975; and Springer, 1977, for some discussion). This is to suggest that experimental evaluations, which typically involve quantitative assessments of performance, reveal lateral asymmetries chiefly when different ways of analyzing input result in different processing efficiencies: given the relative success of an extended programme of such research with normal intact subjects, it is apparent that relative efficiency reflects processing mechanism well enough, in the general case.[15]

The literature reporting right visual field advantages in word recognition has restricted attention, almost without exception, to performances with open class forms; such exceptions as exist (e.g., Hines, 1976) appear to involve errors of item selection. Those properties of an open class recognition device which we have been concerned with are exhibited in the recognition of open class forms presented to either the left or the right visual hemifields: recognition is equally frequency-sensitive in the two cases, both in terms of latencies in lexical decision (Leiber, 1976), and in terms of accuracy of report (Axelrod, Haryadi and Lieber, 1977); and initial segments have primacy in identification in the two cases (Bradshaw, Bradley, Gates and Patterson, 1977). That these general features of word recognition are evident in performances with material directed to each of the hemispheres does not, of course, mean that no qualitative differences exist. Indeed, the manipulation of parameters which have their impact at lower levels of input analysis indicate important divergences: for example, a more script-like typeface biasses relative efficiencies in favor of the right hemisphere in letter identification (Bryden and Allard, 1976). But the evidence at hand suggests that more general properties of an open class recognition system are relatively independent of such pre-processing differences

---

[15] This line of argument suggests that tachistoscopic or dichotic studies with normal subjects which fail to reveal asymmetries (e.g., Hardyck, Tzeng and Wang, 1977, 1978) should perhaps be taken as instances of paradigm failure, rather than as assaults on the (clinically well-supported) hypothesis that the hemispheres differ in important ways.

(e.g., Stanners, Yastremski and Westbrook, 1975; Bradley, 1977).

The recognition performances of Broca's aphasics with open and closed class forms suggest that the left hemisphere plays a crucial role in the support of qualitatively distinct recognition devices, in the sense that the integrity of Broca's area (taking this, now to refer to the tissues supplied by the upper division of the middle cerebral artery) is necessary to the operation of a specialized mechanism whose domain is the closed class vocabulary. Conversely, their apparently normal performance with open class forms (setting aside the word-finding problems which characterize most aphasias) denies a particular significance of those tissues to the processes underlying the operation of a more general recognition system.

In Experiment 8, we evaluate recognition performances for items from the two vocabulary classes, with stimulus presentations which are well-lateralized to the right and left visual hemifields, and those procedures which are standard for establishing field differences in recognition performances for open class items. In conjunction with the analyses of field differences which are typical of research in this area, we look to performance differences over open and closed class sets <u>within visual hemifields</u> for evidence of qualitative divergences between the hemispheres.

EXPERIMENT 8

METHOD

SUBJECTS. Twenty adult native speakers of American English took part (successfully, see below) in the experiment, and were paid for their participation. All were students at the Massachusetts Institute of

Technology. By their own report, subjects were familial right-handers, having no left-handed parents, siblings or children.

APPARATUS. A two-field tachistoscope, Gerbrands model T2B-C, was used to present stimuli. The subject's (self-paced) word identifications were recorded by hand; where a spoken form was homophonous, the subject spelled out his response.

MATERIALS. The item set was made of 32 open class words and 32 closed class words, the word sets being as closely matched as possible. For both item sets, frequencies of occurrence (from the count of Kucera and Francis (1967), expressed here as occurrences per million, in units logarithm base 10) ranged from 2.2 to 3.5, with mean frequencies of 2.8 and 2.9 for open and closed class sets, respectively. Items ranged in length from 3 to 5 letters, mean length being 4.0 letters for both item sets. They were selected from among items previously presented in a lexical decision task (Experiments 1 and 2, above), so that ease of identification, in terms of item reaction times in lexical decision, could be equated between sets: mean reaction times being 469 and 474 for open and closed class sets respectively.

Each item was presented twice to each subject, once in each visual hemifield, for a total of 128 experimental trials which were scored for accuracy of word identification. Two separate pseudo-random orders of 64 trials were prepared, under a constraint prohibiting runs of more than three of the same presentation field (left/right) or item type (open class/closed class). The entire item set was assigned to each sequence. Field of presentation, however, was switched for each item between orders.

The stimulus words were typed horizontally, in Letter Gothic capitals, on 4 x 6 white file cards, and positioned so that (with the 57 cm. viewing distance of the tachistoscope), the nearest approach of a word to fixation subtended 2^{o} of visual angle, and the furthest extent (for five letter items) subtended 3^{o}.

PROCEDURE. All subjects were administered 30 practice trials and 128 scored trials (in two blocks of 64 trials each, corresponding to the two pseudo-random sequences); each trial consisted of a laterally presented word, which the subject attempted to identify. Subjects were alerted prior to each trial, and after 500 msec a central fixation cross gave way to a brief presentation (less than 150 msec) of a stimulus card. The importance of an accurate fixation during a trial was emphasized in the practice session, and again with the start of the two experimental sessions.[16]

Subjects were instructed to report, as best they could, the word shown. During the practice session, which also served to familiarize subjects with the task, exposure duration was reduced from an initial 250 msec. to a level at which the subject reliably identified approximately 50% of the presented words. Only subjects who were able to report at this level, with an exposure duration less than 150 msec, were maintained in the experiment; mean exposure duration over the successful 20 subjects was 115 msec.

---

[16] Some investigators (e.g., Hines, 1977) require report of a centrally presented digit, to ensure fixation; only trials where there is prior correct identification of that symbol at fixation are entered into an analysis. However, there are reports (e.g., McKeever, Subari and van Deventer, 1972) that the presence or absence of such a requirement has little or no effect on the experimental evaluation of field differences.

RESULTS AND DISCUSSION

Recognition scores based on correct identifications of stimulus words are presented in Table 10.

TABLE 10

Percent correct identifications for word types, from left (LVF) and right visual hemifield (RVF) presentations.

Word Type			
Closed Class		Open Class	
LVF	RVF	LVF	RVF
41.4	52.3	44.4	65.8

Both vocabulary classes show the standardly reported superior recognition of words presented in the right visual hemifield. Recognition is more accurate for open class forms presented to the right (65.8%) than to the left (44.4%), minF'(1, 100) = 22.98, $p < .01$; similarly, more closed class items are identified in the right (52.3%) than in the left visual hemifield (41.4%), minF'(1,100) = 6.00, $p < .02$. (Tests are based on simple effects derived from two-way analyses of variance, treating subjects/items as repeated measures).

Of more interest here is an interaction of field differences in recognition performance with vocabulary class: by subject analysis, $F(1,38) = 7.18$, $p < .02$; by item analysis, $F(1,62) = 4.45$, $p < .05$. Figure 7 illustrates the contrast between recognition accuracies for the two vocabulary classes, within field of presentation.

