The Locus of Word-Frequency Effects in the Pronunciation Task: Lexical Access and/or Production?

David A. Balota and James I. Chumbley

University of Massachusetts

Three experiments were conducted to separate the influence of word frequency on lexical access from its influence on production stages in the pronunciation task. A delayed pronunciation task was used in which a word was presented and, after some delay, a cue was presented to pronounce the word aloud. Presumably, subjects should access the word’s lexical representation during the delay interval and any effect of frequency can be attributed to its influence on production rather than on lexical access. In the first two experiments frequency still had a significant effect with a delay interval of 900 milliseconds. At a delay interval of 400 milliseconds, where the subjects should have accessed the word’s lexical representation before the pronunciation cue, the word-frequency effect was 41 milliseconds, only 17 milliseconds less than the 58-millisecond effect obtained with the same words in a normal pronunciation task. A significant word-frequency effect at a delay interval of 2900 milliseconds was found in the third experiment where rehearsal during the delay interval was disrupted by requiring subjects to whisper the alphabet while waiting for the cue to respond. © 1985 Academic Press, Inc.

The process by which printed strings of characters are transformed into meaning, word recognition, has been the center of theoretical concern since the days of Catell. Throughout this period, researchers studying word recognition have relied heavily on the pronunciation task. This emphasis on the pronunciation task has occurred at least partly because it is a simple task and it is intuitively related to processes involved in reading. The task is simple because it only requires that the subject quickly pronounce a visually presented stimulus. Its theoretical relation to reading is quite direct. The prevalent view is that each word is internally represented in a lexicon (by, e.g., a logogen). When there is sufficient overlap between the stimulus word and its internal representation, the word is recognized and its motor code then becomes available to pronounce it aloud.¹ Since pronunciation typically involves lexical access, researchers (e.g., Andrews, 1982; Forster, 1981; Frederiksen & Kroll, 1976) have frequently used the pronunciation task to investigate variables thought to affect the speed of lexical access.

Unfortunately, the simplicity of the pronunciation task may be deceptive. As Rossmeisl and Theios (1982) have noted, re-

¹ In the present paper we are assuming that the pronunciation task actually does involve lexical access. It is possible, on the other hand, that subjects can rely on grapheme-to-phoneme conversion rules in a pronunciation task and bypass lexical access (cf. Coltheart, Davelaar, Jonasson, & Besner, 1977). If this were the case then this would suggest that the effect of word frequency in pronunciation is not on production, rather than on encoding. Although there is currently debate regarding this issue, we are in general agreement with the arguments made by Forster (1979), Theios & Muise (1977), and West and Stanovich (1982) that pronunciation usually involves lexical access.

This research was partially supported by a NIMH Postdoctoral Fellowship (5-25007) to the first author. We acknowledge the assistance of Hayley Arnett throughout this research project. We also thank Charles Clifton, Jr., Kenneth Forster, Susan Lima, and Keith Rayner for comments on an earlier version of the manuscript. Finally, we would like to especially thank Dennis Norris for his direction to the delayed pronunciation task along with many fruitful discussions regarding word recognition. Requests for reprints should be sent to David A. Balota, Department of Psychology, Iowa State University, Ames, Iowa 50011.
searchers have long been concerned (cf. Cattell, 1886) with whether variables that affect pronunciation latency influence the encoding of a word, its production, or both. Because a variable can potentially affect either or both of these components of pronunciation, it is incumbent upon the researcher to determine the locus or loci of the effects being observed. One cannot assume that a variable affects lexical access simply because it affects pronunciation latency. Although in past research this potential difficulty with the task has been implicitly acknowledged, there have yet to be any empirical attempts to tease apart the differential impact of a variable on the two components. The present study begins to provide such evidence.

The principal variable of interest in the present research is printed-word frequency. In numerous studies (e.g., Becker, 1976; Berry, 1971; Forster & Chambers, 1973; Forster, 1981; Frederiksen & Kroll, 1976; Glanzer & Ehrenreich, 1979; James, 1975; Rubenstein, Lewis, & Rubenstein, 1971; Stanners, Jastrzembski, & Westbrook, 1975) it has been demonstrated that subjects respond more slowly to low-frequency words than high-frequency words. The dominant theoretical interpretation of this finding is that frequency of occurrence in print somehow orders the internal lexicon and thereby influences the speed of lexical access. Briefly, some of the theoretical characterizations of the frequency effect have been that (1) low-frequency words are searched only after high-frequency words have been searched (Becker, 1979, 1980; Forster, 1976); (2) low-frequency words are stored in the total lexicon of all words which is searched only after a sublexicon which contains only high-frequency words has been searched (Glanzer & Ehrenreich, 1979); (3) low-frequency words have higher thresholds for word recognition than high-frequency words (Morton, 1969, 1970). The relevance of these theories to word recognition clearly rests on the effect of the word-frequency variable on tasks that can be unequivocally viewed as reflecting lexical access.

Recently, we (Balota & Chumbley, 1984) have argued that the only other speeded task commonly used to investigate lexical access, the lexical decision task, may provide exaggerated estimates of the frequency effect because of the decision stage in that task. That is, in the lexical decision task, subjects are asked to discriminate familiar words from unfamiliar nonwords. Two pieces of information about the letter string that the subject could use to make such discriminations are its visual familiarity and its meaningfulness. In making word/nonword discriminations, subjects should find it more difficult to respond to low-frequency words than to high-frequency words because the former are more similar to the nonwords on a familiarity dimension. Thus, lexical decision task response latencies to low-frequency words may be slower than to high-frequency words because of the discrimination difficulty and not, at least totally, because of differences in lexical access time. This argument, of course, reduces the utility of the lexical decision task for studying the effects of word frequency on lexical access.

