

**Lexical access and frequency sensitivity:
Frequency saturation and open/closed class
equivalence***

BARRY GORDON
ALFONSO CARAMAZZA
The Johns Hopkins University

Abstract

Whether closed-class words use the same lexical access route(s) as open-class words has been intensely debated recently. Differences in frequency sensitivity have been suggested as one manifestation of separable access routes. However, all the lexical decision studies have been limited by floor effects at the higher frequency ranges which could mask theoretically important differences in the behavior of the two classes. We studied lexical decisions to high- and very-high frequency words of both classes using stimulus masking and speeded responses, in order to minimize floor effects, to try to reveal potential differences between the behavior of the two classes, and to contrast theories of lexical access.

We did not find evidence to support the view that closed-class words have a different or special access route. Neither word class showed any appreciable frequency effect for Kučera-Francis frequencies of 400/million or greater, on either reaction time or error analyses. We did find open-class words to have somewhat faster responses than comparable closed-class words, but this may contradict some explanations of the reported word class effect (Bradley et al., 1980). Moreover, our data also show what may be word-specific influences on lexical decision times—effects which may be impossible to factor out of the word class effect in English.

In order to accommodate the frequency insensitivity that we found, logogen-based models of lexical access have to be amended to include a floor on threshold settings. Resonance models (Gordon, 1983), already predict this frequency

*This study was supported by NINCDS Teacher-Investigator Award 1-KO7 NS00721 to the first author and by NINCDS grant NS-14099 to The Johns Hopkins University. The authors thank Dr. Steven G. Lapointe, Barbara Mroz, Joan Sobkov, and Kate Van Dyke for help in creating the stimuli, Andrea Berger, Russ Carstens, and Joan Sobkov for testing subjects, Clara Marin and Antoinette Reed for help with the analysis, and Renee Gordon for editorial refinements. Dr. David Caplan and an anonymous reviewer helped enormously with a number of truly constructive criticisms. Reprint requests should be sent to Barry Gordon, Cognitive Neurology, Meyer 2-222, Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, Maryland 21205, U.S.A.

insensitivity. It should be possible to distinguish between these two accounts by comparing masked and routine lexical decisions, but the unexpected word-specific effects prevented us from doing so.

In languages such as English, two groups of words have been distinguished by a number of criteria. Differences in set size and membership privileges have inspired their nomenclature: The “closed” class is a small set, of about 150–250 words, which only rarely gets new members. The “open” class, in contrast, contains hundreds of thousands of words, with a constantly expanding membership. These classes show a number of other, perhaps more fundamental, differences. They play markedly different grammatical roles: the closed class contains the free grammatical morphemes (articles, prepositions, and the like); the open class contains the nouns, verbs, and adjectives. The two classes also have different acquisition patterns (Bloom, 1970), phonological structure (Kean, 1977), different roles in word formation and in normal speech error patterns (Garrett, 1980), and contrasting impairments in aphasic speech (Caramazza & Berndt, 1985).

Because the two classes have such distinct properties in so many theoretically important aspects of language use, the hypothesis has been entertained that the two classes of words are served by completely different access routes. We will term this the “dual route” hypothesis. A form of this hypothesis has been proposed by Bradley and her colleagues (Bradley, 1978; Bradley & Garrett, 1983; Bradley, Garrett, & Zurif, 1980). They have argued for two separate lexical access systems: One, a general purpose access mechanism serving both open- and closed-class words; the other, a special purpose access mechanism serving only the closed-class words. We will term this view of the organization of the lexical access system the “common plus special route” hypothesis.

Bradley et al. have made some fairly specific claims about how processing differs in the two systems. Their claims have been based on data from a very influential series of lexical decision (and other) tasks with normal and aphasic subjects. One operational marker that they have identified for these different processing modes is the word frequency effect. It is well established that lexical decision time varies with frequency, at least for open-class words (e.g., Gordon, 1983); generally, higher frequency words have faster decision times than lower frequency words. However, Bradley et al. reported that closed-class words did *not* show this effect in normal individuals, although matched open-class words did. Paradoxically, moreover, in agrammatic Broca’s aphasics, closed-class words did show a frequency effect, comparable in magnitude to that of the open-class words. Since normal closed-class access is generally

assumed to be impaired in agrammatic Broca's aphasics, Bradley et al. reasoned that their performance was due to the spared abilities of another lexical access system, one which could clearly make accurate, frequency-sensitive decisions to both open- and closed-class words. Since normal subjects were presumably making use of an intact special-purpose access system, and did not show frequency effects for closed-class words, the special purpose system could not be frequency sensitive. Bradley et al. interpreted these and other data as support for the common plus specific route hypothesis, and as evidence that the two routes had to differ in frequency sensitivity.

Bradley and her colleagues also studied nonword interference effects, using nonwords whose initial segments were closed- or open-class words (Bradley, 1978; Bradley et al., 1980). Their normal subjects did not show any interference with nonword decisions when the initial segment was a closed-class word, but their agrammatics did. To account for this finding, Bradley et al (1980) were led to the further conclusion that access via the special closed-class mechanism must "inhibit the consequences" (p. 283) of the general-purpose access mechanism. They left the nature and site of this inhibition indeterminate. We will therefore have to speculate about possible observable consequences of this inhibition, but we can feel fairly secure about two points. First, Bradley et al. must have felt that the lexical decision task could index lexical access through either system, since they accepted its results in normals (where they were presuming closed-class performance was mediated by the special-purpose system) and in agrammatics (where they presumed that only the general-purpose route was preserved). Second, in normals, decisions about very high frequency closed-class words are no faster than those for lower frequency open-class words, as both Bradley (1978) and Gordon and Caramazza (1982) showed. But Bradley's hypothetical general- and special-purpose processors are both intact in normal subjects. Being frequency sensitive, the general-purpose processor should have been able to process these very high frequency closed-class words faster than the special-purpose processor could, and presumably make a faster decision about them than for lower-frequency words. Since decisions about these words are *not* faster, we must conclude that the special-purpose route is identifying closed-class words (perhaps not individually, but as a class) and then blocking processing in some fashion through the general-purpose route. This is, of course, the same conclusion Bradley et al. came to from their nonword interference results.