INSERT FIGURE 7

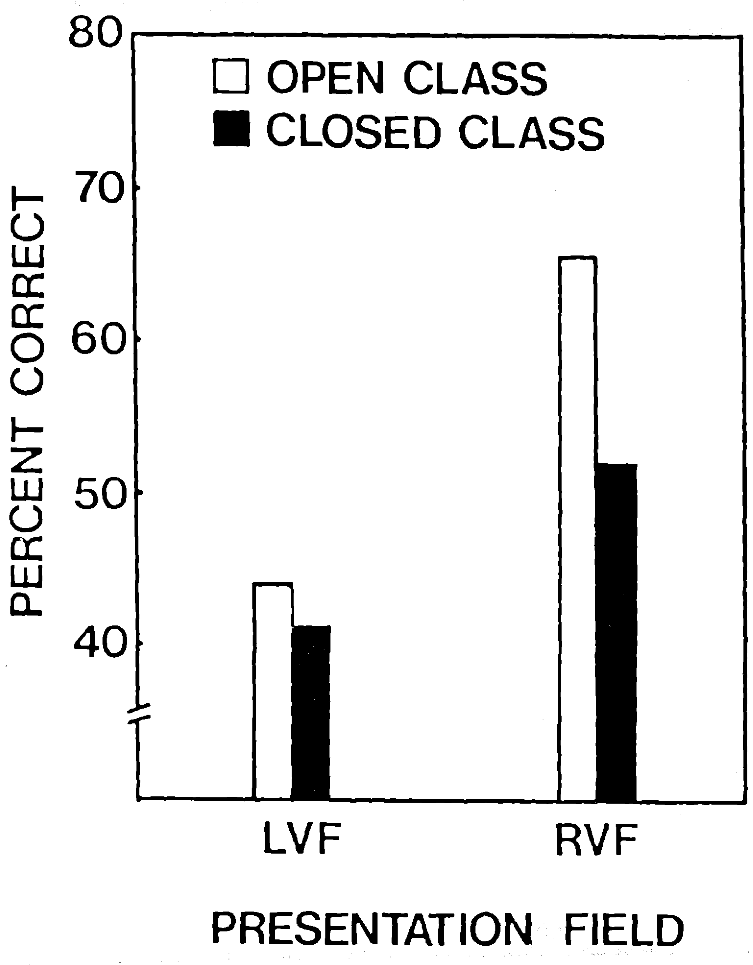


FIGURE : LATERAL ASYMMETRIES IN TACHISTOSCOPIC RECOGNITION

Percent correct identifications for words briefly presented to left and right visual hemifields. Data are for a population of normal speakers.

The interaction can be located in the poorer identification of closed class words (52.3%) relative to open class words (65.8%) in the right visual field, minF'(1,96) = 4.30, $p < .05$; with left visual field presentation, the classes are not distinguished, and the level of identification for closed class forms (41.4%) does not differ from that for open class words (44.4%), minF' < 1. (Tests are based on one-way analyses of covariance, taking as covariate the percent identifications in the opposite field).

The form of the interaction can be summarized as follows: recognition accuracies for closed and open class vocabularies differ when presentation is to the right visual hemifield, in the absence of any such contrast with left visual hemifield presentation. We can speculate that these performance differences reflect a qualitative difference between the left and right hemispheres, in terms of the word recognition devices which are supported: only the left hemisphere supports a specialized retrieval mechanism for the closed class vocabulary. That is, closed class words presented to the right visual hemifield are dealt with by a recognition system which is different from that which operates over the open class vocabulary, while with left visual hemifield stimuli, a single recognition system treats words from both vocabulary classes. This speculation is prompted, in large part, by our previous findings with agrammatic aphasics, who (with a left hemisphere whose integrity is lost) no longer have a specialized closed class recognition system in operation. Clearly, an account along these lines makes a strong claim about the convergence of clinical studies and tachistoscopic studies with normal subjects, and a considerable caution should be exercised. There are many difficulties: on the one hand, the very existence of performance asymmetries in normal intact subjects is somewhat mysterious, given the fact that there are callosal connections between the hemispheres (and a callosal transmission time of the order of 5 milliseconds, Efron (1963)); on the other, there are well-documented (e.g., Gregory, 19) problems in drawing inferences about normal states from abnormal ones.

Granting the (tentative) conclusion that closed and open class forms presented to the right visual hemifield are recognized under separate retrieval systems, we have yet to account for the translation from qualitatively distinct retrieval operations to quantitative differences in recognition accuracy where, moreover, closed class forms are more poorly identified than open class forms. We might point to a general feature of the recognition of closed class words: in many circumstances, these appear to be <u>less available</u> for report than open class forms. Much of the evidence is essentially anecdotal. For example, in sentences where forms are repeated (as in, say, a printer's error), that repetition can be extraordinarily difficult to detect, particularly when the error involves a closed class form. Similarly, when a subject skims over a piece of prose (or even scrambled prose, where structure is obliterated), marking out all occurrences of a given letter, he is much more likely to miss those which occur in closed class forms; so, for example, he will fail to detect instances of the letter <u>t</u> when these occur in the form <u>the</u> (Healy, 1977; Healy and Cutting, 1978). And, in so-called "RSVP" task (Rapid Serial Visual Presentation), where a subject is required to report as many as possible of forms which are presented to him, one at a time, mutually masking, and at a rate that simply overloads the system, report rate for closed class forms is significantly lower than for open class forms, particularly when the serially presented material lacks structure (Garrett and Cutler, 1976). It seems plausible that the fewer correct identifications of closed class forms in the right visual hemifield are another instance of the general phenomenon of relative unavailability of the closed class.[17]

We have suggested that recognition of closed class forms via a specialized retrieval mechanism is to be closely tied to their exploitation as an evidential base for syntactic inference. If the left hemisphere alone supports that recognition device--the strong interpretation of the convergence of clinical studies and the results of Experiment 8--then studies with normal subjects should reveal patterns of performance of the same kind as were found in Experiment 8, where structural variables rather than vocabulary type are manipulated. And this appears to be so. An obvious case, which is very close to the one we have explored, is the enhanced learnability of sequences of nonsense syllables when a minimal basis for structure is permitted by the distribution of closed class forms (e.g. Epstein, 1961). In dichotic listening studies with normal subjects, a superior report of structured nonsense strings (<u>sig the bagee wugs</u>) relative to unstructured strings (<u>wugs bagee sig the</u>) is found only

---

[17] Note that this account requires that unavailability to report be a property of the recognition of closed class forms by a particular retrieval mechanism, rather than a property of closed class forms <u>per se</u>: it is <u>not</u> the case that closed class forms presented in the left visual field are more poorly identified than the corresponding open class forms. This is a claim which is readily tested: agrammatic speakers, who lack a closed class recognition device, should be paradoxically good in reporting closed class forms in, for example, RSVP tasks.

with right ear presentations; with left ear presentations, there is no effect of the structural variable (Zurif and Sait, 1970). Similarly, a manipulation of syntactic complexity has an effect only with right ear presentation so that, for example, progressive tense constructions (e.g., <u>They are fixing benches</u>) which are typically more rapidly verified than adjectival constructions (e.g., <u>They are performing monkeys</u>), fail to exhibit an advantage with (monaural) left ear presentation in some circumstances (Bever, 1971). Other contrasts of structural complexity (e.g. active/passive) provide further examples to the same point (Bever, personal communication). Overall, we see (at least) a convergence between patterns of recognition performance for closed and open class vocabulary items, and effects of structural variables in sentence processing.

EXPERIMENT 1: ITEM DATA FOR OPEN CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E	ITEM	logF	RT	%E
say	3.44	469	0	healthy	2.02	489	0
man	3.33	458	0	thin	2.00	517	3
time	3.28	474	3	wind	1.92	472	0
new	3.23	475	3	sweet	1.90	465	0
year	3.22	454	0	typical	1.81	527	0
state	3.19	454	0	invite	1.72	530	3
go	3.17	472	3	cow	1.66	474	0
give	3.10	478	3	anxiety	1.62	533	0
work	3.07	502	3	desert	1.61	504	0
people	2.96	476	0	ill	1.59	557	3
place	2.92	472	0	monument	1.56	522	0
world	2.92	465	0	dawn	1.47	491	0
life	2.90	484	3	tea	1.45	454	0
turn	2.82	463	0	lamp	1.38	504	3
fact	2.73	468	0	candle	1.36	486	0
general	2.71	497	0	pen	1.30	478	0
night	2.66	459	0	elbow	1.23	525	7
power	2.62	464	3	whistle	1.18	516	0
boy	2.61	463	0	ox	1.12	581	17
national	2.57	519	0	witch	1.11	483	0
white	2.56	498	0	ant	1.11	474	0
action	2.56	465	0	therapy	1.11	624	17
easy	2.46	469	3	pant	1.04	518	0
money	2.42	470	3	addict	0.90	595	13
west	2.37	485	10	*mast	0.90	621	23
education	2.33	518	0	wistful	0.90	650	10
deep	2.30	457	0	thirsty	0.70	490	0
ship	2.20	519	3	almond	0,60	522	0
aid	2.17	517	3	*whim	0.48	660	20
edge	2.10	505	0				
oil	2.08	492	3	MEAN	2.11	497	2.1
balance	2.07	484	3	S.D.	0.78	40	

* Items with error rates greater than 20% are omitted from all calculations.