If the above analysis of the lexical decision task is correct, then data from studies using the pronunciation task are much more important to the study of the influence of printed frequency on lexical access time. However, as noted above, the pronunciation task has been difficult to interpret since

---

2 It is important to note here that the fact that frequency information is available does not provide evidence that this information in some way orders lexical access. We have little doubt that subjects could indicate in a nonspeeded task that they have seen the word "car" more than the word "nip." However, there is a considerable difference between suggesting that individuals have familiarity information available and suggesting that this familiarity slows the lexical access of a low-frequency word by 100 milliseconds, compared to a high-frequency word, the typical size frequency effect found in the LDT.
frequency may be affecting word encoding and/or word production. In fact, it seems intuitively reasonable that oral production frequencies of words may vary at least as much as printed frequencies. Spoken communication usually involves greater time constraints than written communication. There often is not enough time to search for the "best" word in impromptu speech while the good writer will often spend a great deal of time in searching for a word with just the right shade of meaning. Thus, speakers may rely more heavily than do writers on readily available high-frequency words. Assuming that there are considerable differences in the frequency with which particular words are selected to be spoken, it is possible that oral-production-frequency differences are even larger than corresponding printed-frequency differences. Thus, it is unclear why one would not find frequency effects in the production stage of the pronunciation task.

There are two studies (Forster & Chambers, 1973; Theios & Muise, 1977) which provide data relevant to the distinction between frequency effects in lexical access and frequency effects in production. Unfortunately, the results of these studies lead to different conclusions. In Forster and Chambers' study, a word was presented for a full 2 seconds and then a cue for the subject to pronounce the word aloud was presented. The notion was that after a 2-second delay, the word must have been recognized and any remaining frequency effect would be a production-frequency effect. The results suggested that frequency was having its primary effect on lexical access since low-frequency words were only slightly (4 milliseconds by subjects, 7.5 milliseconds by items) and nonsignificantly slower to pronounce than high-frequency words. However, as Rossmeissl and Theios (1982) have pointed out, during a 2-second-delay interval the subject not only can identify the word but also can silently rehearse it in preparation for the pronunciation cue. This rehearsal could attenuate any frequency difference due to articulation. A second concern about the Forster and Chambers study is its generalizability since they used only 15 high-frequency and 15 low-frequency words. The present research addresses both difficulties with the Forster and Chambers study.3

In the second relevant study, Theios and Muise (1977) used homophonic word pairs to equate production frequency while varying word frequency. A set of homophonic word pairs was chosen such that one item had a low printed frequency (heir) and the second item had a high printed frequency (air). If the frequency effect is localized in lexical access, then a large frequency effect should still be observed with such pairs. However, if part of the frequency effect is in the frequency of usage of a particular motor code, then a reduced frequency effect should be observed since both words (air and heir) presumably utilize the same motor code. Theios and Muise found only a small (11 milliseconds), but significant, effect of word frequency. Since this effect was relatively small, they suggested that part of the frequency effect in the pronunciation task is due to production frequency. The Theios and Muise results are suggestive but not conclusive for two related reasons. First, frequency effects in the pronunciation task are often quite small. For example, Forster (1981) recently reported only a 31-millisecond effect. Second, the low-frequency words in the Theios and Muise study had an average value in the Kučera and Francis (1967)

3 In fairness to Forster and Chambers, they were not primarily interested in investigating the frequency effect in production but were simply attempting to ensure that there were no phonological characteristics which would confound their results in a zero-delay pronunciation experiment. Thus, they gave subjects sufficient time to ensure that lexical access was already completed. We simply feel that other processes (such as rehearsal) can occur during a 2-second delay interval.
norms of 30 occurrences per million, a value much higher than the average for low-frequency words, used in most other studies.

Experiments I and IA

Experiment I addressed the issue of production-frequency effects while avoiding some of the difficulties with the past research. The technique utilized was similar to that used by Forster and Chambers (1973); a word was presented and, after some delay, a cue was presented to pronounce the word aloud. In the present experiment, however, six delay intervals ranging (in 250-millisecond increments) from 150 to 1400 milliseconds were used. This procedure made it possible to trace the frequency effect across the different delay intervals. Since subjects should not be able to rehearse the words before the end of some of the short delay intervals, there is a greater likelihood of finding a frequency effect than in Forster and Chambers’ 2-second delay experiment where rehearsal was a real possibility. Also, if lexical access occurs relatively quickly, for example, within about 250 milliseconds, then one may find evidence for a “production”-frequency effect at the longer delay intervals where lexical access should have already been completed. Very simply, if the pronunciation task is a useful method for investigating how frequency influences lexical access, as a number of researchers clearly believe (e.g., Andrews, 1982; Berry, 1971; Forster, 1981; Forster & Chambers, 1973; Frederiksen & Kroll, 1976), then frequency effects should be limited only to the short delay intervals (at most 400 milliseconds) and should disappear after the subject has had time to recognize the stimulus. On the other hand, if there are production-frequency effects then one might expect frequency effects even at the longer delay intervals.

We also conducted a baseline experiment which will be referred to as Experiment IA. In this experiment we obtained an estimate of the frequency effect for the present set of words in a normal pronunciation task. Thus, subjects were simply asked to pronounce the word as soon as it was presented. In this way, we determined a baseline estimate of the size of the total frequency effect, including both lexical access and production-frequency effects, for the present set of words.

Method

Subjects. A total of 48 undergraduate students were recruited from the subject pool at the University of Massachusetts, Amherst, and participated in partial fulfillment of course requirements. Twenty-four subjects participated in Experiment I and 24 subjects participated in Experiment IA. No subject participated in more than one of the experiments reported in this paper.

Apparatus. The experiments were controlled by a North Star Horizon microcomputer. Stimulus words were displayed in uppercase letters on a television monitor driven by an IMSAI memory-mapped video raster generator. In order to maximize legibility, all words were displayed with single spaces between letters. Subjects were seated approximately 50 centimeters from the video monitor. A three-letter word (three letters separated by two spaces) subtended a visual angle of approximately 1.1 degrees while a nine-letter word subtended a visual angle of approximately 3.7 degrees. A voice key connected to the computer detected the onset of the subject’s pronunciation. Response latency and interval timing were both measured with millisecond accuracy.

Materials. A total of 72 high-frequency and 72 low-frequency target words were used in this experiment. These items were selected from the Kučera and Francis (1967) norms such that the low-frequency words had counts less than 7 per million (mean rating = 2.1 occurrences per million) while the high-frequency words had counts greater than 36 per million (mean rating = 169.5 occurrences per million). Word selec-
tion was also constrained so that words ranged in length from three to nine letters and so that each high-frequency word had a corresponding low-frequency word of the same length. The complete list of target words is presented in the Appendix. A total of 75 practice/buffer words (also ranging in length from three to nine letters) were selected from the Kučera and Francis norms.