It should be noted, however, that we may be focusing too narrowly on frequency sensitivity (or lack of it) as a means of distinguishing between possible lexical access mechanisms. One reason for de-emphasizing frequency as a distinctive characteristic of the different processing systems is that the data on frequency dependence that Bradley et al. adduced to support dual-ac-

cess models have been called into question. Lexical decision studies in English (Gordon & Caramazza, 1982, 1983), French (Sequi, Mehler, Frauenfelder & Morton, 1982), Dutch (Kolk & van Grunsven, 1981), and German (Friederici & Heeschen, 1983; Heeschen, Friederici & Drews, 1984) with normal subjects have been interpreted as not showing any appreciable differences in the frequency sensitivity of open- and closed-class words, for normal subjects. Gordon and Caramazza (1983) also failed to find frequency related differences in closed-class lexical decisions between agrammatic Broca's and other aphasic patients. Therefore, frequency may not be the critical variable distinguishing the two hypothetical access systems.

Operationally, the two systems might be distinguished by either their dependence upon lexical access *codes*, or by their lexical access *speeds*. For example, in a reading task, access to the lexical representations of high frequency open-class words might take place directly from the visual representation (cf. McCusker, Hillinger, & Bias, 1981), while access to closed-class word representations might depend upon phonological recoding first. This would actually be an example of a difference in both codes and speed, since it is entirely possible that phonological recoding might take longer than direct access (McCusker et al., 1981). It is also possible to postulate a pure speed advantage for one class over the other. There may be reasons why one access route should be faster and more efficient than the other. Perhaps one route has to be optimized to meet the demands of on-line speech processing. Or perhaps one route could benefit because it need only identify a small number of items, and therefore it can operate with either higher efficiency or greater speed, or both.

We have deliberately not identified which access system we think should have which property, because we wish to emphasize more general points and experimental opportunities: While there could well be a number of differences postulated between the two access systems, many of them might be expressed as differences in speed. Even coding differences might well show up as differences in the speed of stages of coding/access. This being the case, then, a number of hypothetical differences between open- and closed-class processing routes may be testable by two modifications of the straightforward lexical decision task: masked stimulus presentation, and a response deadline (Pachella, Smith, & Stanovich, 1978). These changes might simply exaggerate normal processing differences. Presenting the stimulus string under tachistoscopic, masked, stimulus-limited conditions, for instance, should limit the orthographic information available to make the decision, and so penalize whichever access route is less optimal or whichever is slower. Imposing a response deadline should have a similar effect. It is also possible that these different experimental demands might force the subject to change his/her

lexical decision strategy from what it would normally be. For example, a subject struggling to meet a response deadline might be tempted to monitor the special, rapidly-working closed-class access system Bradley et al. proposed, thus giving those words an advantage.

Clearly, we cannot make strong a priori arguments about what the outcomes from these tasks should be, since we cannot reason from a well-specified base of proposed open- and closed-class differences. Even so, we think it is clear that even a negative result from these types of experimental manipulations could be valuable for setting limits on how empirically distinguishable the different explanations can be, in the absence of any other guides.

In addition, masking and probed response experiments with these high frequency words will address both a practical problem in earlier experiments, and what may be a significant theoretical point about the origin of the frequency effect. Both the Gordon and Caramazza (1982) and the Kolk and van Grunsven (1981) experiments found a flattening or absence of frequency sensitivity for words with frequencies at or above about 400/million (log frequency of about 2.6). Therefore, whether or not there were true differences between open- and closed-class words in this range could not be tested fairly. A masking and/or a speeded response method could test if this limit was due to peripheral rather than central factors. If an upper limit on perceptual appreciation speed had been the limiting factor under normal viewing conditions, limiting the available visual information by masking could eliminate this bottleneck. Similar reasoning would apply for the speeded response situation, if motor response speed had been the floor factor.

If neither manipulation succeeds in removing the floor on reaction times for very high frequency words, then this floor most likely represents a true central saturation of the word frequency effect. This finding would be relevant for theories of the lexical access process itself. The major class of such theories we will consider here are those with parallel access¹ such as Morton's (1970, 1982) logogen model and Gordon's (1981, 1983, in press) resonance model. In the version of Morton's model² developed by Coltheart and his colleagues (Coltheart, Davelaar, Jonasson, & Besner, 1977), lexical access occurs through logogens, counting devices for the perceptual evidence for a word (or other unit). A logogen is incremented whenever there is an input

¹Sequential access theories such as Forster's (1978) are the other major class. However, these theories have not been developed sufficiently to be directly applied to our experimental tasks and concerns.

²We should note that Morton (1982) himself does not contend that logogens are responsible for the frequency effect.

which satisfies one of its defining attributes. The logogen signals its activation to subsequent stages only when the accumulated count exceeds some critical threshold value, which is different for each logogen. Coltheart et al. postulated that high- and low-frequency word logogens accumulate information at the same rate, but that the frequency effect arises from variations in threshold settings. The logogens for the higher frequency words have lower thresholds, and so register the presence of their words more rapidly (with less evidence). However, Morton (1970) has also suggested that the thresholds are under some degree of subject control. Specifically, he argued that subjects might lower all their thresholds to make the system more sensitive, if perceptual factors such as stimulus degradation slowed the rate of information accumulation. If we were to find a flat frequency response, this would imply that the relevant logogens were all at the same, low, threshold setting. If we were to additionally find a flat function for very high frequency words under clear presentation conditions, that remains flat with stimulus degradation, this would imply that there is a lower limit on the threshold settings. Morton (1970) anticipated this possibility, but to our knowledge it has not been tested. In Gordon's (1981, 1983, in press) resonance model, the rate and degree of activation of internal word units is proportional to their frequency. The subject sets a threshold for responding "yes" to this activation in a lexical decision task. Together, these factors imply that a saturation of the frequency effect would occur. But this awaits empirical verification.