EXPERIMENT 1: ITEM DATA FOR CLOSED CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E
the	4.84	482	3
of	4.56	493	7
and	4.46	456	7
with	3.88	556	10
his	3.84	445	0
this	3.71	451	0
but	3.64	482	3
from	3.64	500	0
which	3.55	490	0
you	3.52	449	3
all	3.48	456	0
their	3.43	486	3
more	3.35	468	0
what	3.28	474	0
about	3.26	450	0
than	3.25	513	3
then	3.20	484	0
such	3.11	481	3
our	3.10	468	0
over	3.09	461	0
most	3.06	454	0
also	3.03	491	7
through	2.99	486	0
should	2.95	473	0
those	2.93	488	0
very	2.90	439	3
here	2.88	457	0
both	2.86	481	10
between	2.86	470	0
while	2.83	487	3
without	2.77	496	3

ITEM	logF	RT	%E
however	2.74	466	3
every	2.69	473	0
upon	2.69	484	0
almost	2.64	473	0
less	2.64	482	0
yet	2.62	472	0
often	2.57	469	0
among	2.57	469	0
ever	2.54	466	3
although	2.50	496	3
thus	2.49	487	0
whether	2.46	500	0
either	2.45	479	0
*am	2.36	586	23
therefore	2.31	493	0
soon	2.30	459	0
near	2.30	468	3
*nor	2.29	696	53
else	2.25	490	0
inside	2.24	469	0
below	2.16	481	7
none	2.03	484	0
despite	2.02	504	0
moreover	1.94	496	0
ought	1.83	517	0
besides	1.82	512	3
seldom	1.53	488	0
thereby	1.52	519	3
MEAN	2.88	480	1.6
S.D.	0.70	21	

*Items with error rates greater than 20% are omitted from all calculations

EXPERIMENT 2: ITEM DATA FOR OPEN CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E)

ITEM	logF	RT	%E	ITEM	logF	RT	%E
say	3.44	450	0	hot	2.15	468	0
make	3.37	452	0	key	2.09	525	10
man	3.33	454	0	camp	2.05	500	0
time	3.28	433	0	thin	2.00	489	0
year	3.22	472	0	warm	1.96	487	0
come	3.19	447	0	fresh	1.92	494	0
go	3.17	473	0	hate	1.88	515	0
work	3.07	441	0	safe	1.83	468	0
day	3.04	457	0	fat	1.79	527	0
look	3.01	476	0	meal	1.75	470	5
place	2.92	440	0	bag	1.72	516	0
life	2.90	466	0	egg	1.68	516	0
hand	2.88	462	0	fun	1.64	492	5
thing	2.85	479	0	mad	1.60	501	0
want	2.81	471	0	sad	1.56	477	0
end	2.75	456	0	glow	1.52	528	0
plan	2.72	454	0	tent	1.48	501	5
play	2.68	452	5	harm	1.43	523	5
big	2.62	511	5	clue	1.40	503	5
law	2.59	516	5	dwell	1.36	569	10
pay	2.56	460	0	bake	1.32	478	0
talk	2.52	466	0	hut	1.28	527	15
job	2.48	454	0	bug	1.24	521	0
town	2.44	476	0	cave	1.20	489	0
break	2.41	477	0	toy	1.15	502	0
hard	2.37	489	0	chant	1.11	577	25
drink	2.32	456	0	rust	1.08	517	0
wish	2.29	520	0	hike	1.04	561	5
sit	2.20	480	0	dock	1.00	554	0
farm	2.20	480	0				
throw	2.18	531	0	MEAN	2.18	489	1.8
				S.D.	0.70	33	

EXPERIMENT 2: ITEM DATA FOR CLOSED CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E	ITEM	logF	RT	%E
and	4.46	467	5	did	3.02	527	0
that	4.03	450	0	much	2.97	469	0
was	3.99	462	0	each	2.94	484	0
as	3.86	488	0	those	2.93	445	0
this	3.71	471	0	how	2.92	447	0
are	3.64	493	0	both	2.86	488	0
have	3.60	476	0	same	2.84	462	0
were	3.52	490	5	while	2.83	502	0
all	3.48	501	0	since	2.80	492	0
has	3.39	445	0	few	2.78	503	5
when	3.37	491	0	less	2.64	513	0
who	3.35	492	0	yet	2.62	481	0
more	3.35	467	0	why	2.61	434	0
if	3.34	523	0	next	2.60	446	0
what	3.28	476	5	thus	2.49	484	0
than	3.25	496	10	whose	2.40	459	0
could	3.20	496	0	am	2.36	510	0
then	3.14	470	0	soon	2.30	458	0
now	3.12	469	0	else	2.25	484	0
such	3.11	472	0				
most	3.06	472	0	MEAN	3.11	479	0.8
				S.D.	0.50	22	

EXPERIMENT 2: ITEM DATA FOR OPEN CLASS WORDS, SUBJECT REPLICATION. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time for subject group #1 (RT^1, in msec), and mean reaction time for subject group #2 (RT^2, in msec).

ITEM	logF	RT^1	RT^2	ITEM	logF	RT^1	RT^2
say	3.44	458	442	hot	2.15	478	455
make	3.37	447	457	key	2.09	527	540
man	3.33	451	457	camp	2.05	468	516
time	3.28	447	418	thin	2.00	498	479
year	3.22	499	444	warm	1.96	489	485
come	3.19	441	453	fresh	1.92	482	492
go	3.17	479	467	hate	1.88	505	524
work	3.07	433	449	safe	1.83	474	462
day	3.04	474	439	fat	1.79	521	502
look	3.01	480	472	meal	1.75	454	483
place	2.92	469	439	bag	1.72	526	506
life	2.90	482	450	egg	1.68	522	520
hand	2.88	449	474	fun	1.64	487	502
thing	2.85	485	473	mad	1.60	498	528
want	2.81	505	437	sad	1.56	482	473
end	2.75	449	464	glow	1.52	546	511
plan	2.72	456	452	tent	1.48	510	497
play	2.68	419	450	harm	1.43	545	508
big	2.62	495	507	clue	1.40	483	570
law	2.59	559	483	dwell	1.36	599	559
pay	2.56	482	437	bake	1.32	456	501
talk	2.52	457	475	hut	1.28	486	572
job	2.48	463	445	bug	1.24	538	505
town	2.44	459	480	cave	1.20	496	483
break	2.41	479	475	toy	1.15	502	507
hard	2.37	475	504	chant	1.11	683	544
drink	2.32	489	438	rust	1.08	506	523
wish	2.29	525	514	hike	1.04	573	564
sit	2.23	481	480	dock	1.00	569	542
farm	2.20	485	475				
throw	2.18	552	510	MEAN	2.18	494	487
				S.D.	0.70	44	37

EXPERIMENT 2: ITEM DATA FOR CLOSED CLASS WORDS, SUBJECT REPLICATION. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time for subject group #1 (RT^1, in msec), and mean reaction time for subject group #2 (RT^2, in msec).