Procedure. Each subject was presented two blocks of 15 practice trials. These practice blocks were followed by six test blocks of 34 trials each. The first 10 trials in each test block were buffer trials with the remaining 24 trials being target trials. In Experiment 1 each subject responded to a total of 144 target words, 12 high-frequency target words and 12 low-frequency target words at each of the six delay intervals. Across the first six subjects a given target word was paired once with each of the delay intervals. Furthermore, after each group of six subjects was completed, the total list of words was rerandomized with each word being assigned to a different position in a Latin square to again be cycled through all six delay intervals. Thus, with this procedure, each subject saw each word once at only one delay interval but across the 24 subjects each word was presented four times at each of the six delay intervals. After the target words were selected for a given test block they were randomly ordered. Within the practice block, all trials were randomly assigned to delay interval. However, within a test block all 34 trials had the same delay interval. Thus, delay interval was a blocked variable in Experiment 1. Furthermore, within every group of six subjects the delay intervals for a given block were sequenced through a Latin square across test blocks, with a different Latin square being utilized for each group of six subjects.

On each trial in Experiment 1 the following sequence occurred: (a) a 500-Hertz warning tone was presented for 250 milliseconds; (b) a 250-millisecond interstimulus interval; (c) the stimulus word was presented at the center of the monitor; (d) after a delay interval ranging from 150 to 1400 milliseconds, parentheses were presented which enclosed the word, for example (FLOOR); (e) the word enclosed by parentheses was presented until the computer detected the onset of the subject’s pronunciation; (f) the word and parentheses were immediately erased and after a 1-second delay interval the message “Response OK?” was presented; (g) the subject either pulled the right lever to indicate that a correct pronunciation of the word triggered the computer or the left lever to indicate that an incorrect pronunciation of the word or some other sound (such as a cough) triggered the computer; and (h) after one of these levers was pulled there was a 4-second blank interval before the warning tone for the next trial was presented.

All subjects were tested individually. Upon their arrival at the experiment subjects were seated in a sound-deadened room and were given instructions to read. It was emphasized in both the written, and later oral, instructions that the subject should resist the tendency to pronounce the word before the presentation of the parentheses. Subjects were told that it was crucial that they wait until the parentheses were presented and then pronounce the word as quickly as possible without mispronouncing it. The experimenter used an intercom, a monitor that displayed the same stimulus the subject was viewing, and a second monitor which displayed the subject’s response latency to ensure that subjects were following instructions.

Two response levers, one designated “Response OK” and the other “Error,” were placed in front of the subject along with a microphone. Feedback regarding both the subject’s reaction time (RT) and the percentage of trials in which the subject pulled the “Response OK” lever was given after each block of trials. Subjects were required to rest 10 seconds between trial blocks before a message on the monitor indicated that they could press a button when
they were ready to continue the experiment.

The same procedure was used in Experiment 1A except that there were no parentheses presented and the subjects were instructed to pronounce the word as soon as it appeared on the screen.

Results

A mean response latency for each Frequency x Delay Interval condition was computed for each subject. Any RT which was greater than 850 milliseconds or less than 125 milliseconds (possible anticipations) was treated as an outlier. The mean error and outlier rates were 2.3 and 3.8%, respectively, for Experiment 1, and 3.4 and 4.7%, respectively, for Experiment 1A.

Two two-way analyses of variance treating frequency and delay interval as variables were conducted on the data from Experiment 1. One analysis was based on item means and the second was based on subject means. The statistic based on the subject means will be presented first and then will be followed in brackets by the statistic based on the item means. Unless otherwise specified all significant effects have p values less than .05.

Figure 1 displays the mean pronunciation latencies for the delay intervals of Experiment 1 along with the baseline-frequency effects found in Experiment 1A. The mean differences between the high-frequency and low-frequency conditions are also provided along the abscissa of Figure 1. There are four points to note about these results. First, high-frequency words were pronounced faster than low-frequency words in the baseline conditions of Experiment 1A, F(1,23) = 159.58, MSe = 256.7 [F(1,142) = 93.34, MSe = 1372.8]. Second, high-frequency words were pronounced faster than low-frequency words across the delay intervals of Experiment 1, F(1,23) = 49.38, MSe = 1222.2 [F(1,142) = 27.92, MSe = 5629.1]. Third, response latency decreased across the delay intervals, F(1,23) = 10.59, MSe = 6601.8 [F(5,710) = 51.49, MSe = 3723.4]. Fourth, the frequency effect decreased across the delay intervals. This decrease is reflected by a Delay Interval x Frequency interaction which reached significance by subjects, F(5,115) = 2.37, MSe = 1279.1, but not by items, F(5,710) < 1, MSe = 3723.4.

Planned comparisons at each delay interval indicated that there were significant frequency effects at the 150-millisecond delay interval, t(23) = 4.57 [t(142) = 2.98]; the 400-millisecond delay interval, t(23) = 5.40 [t(142) = 3.98]; the 900-millisecond delay interval, t(23) = 3.09 [t(142) = 2.50]; the 1150-millisecond delay interval, t(23) = 3.09 [t(142) = 2.68]; and, the 1400-millisecond delay interval, t(23) = 2.21 [t(142) = 2.08]. Although the frequency difference was significant at the 650-millisecond delay interval by the subject analysis, t(23) = 2.21, this difference did not reach significance by items, t(142) = 1.12.

Discussion

The results of Experiments 1 and 1A are quite clear. There were large frequency effects at both the zero-delay interval of Experiment 1A and at the 150- and 400-millisecond intervals of Experiment 1. In fact,
there was only a 14-millisecond drop in the
correction delay interval to the 150-millisecond delay
interval and there was only a 1-millisecond
difference between the frequency effect ob-
tained at the 150- and 400-millisecond delay
intervals. This is especially interesting
when one considers that overall pronunci-
ation latency dropped by 160 milliseconds
between the baseline overall mean and the
400-millisecond delay interval. Thus, sub-
jects were clearly utilizing the 400-millise-
cond delay interval to process the word.
Also, it is noteworthy that there were sig-
nificant frequency effects even at the longest,
1400-millisecond, delay interval.