To summarize, the primary motivation for this study is that, if the notion of separate open- and closed-class access routes has any substance, we expect to see differences emerge between the classes under either tachistoscopic, masked presentation, or with speeded responses. Furthermore, finding a saturation of the frequency effect, particularly one independent of masking or response speeding, would be of additional importance for the logogen and resonance accounts of lexical access mechanisms.

Experiment 1: Tachistoscopic, masked presentation

Subjects

Twenty-six native speakers of English, with normal or corrected-normal visual acuity, aged 16 to 49, were recruited through hospital advertising. Nineteen were female; three were left-handed. Educational background ranged from 11 to 20 years. Each was paid 10 dollars for participation.

Materials

Two stimulus lists were constructed:

The *critical* list was made up of 305 words and 255 nonwords, for a total of 560 trials. The words consisted of closed- and open-class items. The selection of closed-class words was exhaustive: all those with log summed Kučera-Francis (1967) frequencies of 1.75 or greater, 1–8 letters long (mean = 4.16), and with 1–2 syllables (mean = 1.3) were included; this amounted to 146 words.

Sixty-six of the closed-class words could be very closely paired with open-class words, matching first by frequency (± 0.1 log frequency unit), next by length and number of syllables (\pm one letter or one syllable for the longer words, but not both). This is the set we will subsequently refer to as the “matched pairs”. The rest of the critical open-class words were chosen to achieve an even frequency distribution, while still trying to maintain a match by length and syllables. A total of 161 open-class words were used, of 2–8 letters (mean = 4.3) and 1–2 syllables (mean = 1.2). The overall matching was fairly good, but there were some inevitable differences of potential interpretive importance. For the log frequency range less than 2.6, closed-class words were on the average slightly longer than the open-class ones (5.5 vs. 4.2 letters, $t = 5.7$, $df = 126$, $p < .001$) and had more syllables (1.7 vs. 1.2, respectively, $t = 5.6$, $p < .001$). In the middle frequency interval between 2.6 and 3.44 (inclusive)—the highest end of the open-class frequency spectrum—the 74 closed-class words were nonsignificantly shorter than the 67 open-class words (mean length of 4.2 vs. 4.4, $t = -0.85$, ns), and had nonsignificantly more syllables (1.3 vs. 1.2, $t = 1.0$, ns). For the entire group with log frequency 2.6 or greater, closed-class words were slightly shorter (3.7 vs. 4.4, $t = -3.3$, $df = 175$, $p = .001$) but equal in syllables (1.18 vs. 1.19, $t = -0.2$, ns).

Nonwords were selected to be phonologically and orthographically legal letter combinations, matching the word pool in length, number of syllables, and initial letter or letter combination.

The masking stimuli were pseudo-randomly selected letter strings, eight characters long. All the words selected for the experiment were divided by position into their constituent letters; these letters were then incorporated into a pool for each position, with the number of occurrences of a letter in the pool matching the number of occurrences of that letter in that position in the set of actual words. Pseudo-random selection without replacement then produced a letter for each position of the mask, with meaningful letter combinations and more than two of the same letters in a row excluded.

A 330-trial *practice* list was made from 167, 1–2 syllable, open-class words

with frequencies above 1.75, which had *not* been used in the critical list because they did not match individual closed-class words closely enough. Their lengths ranged from 4 to 7 (mean = 5.25 ± 1.1). Nonwords and mask strings were made as described above (the letter pool for the eighth letter position was a duplicate of that for the seventh). Lists of the stimuli are available from the authors.

Apparatus

Stimulus presentation and response recording were controlled by an Apple microcomputer modified so that stimulus presentation could be accurately synchronized with the video sweep (Reed, 1979). Stimuli were shown in upper-case characters on a high resolution video screen. Items subtended no more than 3.0° horizontally and 0.5° vertically at a comfortable reading distance.

Procedure

Subjects were instructed verbally to decide whether the letter strings they saw before the mask were words or not. They were warned that the nonwords might look and sound like real English words. They were asked to make as rapid but accurate a response as possible by pushing finger-actuated micro-switches held in both hands: the switch held in the dominant hand for "yes", the one in the nondominant hand for "no". Each trial was initiated by the subject.

Each subject was first tested using the practice list to determine the stimulus-mask interval that gave 75% correct responses. After initiating a trial, an asterisk which served as a fixation point disappeared from the center of the screen. It was replaced 400 ms later by the centered test string. This was then replaced by the mask string, which was displayed for 2000 ms. The test string presentation time therefore corresponded to the string-to-mask interval. The subject was given feedback on the screen as to whether he had made the correct choice and whether the item had been a word or nonword.

For each subject, the initial test string presentation intervals were set to 167 ms. In the course of the practice trial block, this presentation interval was adjusted on the basis of the subject's accuracy (using a moving average of the previous 20 trials), in steps of 16.7 ms, to give 75% accuracy. For most subjects, this required intervals of 16.7–50.1 ms. When the accuracy level had been stable at 75% for 40 trials, the practice session was ended.

This mask delay giving 75% accuracy was then used for the entire series of 560 critical trials. The instructions and procedure were otherwise identical

to those for the practice trial, except that there was no feedback given to the subjects. Responses and reaction times were again recorded.

Data analysis

Errors due to extraneous factors (e.g., equipment failure) were excluded from all analyses. Individual reaction times were log transformed before analysis, and mean reaction times transformed back into milliseconds for group comparisons. Group proportions correct were arcsin transformed before averaging and analysis (Myers, 1972); mean values were converted back to proportions correct only for explication. For each subject, word and nonword reaction times which exceeded the mean ± 3 SD for that class were treated as errors and dropped from analysis. Log frequencies per million (summed over all regular derivational forms) were calculated as described in Gordon and Caramazza (1982). Averaged over all subjects, the mean percentage of words correct was 75%, the target value, but with a range from 47% to 98% (SD = 13%). Therefore, data from both the entire set of subjects and from the set who were between 65% and 85% correct (which we will term the "restricted set") were analyzed separately. Differences and trends for the restricted set of subjects essentially paralleled those of the entire group. A comparison of the first and second halves of each experiment, in the upper and lower frequency ranges, likewise did not show apparent effects of the practice session's use of only open-class words, so all critical trials were considered together.