ITEM	logF	RT^1	RT^2
and	4.46	487	450
that	4.03	459	437
was	3.99	483	441
as	3.86	502	475
this	3.71	481	457
are	3.64	497	488
have	3.60	501	456
were	3.52	491	501
all	3.48	502	501
has	3.39	460	433
when	3.37	508	473
who	3.35	496	489
more	3.35	479	455
if	3.34	520	515
what	3.28	486	450
than	3.25	502	490
could	3.20	503	489
then	3.14	518	440
now	3.12	488	450
such	3.11	506	438
most	3.06	485	456

ITEM	logF	RT^1	RT^2
did	3.02	575	479
much	2.97	450	469
each	2.94	473	474
those	2.93	435	454
how	2.92	452	442
both	2.86	484	493
same	2.84	467	456
while	2.83	497	508
since	2.80	499	485
few	2.78	541	472
less	2.64	541	506
yet	2.62	503	458
why	2.61	456	412
next	2.60	442	444
thus	2.49	506	457
whose	2.40	449	470
am	2.36	527	492
soon	2.30	467	451
else	2.25	497	470
MEAN	3.11	490	467
S.D.	0.50	29	24

EXPERIMENT 3: ITEM DATA FOR MIXED NOUNS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E
price	2.27	477	0
leg	2.11	513	0
yard	2.00	470	0
wind	1.92	475	0
stone	1.87	480	0
joy	1.67	454	0
cup	1.77	457	5
tea	1.43	483	0
coin	1.34	490	0
shelf	1.32	492	10
straw	1.30	508	0
gossip	1.20	545	0
silk	1.18	551	0
toy	1.15	488	0
bean	1.11	547	0
ankle	1.15	499	0
dock	1.00	543	15
wool	1.00	533	5
grief	1.00	516	5
armor	0.95	556	0
loaf	0.90	535	0
mint	0.85	525	5
ore	0.78	578	45
raft	0.70	579	0
weed	0.85	524	0
ulcer	0.70	561	15
keg	0.60	601	30
pastry	0.60	537	0
chalk	0.60	592	0
kennel	0.48	574	0
ale	0.00	618	5
elm	0.60	559	0
calico	0.30	619	10
talon	0.00	608	45
raisin	0.00	544	0
sleet	0.00	578	5
MEAN	1.02	534	5.6
S.D.	0.60	46	

EXPERIMENT 3: ITEM DATA FOR ANIMAL NAMES. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E
horse	2.29	504	5
dog	2.15	469	0
bird	1.96	466	5
bear	1.92	481	0
snake	1.87	454	0
cow	1.66	472	0
cat	1.62	458	0
bee	1.43	510	0
lion	1.41	503	0
sheep	1.36	506	0
mouse	1.30	508	5
rabbit	1.20	475	0
hawk	1.18	535	0
fox	1.15	460	0
deer	1.11	492	0
eagle	1.08	504	0
wolf	1.00	502	5
worm	1.00	506	0
tiger	0.95	481	0
goat	0.90	481	0
goose	0.85	529	0
mule	0.85	536	0
owl	0.78	510	5
mink	0.78	510	5
wasp	0.78	520	5
otter	0.70	615	0
hog	0.70	551	5
pigeon	0.60	514	5
shark	0.60	527	0
beaver	0.48	521	0
ape	0.48	546	0
elk	0.30	544	0
coyote	0.30	558	5
bison	0.00	639	15
walrus	0.00	592	0
skunk	0.00	650	5
MEAN	1.02	518	1.8
S.D.	0.59	47	

EXPERIMENT 3: SUBJECT DATA FOR ANIMAL NAMES AND MIXED NOUNS. Mean reaction time (RT, in msec) as a function of word type.

	ANIMAL NAMES	MIXED NOUNS
Subject #	RT	RT
1	505	523
2	489	512
3	509	523
4	527	538
5	570	603
6	484	511
7	560	574
8	512	551
9	548	524
10	478	511
11	672	650
12	516	559
13	549	558
14	583	602
15	517	559
16	413	463
17	460	485
18	482	477
19	450	463
20	545	572
MEAN	518	538
S.D.	56	49

EXPERIMENT 4: ITEM DATA FOR NONWORDS. Frequency of occurrence of initial string (per million, from Kucera and Francis (1967), expressed as units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

CLOSED CLASS NONWORDS				OPEN CLASS (HIGH) NONWORDS				OPEN CLASS (LOW) NONWORDS				BASELINE NONWORDS		
ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	RT	%E
mostner	3.06	698	5	lostner	2.25	579	0	mistler	1.15	616	10	vostner	576	0
lessipen	2.64	624	5	missalen	2.41	604	0	fussitan	0.60	601	0	ressipen	591	0
whendour	3.37	576	0	turndour	2.37	702	5	shindour	0.48	596	5	treshour	583	0
thanage	3.25	582	0	thinage	1.96	628	15	chopage	0.48	647	5	scutage	576	0
frominy	3.64	592	5	formony	2.57	626	5	drumony	1.04	596	5	prumeny	648	0
fewlet	2.78	585	10	newlet	3.21	613	5	cowlet	1.46	644	20	bewlet	569	0
herege	2.88	686	10	surege	2.42	614	5	marege	1.20	579	0	birege	580	0
thisire	3.71	580	0	factire	2.65	736	10	cornire	1.53	584	25	garnire	611	5
bothrane	2.86	577	0	pathrine	1.66	606	0	mothrane	0.00	586	0	suthrane	574	5
thattace	4.03	599	5	shuttice	1.66	652	0	chattice	0.70	671	25	thiddace	566	0
sucherty	3.11	582	0	worderty	2.44	635	0	casterty	1.65	705	10	nacherty	616	5
yetitude	2.62	536	0	setitude	2.62	602	0	petitude	0.90	648	20	dititude	581	0
forald	3.98	597	5	sayeld	2.70	599	0	vowald	0.00	575	0	lereld	570	0
didrin	3.02	547	0	kidrin	1.79	602	5	saprin	0.00	586	0	picrin	608	5
whatent	3.28	581	0	costent	2.36	726	20	hintent	0.95	627	0	brisent	573	0
notrome	3.66	571	5	lotrume	2.10	533	0	matrume	0.70	625	0	naltome	584	0
MEAN	3.24	595	3.1		2.32	629	4.4		0.82	618	7.8		588	1.3
S.D.	0.45	43			0.41	41			0.51	37			22	

EXPERIMENT 4: SUBJECT DATA FOR NONWORDS. Mean reaction time (RT, in msec) as a function of nonword type.

	CLOSED CLASS NONWORDS	OPEN CLASS (HIGH) NONWORDS	OPEN CLASS (LOW) NONWORDS	BASELINE NONWORDS
Subject #	RT	RT	RT	RT
1	552	595	574	517
2	616	629	642	624
3	559	606	579	574
4	486	524	526	464
5	596	596	600	568
6	558	617	577	599
7	594	654	644	612
8	614	694	630	602
9	648	649	606	575
10	599	608	615	591
11	730	672	688	682
12	515	574	557	506
13	573	599	628	576
14	623	636	592	592
15	514	552	570	550
16	615	698	701	664
17	565	609	604	566
18	641	659	642	609
19	687	727	728	640
20	578	697	656	625
MEAN	593	630	618	587
S.D.	58	52	51	52

EXPERIMENT 4': ITEM DATA FOR WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E). Items marked * are not included in the analysis.

CLOSED CLASS				OPEN CLASS (HIGH)				OPEN CLASS (LOW)			
ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	logF	RT	%E
most	3.06	476	0	lost	2.73	449	0	mist	1.32	546	0
less	2.64	479	5	miss	2.52	480	0	*fuss	1.00	581	20
when	3.37	479	0	turn	2.82	459	10	*shin	0.48	770	60
than	3.25	528	5	thin	2.00	484	0	chop	1.18	535	0
from	3.64	494	0	form	2.77	490	0	drum	1.56	511	10
few	2.78	472	0	new	3.23	470	5	cow	1.66	471	0
here	2.88	482	10	sure	2.42	479	0	*mare	1.28	608	25
this	3.71	505	0	fact	2.73	477	5	corn	1.53	470	0
both	2.86	461	5	path	1.76	508	0	moth	0.60	556	0
that	4.03	489	0	shut	1.69	516	10	chat	0.95	521	5
such	3.11	481	0	*word	2.75	443	0	cast	1.71	532	0
yet	2.62	506	0	set	2.68	477	5	pet	1.15	469	0
for	3.98	460	10	say	3.44	454	0	vow	1.15	600	15
did	3.02	495	15	kid	2.05	511	5	sap	0.78	639	15
what	3.28	458	0	cost	2.62	471	0	hint	1.43	479	0
not	3.66	480	0	lot	2.23	492	0	mat	0.85	594	5
MEAN	3.24	484	3.1		2.51	481	2.7		1.22	533	3.8
S.D.	0.45	19			0.50	20			0.35	54	

EXPERIMENT 5: ITEM DATA FOR NONWORDS. Frequency of occurrence of initial string (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