These results suggest that a large com-
ponent of the word-frequency effect in the
pronunciation task involves production in-
stead of simple lexical access. It is unclear
how subjects could still be searching the
lexicon after a 400-millisecond delay in-
terval since normal reading rates are on the
order of 250 milliseconds per word. Thus,
frequency in print appears to be related to
the time taken to pronounce a word inde-
pendent of the time taken to access the
word’s lexical representation. In fact,
based on the results of Experiment 1, it ap-
ppears that frequency has a relatively small
effect upon the time taken to search the
lexicon for a letter string’s representation.
Again, the frequency effect was only 14 mil-
liiseconds smaller at the 400-millisecond
delay interval than at the zero-millisecond
delay interval. We shall return to the issue
of how much time is needed for lexical ac-
cess later in the paper.

To our surprise the results of Experiment
1 yielded significant frequency effects even
at the 1150- and 1400-millisecond delay in-
tervals. One concern that arises regarding
these results is that subjects may have pro-
cessed the stimulus in a different manner at
these long delay intervals since the trials
were blocked by delay interval. Thus, the
frequency effect in the longer delay inter-
vals may have been produced because sub-
jects knew approximately when the pro-
nunciation cue would be presented and
waited to process the stimulus until slightly
before the parentheses were presented.

This possibility was addressed in Experi-
ment 2 in which the delay interval ran-
domly varied from trial to trial. With this
procedure it was not possible for the sub-
jects to anticipate when the cue would be
presented and they should have been forced
to begin processing the word as soon as it
appeared. If the significant frequency ef-
effects at the longer delay intervals in Exper-
iment 1 were simply due to subjects not
processing the word until slightly before the
parentheses appeared, then a frequency ef-
effect should not be found at the longer delay
intervals in Experiment 2.

Experiment 2

Method

Subjects. Twenty-four undergraduate
students were recruited from the subject
pool described for Experiment 1.

Materials. The target words used in Ex-
periment 1 were again used in Experiment
2. Only 50 of the 75 buffer/practice words
used in Experiment 1 were utilized in Ex-
periment 2.

Procedure. Each subject was presented
two blocks of 15 practice trials which were
followed by four test blocks of 41 trials
each. The first five trials of each test block
were buffer trials. Each word was assigned
to a delay interval as in Experiment 1.
Within a block of trials each delay interval
appeared with six different target words at
each frequency level and, within a block of
trials, the delay intervals were randomly or-
dered. All other aspects of the procedure
were identical to those of Experiment 1.

Results

The analyses conducted on the results
from Experiment 2 were the same as those
conducted for Experiment 1. The mean
error and outlier rates were 3.1 and 4.2%,
respectively.

Figure 2 displays the mean pronunciation
latencies for the delay intervals of Experiment 2 along with the baseline-frequency effects obtained in Experiment 1A. Also, shown in Figure 2 are the mean differences between the high-frequency and low-frequency conditions. There are three points to note about Figure 2. First, high-frequency words were consistently pronounced faster than low-frequency words, $F(1,23) = 25.29$, $MS_e = 1470.7$ [$F(1,142) = 34.90$, $MS_e = 3820$]. Second, response latencies decreased across the delay intervals, $F(1,115) = 125.21$, $MS_e = 1279.1$ [$F(5,710) = 181.61$, $MS_e = 2633.3$]. Third, the frequency effect also decreased across the delay intervals which is reflected by a Delay Interval x Frequency interaction, $F(5,115) = 6.77$, $MS_e = 1279.1$ [$F(5,710) = 2.93$, $MS_e = 2633.3$].

Planned comparisons at each delay interval indicated that there were significant frequency effects at the 150-millisecond delay interval, $t(23) = 6.06$ [$t(142) = 5.04$]; the 400-millisecond delay interval, $t(23) = 4.50$ [$t(142) = 4.05$]; the 650-millisecond delay interval, $t(23) = 2.92$ [$t(142) = 2.52$]; and, the 900-millisecond delay interval, $t(23) = 2.71$ [$t(142) = 2.11$]. On the other hand, the frequency effects did not reach significance at either the 1150-millisecond delay interval, $t(23) = 1.52$ [$t(142) = 1.64$], or the 1400-millisecond delay interval, $t(23) = 1.10$ [$t(142) = 1.28$].

Discussion

The results of Experiment 2 were very similar overall to the results of Experiment 1. There were again large frequency effects found for the 150- and 400-millisecond delay intervals. Also, there were significant frequency effects at both the 650- and 900-millisecond delay intervals. However, the frequency effects found in Experiment 1 at the 1150- and 1400-millisecond delay intervals did not reach significance in Experiment 2. Thus, it may indeed be the case that the frequency effects found at the 1150- and 1400-millisecond delay intervals in Experiment 1 were, in part, due to the fact that subjects knew when the parentheses would be presented and waited until just prior to the presentation of the parentheses to prepare the response.

On the other hand, there is an alternative account of the results of Experiment 2. It is possible that during the long delay intervals subjects were engaging in a strategy which overrides any production-frequency effect. For example, it is possible that the frequency effect disappeared at the 1400-millisecond delay interval of Experiment 2 because subjects were silently rehearsing the words during the delay interval. In fact, if the present arguments for a production-frequency effect are correct, then it is unclear why one would not find the effect at longer delay intervals as long as subjects were not engaging in a process which eliminated the effect. However, if subjects engaged in some rehearsal during the delay interval, then the repetition of the motor code for pronunciation may have led to a reduction in any observed frequency effect. That is, one might expect motor code repetition to have a larger impact on the production of a low-frequency motor code than on the production of a high-frequency motor code. Such a differential impact of
repetition on low- and high-frequency words has been demonstrated in lexical decision performance (Scarborough, Cortese, & Scarborough, 1977; Scarborough, Gerard, & Cortese, 1979).

**Experiment 3**

In Experiment 3 we addressed the possibility that subjects in Experiment 2 were silently rehearsing the words during the silent delay interval thereby diminishing any production-frequency effect. During the delay interval in Experiment 3, subjects were required to whisper the letters in the alphabet forward from a randomly determined starting letter. If the production-frequency effect disappeared at the longer delay intervals in Experiment 2 because subjects were silently rehearsing the words during the delay intervals and, if such rehearsal can be suppressed by whispering the alphabet, one would expect a production-frequency effect to reappear at the longer delay intervals in Experiment 3. Furthermore, as long as subjects rehearse the alphabet continuously during the delay interval, such production-frequency effects should be found at much longer delay intervals. Thus, in Experiment 3 there were six delay intervals which randomly varied from 400 to 2900 milliseconds in 500-millisecond increments.