Results

The data for all subjects are plotted in Figure 1 (mean reaction time per word as a function of class and frequency) and in Figure 2 (proportion correct for each word across subjects, by class and frequency).

Masking prolonged response times increased error rates, as intended. Mean reaction times for words with frequencies of 2.6 or greater (where Gordon & Caramazza, 1982, and Kolk & van Grunsven, 1981, had found a flattening of the frequency effect) were over 100 ms greater than those of an approximately comparable experiment without masking (Experiment 3 of Gordon & Caramazza, 1982), and 53% of the words in this masking experiment had error rates of 20% or more, whereas only 1% of the words in the earlier experiment had so many errors.

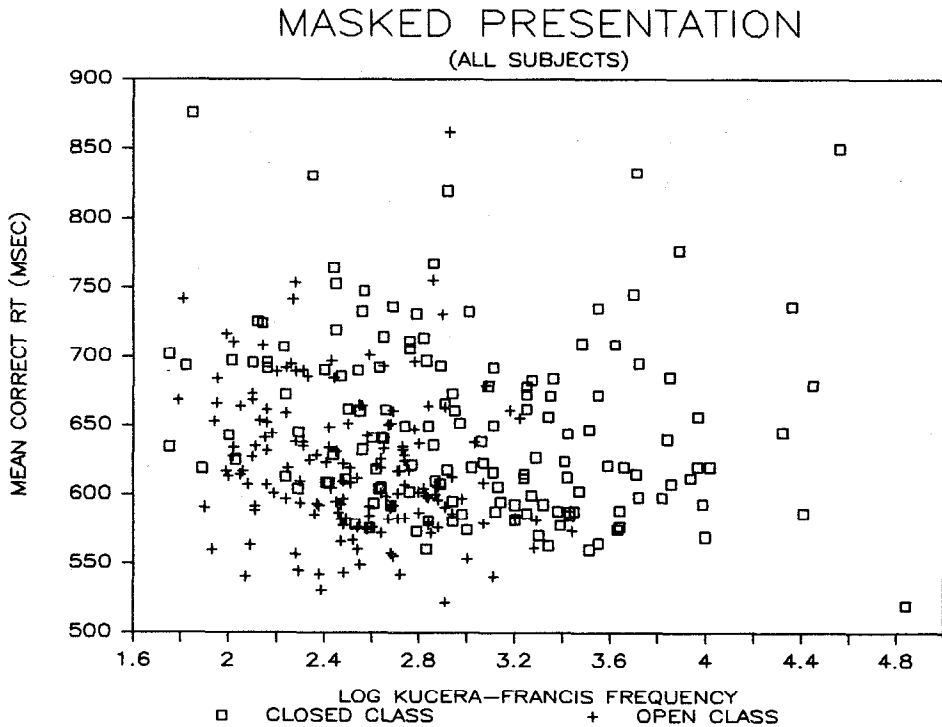


Figure 1. *Experiment 1. Mean correct lexical decision times as a function of summed Kučera-Francis frequency, for all subjects.*

Frequency sensitivity

In the lower frequency range (below log frequency of 2.6), the open-class words show a frequency effect (for all subjects, $r = -.30$, $df = 91$, $p = .002$, length partialled out; for the restricted set, $r = -.16$, $p = .06$). The closed-class words have negative but nonsignificant correlations in this range (for all subjects, $r = -.12$, $df = 31$, ns; for the restricted set, $r = -.09$, ns).

Neither word type shows any appreciable frequency effect for frequencies of 2.6 or greater, with nonsignificant r 's ranging from $-.04$ to $+.02$ (for the open class, $df = 64$; for the closed, $df = 107$). This was also generally true when sensitivity was studied post hoc on the sets of words with homogeneous length and syllable counts having enough members (10 or more) to test both sides of the null hypothesis. Of the twelve measures of sensitivity (reaction time and error rate for the four closed-class sets (monosyllabic, lengths = 2

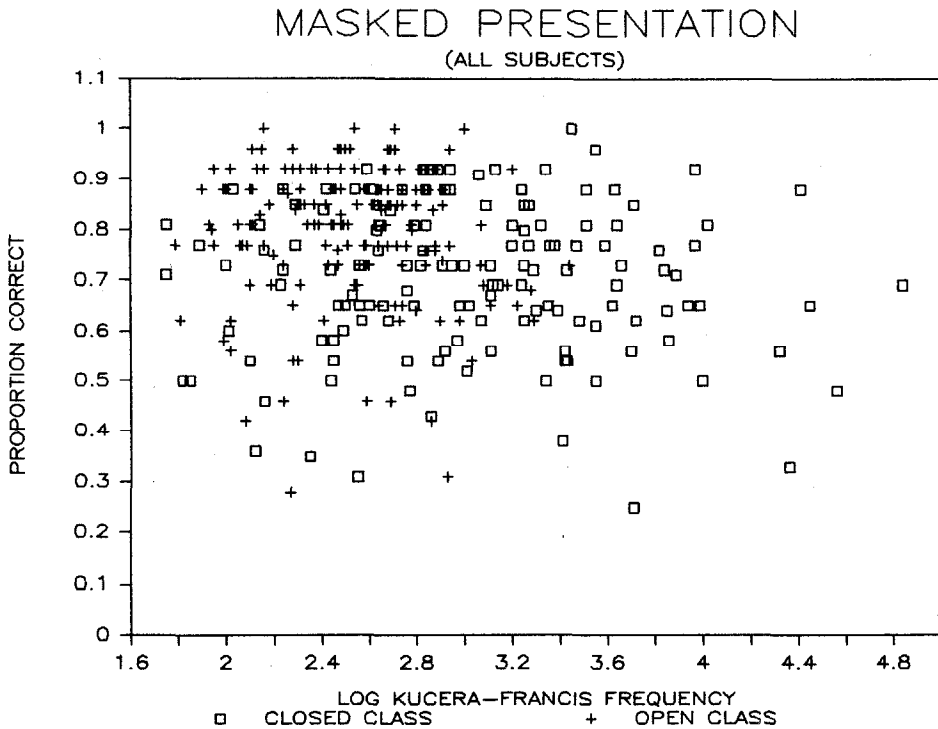


Figure 2. *Experiment 1. Proportion of correct lexical decisions for words as a function of summed Kučera-Francis frequency, for all subjects.*

($n = 20$), 3 ($n = 27$), 4 ($n = 26$), and 5 ($n = 10$)) and two open-class ones (monosyllabic, lengths = 4 ($n = 28$) and 5 ($n = 12$)), only two were significant, and both of these effects were the opposite of those usually expected for frequency.