OPEN CLASS ACTIVE NONWORDS				OPEN CLASS INACTIVE NONWORDS				BASELINE NONWORDS		
ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	RT	%E
checkose	1.94	610	0	reachose	2.03	562	0	nerchose	566	0
chessiate	0.48	615	0	thiefiate	0.90	637	5	chirdiate	562	5
drumacle	1.04	574	0	dawnicle	1.45	579	5	drimacle	567	0
folkard	1.53	601	0	riskard	1.73	583	5	nslkard	550	0
postilude	1.92	660	10	lostitude	2.24	642	0	tostilude	549	5
handure	2.63	574	0	helpure	2.49	605	0	hesture	593	0
dustarn	1.84	597	5	plotarn	1.57	606	5	blitarn	561	0
farmesty	2.10	603	5	foamesty	1.57	598	0	feamesty	572	5
railont	1.45	609	0	loadant	1.65	591	0	daulent	568	0
raintice	1.94	560	0	grintice	1.72	558	0	bointice	565	0
bathect	1.41	559	0	auntect	1.34	578	0	peshect	563	5
flagify	1.20	591	0	gripify	1.30	605	0	tregify	544	0
earilege	1.46	565	0	eatirege	1.79	584	5	eaderage	573	5
airlat	2.41	621	0	aidret	2.11	560	0	aitlet	551	0
dayrin	2.84	576	0	sayrin	2.70	547	0	tayrin	527	0
guneld	2.07	596	5	jobeld	2.38	596	0	pobeld	563	0
MEAN	1.77	594	1.6		1.81	589	1.6		561	1.6
S.D.	0.60	27			0.48	27			15	

EXPERIMENT 5: SUBJECT DATA FOR NONWORDS. Mean reaction time (RT, in msec) as a function of nonword type.

Subject #	OPEN CLASS ACTIVE NONWORDS RT	OPEN CLASS INACTIVE NONWORDS RT	BASELINE NONWORDS RT
1	544	559	565
2	493	490	480
3	618	618	607
4	571	549	533
5	580	556	555
6	477	466	473
7	671	645	651
8	731	772	681
9	584	590	520
10	551	575	512
11	600	580	585
12	589	576	525
13	506	500	457
14	667	652	602
15	643	640	615
16	573	600	548
17	643	609	593
18	722	701	700
19	590	567	531
20	544	534	493
MEAN	595	589	561
S.D.	69	72	68

EXPERIMENT 6: ITEM DATA FOR OPEN CLASS WORDS, NORMAL PATIENTS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E)

ITEM	logF	RT	%E	ITEM	logF	RT	%E
say	3.44	765	0	healthy	2.02	768	0
man	3.33	829	0	thin	2.00	856	0
time	3.28	815	0	wind	1.92	746	0
new	3.23	854	0	sweet	1.90	688	0
year	3.22	755	20	typical	1.81	822	0
state	3.19	751	20	invite	1.72	744	0
give	3.10	792	0	cow	1.66	798	0
work	3.07	712	0	anxiety	1.62	834	20
people	2.96	748	0	desert	1.61	759	0
place	2.92	802	0	ill	1.59	884	0
world	2.92	738	0	monument	1.56	866	0
life	2.90	801	0	dawn	1.49	813	0
turn	2.82	839	0	tea	1.45	770	0
fact	2.73	795	0	lamp	1.38	859	0
general	2.71	773	0	candle	1.36	800	0
night	2.66	704	0	pen	1.30	770	0
power	2.62	752	0	elbow	1.23	836	0
boy	2.61	678	0	whistle	1.18	858	0
national	2.57	846	0	ox	1.12	1013	0
white	2.56	690	0	witch	1.11	704	0
action	2.56	798	0	ant	1.11	858	20
easy	2.46	776	0	therapy	1.11	839	0
money	2.42	735	0	addict	0.90	848	0
west	2.37	749	0	mast	0.90	815	20
education	2.33	869	20	wistful	0.78	877	20
deep	2.30	708	0	thirsty	0.70	795	0
ship	2.20	806	0	almond	0.60	782	0
aid	2.17	831	0	whim	0.48	944	20
edge	2.10	849	0	pant	1.04	810	0
oil	2.08	738	0				
balance	2.07	775	0	MEAN	2.11	797	2.6
				S.D.	0.78	63	

EXPERIMENT 6: ITEM DATA FOR CLOSED CLASS WORDS, NORMAL PATIENTS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E
the	4.84	799	0
of	4.56	902	20
and	4.46	766	0
with	3.88	770	0
his	3.84	793	0
this	3.71	772	0
but	3.64	915	0
from	3.64	711	0
which	3.55	954	0
you	3.52	731	0
all	3.48	762	0
their	3.43	752	0
more	3.35	789	0
what	3.28	729	0
about	3.26	722	0
than	3.25	810	0
then	3.20	807	0
such	3.11	845	0
our	3.10	772	0
over	3.09	834	0
most	3.06	861	0
also	3.03	826	20
through	2.99	860	20
should	2.95	788	0
those	2.93	732	0
very	2.90	816	0
here	2.88	827	0
both	2.86	752	20
between	2.86	853	0
while	2.83	743	0
without	2.77	728	0

ITEM	logF	RT	%E
however	2.74	783	0
every	2.69	674	0
upon	2.69	780	0
almost	2.64	751	20
less	2.64	708	0
yet	2.62	783	0
often	2.57	734	0
among	2.57	751	0
ever	2.54	730	0
although	2.50	858	0
thus	2.49	859	0
whether	2.46	744	0
either	2.45	781	0
am	2.36	765	0
therefore	2.31	803	0
soon	2.30	615	0
near	2.30	734	0
nor	2.29	907	20
else	2.25	779	20
inside	2.24	705	0
below	2.16	697	0
none	2.03	788	0
despite	2.02	754	0
moreover	1.94	876	0
ought	1.83	894	0
besides	1.82	852	0
seldom	1.53	809	0
thereby	1.52	862	0
MEAN	2.86	788	2.4
S.D.	0.69	65	

EXPERIMENT 6: ITEM DATA FOR OPEN CLASS WORDS, APHASIC PATIENTS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E	ITEM	logF	RT	%E
say	3.44	903	0	healthy	2.02	945	0
man	3.33	942	0	thin	2.00	1124	0
time	3.28	808	0	wind	1.92	767	0
new	3.23	877	0	sweet	1.90	874	0
year	3.22	866	0	typical	1.81	1080	0
state	3.19	848	0	invite	1.72	980	20
go	3.17	865	20	cow	1.66	877	0
give	3.10	903	20	anxiety	1.62	986	20
work	3.07	1015	0	desert	1.61	964	0
people	2.96	809	0	ill	1.59	972	0
place	2.92	919	0	monument	1.56	921	0
world	2.92	907	0	dawn	1.49	1139	0
life	2.90	945	0	tea	1.45	992	0
turn	2.82	837	0	lamp	1.38	838	0
fact	2.73	1109	0	candle	1.36	1080	0
general	2.71	980	0	pen	1.30	801	0
night	2.66	866	0	elbow	1.23	1017	20
power	2.62	908	0	whistle	1.18	937	0
boy	2.61	873	20	ox	1.12	1034	20
national	2.57	1019	0	witch	1.11	924	0
white	2.56	984	0	ant	1.11	910	20
action	2.56	989	0	therapy	1.11	896	20
easy	2.46	934	0	pant	1.04	950	0
money	2.42	928	0	addict	0.90	896	20
west	2.37	1038	0	mast	0.90	1177	0
education	2.33	905	0	wistful	0.78	941	20
deep	2.30	845	0	thirsty	0.70	1037	0
ship	2.20	939	0	almond	0.60	1011	20
aid	2.17	1077	0	whim	0.48	1038	40
edge	2.10	969	0				
oil	2.08	954	0	MEAN	2.11	951	4.6
balance	2.07	982	0	S.D.	0.78	88	

EXPERIMENT 6: ITEM DATA FOR CLOSED CLASS WORDS, APHASIC PATIENTS. Frequency of occurrence (per million, from Kucera and Francis (1967) expressed in units logarithm base 10), mean reaction time (RT, in msec), and percentage error (%E).