**Method**

**Subjects.** Thirty subjects were recruited from the subject pool described in Experiment 1.

**Materials.** The materials were precisely the same as those used in Experiment 2.

**Procedure.** The procedure was the same as that of Experiment 2 with the following exceptions. First, on each trial a randomly chosen letter of the alphabet was presented in the center of the video screen. This letter appeared for 500 milliseconds before being erased simultaneously with the presentation of the warning tone which was used as in the previous experiments. Second, subjects were instructed to whisper the alphabet forward starting from the randomly chosen letter. The experimenter instructed the subject to whisper the alphabet as quickly as possible until the parentheses were presented at which point she/he should pronounce the word aloud as quickly as possible. It is important to note here that, just as in the previous experiments, the target word was displayed on the screen during the interval between the warning tone and the subject's pronunciation of the word aloud. Third, because of the importance of having the subjects follow instructions, the experimenter sat in the testing room slightly behind and to the right of the subject. If the subject was not whispering the alphabet quickly enough the experimenter encouraged the subject to pick up the pace. This control was necessitated by the fact that this experiment hinges on the assumption that the subjects were actually following instructions.

**Results**

Since response latencies were considerably longer in this experiment, a 1250-millisecond upper criterion was used to define an outlier. The lower criterion remained at 125 milliseconds. The mean error and outlier rate were 5 and 1.5%, respectively. The analyses conducted on the results from Experiment 3 were the same as those conducted on the earlier experiments.

Figure 3 displays the mean pronunciation latencies for the delay intervals of Experiment 3, along with the mean differences between the high- and low-frequency words. There are three points to note about Figure 3. First, high-frequency words were pronounced faster than low-frequency words, \( F(1,29) = 21.45, M_{S_e} = 1769.1 \) \[ F(1,142) = 14.71, M_{S_e} = 6925.9 \]. Second, response latencies appear to decrease from, at least, the 400-millisecond interval to the longer intervals, \( F(5,145) = 3.50, M_{S_e} = 1769.1 \) \[ F(5,710) = 4.31, M_{S_e} = 4793.1 \]. Third, there is little difference between the frequency effects across the different delay intervals as indicated by the lack of a Fre-
to trial, it is unlikely that subjects were allowing their attention to drift during the delay interval and rerecognizing the word after the parentheses were presented. In fact, the relatively fast response latencies (mean RT = 603 milliseconds) for such a complex task indicates that subjects were following instructions and attempting to pronounce the word very quickly upon presentation of the parentheses (possibly due to the demand characteristics of having an experimenter monitoring their performance in the room).

Second, if subjects were allowing their attention to drift, then one would expect an increase in response latency at the longer delay intervals. However, as shown in Figure 3, overall response latency was flat from the 900-millisecond delay interval to the 2900-millisecond delay interval. On the other hand, it is noteworthy that overall RT did decrease from the 400-millisecond delay interval to the 900-millisecond delay interval. Similar patterns were found in Experiments 1 and 2 where it can be seen that there was some decrease in response latency from the 400-millisecond delay interval to the 900-millisecond delay interval (see Figures 1 and 2). This decrease in latency in Experiment 3 is important because it also suggests that subjects were attending to the word on the screen during the delay interval since RT is decreasing in a similar manner across all three experiments and then reaching asymptote.

A third point worth noting in Figure 3 is that the frequency effect was only 25 milliseconds at the 400-millisecond delay interval. In Experiment 1 it was 44 milliseconds and in Experiment 2 it was 38 milliseconds. It is possible that there was a reduced frequency effect at the 400-millisecond delay interval in Experiment 3 because the "actual" delay interval until pronunciation was 200 milliseconds longer in Experiment 3. This longer delay interval occurred because subjects needed to stop whispering the alphabet and program their response; processes which could occur ei-

Discussion

The results of Experiment 3 are quite clear. If one disrupts rehearsal during the delay interval, then a significant frequency effect is found even though the subject has had close to 3 seconds to access a visually presented word's lexical representation. Obviously, such a frequency effect is not simply occurring in lexical access.

There are a number of points to note about this pattern of data. First, since the delay intervals randomly varied from trial

**Fig. 3.** Mean pronunciation latencies as a function of frequency and delay interval for Experiment 3. The numbers in parentheses indicate the differences between the high-frequency and low-frequency words for that delay interval.

Frequency × Delay Interval interaction $F(1,145) < 1, MSe = 1248.2 \ [F(5,710) < 1, MSe = 4793.1]$.  

Planned comparisons at each delay interval indicated that there were significant frequency effects at the 400-millisecond delay interval, $t(29) = 2.34 \ [t(142) = 1.85, p < .05$, directional]; the 900-millisecond delay interval, $t(29) = 2.36 \ [t(142) = 1.92, p < .05$, directional]; the 1400-millisecond delay interval, $t(29) = 2.22 \ [t(142) = 2.00]$; the 1900-millisecond delay interval, $t(29) = 2.91 \ [t(142) = 2.31]$; and, the 2900-millisecond delay interval, $t(29) = 2.22 \ [t(142) = 2.04]$. The only exception to this pattern was the 2400-millisecond delay interval in which the frequency effect did not approach significance, both $t$'s < 1.
ther in cascade or in parallel (cf. McClelland, 1979). Of course, the reduced frequency effect at the 400-millisecond delay interval which may be attributable to the longer "actual" delay interval does not reduce the importance of finding significant frequency effects at the very long delay intervals.

**Further Analyses of the Results**

The results of the present experiments indicate that there are quite large frequency effects in the pronunciation task which are traceable to the output of the stimulus word rather than only to its encoding. Because of the importance of these results with respect to the locus of the word-frequency effect in the pronunciation task, it is necessary to consider some alternative accounts of the data.

One possibility is that some other variable covaried with frequency in our list of materials and it produced the delayed pronunciation "frequency" effects. There are a number of points to note here. First, such a variable, unless it is specific to our materials, most likely varied in other studies which have utilized the pronunciation task. The set of high- and low-frequency words used in our study were equated on the principal variables (such as length and concreteness) equated in other pronunciation studies. Second, since the effects generalize not only across subjects but also across the set of items utilized in the experiments, the effects are not simply being produced by a few peculiar words. Third, we used a considerably larger sample (144 words) than has typically been used in studies of this type. For these reasons, our results should be more likely to generalize to the population of words instead of being restricted to the particular sample set [see, e.g., the earlier discussion of Forster & Chambers' (1973) study].