Class reaction time differences

Reaction times (RTs) for the closed-class words were consistently slightly longer than for the open-class words: this difference can be seen in Figure 1. The differences found for the matched pairs are representative of the differences found for the larger groups: For the log frequency range less than 2.6, mean closed-class RT = 662 ms versus 620 ms for the open-class ($df = 22$, $t = 3.5$, $p < .002$, two-tailed); for the overlap range between log frequency

2.6 and 3.44 (inclusive), closed-class RT = 640 ms versus 622 for the open-class ($t = 1.89$, $df = 41$, $p < .04$). For the entire set of words with frequencies of 2.6 or greater, closed-class mean RT = 640 ms versus open-class mean RT = 614 ms, $t = 3.02$, $df = 155.4$, $p < .002$, two-tailed). A post hoc analysis of more directly commensurate (but not necessarily paired) monosyllabic words of three or more letters showed the same trend for faster open class reaction times (614 vs. 648), although this was not significant. For all comparisons, the proportions correct tended to vary in the same direction as the reaction times, but these differences were small and not significant.

Further analysis of these results will be given as part of the overall discussion. In keeping with the logic of the introduction, we also tested the same word stimuli under speeded response conditions.

Experiment 2: Speeded responses

Subjects

Thirty native speakers of English who had not participated in the previous study were tested. These were largely undergraduates who received course credit or payment for participation. Ages ranged from 17 to 23; 19 were male; 4 were left-handed.

Materials and apparatus

As in Experiment 1, except that a masking string was not used.

Procedure

Subjects were given a training session and then a critical session as in Experiment 1.

Subjects were instructed verbally to make a lexical decision as soon as possible after the onset of a loud 1000 Hz tone. For the training session, tone onset was initially set at 300 ms after the onset of the letter string display. Using a moving average over the previous 20 trials as before, this onset was adjusted to achieve a 75% accuracy of response. The experimenter monitored how well the subject was being guided by the tone; subjects whose RTs were outside the time band from tone onset to tone onset +400 ms were encouraged to respond more quickly. When a subject had been responding at a 75% accuracy level for 40 trials, the critical list was begun with tone onset fixed at that individual value. The necessary tone onset times ranged from 25 to 300 ms.

Subjects did not receive any error feedback during the critical list, but their adherence to the tone-signalized RT band was monitored and they were encouraged to respond faster or slower as necessary.

Data analysis

As for Experiment 1. Averaged over all subjects, the mean percentage of words correct was 78% (SD = 5.3%, range 67% to 92%), slightly better than the target value of 75% achieved on the practice trials. No subject subgroups were separated out.

Results

As desired, mean RTs with prompted responses were far faster than otherwise expected with the standard speed-accuracy instructions (over 230 ms faster than those of Experiment 3 in Gordon & Caramazza, 1982). Since the speeded response method tends to equalize reaction times, error rate is the primary variable of interest (cf. Santee & Egeth, 1982). Figure 3 shows these error rates as a function of frequency and word class for Experiment 2.

For the matched pairs with log frequency below 2.6, closed-class words had a slight, nonsignificant tendency for a lower proportion correct than open-class words: 0.77 versus 0.80, respectively ($t = -1.35$, $df = 22$, ns). This difference was not seen with the complete word set in this frequency range (proportion correct = 0.78 for both); monosyllabic words of length 3 or greater behaved similarly. As expected, reaction times were quite close for all the groups: for the matched pairs, 288 ms and 287 ms, respectively ($t = 0.11$, ns); for the entire set in this range, 288 and 285 ($t = 0.59$, ns); for the monosyllabic, length 3 or greater words, 284 and 286 ms ($t = -0.15$, ns).

For the matched pairs in the 2.6–3.44 frequency range, closed-class words had a minimally, nonsignificantly greater proportion correct: 0.81 versus 0.79 ($t = 1.11$, $df = 41$, ns). For all words with frequencies of 2.6 or greater, this difference disappeared (for both, proportion correct = 0.79, $t = 0.03$, $df = 104$, ns); similar values were obtained for the monosyllabic, length 3 or greater, group. Again, reaction times were quite similar: For the matched pairs, 281 and 284 ($t = -0.57$, ns); for the entire set, 282 and 281 ($t = 0.39$, ns); for the monosyllabic words, 280 and 276 ($t = 0.88$, ns).

There is a frequency effect for open-class words in the frequency segment below 2.6 ($r = .30$, $df = 91$, $p = .002$, length partialled out), but not for the closed-class words ($r = -0.08$, $df = 31$, ns). Above log frequency 2.6, neither class shows any appreciable frequency effects (r 's of .03 and -0.04 for $df =$