ITEM	logF	RT	%E	ITEM	logF	RT	%E
the	4.84	873	20	however	2.74	956	0
of	4.56	1084	20	every	2.69	855	0
and	4.46	902	0	upon	2.69	840	40
with	3.88	845	0	almost	2.64	989	0
his	3.84	896	0	less	2.64	812	20
this	3.71	910	0	yet	2.62	1066	20
but	3.64	809	0	often	2.57	981	0
from	3.64	1048	0	among	2.57	939	0
which	3.55	889	0	ever	2.54	920	20
you	3.52	865	0	although	2.50	998	20
all	3.48	766	0	thus	2.49	1082	20
their	3.43	907	0	whether	2.46	1015	0
more	3.35	1092	0	either	2.45	958	0
what	3.28	900	0	am	2.36	987	0
about	3.26	879	0	therefore	2.31	1028	20
than	3.25	924	0	soon	2.30	878	0
then	3.20	901	0	near	2.30	979	0
such	3.11	853	0	nor	2.29	965	60
our	3.10	944	0	else	2.25	920	40
over	3.09	1004	0	inside	2.24	989	0
most	3.06	1007	0	below	2.16	988	0
also	3.03	885	20	none	2.03	no data	
through	2.99	957	0	despite	2.02	1032	40
should	2.95	881	0	moreover	1.94	952	20
those	2.93	866	0	ought	1.83	1094	20
very	2.90	901	0	besides	1.82	910	0
here	2.88	986	0	seldom	1.53	1098	20
both	2.86	904	0	thereby	1.52	1111	0
between	2.86	874	0				
while	2.83	1071	0	MEAN	2.88	946	7.6
without	2.77	902	0	S.D.	0.70	82	

EXPERIMENT 7: ITEM DATA FOR NONWORDS, NORMAL PATIENTS. Frequency of occurrence of initial string (per million, from Kucera and Francis (1967), expressed as units logarithm base 10), mean reaction time (RT, in msec) and percentage error (%E).

CLOSED CLASS NONWORDS				OPEN CLASS (HIGH) NONWORDS				OPEN CLASS (LOW) NONWORDS				BASELINE NONWORDS		
ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	RT	%E
mostner	3.06	1017	0	lostner	2.25	1039	0	mistler	1.15	964	0	vostner	877	0
lessipen	2.64	858	0	missalen	2.41	1008	0	fussitan	0.60	912	0	ressipen	1039	0
whendour	3.37	905	0	turndour	2.37	1118	20	shindour	0.48	971	0	treshour	984	0
thanage	3.25	916	0	thinage	1.96	970	20	chopage	0.48	1016	0	scutage	849	0
frominy	3.64	997	0	formony	2.57	1054	20	drumony	1.04	1093	0	prumeny	1024	0
fewlet	2.78	1019	0	newlet	3.21	1004	0	cowlet	1.46	1136	0	bewlet	947	0
herege	2.88	969	0	surege	2.42	1063	0	marege	1.20	975	0	birege	900	0
thisire	3.71	878	0	factire	2.65	1002	0	cornire	1.53	959	20	garnire	1019	0
bothrane	2.86	920	0	pathrine	1.66	1041	0	mothrane	0.00	1116	0	suthrane	998	0
thattace	4.03	894	0	shuttice	1.66	1063	0	chattice	0.70	1057	20	thiddace	883	0
sucherty	3.11	943	0	worderty	2.44	985	0	casterty	1.65	1103	0	nacherty	970	0
yetitude	2.62	1024	0	setitude	2.62	1082	0	petitude	0.90	1029	20	dititude	1021	0
forald	3.98	977	0	sayeld	2.70	928	0	vowald	0.00	925	0	lereld	897	0
didrin	3.02	878	0	kidrin	1.79	1118	0	saprin	0.00	938	20	picrin	887	0
whatent	3.28	1108	0	costent	2.36	938	20	hintent	0.95	1022	0	brisent	1047	0
notrome	3.66	951	40	lotrume	2.10	966	0	matrume	0.70	1149	0	naltome	905	0
MEAN	3.24	953			2.32	1024			0.82	1023			953	
S.D.	0.45	68			0.41	58			0.51	78			67	

EXPERIMENT 7: SUBJECT DATA FOR NONWORDS, NORMAL PATIENTS. Mean reaction time (RT, in msec) as a function of nonword type.

Subject #	CLOSED CLASS NONWORDS RT	OPEN CLASS (HIGH) NONWORDS RT	OPEN CLASS (LOW) NONWORDS RT	BASELINE NONWORDS RT
1	1111	1164	1216	1121
2	680	780	754	720
3	905	933	949	924
4	852	904	874	850
5	1219	1338	1322	1148
MEAN	953	1024	1023	953
S.D.	214	224	238	182

EXPERIMENT 7: ITEM DATA FOR NONWORDS, APHASIC PATIENTS. Frequency of occurrence of initial string (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), mean reaction time (RT, in msec) and percentage error (%E)

CLOSED CLASS NONWORDS				OPEN CLASS (HIGH) NONWORDS				OPEN CLASS (LOW) NONWORDS				BASELINE NONWORDS		
ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	logF	RT	%E	ITEM	RT	%E
mostner	3.06	1193	60	lostner	2.25	1146	20	mistler	1.15	1201	40	vostner	984	0
lessipen	2.64	1343	20	missalen	2.41	1067	0	fussitan	0.60	1131	0	ressipen	1104	20
whendour	3.37	1291	0	turndour	2.37	1411	40	shindour	0.48	1002	0	treshour	1139	0
thanage	3.25	1074	0	thinage	1.96	1266	20	chopage	0.48	1292	0	scutage	1062	0
frominy	3.64	1387	0	formony	2.57	1154	20	drumony	1.04	1419	20	prumeny	1168	40
fewlet	2.78	1256	0	newlet	3.21	1403	20	cowlet	1.46	1270	80	bewlet	1216	0
herege	2.88	1134	60	surege	2.42	1073	20	marege	1.20	1114	20	birege	1011	0
thisire	3.71	1311	40	factire	2.65	1307	40	cornire	1.53	1164	20	garnire	1286	40
bothrane	2.86	1170	0	pathrine	1.66	1436	0	mothrane	0.00	1048	0	suthrane	1181	20
thattace	4.03	1143	20	shuttice	1.66	1173	0	chattice	0.70	1268	40	thiddace	988	20
sucherty	3.11	1208	0	worderty	2.44	1278	0	casterty	1.65	1308	20	nacherty	993	0
yetitude	2.62	971	0	setitude	2.62	1088	0	petitude	0.90	1249	60	dititude	1221	40
forald	3.98	1302	20	sayeld	2.70	977	20	vowald	0.00	967	0	lereld	1050	20
didrin	3.02	1222	0	kidrin	1.79	1199	0	saprin	0.00	1146	40	picrin	1202	0
whatent	3.28	1224	40	costent	2.36	1167	40	hintent	0.95	1256	20	brisent	1145	20
notrome	3.66	1021	0	lotrume	2.10	1156	20	matrume	0.70	1252	20	naltome	1025	0
MEAN	3.24	1203			2.32	1206			0.82	1193			1112	
S.D.	0.45	115			0.41	134			0.51	121			97	

EXPERIMENT 7: SUBJECT DATA FOR NONWORDS, APHASIC PATIENTS. Mean reaction time (RT, in msec) as a function of nonword type.

Subject	CLOSED CLASS NONWORDS RT	OPEN CLASS (HIGH) NONWORDS RT	OPEN CLASS (LOW) NONWORDS RT	BASELINE NONWORDS RT
1	1019	1124	1004	1009
2	1468	1338	1387	1223
3	1256	1237	1259	1181
4	1273	1214	1181	1177
5	1000	1123	1134	967
MEAN	1203	1206	1193	1111
S.D.	196	88	143	115

EXPERIMENT 8: ITEM DATA FOR CLOSED CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), and correct recognition percentages, for left (LVF) and right visual hemifield (RVF) presentations.