Multiple regression analyses were also used in an effort to detect effects of variables covarying with frequency. The five predictor variables selected for these analyses were frequency; length in letters; number of syllables (cf. Sternberg, Monsell, Knoll, & Wright, 1980); the log of the number of dictionary meanings for the word (cf. Jastrzembski, 1981); and a measure of the phonological characteristics of the beginning of the word. The latter variable simply divided the words into five classes ranked according to each word's beginning phoneme. The five ordered categories were (as represented by the letter of the alphabet corresponding to the phonemes) (1) P, T, K, C, and Q; (2) D, G, and B; (3) A, E, I, O, and Y; (4) F, L, M, N, R, W, and V; and (5) H, J, S, CH, SH, TH. The rationale for these five categories was that words with beginning phonemes from the first category might be produced more quickly or might activate the voice key more quickly than words which have a beginning phoneme from one of the latter categories.

A series of regression analyses were conducted using the above five predictor variables and the mean response latency for each word, across subjects, as the criterion variable. A regression analysis was conducted on the results for each of the six delay intervals for each experiment. With the exception of frequency and length, the results of these regression analyses did not yield any consistent pattern of significant effects nor was there a consistent pattern of regression coefficients.

Actually, the lack of a syllable effect in the pronunciation task is consistent with a number of other studies (Balota & Chumbley, 1984; Chumbley & Balota, 1983; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Klapp, 1971; see, however, footnote 4). Also, the lack of an influence of the log number of meanings on pronunciation performance is consistent with a re-

---

2 Sternberg et al. (1980) found a small, but significant, 4.5-millisecond effect of syllables in the production of word strings. Since length of the word in letters appeared to be confounded with length of the word in syllables it is difficult to interpret this result.
cent study conducted in our own laboratory (Chumbley & Balota, 1983). It is of course possible that the scaling for phonological onset that we utilized was not a sensitive measure. However, in closer examination of our materials, we were unable to find any further phonological characteristics that consistently varied between the high- and low-frequency sets of words.

One variable, in addition to frequency, which did significantly influence pronunciation latency was length. A number of researchers have also found that length has a large impact on pronunciation latency (e.g., Forster & Chambers, 1973; Frederiksen & Kroll, 1976). In these earlier studies the impact of length, like the impact of word frequency, could have occurred at a number of stages in pronunciation. For example, length could be influencing (1) a letter-by-letter readout from sensory memory (see Gough, 1972; Gough & Cosky, 1977; Sperling, 1963); (2) the comparison of the visual stimulus with its lexical representation; (3) retrieval of phonologically based subprograms to pronounce the letter string; and (4) possibly, the readout from some phonologically based buffer. With regard to the latter possibility, Sternberg et al. (1980) found that the number of words in a to-be-pronounced word string has a potent effect on latency to start production of the string in a delayed pronunciation procedure. Fortunately, the present data can provide some useful information regarding these possibilities. If length is influencing production then one should find length effects (like the obtained frequency effects) at the longer delay intervals. On the other hand, if the length effect is primarily affecting an earlier stage in processing, then one might expect that the length effect would occur only at the very short delay intervals.

In an attempt to discriminate between the above two hypotheses regarding the locus of the length effect, the high-frequency words and the low-frequency words were split into two subgroups of 32 words each. (In order to prevent overlapping groups of long and short words only 128 words could be included in this analysis instead of all 144 words.) The words in the long group ranged from 6 to 9 letters in length (mean length = 6.97 letters); the words in the short group ranged from 3 to 5 letters in length (mean length = 4.53 letters).

The mean RTs for each word for each experiment at each Delay Interval X Length X Frequency condition were submitted to separate analyses of variance. Since for this post hoc partitionings of words there was not an equal number of observations for a given subject, only the item analyses will be presented.

Figure 4 displays the mean pronunciation latencies as a function of length and delay interval for Experiments 1, 2, and 3 along with the baseline length effect found at the zero-delay interval in Experiment 1A. First, note the graphs for Experiments 1, 1A, and 2. As expected, the main effect of frequency was again highly significant in Experiment 1, $F(1,124) = 26.03$, $MSE = 5656.6$, Experiment 1A, $F(1,124) = 94.66$, $MSE = 1164.6$, and in Experiment 2, $F(1,124) = 25.36$, $MSE = 4103.5$. Also, the main effect of delay interval was significant in Experiment 1, $F(5,620) = 44.53$, $MSE = 3727.8$, and in Experiment 2, $F(5,620) = 164.14$, $MSE = 2616.8$. More importantly, consider the length effect displayed in Figure 4. As can be seen, the only substantial length effect is at the zero-delay interval baseline condition of Experiment 1A, $F(1,124) = 29.86$, $MSE = 1164.6$. The main effect of length did not reach significance in either Experiment 1, $F(1,124) = 2.71$, $MSE = 5656.6$, or Experiment 2, $F(1,124) = 1.63$, $MSE = 4103.5$. Furthermore, planned comparisons indicated that only the 150-millisecond condition in Experiment 2, $t(126) = 2.47$, produced a significant length effect. In fact, if one considers the mean difference between long and short words at delay intervals greater than 150 milliseconds in both Experiments 1 and 2, this difference is only 4 milliseconds.

Turning to the results of Experiment 3,
both frequency, $F(1,124) = 13.33$, $MS_e = 7113$, and delay interval, $F(5,620) = 5.32$, $MS_e = 4878$, were again highly significant. However, as can be seen in Figure 4, the results of Experiment 3 appear to show only a small impact of length. In fact, neither the main effect, $F(1,124) = 2.23$, $MS_e = 7113$, nor any of the $t$ tests at any of the delay intervals yielded significant length effects. It is interesting to note that since the shortest delay interval in Experiment 3 was 400 milliseconds, the lack of a length effect in this experiment is consistent with the fact that there was no length effect found at any of the delay intervals longer than 150 milliseconds in the results of Experiments 1 and 2.