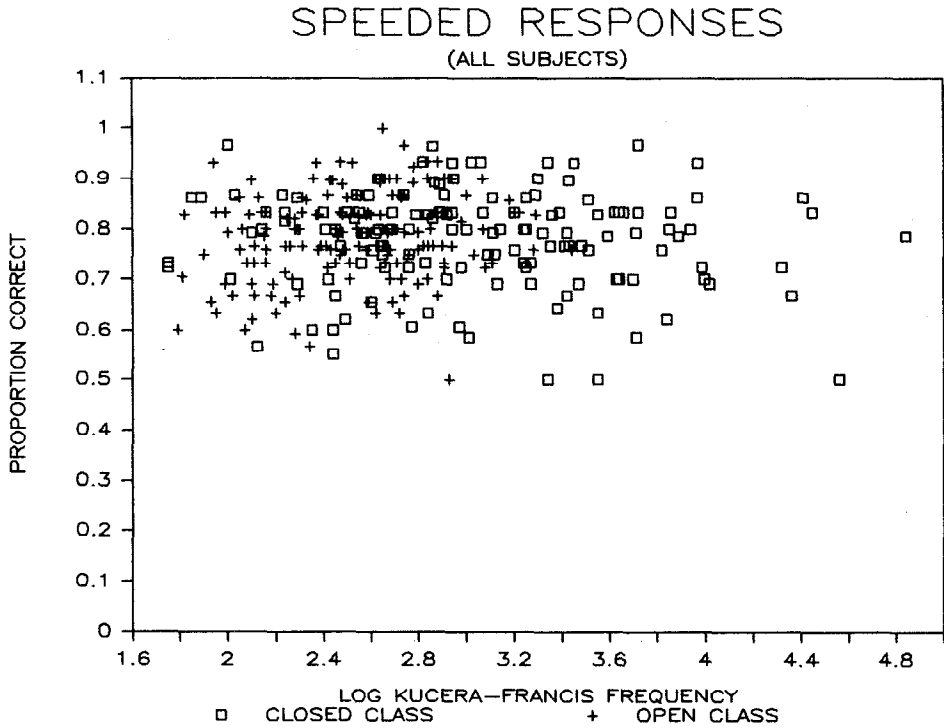


Figure 3. *Experiment 2. Proportion of correct lexical decisions for words as a function of summed Kučera-Francis frequency, for all subjects.*

64 and 107, respectively). Again, this was also true on a post hoc analysis of the same homogeneous subsets of open- and closed-class words described earlier. Of the twelve possible measures, only one reached significance, and this was again against the direction the frequency effect would have been expected to take.

Discussion

We will first consider the implications of our results for possible processing distinctions between open- and closed-class words.

These experiments eliminate the possibility that a simple floor on perceptual appreciation speed or motor response speed is a barrier to comparing very high frequency open- and closed-class words. RTs were appreciably

lengthened by stimulus masking (Experiment 1) without eliminating the flat frequency function for very high frequency words of either class; with RTs equalized (Experiment 2), error rates also did not show any frequency effects for these items. The pattern of lexical decision times they reveal is therefore not due to simple perceptual or motor limitations.

On the surface, these patterns seem to confirm some of Bradley's (1978) original findings: We did not find a frequency effect for closed-class words with log frequencies of 1.75 or greater, whether measured by reaction times or error rates. Yet we did show a frequency effect for fairly well matched open-class words over this range, particularly over its lower half. However, we do not think these results necessarily contradict our earlier ones (Gordon & Caramazza, 1982), nor those of others (Friederici & Heeschen, 1983; Kolk & van Grunsven, 1981; Segui et al., 1982) on the equivalent frequency sensitivity of the open- and closed-class, because we found word-specific influences on lexical decision speed which make it difficult if not impossible to attribute class-specific influences in this frequency range.

We had not originally assumed that any word-specific effects would be important. In selecting our stimuli, we had controlled for the usual variables known to affect lexical decision times (frequency, length and syllabicity) and ignored others which have not affected performance with high frequency words (e.g., bigram frequency, orthographic and phonologic regularity, concreteness, and polysemy; cf. Gernsbacher, 1984). If these standard precautions were adequate, then any differences in reaction times and error rates after accounting for class, length, and syllabicity in the frequency-insensitive range should have been just random noise. Yet the individual word latencies were not random, even for these closely matched words. We developed this impression in the course of our experiments; we tested it by a post hoc analysis using Kendall's coefficient of concordance (W) (Siegel, 1956) to measure the consistency of rankings of decision latencies, for words with log frequencies above 2.6 that had appeared in all four of our experiments with normal subjects using reaction time as the dependent variable (all three experiments of Gordon & Caramazza, 1982, and the masking experiment here). Only sublevels with ten or more words were tested, in order to guarantee enough power to test the consistency hypothesis in both directions. Three sets of words (two closed-class, one open-class) met these criteria. All showed highly significant ranking patterns. These words and their mean rankings are given in Table 1. (For one set of monosyllabic closed-class words, length = 3, $W = 0.45$, $n = 19$, $p = .02$; for the other monosyllabic closed-class set of length = 4, $W = 0.49$, $n = 18$, $p = .01$; for the one set of monosyllabic open-class words, of length 4, $W = 0.54$, $n = 13$, $p = .01$. The same set of length 3, closed-class words showed a borderline significant frequency effect,

even though the larger set of monosyllabic, length 3, closed-class words did not; the other two sets did not show frequency effects despite significant rank concordance.) This consistency was also apparent for many other subsets, although not statistically testable.

So even after controlling for class membership, length, and syllabicity, and in a range where frequency has no appreciable effect, there are properties of individual words which help determine their decision latencies across experiments, subjects, viewing conditions (masked in the current Experiment 1, standard in the other three), and response methodology (go–no go in Gordon & Caramazza's, 1982, Experiments 1 and 2; standard yes–no in the other two). We can speculate that these properties correspond to visual word forms (Warrington & Shallice, 1980) or the configurational properties of the words (Besner, 1983), but this will have to await a further exploration. Their significance now is that there have been uncontrolled, fixed determinants of RT in our experiments (and presumably others in English with these words).

These unanticipated effects are probably one of the reasons these experiments failed to find an appreciable frequency effect for the closed-class words