ITEM	logF	LVF	RVF
all	3.48	90	100
has	3.39	35	35
when	3.37	35	55
more	3.35	40	50
what	3.28	55	50
than	3.25	15	30
could	3.20	60	75
then	3.14	35	65
now	3.12	10	30
such	3.11	10	20
most	3.06	45	65
also	3.03	95	80
much	2.97	5	45
each	2.94	35	75
those	2.93	25	45
how	2.92	40	25
very	2.90	35	60
here	2.88	20	45
both	2.86	25	55
same	2.84	65	65
while	2.83	60	50
since	2.80	10	30
few	2.78	60	70
every	2.69	75	50
less	2.64	70	95
yet	2.62	65	65
next	2.60	5	25
ever	2.54	30	40
thus	2.49	5	5
whose	2.40	50	10
soon	2.30	25	90
else	2.25	95	75
MEAN	2.90	41	52
S.D.	0.32	27	24

EXPERIMENT 8: ITEM DATA FOR OPEN CLASS WORDS. Frequency of occurrence (per million, from Kucera and Francis (1967), expressed in units logarithm base 10), and correct recognition percentages, for left (LVF) and right visual hemifield (RVF) presentations.

ITEM	logF	LVF	RVF
say	3.44	80	90
make	3.37	10	50
man	3.33	25	50
time	3.28	75	85
new	3.23	35	20
year	3.22	35	75
state	3.19	40	90
give	3.10	40	50
work	3.07	50	60
look	3.01	65	95
place	2.92	65	90
world	2.92	35	15
life	2.90	65	85
hand	2.88	40	45
thing	2.85	20	65
turn	2.82	25	65
end	2.75	90	90
want	2.81	20	45
fact	2.73	35	80
plan	2.72	35	45
play	2.68	80	95
night	2.66	15	35
law	2.59	75	100
talk	2.52	60	100
job	2.48	75	90
easy	2.46	65	95
town	2.44	5	40
break	2.41	20	75
hard	2.37	20	20
deep	2.30	40	40
wish	2.29	35	40
farm	2.20	45	80
MEAN	2.81	44	66
S.D.	0.35	23	26

EXPERIMENT 8: SUBJECT DATA. Percentage of open and closed class words correctly recognized, for left (LVF) and right visual hemifield (RVF) presentations.

	OPEN CLASS		CLOSED CLASS	
Subject #	LVF	RVF	LVF	RVF
1	56	72	38	50
2	62	69	47	56
3	50	75	72	75
4	44	66	47	59
5	22	53	38	50
6	31	59	53	56
7	53	62	47	31
8	25	72	34	72
9	47	78	28	34
10	56	62	38	38
11	28	66	19	56
12	53	75	47	44
13	47	66	50	44
14	41	50	13	38
15	56	69	47	72
16	44	59	34	50
17	59	81	62	69
18	41	66	47	66
19	34	59	34	56
20	38	56	34	31
MEAN	44.4	65.8	41.4	52.3
S.D.	11.7	8.3	13.6	13.9

REFERENCES

ANDREEWSKY, A. & SERON, X. Implicit processing of grammatical rules in a classical case of agrammatism. Cortex, 1975, 11, 379-390.

AXELROD, S., HARYADI, T. & LEIBER, L. Oral report of words and word approximations presented to the left or right visual field. Brain and Language, 1977, 4, 550-557.

BATTIG, W. F. & MONTAGUE, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, 1969, 80, 1-46.

BEDNALL, E. S. The interaction of frequency and semantic relatedness in lexical access. Unpublished paper, Monash University, 1976.

BEVER, T. G. The nature of cerebral dominance in the speech behaviour of the child and adult. In R. Huxley and E. Ingram (Eds.) Language Acquisition: Models and Methods. New York: Academic Press, 1971.

BRADLEY, D. C. Frequency effects in lexical decision and stimulus quality. Unpublished manuscript, M.I.T., 1977.

BRADSHAW, J. L., BRADLEY, D., GATES, A. & PATTERSON, K. Serial, parallel, or holistic identification of single words in the two visual fields? Perception and Psychophysics, 1977, 21, 431-438.

BROADBENT, D. E. The role of auditory localization in attenuation and memory span. Journal of Experimental Psychology, 1954, 47, 191-196.

BROADBENT, D. E. Word-frequency effect and response bias. Psychological Review, 1967, 74, 1-15.

BROWN, R. LInguistic determinism and parts of speech. Journal of Abnormal and Social Psychology, 1957, 55, 1-5.

BRYDEN, M. P. & ALLARD, F. Visual hemifield differences depend on typeface. Brain and Language, 1976, 3, 191-200.

CAPLAN, D., HOLMES, J. M. & MARSHALL, J. C. Word classes and hemispheric specialization. Neuropsychologia, 1974, 12, 331-337.

CARAMAZZA, A. & ZURIF, E. G. Dissociation of algorithmic and neuristic processes in language comprehension: Evidence from aphasia. Brain and Language, 1976, 3, 572-582.

CAREY, S. The child as word learner. In J. Bresnan, G. Miller and M. Halle (Eds.) Linguistic Theory and Psychological Reality. Cambridge: M.I.T. Press, 1978.

CHAMBERS, S. M. & FORSTER, K. I. Evidence for lexical access in a simultaneous matching task. Memory and Cognition, 1975, 3, 549-559.

COLTHEART, M., DAVELAAR, E., JONASSON, J. T. & BESNER, D. Access to the internal lexicon. In S. Dornic (Ed.) Attention and Performance VI. New Jersey: Lawrence Erlbaum Associates, 1977.

DUNNETT, C. W. A multiple comparison procedure for comparing several treatments with a control. Journal of the American Statistical Association, 1955, 50, 1096-1121.

EFRON, R. The effect of handeness on the perception of simultaneity and temporal order. Brain, 1963, 86, 261-284.

ELLIS, H. D. & SHEPARD, J. W. Recognition of abstract and concrete words presented in the right and left visual fields. Journal of Experimental Psychology, 1974, 103, 1035-1036.

EPSTEIN, W. The influence of syntactical structure on learning. American Journal of Psychology, 1961, 74, 80-85.

ERICKSEN, G. A. & ERICKSEN, C. W. The importance of being first: A tachistoscopic study of the contribution of each letter to the recognitxion of four-letter words. Perception and Psychophysics, 1974, 15, 66-72.

FISCHLER, I. Semantic facilitation without association in a lexical decision task. Memory and Cognition, 1977, 5, 335-339.

FORSTER, K. I. Separation of effects of meaning on word recognition and sentence processing. Paper presented at the First Melbourne Experimental Psychology Conference, September, 1974.

FORSTER, K. I. Accessing the mental lexicon. In R. J. Wales and E. Walker (Eds.) New Approaches to Language Mechanisms. Amsterdam: North Holland, 1976.

FORSTER, K. I. Levels of processing and the structure of the language processor. In W. E. Cooper and E. Walker (Eds.) Sentence Processing: Psycholinguistic Studies presented to Merrill Garrett. Forthcoming.

FORSTER, K. I. & CHAMBERS, S. M. Lexical access and naming time. Journal of Verbal Learning and Verbal Behaviour, 1973, 12, 627-635.

FREDERIKSEN, J. R. & KROLL, J. F. Spelling and sound: Approaches to the internal lexicon. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 361-379.

FRIES, C. C. The Structure of English. New York: Harcourt, Brace & World, Inc., 1952.

FRIES, C. C. Linguistics and Reading. New York: Holt, Rinehart & Winston, 1962.

GARRETT, M. The analysis of sentence production. In G. Bower (Ed.) The Psychology of Learning and Motivation: Advance in Research and Theory, vol. 91. New York: Academic Press, 1975.

GARRETT, M. Syntactic processes in sentence production. In E. Walker and R. Wales (Eds.) New Approaches to Language Mechanisms. Amsterdam: North Holland, 1976.

GARRETT, M. F. & CUTLER, E. A. Report levels in RSVP: open vs. closed class forms. Unpublished M.I.T. manuscript, 1975.

GLANZER, M. Grammatical category: A rote learning and word association analysis. Journal of Verbal Learning and Verbal Behaviour, 1962, 1, 31-41.