The above analyses suggest that the length effect in normal pronunciation may be localized in a very early component of the pronunciation process. This component is apparently completed within approximately 150 to 400 milliseconds. There were only small nonsignificant differences between long and short words at the longer delay intervals. This pattern of length effects is in direct contrast to the frequency effects which were found at delay intervals reaching 2900 milliseconds. Thus, it appears that frequency and length may be influencing different stages of word analysis. The stage in word pronunciation which is influenced by length is completed after 150 milliseconds and before 400 milliseconds, whereas, at least one stage in word pronunciation which is influenced by frequency is not completed until the word is actually produced. In this light, these results suggest that length may be affecting the encoding component of word analysis. As noted earlier, a number of researchers have suggested that length is playing a role in sensory readout. Both Spelke (1963) and Gough (1972; Gough & Cosk, 1977) have suggested that such readout involves approximately 10 milliseconds per letter. In the present set of words, the mean latency difference between the long and short words at the zero-delay interval was 33 milliseconds and the mean difference in length between these two sets of words was 2.44 letters. Thus, the increase in RT found in our studies (13.5 milliseconds per letter) is relatively consistent with the estimates noted earlier. Furthermore, the present results suggest that the length effect does indeed appear to be localized early in word processing. Further research is needed to
corroborate this post hoc analysis of the impact of length (in letters) on pronunciation latencies.

**General Discussion**

The present series of experiments was motivated by the need to discriminate between the impact of word frequency upon lexical access and its impact on production in the pronunciation task. The results of Experiment 1 yielded significant frequency effects even when subjects were given 1400 milliseconds to access the word's lexical representation. This result suggests that frequency does indeed have a postaccess influence in word pronunciation.

In Experiment 2 we addressed the possibility that since the delay intervals were blocked in Experiment 1, subjects may have anticipated when the pronunciation cue would be presented and waited to process the letter string until just prior to the cue's presentation. In order to prevent such anticipations, the delay intervals were randomly ordered in Experiment 2. The results of this experiment yielded significant frequency effects at delay intervals as long as 900 milliseconds.

Experiment 3 addressed the concern that subjects were possibly rehearsing the words during the longer delay intervals of Experiment 2 thereby eliminating the significant frequency effect found in Experiment 1 at the 1150- and 1400-millisecond delay intervals. Rehearsal was prevented in Experiment 3 by having subjects whisper the alphabet during the delay interval. This experiment yielded significant frequency effects for delay intervals up to 2900 milliseconds, the longest delay interval used.

Before discussing the implications of these results it is important to address an alternative account of the frequency effect at longer delay intervals. This account is based on the premise that lexical access sometimes occurred twice in a given trial; once when the word was presented and again upon presentation of the parentheses used as a pronunciation cue. If a second access occurred, it is reasonable to expect that frequency effects were introduced by the process of reaccessing the word's lexical representation.

Although this reaccess account is a possibility, it is inconsistent with a number of our experimental results. First, overall latencies dropped 160 milliseconds from the zero-millisecond delay interval to the 400-millisecond delay interval and then asymptoted at approximately 370 milliseconds. If the subjects were reaccessing the word when the cue was presented then it is unclear why there was such a large drop in response latency. Clearly, subjects have completed one aspect of the pronunciation task which lasts 150 milliseconds; an amount of time within the range of other estimates of lexical access (see discussion below). A second problem is that one might expect that subjects would be more likely to reaccess an item at the longer delay intervals thereby producing a rise in latency at these delay intervals. In fact, none of the results from the three experiments provide any evidence for an increase in latency-with-delay interval. Third, we feel that if reaccess were to occur it would be more likely to occur when subjects could anticipate the cue than when the cue randomly varied between the six delay intervals. On the contrary, there was very little difference between the results obtained in the blocked conditions of Experiment 1 and the randomized conditions of Experiment 2. Finally, if reaccess were occurring throughout the delay intervals, it is unclear why there was evidence for a length effect only at the zero-delay and 150-millisecond delay intervals. As noted above, the pattern of length effects seems most consistent with an effect on lexical access and if reaccessing occurred, there should have been length effects at the longer delay intervals.

While we do not believe that a second lexical access account is compatible with our data, there is one form of a reaccess argument which seems plausible to us. Once the subject has recognized the word
to be pronounced and prepares to pronounce it when cued, this "preparation" must be maintained in some form. There seem to be two obvious candidates for the form of storage: (1) a motor program which is the basis for output, or (2) a more abstract code which must retrieve and implement a motor program when the pronunciation cue is presented. Since the asymptotic RT difference between low-frequency words and high-frequency words was about 20 milliseconds in all three experiments, it is doubtful that a motor program was the form of storage. That is, it is unlikely that subjects could have maintained a motor program for the stimulus word while rehearsing the alphabet in Experiment 3.

Assuming that subjects were storing an abstract code there are at least two forms it could take. First, it could be a "visual" code which would allow reaccessing the lexicon and production of a response. This type of code does not seem reasonable for all of the reasons given above in rejecting the "two-access" account of our data. A more likely candidate would be an abstract, possibly semantic, code which could be used to retrieve and implement a motor program for pronunciation of the word, just as is required in normal conversation. The time required for retrieval and for implementation of this motor program could very well depend upon frequency of usage.

At this point we can only speculate about whether the production frequency effect we observed is due to longer retrieval times or longer implementation times or both. Obviously, more research is needed to settle the issue. What is clear, however, is that differences between low- and high-frequency words in either of these processes should affect performance in a normal undelayed pronunciation task. There is no reason to believe that the connection from a lexical entry to a motor program is more direct and less subject to these effects than is the connection from an abstract code.

There is one last issue which needs to be addressed in the present study. That is, what are the relative impacts of frequency on lexical access and on postaccess processes? The force of the present study decreases considerably if only a very small proportion of the frequency effect in the pronunciation task can be attributed to postaccess processes. The data from the current experiments can be used to address this question but an estimate of the maximal time needed to complete lexical access is needed. By comparing the frequency effect at the delay interval corresponding with this time with the frequency effect obtained at the zero-delay interval, a rough estimate of the impact of frequency on lexical access can be obtained.5

Lexical access time can be estimated from a number of different situations. First if one makes the reasonable assumption that subjects are recognizing words in normal reading then one must consider the average reading rate of 200 to 250 milliseconds per word (Rayner, 1978) as one estimate of lexical access time.6 Second, the

5 In the present paper we assume a sequential stage model (Sternberg, 1969) of pronunciation in which lexical access is first completed and then some postaccess motor retrieval and production occur. An alternative to this view is the cascade framework developed by McClelland (1979). Such an alternative would assume that lexical access overlaps with transmission of information to the postaccess stages. We have assumed a discrete framework because the relevant models (Becker, 1980; Forster, 1979; Morton, 1970) assume that lexical access is completed before pronunciation and meaning information become available. Furthermore, we feel that the present data are equally problematic for a cascade model of word recognition.