Table 1. *Mean rank across experiments for common word sets*

Closed-class, length = 3		Closed-class, length 4		Open-class, length = 4	
Mean rank	Word	Mean rank	Word	Mean rank	Word
18.25	DID	14.88	SUCH	11.50	FACT
15.88	HER	14.75	LESS	10.50	PLAN
13.50	FEW	14.75	WERE	9.50	LIFE
13.50	HAD	12.25	THAN	9.50	WANT
13.00	OUR	11.75	WHAT	8.50	YEAR
11.50	ALL	11.50	THEY	7.25	MAKE
11.50	YET	11.38	WHEN	6.75	HEAD
11.25	FOR	11.00	THAT	6.50	TURN
10.50	ARE	10.50	WITH	6.50	WORK
10.50	WHO	10.25	FROM	5.25	PLAY
9.25	HOW	9.50	BOTH	4.00	HAND
7.75	THE	8.50	MUCH	2.75	TIME
7.75	YOU	6.75	MORE	2.50	LOOK
7.63	HIS	5.75	HAVE		
7.50	AND	5.75	THEN		
5.75	BUT	4.50	THIS		
5.50	HAS	4.00	MOST		
5.25	WAS	3.25	EACH		
4.25	NOW				

in the 1.75 to 2.6 log frequency range, since other studies (except for Bradley's, 1978) have consistently found closed-class frequency sensitivity (Friederici & Heeschen, 1983; Gordon & Caramazza, 1982, Kolk & van Grunsven, 1981; Segui et al., 1982). Our current experiments could have been expected to be much less reliable than these: we studied only a relatively narrow frequency band, and used a selective and comparatively small set of words (only one-third as many closed-class as open-class items). We now also know that word-specific effects could have aggravated these selection problems. So we cannot regard our negative finding as valid. But could these word-specific effects have also influenced the previous studies, perhaps even to the point of artifactually creating an apparent closed-class frequency effect? We do not think so. For the Gordon and Caramazza (1982) study, the word-specific effects were relatively small compared to the range of decision times for these lower-frequency words (the standard deviations of the word-specific effects were 23 to 38 ms for the three sets of words we isolated, compared to 150–200 ms or greater latency differences across this range). It seems reasonable to assume that these word-specific effects are independent of frequency, randomly distributed with respect to both frequency and class, and similar across languages. If these are correct, then the word-specific effects should not have appreciably biased the data of the Gordon and Caramazza (1982) or other studies, so their conclusion that the classes show equal and appreciable frequency sensitivity would still be valid. Of course, a more definite judgment will have to wait until the rules governing the word-specific effects are determined, and these assumptions checked.

While the individual word effects we uncovered prevent some conclusions, they do not detract from our major observation: We found no appreciable class-specific differences despite our experimental manipulations. The two classes show the same lack of frequency sensitivity at the upper end of the frequency spectrum; they show a similar variability as well, with no evidence that the closed-class receives the special treatment we would have expected a special processor to provide. The only possible class difference we did find was an advantage for open-class words (under masked conditions), which ran counter to our expectations that the closed-class would be the one to use and benefit from specialized processing. Therefore, even though our chains of reasoning were individually tenuous, the predictions of dual- or general-plus-dual route accounts of closed-class access seem to be collectively disproved. Our current findings are therefore in accord with those of Gordon and Caramazza (1983), who failed to find any difference in closed-class lexical decisions across agrammatic and nonagrammatic aphasics, nor between open- and closed-class words. So even under pathologic conditions, where access differences should have been accentuated, no differences have been found.

We are not rejecting the position that there are processing distinctions between closed- and open-class words. There is compelling evidence in the literature on language processing both in normal speakers (Garrett, 1980) and in aphasics and dyslexics (Caramazza & Berndt, 1985) to support the claim that open- and closed-class words are served by different cognitive mechanisms. The real question is at what level of access and/or post-access processing these differences emerge. Insofar as the lexical decision tasks we have used tap stages of lexical access (cf. Balota & Chumbley, 1984, for reservations about this task), we cannot demonstrate differences between the classes. Therefore, the basic access processes for the two classes of words are most probably identical; their higher-order consequences are not.

Frequency sensitivity

Our Experiments 1 and 2 were not intended as replications of the already well-documented frequency effect in lexical decision for words with frequencies below about 400/million (e.g., Scarborough, Cortese, & Scarborough, 1977). Nevertheless, the other experiments in the literature have typically used only open-class words, and the highest frequencies they explored have therefore been up to about 3.0–3.44 (log frequency units). Including closed-class words extends this frequency range by another order of magnitude or more. In this range, there are no appreciable frequency effects, as Gordon and Caramazza's (1982) and Kolk and van Grunsven's (1981) experiments had suggested, and as the current experiments help confirm. (In the Segui et al., 1982, experiments, stimuli ranged up to log 3.4 in frequency, so they may not have been expected to show this effect. Nevertheless, their published plots are consistent with a flattening of the frequency sensitivity curve.) What we have shown in addition is that this flattening of RTs occurs even when possible perceptual and motor response limits on speed of response are removed. Therefore, it appears to be a more central limit. Other data suggests that this limit is reached gradually (Gordon, in press).

As we discussed earlier, this frequency saturation implies that, if Coltheart et al.'s (1977) logogen-based account for the frequency effect is accepted, then there must be an absolute lower limit on the threshold settings. The same saturation is a fairly predictable consequence of Gordon's resonance model (Gordon, 1981; 1983; in press), which can quantitatively fit this and other frequency data (Gordon, in press). These data, then, warrant modification of the logogen theory, but are consistent with the resonance model.

In addition, it should be possible to differentiate between the two accounts by examining the behavior of words with frequencies just below the limits of complete saturation. In the logogen account, the slower activation caused by

perceptual degradation would induce a lowering of the logogens' thresholds, until the lowest limit is reached. So reaction times for these words would approach those of words already at the lower limit, flattening the reaction time versus frequency function in this border zone. This prediction is opposite the one made by the resonance model. Although the resonance model also assumes that perceptual degradation could slow the activation of word-units, it predicts that the correct subject strategy would be to *raise* the decision threshold. This should accentuate the frequency effect for borderzone words (and also reveal one for words of slightly higher frequency which were at saturation under clear viewing conditions). Unfortunately, in our experiments the existence of word-specific latency differences precluded the group comparisons we had planned. Post hoc comparison of borderzone words that were identical across experiments showed only a nonsignificant trend for masking to flatten the frequency effect in the borderzone region. It appears, then, that differentiating the two accounts in this fashion should be feasible, but this will have to be a future project.

In summary, we have not found word-class specific effects in lexical decision despite manipulations which we had reason to believe would elicit any differences that do exist. We conclude that the open- and the closed-class are accessed through the same processes, at least at the level(s) tapped by the lexical decision tasks we used. The very highest frequency words of both classes show a saturation of the frequency effect, which needs to be taken into account in theories of the lexical access mechanism. Finally, we found significant individual regularity in decision latencies for these very high frequency words, without any obvious factors which could determine such patterns. For now, this was an experimental obstacle. But this finding may warrant pursuit in its own right.