GOODGLASS, H. Agrammatism. In H. Whitaker and H. A. Whitaker (Eds.) Studies in Neurolinguistics, Vol. 1. New York: Academic Press, 1976.

GOODGLASS, H., FODOR, I. & SCHULHOFF, C. Prosodic factors in grammar: Evidence from aphasia. Journal of Speech and Hearing Research, 1967, 10, 5-20.

GOODGLASS, H. & KAPLAN, E. The Assessment of Aphasia and Related Disorders. Philadelphia: Lea & Febiger, 1972.

HARDYCK, C., TZENG, O. J. L., & WANG, W. S.-Y. Cerebral lateralization effects in visual half-field experiments. Nature, 1977, 269, 705-707.

HARDYCK, C., TZENG, O. J. L., & WANG, W. S.-Y. Cerebral lateralization of function and bilingual decision processes: Is thinking lateralized? Brain and Language, 1978, 5, 56-71.

HEALY, A. F. Detection errors on the word the: Evidence for reading units larger than letters. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 235-242.

HINES, D. Differences in tachistoscopic recognition between abstract and concrete words as a function of visual half-field and frequency. Cortex, 1977, 13, 66-73.

HOWES, D. A word count of spoken English. Journal of Verbal Learning and Verbal Behaviour, 1966, 5, 572-606.

HOWES, D. & GESCHWIND, N. Quantitative studies of aphasic language. In D. Rioch and E. A. Weinstein (Eds.) Disorders of Communication. London: Churchill, 1964.

ISSERLIN, M. Uber Agrammatismus. Zeitschrift für die Gesamte Neurologie und Psychiatrie, 1922, 75, 332-416.

JOHNSON, S. C. Hierarchical clustering schemes. Psychometrika, 1967, 32, 241-254.

KATZ, B., BAKER, E. & McNAMARA, J. What's in a name? On the child's acquisition and development of proper and common nouns. Child Development, 1974, 45, 269-273.

KEAN, M.-L. The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. Cognition, 1977, 5, 9-46.

KIMURA, D. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 1961, 15, 166-171.

KUCERA, H. & FRANCIS, W. N. Computational Analysis of Present-Day American English. Providence, R. I.: Brown University Press, 1967.

LANDAUER, T. K. & STREETER, L. A. Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. Journal of Verbal Learning and Verbal Behaviour, 1973, 12, 119-131.

LEIBER, L. Lexical decisions in the right and left cerebral hemispheres. Brain and Language, 1976, 3, 443-450.

LEVELT, W. J. M. Hierarchical clustering algorithm in the psychology of grammar. In G. B. Flores and W. J. M. Levelt (Eds.) Advances in Psycholinguistics. Amsterdam: North Holland, 1970.

MARSLEN-WILSON, W. Sentence perception as an interactive parallel process. Science, 1975, 189, 226-228.

MARSLEN-WILSON, W. D. & WELSH, A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology, 1978, 10, 29-63.

McKEEVER, W. F., SUBARI, M. & VAN DEVENTER, A. D. Fixation control in tachistoscopic studies of laterality effects: Comments and data relevant to Hines' experiment. Cortex, 1972, 8, 473-479.

MEYER, D. E. & SCHVANEVELDT, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal of Experimental Psychology, 1971, 90, 227-234.

MEYER, D. E., SCHVANEVELDT, R. W. & RUDDY, M. G. Activation of lexical memory. Paper presented at the meeting of the Psychonomic Society, St. Louis, November, 1972.

MEYER, D. E., SCHVANEVELDT, R. W. & RUDDY, M. G. Functions of graphemic and phonemic codes in visual word-recognition. Memory and Cognition, 1974, 2, 309-321.

MEYER, D. E., SCHVANEVELDT, R. W. & RUDDY, M. G. Loci of contextual effects in visual word recognition. In P. M. A. Rabbitt and S. Dornic (Eds.) Attention and Performance V. New York: Academic Press, 1975.

MILLER, G. A., NEWMAN, E. B. & FRIEDMAN, E. A. Length-frequency statistics for written English. Information and Control, 1958, 1, 370-389.

MISHKIN, M. & FORGAYS, D. G. Word recognition as a function of retinal locus. Journal of Experimental Psychology, 1952, 43, 43-48.

MOHR, J. P. Broca's area and Broca's aphasia. In H. Whitaker and H. A. Whitaker (Eds.) Studies in Neurolinguistics, Vol. 1. New York: Academic Press, 1976.

MORTON, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

MORTON, J. A functional model for memory. In D. A. Norman (Ed.) Models of Human Memory. New York: Academic Press, 1970.

O'CONNELL, D. C., TURNER, E. A. & ONUSKA, L. A. Intonation, grammatical structure and contextual association in immediate recall. Journal of Verbal Learning and Verbal Behaviour, 1968, 7, 110-116.

OLDFIELD, R. C. Things, words and the brain. Quarterly Journal of Experimental Psychology, 1966, 340-353.

PATTERSON, K. & BRADSHAW, J. L. Differential hemispheric mediation of non-verbal stimuli. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 246-252.

RUBENSTEIN, H. R., LEWIS, S. S. & RUBENSTEIN, M. A. Evidence for phonemic recoding in visual word recognition. Journal of Verbal Learning and Verbal Behaviour, 1971, 10, 645-657.

SAVIN, H. B. Word-frequency effect and errors in the perception of speech. Journal of the Acoustical Society of America, 1963, 35, 200-206.

SCHVANEVELDT, R. W. & MEYER, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.) Attention and Performance IV. New York: Academic Press, 1973.

SPRINGER, S. P. Tachistoscopic and dichotic listening investigations of laterality in normal human subjects. In S. Harnad (Ed.) Lateralization in the Nervous System. New York: Academic Press, 1977.

STANNERS, R. F., YASTREMSKI, J. E. & WESTBROOK, A. Frequency and visual quality in a word-nonword classification task. Journal of Verbal Learning and Verbal Behaviour, 1975, 14, 259-264.

TAFT, M. Morphological and syllabic analysis in word recognition. Unpublished Ph. D. thesis, Monash University, 1976.

TAFT, M. Recognition of words via an orthographic access code: the Basic Orthographaphic Syllabic Structure (BOSS). Forthcoming.

TAFT, M. & FORSTER, K. I. Lexical storage and retrieval of prefixed words. Journal of Verbal Learning and Verbal Behaviour, 1975, 14, 638-647.

TAFT, M. & FORSTER, K. I. Lexical storage and retrieval of polymorphic and polysyllabic words. Journal of Verbal Learning and Verbal Behaviour, 1976, 15, 607-620.

THORNE, J. P., BRATLEY, P. & DEWAR, H. The syntactic analysis of English by machine. In D. Michie (Ed.) Machine Intelligence, vol. 3. New York: American Elsevier, 1968.

TWEEDY, J. R., LAPINSKI, R. H. & SCHVANEVELDT, R. W. Semantic-context effects on word recognition: Influence of varying the proportion of items presented in an appropriate context. Memory and Cognition, 1977, 5, 84-89.

VON STOCKERT, T. R. & BADER, L. Some relations of grammar and lexicon in aphasia. Cortex, 1976, 12, 49-60.

WHALEY, C. P. Word-nonword classification time. Journal of Verbal Learning and Verbal Behaviour, 1978, 17, 143-154.

WHITE, M. Laterality differences in perception: a review. Psychological Review, 1969, 72, 387-405.

ZANGWILL, O. L. Dyslexia in relation to cerebral dominance. In J. Money (Ed.) Reading Disability: progress and research needs in dyslexia. Baltimore: The Johns Hopkins Press, 1962.

ZURIF, E. B. & CARAMAZZA, A. Psycholinguistic structures in aphasia: Studies in syntax and semantics. In H. Whitaker and H. A. Whitaker (Eds.) Studies in Neurolinguistics, Vol. 1. New York: Academic Press, 1976.

ZURIF, E., CARAMAZZA, A. & MYERSON, R. Grammatical judgements of agrammatic patients. Neuropsychologia, 1972, 10, 405-417.

ZURIF, E. B. & SAIT, P. E. The role of syntax in dichotic listening. Neuropsychologia, 1970, 8, 239, 243.