6 Obviously, there may be contextual constraints in reading which are not present in isolated word pronunciation and these constraints may decrease fixation duration in reading. However, it is important to note two considerations. First, the size of strong contextual manipulations (e.g., high associates) on fixation duration are relatively small. Zola (1982) found only a 16-millisecond effect for high associates, and Ehrlrich and Rayner (1981) found a 55-millisecond effect for very strong contextually biasing paragraphs. Furthermore, a number of researchers have suggested that the predictive power of such contextual constraints in reading are quite small (cf. Stanovich & West, 1983, for a discussion of such effects). Thus, even if context
estimates of semantic processing rates obtained from research using the rapid sequential visual presentation (RSVP) method (Fischler & Bloom, 1980; Potter, Kroll, & Harris, 1980) are less than 100 milliseconds per word. Third, estimates from more direct attempts to measure lexical access time have ranged from 183 (Sabol and DeRosa, 1976) to 210 milliseconds (Neisser & Beller, 1965). Finally, Gough and Cosky (1977) have estimated that the extraction of semantic category information occurs in less than about 300 milliseconds. Thus, the available estimates of lexical access time appear to consistently converge on a time of between 100 and 300 milliseconds.

Interestingly, this estimate is quite consistent with a number of aspects of the present data. One such aspect is that significant length effects were not found for delay intervals over 150 milliseconds. This suggests that the component of the pronunciation response which was influenced by length required at least 150 milliseconds but less than 400 milliseconds. Furthermore, if lexical access is taking 150 milliseconds of the total response time in the zero-delay condition, then RT should asymptote at a value which is approximately 150 milliseconds less than overall response latency in the zero-delay condition. This argument is again based on the notion that subjects should have completed lexical access by the longer delays and therefore at those delay conditions, where performance has asymptomed, response latency should be 150 milliseconds less than the zero-delay condition which includes lexical access. As noted above response latency asymptoted at approximately 150 milliseconds less than at the zero-delay condition. Thus, we have a remarkably consistent, 150 millisecond estimate of lexical access time.

The present data and the review of earlier studies suggest that 150 milliseconds, and at most 400 milliseconds, represent delays in which the subject should have accessed the word's lexical representation. If one considers the frequency effect at the 400-millisecond delay interval, it is quite surprising that this frequency effect (41 milliseconds across both Experiments 1 and 2) is only 17 milliseconds less than at the zero-delay (58 milliseconds) condition. This clearly suggests that a major portion of the word-frequency effect in the pronunciation task cannot be unequivocally attributed to access, but rather, appears more likely to be due to processes occurring after lexical access.

**CONCLUDING REMARKS**

The present study provides firm evidence that frequency influences more than lexical access in the pronunciation task. We have argued that lexical access should have occurred quite early in the delay intervals utilized in the present studies. If these arguments are correct, then one must note that there is a drop of only 17 milliseconds from a condition including lexical access, the zero-delay interval of Experiment 1A, to a condition where lexical access effects should be minimal, the 400-millisecond delay interval of Experiments 1 and 2.

It is important to note that we are not arguing that word frequency has no impact on lexical access, but rather, that one must be very cautious in unequivocally attributing the frequency effect found in this task to lexical access. We have argued elsewhere (Balota & Chumbley, 1984) that the decision process in the lexical decision task, the other task commonly used to study speeded lexical access, exaggerates the role of word frequency. Now we have provided evidence that the pronunciation task also includes exaggerated effects of frequency because of the production stage.
In this light, we feel it is crucial that researchers be especially cautious in interpreting the results of tasks which have been viewed as relatively pure indices of lexical access. There may be hidden components (e.g., production in pronunciation and decision in lexical decision) which may produce effects which in turn could potentially misdirect our theoretical characterizations of the processes we intend to study.

APPENDIX

High-frequency words

MAN CAR DOG KEY OIL HOME FEET DOOR TOWN FIRE PAPER GLASS HORSE RADIO WORLD WATER MONEY BLOOD GRASS BIRDS TEETH TREES BONES FLOOR BREAD HEART WALLS SHOES MOUTH SEEDS PLANE LETTER ENGINE CATTLE OFFICE WINDOW SQUARE STREET MOTHER PLANTS PALACE COFFEE GARDEN COTTON FOREST MARKET BILLS ISLAND YELLOW COUSIN VALLEY ANIMALS MACHINE FRIENDS COLLEGE TEACHER VILLAGE KITCHEN CAPTAIN CHICKEN TONGUE LICENSE SENATOR GROWTHS SOLDIERS MAGAZINE MERCHANT BASEBALL STUDENTS MOUNTAINS PRESIDENT APARTMENT

Low-frequency words

PEW ELK SOD GEM PAN LAVA KELP SILO MALT DIME BARON LADLE WHARF NAVAL SNAIL CLEAT SHACK TUNIC FERNBANJO THUGS LAPEL FELON RIVET PECAN NYLON BEETS MELON CRATE CLOAK REELS SABLE RETINA HEARSE BEAKER SKATES RACOON CASHIWS ICICLE GASKET SEQUIN HANGER RAFTER RUDDER ROBOTS WEASEL PARCEL QUARTZ PELVIS ARMPIT RAVINE BEGGER JEWELS SCALLOP HORNETS PEBBLES FRECKLE GAZELLE HADDOCK TORPEDO LOBSTER MAGGOTS TRESTLE TRINKET NOSTRILS MINSTREL LIGAMENT PARAKEET

REFERENCES


FORSTER, K. I. (1981). Frequency blocking and lexical access: One mental lexicon or two? Journal of
Verbal Learning and Verbal Behavior, 20, 190–203.


(Received October 20, 1983)
(Revision received April 17, 1984)