References

- Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340-357.
- Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology*, 37, 429-438.
- Bloom, L. (1970). *One word at a time*. The Hague: Mouton.
- Bradley, D.C. (1978). Computational distinctions of vocabulary type. Unpublished doctoral dissertation, MIT, Cambridge, MA.
- Bradley, D.C., & Garrett, M.E. (1983). Hemispheric differences in the recognition of closed and open class words. *Neuropsychologia*, 21, 155-159.
- Bradley, D.C., Garrett, M.E., & Zurif, E.B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes*. Cambridge, MA: MIT Press.

- Caramazza, A., & Berndt, R.S. (1985) A multi-component deficit view of agrammatic Broca's aphasia. In M.L. Kean (Ed.), *Agrammatism*. New York: Academic Press.
- Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale: Erlbaum.
- Forster, K.I. (1978). Accessing the mental lexicon. In E.W. Walker (Ed.), *Explorations in the biology of language*. Montgomery, VT: Bradford Books.
- Friederici, A., & Heeschen, C. (1983). *Lexical decision of inflected open-class items and inflected closed-class words*. Paper presented at the Annual Meeting of the Academy of Aphasia, Minneapolis, MN.
- Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 1). London: Academic Press.
- Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*, 256–281.
- Gordon, B. (1981). *Lexical access and lexical decision: Mechanisms of frequency sensitivity*. Unpublished doctoral dissertation, Johns Hopkins University, Baltimore, MD.
- Gordon, B. (1983). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior*, *22*, 24–44.
- Gordon, B. (in press). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory and Language*.
- Gordon, B., & Caramazza, A. (1982). Lexical decision for open- and closed-class words: Failure to replicate frequency sensitivity. *Brain and Language*, *15*, 143–160.
- Gordon, B., & Caramazza, A. (1983). Closed- and open-class lexical access in agrammatic and fluent aphasics. *Brain and Language*, *19*, 335–345.
- Heeschen, C., Friederici, A., & Drews, E. (1984). Lexical access in agrammatism: An inflected language as a testing ground. Paper presented at the Annual Meeting of the Academy of Aphasia, Los Angeles, CA.
- Kean, M.L. (1977). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition*, *5*, 9–46.
- Kolk, H.H.J., & van Grunsven, M. (1981). *Nonsyntactic sources of agrammatism*. Paper presented at the Academy of Aphasia Annual Meeting, London, Ontario, and personal communication (1984).
- Kučera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- McCusker, L.X., Hillinger, M.L., & Bias, R.G. (1981). Phonological recoding and reading. *Psychological Bulletin*, *89*, 217–245.
- Morton, J. (1970). A functional model for memory. In D.A. Norman (Ed.), *Models of human memory*. New York: Academic Press.
- Morton, J. (1982). Disintegrating the lexicon: An information processing approach. In J. Mehler, E.C.T. Walker, & M. Garrett (Eds.), *Perspectives on mental representation*. Hillsdale, NJ: Erlbaum.
- Myers, J.L. (1972). *Fundamentals of experimental design*. Boston: Allyn and Bacon.
- Pachella, R.G., Smith, J.E.K., & Stanovich, K.E. (1978). Qualitative error analysis and speeded classification. In N.J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, NJ: Erlbaum.
- Reed, A.V. (1979). Microcomputer display timing: Problems and solutions. *Behavior Research Methods & Instrumentation*, *11*, 572–576.
- Santee, J.L., & Egeth, H.E. (1982). Do reaction time and accuracy measure the same aspects of letter recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 489–501.
- Scarborough, D.L., Cortese, C., & Scarborough, H.S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 1–17.
- Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia*, *20*, 615–627.
- Siegel, S. (1956). *Nonparametric statistics*. New York: McGraw-Hill.
- Warrington, E., & Shallice, T. (1980). Word-form dyslexia. *Brain*, *103*, 99–122.

Résumé

La question de savoir si les mots de la classe fermée et les mots de la classe ouverte suivaient la même voie d'accès au lexique a été récemment l'objet d'un vif débat. On a suggéré que les différences dans la sensibilité aux fréquences pouvaient indiquer des voies d'accès séparées. Cependant les études sur les décisions lexicales ont souffert des effets-plancher pour les fréquences élevées. Ces effets peuvent masquer des différences théoriques importantes dans le comportement des deux classes. Nous avons étudié les décisions lexicales pour des mots ayant une grande ou une très grande fréquence dans chacune des classes. Nous avons utilisé des stimuli masqués et des réponses accélérées dans le but de minimiser les effets-plancher, d'essayer de voir les différences potentielles entre le comportement des deux classes et de d'opposer les théories d'accès au lexique.

Nous n'avons pas trouvé de preuves à l'appui de l'hypothèse que les mots de la classe fermée ont une voie d'accès lexical différente ou spéciale. On ne trouve pour aucune des deux classes de mots d'effet de fréquence pour les fréquences de Kučera-Francis de 400 million ou plus et ceci aussi bien pour les temps de réaction que dans l'analyse des erreurs.

Les mots de la classe ouverte donnent parfois lieu à des réponses plus rapides que les mots comparables de la classe fermée mais ceci peut être en contradiction avec des interprétations sur l'effet de la classe des mots (Bradley et al 1980). En outre nos données montrent également quelles pourraient être les influences spécifiques aux mots sur les temps de décision lexicale – effets qui pourraient être impossibles à séparer de l'effet de la classe des mots en anglais.

Pour interpréter la non sensibilité aux fréquences trouvée dans ces expériences on doit aménager les modèles d'accès lexical basés sur le logogen et y inclure une limite inférieure pour les seuils de positionnements. Les modèles de résonance (Gordon, 1983) prédisent déjà cette insensibilité aux fréquences. Il aurait du être possible de distinguer entre les deux modèles lors des tâches de décision lexicale avec masquage mais des effets inattendus inhérents aux mots nous ont empêché de le faire.